Center for Advanced Multimodal Mobility

Solutions and Education

Project ID: 2018 Project 02

USING GENERAL TRANSIT FEED SPECIFICATION (GTFS) DATA AS A BASIS FOR EVALUATING AND IMPROVING PUBLIC TRANSIT EQUITY

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 2019

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

IPTG – Index of Public Transport Gaps

IPTN – Index of Public Transport Needs

IPTP – Index of Public Transport Provision

IVR – Interactive Voice Response

MOE – Measure of Effectiveness

NCDOT – North Carolina Department of Transportation

RUC – Residential Units Covered by Transit Stop/Station

RUT –Residential Units (Total) in Block groups

SQL – Structured Query Language

TGI – Transit Gap Index

TD – Transportation Disadvantaged/Transit Dependent

TDS – Transit Dependent Score

TDS' – Standardized Transit Dependent Score

TOI – Transit Opportunity Index

TSCR – Transit Service Coverage Ratio

TSS – Transit Service Score

TSS' – Standardized Transit Service Score

TTAS – Travel Time between All Stations

TTP – Total Population

WATT – Weighted Average Travel Time

EXECUTIVE SUMMARY

As a critical part of economic and social fabric of metropolitan areas, public transit is necessary to provide mobility for users. A crucial task of transit planning is to better assess the equity and accessibility of public transit. The basic concept of transit equity refers to the degree to which transportation systems enable people to reach desired activity locations with fair and appropriate distribution of impact (benefits and costs), which explains the complex relationship between transportation, human activity and land use. Although years of research efforts have been done for better quantifying, analyzing, and planning for transit accessibility and equity, they are still challenging due to many types of barriers (including spatial, temporal, financial, and social, etc.), all of which can limit accessibility and equity. Meanwhile, the development of General Transit Feed Specification (GTFS), an open standard format, provides new opportunities for transit performance measurement, benchmarking and research, especially in the field of transit equity and accessibility assessment. This standard transit feed data format has been demonstrated to be extremely useful, due to its contents associated with spatial and temporal characteristics. However, the progress of studies combining those two together is still relatively slow and modest. To improve such studies, more spatially disaggregated, individualized and temporally-aware accessibility metrics, and more sophisticated spatial computational tools to operationalize such metrics and improve measurement of transit accessibility and equity in empirical research, are required.

This research develops and recommends an advanced and practical method to better evaluate and improve the equity and accessibility of public transit for people. In such sense, the transit gap index (TGI) is developed by taking demographic features, spatial and temporal transit service characteristics into consideration. A case study in the City of Charlotte is conducted and the associating comprehensive gap analysis based on the proposed methodology is provided. This research also develops guidelines and recommends best practices for the use of GTFS data as a main data source to better understand and assess public transit equity and accessibility for public transportation planning and operation. Summary and conclusions are made, and further research directions are also given.

Chapter 1. Introduction

1.1 Problem Statement

As a critical part of economic and social fabric of metropolitan areas, public transit is necessary to provide mobility for users. A crucial task of transit planning is to better assess the equity and accessibility of public transit. Due to the complex characteristics of transportation equity, the analysis should consider various perspectives and relevant impacts. Litman (2002) presented a modeling framework for transportation equity analysis by categorizing it by type, impact, measurement unit, and group of people. Since the equity assessment is a crucial part of transit performance evaluation, many standards, guidebooks and reports (i.e., *TCQSM*; *Canadian Transit Handbook 3rd Edition, 1993; Bus Service Evaluation Methods: A Review, USDOT 1984; RTD Service Standards*; *TCRP Reports and TCRP Synthesis; NCHRP Reports; FDOT Final Report: Best Practices in Evaluating Transit Performance, 2014; Establishing a Framework for Transit and Rail Performance Measures, CDOT 2012*) have been developed in which several related measurements and metrics have been proposed. Following all those guiding documents, many studies have been conducted to explore the topics. However, few research efforts have been made to develop comprehensive measures to assess the public transit equity/accessibility for the transit dependent population. In addition, previous relevant studies present some limitations. For instance, most previous studies have used census blockgroup or census tract as the basic spatial unit of analysis to measure transportation accessibility, assuming uniform distribution of the population, which is not the case in reality. Besides, the common use of demographic data to single out transit disadvantage population (i.e., elder, people with disability, and people with low-income) in many studies results in unavoidable overlaps. Thus, in order to address the limitations, it is still worthwhile to enrich the set of public transit equity/accessibility evaluation in the field of transportation.

With the rapid development of GTFS over the past few years, research efforts have shown both the effectiveness and efficiency of using the GTFS data for measuring the equity and accessibility of public transit by using different indexes and indicators. Gandavarapu (2012) developed a framework to use GTFS in computing transit accessibility measures based on the population and employment for each of the traffic analysis zones by constructing a shortest path tree of any location. Wong (2013) proposed a method to conduct a comparison between measures that were calculated based on the GTFS data and measures that were observed by agencies. Jiao and Nichols (2015) used the GTFS data to locate “transit deserts” by identifying the transit dependent populations by estimating the transit demand, calculating the transit supply, and then subtracting the supply from the demand to measure the service gap (as a measure of accessibility and equity). Ma and Jan-Knaap (2014) employed the GTFS data and OpenStreetMap data to model the employment accessibility at the neighborhood level through the time-space combined mapping, and they also applied the method to the Purple Line in the State of Maryland as an example to analyze the employment accessibility change for planning purpose. Porter et al (2014) developed a prototype model using GTFS data to help the Oregon Department of Transportation optimize their transit network. Bertolaccini and Lownes (2015) developed the Transit Opportunity Index (TOI), a comprehensive accessibility measure, to quantify the transit accessibility using only the GTFS and population data. Sarker et al. (2016)

developed and used a graph theory-based methodology to measure transit connectivity (which did not require the use of transit ridership data and transit assignment models) using the GTFS, demographic and socio-economic data. Bejleri et al (2018) used spatial overlays and conducted a network analysis to identify transit dependent (TD) population areas with major gaps in alternative transportation services. The GTFS data were utilized to measure access to public transportation as one of the three alternative transportation services. In this regard, the topic to be explored in this study is highly relevant to and in line with the FAST ACT priority area of "Improving Mobility of People and Goods".

1.2 Objectives

The main goal of this research project is to develop and recommend an advanced and practical method to better evaluate and improve the equity and accessibility of public transit for people. To achieve the goal, 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 assessment, with a focus on those using the GTFS based data as a major data source; (2) identify and develop suitable public transit equity measures; (3) develop a unique approach to measuring public transit equity mainly using the GTFS data; and (4) analyze the performance of the method developed and provide recommendations for future research directions.

1.3 Expected Contributions

To accomplish these objectives, several tasks have been undertaken. A comprehensive review of public transit equity and accessibility assessment associated with using GTFS data as main data source has been conducted. Based on the literature review, a sophisticated measure for better assessing the public transit equity has been developed. A GIS-based methodology has been developed to evaluate the equity-related public transit performance by conducting a case study in the City of Charlotte. It should be noted that both the developed metric and methodology are very general and implementable with easily accessible sources of data.

All those outcomes can be easily integrated into current practices so as to better assess the public transit equity and accessibility, help provide guidance and identify further opportunities to develop advanced methods that can be used to optimize public transit infrastructure investments to maximize equity and accessibility related impacts in future research.

1.4 Report Overview

The remainder of this report is organized as follows: Chapter 2 presents a comprehensive review of the state-of-the-art and state-of-the-practice literature on the general transit feed specification (GTFS), public transit equity assessment metrics (measurements) and methods, and also the use of GTFS in public transit equity evaluation. Chapter 3 provides a detailed explanation of and formulation for the developed metric (measurement) for public transit equity assessment. Chapter 4 describes the solution methodology for the developed metric (measurement) to evaluate the public transit equity. 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 (GTFS data, transportation network data, and demographic data) 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 of general transit feed specification (GTFS), public transit equity assessment metrics (measurements) and methods, and also the use of GTFS in public transit equity evaluation. This should give a clear picture of public transit equity assessment methods, available and potential use of GTFS data in public transit equity evaluation.

The following sections are organized as follows. Section 2.2 presents several definitions of public transit equity, followed by the discussion of a list of different classifications of public transit equity in section 2.3. Section 2.4 provides detailed information about the GTFS data. Section 2.5 gives brief descriptions of previous studies that mainly used GTFS data as the basis to assess public transit equity. Section 2.6 shows some frequently used measures of effectiveness (MOEs) for quantifying public transit equity. Finally, section 2.7 concludes this chapter with a summary.

2.2 Definition of Public Transit Equity

As a critical part of economic and social fabric of metropolitan areas and an important avenue of research that is in line with the Fixing America's Surface Transportation (FAST) Act priority area, public transit is necessary to provide mobility for users and also attracts much attention from researchers in the transportation field. Prior to conducting evaluations of public transit equity and developing potential solutions to improve public transit equity, it is necessary to clearly define what public transit equity is and understand its features.

There are a myriad of definitions for public transit equity. However, it is a consensus that transit equity refers to the degree to which transportation systems enable people to reach desired activity locations with the fair and appropriate distribution of impact (benefits and costs). Clearly, it reveals the complex relationship between transportation, human activity, and land use. Despite the common point, there are still some slight differences in describing public transit equity. At times, the equity of public transit might also be referred to as accessibility, justice or fairness (Litman, 2002).

Even though not directly, the significance of public transit equity was firstly and formally put forward and addressed in the Civil Rights Act of 1964. Title VI of this Act regulated that the distributions of federal resources by federal agencies must be in the fairest and least discriminatory manner, in order to provide and maintain equitable services (Colopy, 1994). In the act, transportation (including the public transit) was included as one of the services. This concept is very important. It was the very first time to combine the anti-discrimination and service quality of public transportation together, which certainly has a great impact on later acts and policies.

Krumholz and Forrester (1990) defined the equity in transit planning in a book, entitled "Making Equity Planning Work: Leadership in the Public Sector", as an effort to render multitudinous choices for people who have fewer ones. Planners always seek to balance the maximization of ridership and improvement of coverage of less-populated area. One of the root-causes of excluding and isolating the poor from the more developed areas was the lack of adequate public transit (Garrett and Taylor, 1999).

The definitions above come more from a high level of policy makers and regulations, which might seem to lack practical applications and relevant quantification implications. Hansen defined accessibility (equity) as "a measurement of the spatial distribution of activities about a point, adjusted for the ability and the desire of people or firms to overcome spatial separation" (Hansen, 1959). Furthermore, based on this definition, he proposed a gravity-based method to calculate a travel time-based accessibility index, which has been widely used in various ways in the field of transportation research.

Built on this concept, some researchers defined equity as evenly spatial transit opportunity where the distances from each resident to public transit facilities are same or have uniform spatial distribution in a geographical region (Chang and Liao, 2011; Tsou et al., 2005). This definition is closely related to another important concept, "horizontal equity", in the public transit equity assessment, which will be discussed in the following section in details. Unlike previous vague definitions, it gave a clear picture on how to quantify the equity. However, the limitation of such definition lies in the ignorance of varying population densities in regions and different levels of requirements to distribute the public transit benefit among all residents. After all, it is almost impossible to offer perfect even distribution of public transit services in space.

To overcome the limitation of this too idealistic definition, several studies were undertaken to make some modifications to the previous concept and suggested to define equity as providing certain quality of service, benefit and coverage of public transit upon different populations, while allowing a range of acceptable distribution (Martens et al., 2012). Thus, it is well understood that an equitable distribution of transportation benefits, particularly public transit, should firstly maintain a decent level of benefits for socially disadvantaged groups, and then maximize the average for all.

2.3 Classification of Public Transit Equity

2.3.1 Horizontal vs. Vertical Equity

Horizontal equity treats different individuals and groups in even ability and need while considering the distributions of transit services. Thus, no difference in cost/benefit between individuals and groups will appear, meaning that favoring one individual or group over others should not happen. A simply common interpretation of such type of public transit equity is that particular individuals or groups "get what they pay for and pay for what they get" with fares and other costs unless there is any exemption (e.g., specific subsidies).

Vertical equity is where benefits (transit services) are provided on purpose for one specific group (in this case, often low income) to ensure them receive a relatively equal level of transportation services, comparing to those who can afford to choose different efficient transportation modes other than public transit. In other words, in order to achieve equity, compensations will be made by favoring economically and socially disadvantaged groups (Rawls 1971).

Despite the basic considerations of vertical equity, there is an expansion to it. Such expansion takes the mobility need and ability into account. Therefore, the needs of travelers with mobility impairments or other special needs are met.

2.3.2 Equity Based on Access to Different Opportunities (Accessibility to Food, Employment, and Health Services, etc.)

Generally speaking, the ultimate goal of providing public transit services is to enable people to reach desired activity locations with fair and appropriate distribution of impact (both benefits and costs). However, not all activities have the same priority level that require full attention and efforts to support traveling by public transit systems. As such, particular equity consideration based on specific trip purpose (e.g., education, food, employment and health services) has been put forward to further examine the aspects that are associated with survival but might have negative impacts on and pose a significant threat to particular populations. The major focus is on the job/employment accessibility via the transit system.

Sanchez (1999) found that labor participation rate is significantly influenced by accessibility to transit in Portland, Oregon, and Atlanta, Georgia, and similar results were also concluded by Kawabata (2002) in his study in Los Angeles. Many researchers devoted to this specific transit equity issue (Ma, 2014; Owen and Levinson, 2015). Neutens (2015) conducted a literature review on the impact of the transport related equity issues on the health care services and further identified the knowledge gaps in this field for future research. The unequal accessibility to healthy foods is one of the most pressing health issues in the U.S. Farber et al. (2014) made a comprehensive review of this topic and a method was developed to measure the variability in spatial access to supermarket dependent on the time of departure. Attentions were also paid to other aspects, such as accessibility to educational facilities (Chen, et al, 2011).

These different classifications and types of equity often overlap or conflict. For example, horizontal equity requires that users bear the costs of their transport facilities and services. On the other hand, vertical equity often requires subsidies for disadvantaged people. Therefore, transportation planning often involves making tradeoffs between different equity objectives.

2.4 General Transit Feed Specification (GTFS)

2.4.1 History and Development of GTFS

Prior to 2005, the public transit data were not readily available and due to this reason, planning for and relevant research efforts in regional transit networks were quite difficult. This situation was alleviated thanks to the emergence of google transit feed specification, which is later known as general transit feed specification (GTFS). GTFS defines a common format for public transportation schedules and associated geographic information. GTFS was firstly created by TriMet and Google in 2005 for the Google Transit Web-based trip planner, including schedules, trips, routes, and stops data in an open-source format of transit agencies. Due to the collaborative nature of this project between TriMet and Google, the major characteristics of GTFS data format were its simplification of being created by agencies and being used and accessed by developers, while containing sufficient information about a transit system. GTFS identifies a series of comma separated files which together describe the stops, trips, routes and fare information about an agency's service. After the first agency, TriMet, as of January 15, 2016, there was an estimated number of 1026 transit agencies worldwide, including 864 transit agencies in the U.S., who share their GTFS data openly with the general public (Front Seat Management, LLC., 2016). Nowadays, GTFS is the only worldwide standard format for public transit stops, routes, and schedules. "Open source" is another important property of GTFS data. Despite the primary purpose of creating such data benefiting the free Google Transit trip planners, many other types of services could be achieved by GTFS. As a result of third-party developer innovation, GTFS data are now being used by a variety of third-party software applications for many different purposes (i.e., trip planning, maps, timetable creation, mobile data, visualization, accessibility, and analysis tools for planning, and real-time information systems). In 2010, the GTFS format name was changed to the General Transit Feed Specification to accurately represent its use in many different applications beyond Google products (TransitWiki). Despite the popular usages by many business programmers and its quick development, much attention was paid by many researchers, making GTFS a recognized standard within industry (Nassir et al, 2011; Nazem et al, 2013; Wong, 2013; Bertolaccini and Lownes, 2014; Liu and Cirillo, 2015).

2.4.2 Contents of GTFS

A common GTFS data are compressed in a zip file, which describes 13 unique text files from a GTFS feed. Though each file is stored as a text file, its format is comma-separated-value and the particular header fields in the file are strictly prescribed by the specification. Generally speaking, a set of GTFS files contain information about network topology, vehicle frequencies and headways, in-vehicle travel time, and stop locations, etc. Figure 1 shows a typical set of GTFS files and one sample text file of the 13 files.

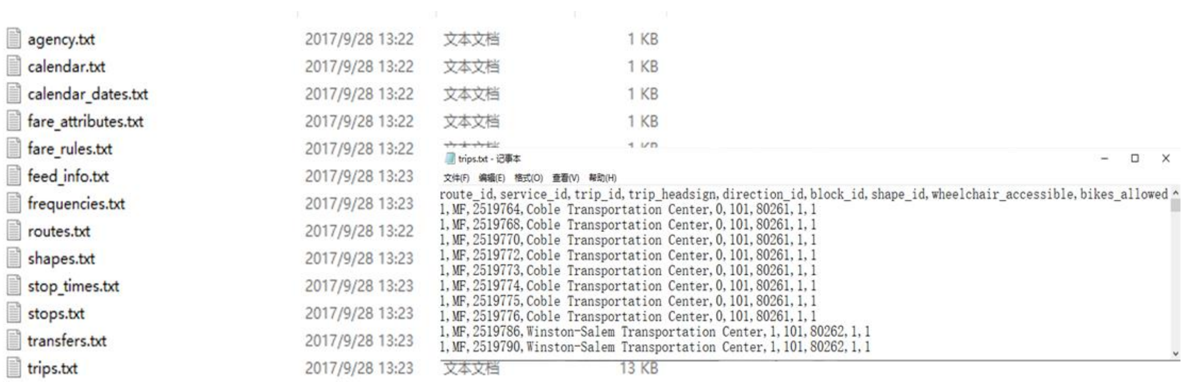


Figure 2.1 GTFS Set and Sample Text File from a GTFS Feed

The structure of GTFS data is similar to a relational database since certain shared values are used to relate files to one another (e.g., unique stop, trip, and route IDs) by creating cascading such that there will not be any duplicative information when processing the GTFS data (Wong, 2013). There are two main categories of files in one set of GTFS files: 1. required files, and 2. optional files. Both contain required and optional fields which present specific attributes of the transit system. Figure 2 provides the GTFS data model structure and diagram, also shows whether the file or field is required or not. As one can see, only 6 out of 13 files are basically required for the GTFS data to function properly.

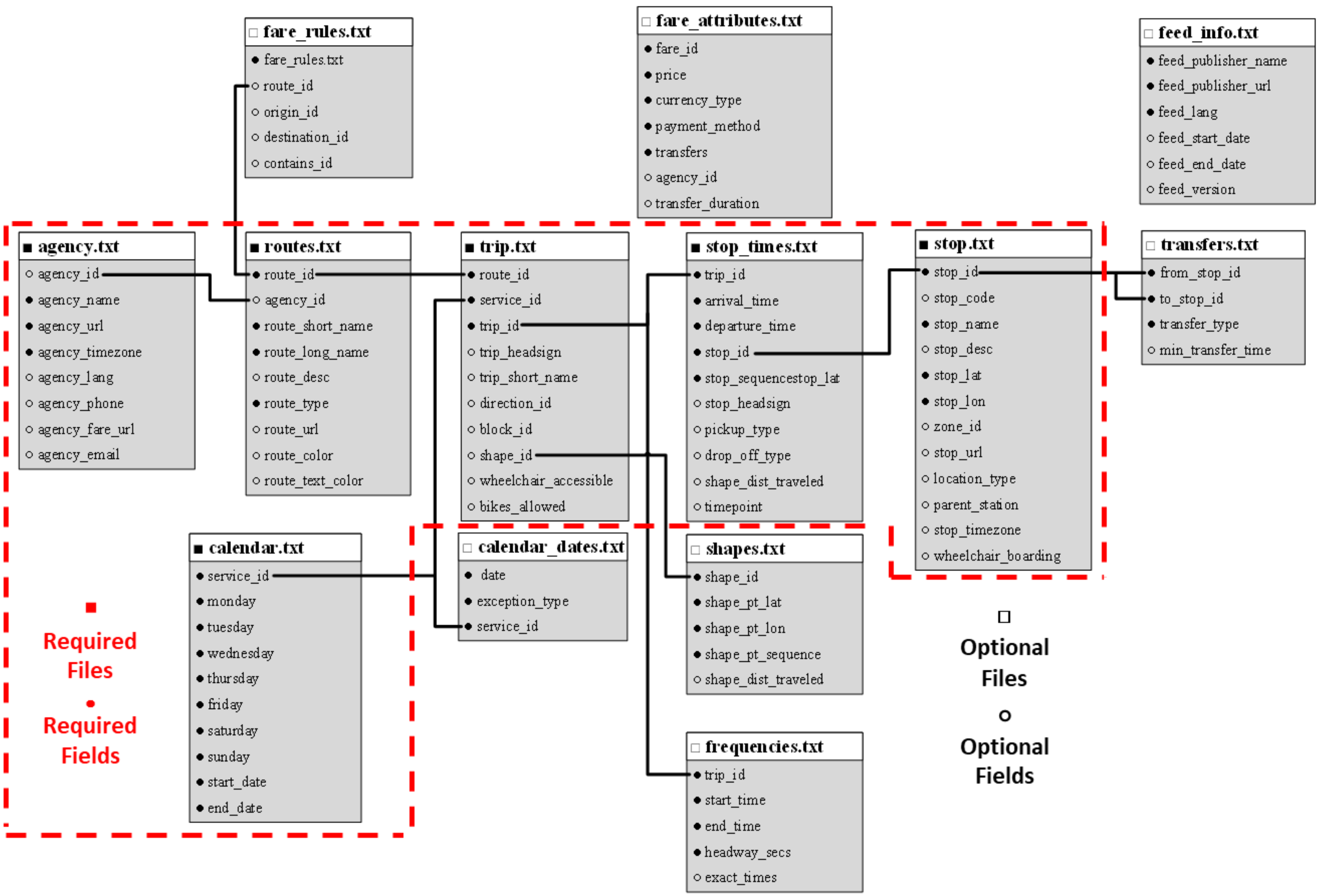


Figure 2.2 GTFS File Structure and Diagram

2.4.3 Different GTFS-Consuming Applications

Creating and maintaining GTFS data are burdensome tasks for agencies, which may require a good understanding of the GTFS structure and a lot of relevant efforts. However, benefits of publishing and utilizing such standardized format feed data are far more obvious. This subsection presents some typical examples of the types of applications and names of existing applications that use GTFS, in a categorical manner (Antrim and Barbeau, 2010; TransitWiki):

1. Trip planning and maps: These are the major applications that consume the GTFS data and there are a bunch of applications within this category. The main function of these applications is to assist a transit customer in planning a trip from one location to another using public transit.

2. Timetable creation: These applications use GTFS data to create the agency's schedule in a timetable format.

3. Data visualization: This category also largely consumes the GTFS data. A wide variety of applications utilize GTFS data to show information about transit routes, stops, and schedule data in a visual manner. Detailed information about the walkability, the quality of public transit serving the area, and relating those factors to a third criterion associated with the service (i.e., apartments available in the area) could also be provided.

4. Accessibility devices and applications: These include applications that assist transit riders with disabilities in using public transit.

5. Real-time transit information: By combining GTFS data and real-time information together, applications of this category can deliver estimated departure or arrival information to public transit riders. Recently, newer formats, such as GTFS-realtime and SIRI, were developed as extensions to the basic GTFS to support such function.

6. Frameworks and database tools: These applications are developed to support forming and organizing GTFS files and their relevant databases.

7. Interactive Voice Response (IVR): Interactive Voice Response (IVR) telephone systems can provide travel directions by phone. At least two vendors offer phone-base trip planning using voice recognition. Pricing for these products depends on features selected, agency and region size, and call volume. Ontira Communcations, Inc. offers BusLine. LogicTree offers TransitSpeak and TravelSpeak. The systems with voice recognition can be very expensive and have been noted to provide a frustrating user experience (Lee, 2009).

8. Transit network planning: most of the applications that fall into this category are used by Departments of Transportation (DOTs) or researchers, in order to accomplish planning and operational optimization for the regional public transit network. This category also includes the purpose of research studies, which use GTFS to assess and improve public transit equity.

Other than the application categories as listed above, there are still some variations. However, the above eight categories cover all the major applications of utilizing GTFS data in the field of public transit. Despite the recently emerged new format that will contain real-time information, overall, the basic GTFS data are for the static exchange of public transportation stop and schedule data. This standard transit feeds data format has been demonstrated to be extremely useful, due to its contents associated with spatial and temporal characteristics and the current valid various usages. This subsection aims to present a general outline of how GTFS data can be utilized. As mentioned, this study will focus on using GTFS data for public transit equity assessment, and therefore more detailed discussions will be made in later sections.

2.5 Previous Studies Using GTFS Data for the Assessment of Public Transit Equity

Previous sections were presented by providing details of public transit equity definitions, classifications, and also the GTFS data, respectively. Public transit equity evaluation or assessment itself indeed is an important topic in the research field of transportation. The major focus of this research is to identify the immediate opportunity of the available GTFS data and use them as a basic data source for assessing public transit equity. With the rapid development of GTFS over the past few years and its relative convenient and powerful nature in network analysis, research efforts have shown both the effectiveness and efficiency of using GTFS data for assessing accessibility and equity of public transit by developing and using a variety of methodologies, measures, and indicators. This section presents and lists some of the most relevant previous studies on this topic.

In 2011, a nationwide study that was conducted by Brookings Study of Transit and Jobs in America (Adie, et al., 2011) used GTFS data to examine the level of service of the public transit system that connects the people to jobs (which can also be interpreted as “job accessibility”). Additional data including the Census 2000 block population data and data on the working-age population (18 to 64 years old) and the neighborhood income from the Nielsen Pop-Facts 2010 Database, were used with the GTFS data. By using a specialized GIS extension (called Traffic Analyst for modeling and analyzing), results were presented through three primary metrics, including: 1. public transit service coverage, 2. service frequency, and 3. job access. In the same year, a research project conducted by the National Center for Transit Research examined opportunities of using GTFS data for evaluating the service planning and operational activity. Though it was not directly related to public transit equity evaluation, it was found that by combining location and time elements, the GTFS data provided new opportunities to evaluate transit service by measuring accessibility based on time and location (Catala et al., 2011).

Gandavarapu (2012) developed a framework to use GTFS in computing transit accessibility measures for the population and employment in each of the traffic analysis zones by constructing a shortest path tree of specific location. In the paper, he briefly introduced the GTFS data and conducted a simple literature review on earlier applications of GTFS in transportation modeling. Discussions were made in the following areas of applications: 1. transit accessibility measures for peak and off-peak hours, 2. auto ownership, and 3. trip attraction (generation) models.

Since 2012, Owen and Levinson undertook a series of research projects on transit accessibility to jobs and presented detailed accessibility values for each metropolitan area, as well as block-level maps that illustrated the spatial patterns of accessibility within each area. A U.S. Census tract-level map was used to show accessibility patterns at a national scale (Owen and Levinson, 2012-2015). By integrating with GTFS data, this study mainly focused on estimating the accessibility to jobs by transit and walking for each of the United States' 11 million census blocks and analyzed these data in 49 of the 50 largest (by population) metropolitan areas using transit schedules (i.e., GTFS data) from 2015. Based on the U.S. Census data (blocks, core-based statistical areas), the study firstly divided the geographical U.S.A. into 4879 "analysis zones" (with defining both origins and destinations inside), each including no more than 5,000 Census blocks. All the accessibility relevant measurements as listed above were calculated to rank the accessibility opportunities for each metropolitan area. After the calculations, conclusions and discussions were made on some issues related to the transit service and land use effects. The main calculation methodology was based on a gravity-based method (Ingram, 1971; Morris et al, 1979), which will be introduced in the later section of the measure of effectiveness.

Fransen, et al (2015) used GTFS data and demographic data along with the opportunity data (i.e., facility locations [supermarkets, physicians, day-care centers, administrative centers, etc.] data, educational capacity, and number of jobs) to examine the public transportation gaps using time-dependent accessibility levels. The study developed a methodology to identify the public transit gaps (i.e., the gap between the public transit service supply and the public transit demand derived from the true need of the population) while accounting for the temporal variability. The compensations from local events to the accessibility to public transit were discussed. The research also constructed an Index of Public Transport Needs (IPTN) and an Index of Public Transport Provision (IPTP) based on the studies of Currie (2010) and Jaramillo et al. (2012), to compute the index of public transport gaps (IPTG = IPTN–IPTP).

Jiao et al. (2012, 2015) used GTFS data to estimate the number of "transit dependent population", who are unable to drive because of age (too young or too old), poverty or physical disability. The results were also used to identify the "transit deserts", which are defined as areas that lack sufficient public transit services to such population.

As mentioned in section 2.3.2, Farber et al. (2014) specifically focused on the unequal accessibility to healthy foods, which is one of the most pressing health issues in the U.S. A transit time-dependent analysis was conducted in this study. Such analysis calculated the transit travel time from each Cincinnati census block to its nearest supermarkets at different times of the day. Meanwhile, by associating this time-dependent analysis with census demographic data (i.e., race, income, and age), the authors identified areas with the lowest accessibility to healthy food stores by factoring schedule-dependent public transit into the measures of accessibility.

Ma and Jan-Knaap (2014) used GTFS data and OpenStreetMap data to model the employment accessibility at the neighborhood level through time-space combined mapping, and also applied the method to the Purple Line in the State of Maryland as an example to analyze employment accessibility change for planning.

Bertolaccini and Lownes (2015) applied the Transit Opportunity Index (TOI) (Mamun, et al, 2013), a comprehensive accessibility measure to quantify the transit accessibility using only GTFS and population data by undertaking the following tasks:

1. Quantifying the accessibility using TOI, which considers three major aspects of access: spatial, temporal, and trip coverage;

2. Using TOI as measure of accessibility to show the changes through the day;

3. Creating a Python script which automates the calculation of the TOI for a transit service area, only using GTFS and census data.

In order to reduce the workload involved in quantifying the transit accessibility, this study focused on creating a script which automates the calculation of TOI, using only publically available data. Transit system data were obtained exclusively from General Transit Feed Specification (GTFS) data. The only other data required were basic population count at the block group level. Such data were available from the either Census Bureau or the relevant state data center. The authors also applied the developed script to the six Connecticut Transit operated bus systems. The script was also modified to explore how transit accessibility changes throughout the day in the Connecticut Transit Hartford bus system. The TOI script using public data substantially reduced the amount of time required to calculate and map transit accessibility. The details of this measurement were also included in this study, as will be shown in section 2.6.

Sarker et al. (2016) developed a graph theory-based methodology to measure transit connectivity (which did not require transit ridership data and transit assignment models) using GTFS, demographic and socio-economic data. This study mainly focused on estimating the accessibility to jobs by transit and walking for each of the United States' 11 million census blocks. 49 of the 50 largest (by population) metropolitan areas were analyzed using transit schedules from 2015. For accessibility calculation, simple steps were designed and followed:

1. For each Census block, calculating travel time to all other blocks within 60km for each departure time at 1-minute intervals, over 7 – 9 AM period.

2. Calculating cumulative opportunity accessibility (Ingram, 1971; Morris et al., 1979).

3. Calculating average accessibility for each block over 7 – 9 AM period.

4. Calculating average accessibility for each CBSA over all blocks, weighted by the number of workers in each block.

5. Calculating weighted ranking for each metropolitan area.

Another graph-oriented method was proposed by Fortin et al. (2016). This study was conducted to overcome the difficulty in incorporating the classic indicators with the dynamic elements of transit service (i.e., transfers between routes or stops or buses following a specific route). Three indicators were developed which were adopted from the graph theory, including:

1. The dynamic connectivity between pairs of stops throughout the day (stop level).

2. The extent of the service [i.e., departure and arrival opportunities] offered at each stop (stop level).

3. The service speed (route level).

They used a time-expanded model to evaluate the Classical Transit Indicators (such as transit system length, and number of stops) and the graph-oriented indicators. Nodes in such model represent events (arrival, transfer, and departure). Six types of edges (1. Departure-Edges [T=>D]; 2. Connection-Edges [D=>A]; 3. Station-Edges [T=>T]; 4. Transfer-Edges [A=>T]; 5. Vehicle-Edges [A=>D]; 6. Overnight-Edges) were used to forward the progresses in the graph.

2.6 Measure of Effectiveness for Assessing Public Transit Equity

Since this project seeks to assess public transit equity using GTFS data for better evaluating the current public transit system to support equitable planning and operations, this section mainly focuses on presenting the measure of effectiveness used to assess public transit equity utilizing the GTFS data as the major input. Meanwhile, this section also intends to explore any feasible methodologies that could be well suited to assess the public transit equity, but still yet to use GTFS data as the major input.

Wong (2013) examined the metrics in the "*Transit Capacity and Quality of Service Manual*" (TCQSM). There were six different performance measures used for fixed-route transit pertaining to the availability of transit services and the comfort/convenience of those services, in which GTFS data were applicable solely or associated with other data sources, as shown in Table 2.1:

Table 2.1 GTFS Data in TCQSM Analyses

Measure	GTFS Applicable	Additional Data Required
Average headway	Yes	None
Hours of service	Yes	None
Percentage of transit-supportive areas covered	Yes	Employment, residential densities
Passenger load	No	Passenger counts
On-time performance	Yes	Archived actual arrival times
Travel time difference	Yes	Traffic network

Even though not all the metrics are directly related to equity assessment, TCQSM is the leading resource on analytic methods developed for evaluating transit in the United States. It is unavoidable to convert and combine some of the six measures to further develop suitable metrics utilizing GTFS data to evaluate the public transit equity. Relevant work that used average headway to assess public transit equity can be found in Tribby and Zandbergen (2012) and Welch and Mishra (2013).

As mentioned previously, Jiao et al. (2012, 2015) used GTFS data to estimate the number of "transit dependent population" (who are unable to drive because of age (too young or too old), poverty or physical disability) and to identify the "transit deserts" (which were defined as areas that lack of sufficient public transit services to such population). The concept and calculation were relatively simple and straightforward, which can be interpreted as follows:

A. transit dependency:

Household drivers = (population age 16 and over) – (persons living in group quarters).

Transit-dependent household population = (household drivers) – (vehicles available).

Transit-dependent population = (transit-dependent household population) + (population ages 12–15) + (non-institutionalized population living in group quarters).

B. transit supply:

1. Number of bus and rail stops in each block group.

2. Frequency of service for each bus and rail stop per day (weekday service) in each block group.

3. Number of routes in each block group.

4. Length of bike routes and sidewalks (miles) in each block group.

The supply was then subtracted from the demand to measure the service gap, as a measure of accessibility and equity. It should be mentioned that the calculation of "transit dependency" changes the focus from why individuals may not drive (age, income, mobility) to examining where there are limited vehicles available for individuals to use. Since census data on the topics of age, income, and mobility do not account for the fact that these groups often overlap, this formula can effectively eliminate the overlapping by simply counting each criterion and adding them together. Such calculation can also be interpreted as "maximum potential transit dependent population".

Bertolaccini and Lownes's work (2015) was mainly built up on a study of Mamun (2013), in which a new method was developed to quantify public transit performance, the Transit Opportunity Index (TOI), by combining measures of spatial coverage, temporal coverage, and trip coverage. This index quantified public transit opportunity or the ease of reaching a destination from a given location using public transit by integrating transit accessibility (spatial and temporal coverage) via topological network connections and travel time (trip coverage) in a new transit service performance measure, while also accounting for O–D pair-wise transit connectivity with binary connectivity and decay factors. The procedures of developing the TOI is shown as follows:

1. Estimating the transit accessibility ($A_{ijl} = R_{il} S_{ijl}$, where $R_{il} = \frac{B_{i,l,buffer}}{B_{i,total}}$ [the spatial coverage area of a transit line (l) / the total area], is the spatial coverage; and $S_{ijl} = \frac{V_{ijl} U}{P_i}$ [daily available seats per capita; represented by vehicle runs * bus capacity / total population] is the service frequency for each O-D pair, and

2. A binary connectivity parameter (δ_{ijl}) is then aggregated over all transit lines based on the existence of a direct route from an origin to a destination on a particular transit line.

3. Developing a decay function ($f_{ijl} = \frac{L}{1+\alpha e^{-\beta T}}$, $T_{ijl} = T_{access} + T_{wait} + T_{in-vehicle} + T_{egress}$ represents the travel time in minutes; L is the upper limit of the connectivity factor which is assumed to be 1.0 in this study (and represents no decay in the connectivity of an O–D pair)) to reflect decreasing connectivity (↓) with increasing travel time (↑).

4. Calculating the Transit Opportunity Index for each O–D pair (TOI_{ij}) using the parameters mentioned above.

The evaluation of transfer's capability (penalty) could be further extended. The original formulation of the TOI assumed that the vehicle capacity is constant across routes, trips, and times of day, which would be relaxed for further development. In addition, it only accounted for physical connectivity and was not associated with socioeconomics.

According to Fayyaz's work (2017), the weighted Average Travel Time (WATT) was also developed and used, which was actually derived from a gravity-based method. Its simple mathematical form can be shown as follows:

$$WATT_i = \frac{\sum_{j=1}^{J} M_j * tt_{ij}}{\sum_{j=1}^{J} M_j}$$

In this equation, $WATT_i$ represents Weighted average travel time of station i, M_j is the population in the 700-meter radius of the station j, tt_{ij} means the travel time (including egress, ingress, and transfer time) using public transit from station i to station j, and J denotes the total number of stations in transit network. Any increase in population (gravity) and decrease in travel time (distance) will increase the accessibility (gravity force) between two stations (masses). It provides a different method to assess the public transit equity by using a measure of time which is understandable and tangible. The efficiency of the algorithm made this method computationally efficient to calculate the Travel Time between All Stations (TTAS or WATT) for all times of a day, giving more insight to the public transit equity. However, similar to Bertolaccini and Lownes's work (2015), it did not consider any socioeconomic characteristics of the population around the stations (including age, gender, and average salary).

Bejleri et al. (2018) used spatial overlays and network analysis to develop a methodology that was aimed at examining transportation disadvantaged (TD) population (elderly, people with disabilities and housing units without vehicles) areas with major gaps in alternative transportation services (public transportation, on-demand service, and taxi service). The GTFS data were utilized to measure access to public transportation services, one of the three alternative transportation services. For calculations of transportation supplies, the authors measured the following accessibilities for all three alternative transportation services:

1. Measuring Access to Public Transportation Services: the authors utilized GIS network analysis to determine the area of five-minute walk distance around each transit stop, and then calculated the public transportation accessibility ($TA_{pt} = \frac{RU_S}{RU_T}$, where RU_S is number of covered residential units by the transit stop within area of five-minute walk distance, and RU_T is the total number of residential units within the blockgroup).

2. Measuring Access to On-Demand Services: (1) Flat fee: the measurement was the same as public transportation services; (2) Variable fee: the authors used GIS OD matrix analysis to determine the OD distances and then calculated the "cumulative opportunity score" ($S_{CO} = \sum \frac{W_i}{d_{ij}}$, where S_{co} is the cumulative opportunity score, d_{ij} represents the distance between each blockgroup and W_i denotes the relative trip frequency to each destination category, which is based on NHTS).

3. Measuring Access to Taxi Services: The similar measurement was used and described in the On-Demand Services.

4. The next step was to standardize them in a scale of 0 to 1 and make a combination as the "comprehensive transportation accessibility", with different weights (W_p=0.55, W_o=0.35, and W_t=0.1): $CTA = (SS_p \times W_p) + (SS_o \times W_o) + (SS_t \times W_t)$

5. For transportation demand, the authors adopted the method from Currie (2010) to compute the "demand score" ($DS = (SS_e \times W_e) + (SS_d \times W_d) + (SS_v \times W_v)$, where SS_i and W_i are standardized score and weight (same weight for all three population) for the number of transportation disadvantaged population *i*, respectively, and they correspond to the populations of elderly, people with disabilities and housing units without vehicles.)

6. Finally, the authors compared the supply and demand score for all the combined populations to examine the transportation gaps.

One improvement should be mentioned of the "public transportation accessibility", is that it considered the actual spatial coverages of the residential units instead of measuring the served areas as a ratio of the service area to the total area of the blockgroup. Despite the coverages of the services, other features such as frequencies and capacities of the services were not included in the public transit accessibility. Another drawback is that the potential for "double-counts" in computing transportation demand by simply adding each criterion (elderly, people with disabilities and housing units without vehicles) together.

2.7 Summary

A comprehensive review and synthesis of the current and past research efforts related to general transit feed specification (GTFS), public transit equity assessment metrics (measurements) and methods, and also the use of GTFS in public transit equity evaluation have been discussed and presented in the preceding sections. This is intended to provide a solid reference and assistance in formulating public transit equity assessment methods and developing effective improvement strategies for future tasks.

Chapter 3. Formulation of the Transit Gap Index

3.1 Introduction

As can be seen in the literature review conducted in Chapter 2, a "gap analysis" between transit supply and demand is generally the most common form performed 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. Section 2.5 and Section 2.6 provide clear indications of the popular utilization of ArcGIS tools in such analysis. On the other hand, when considering the availability of the features in GTFS data, the definition of equity for this research is designed around those components available through GTFS. This chapter formulates the transit gap index from both supply and demand sides of the public transit service system.

The following sections are organized as follows. Section 3.2 provides a general overview of the criteria for composing the transit gap index. The process of formulating components of the transit gap index from both supply and demand sides will be discussed in Section 3.3 and Section 3.4, respectively. Section 3.5 presents the transit gap index. Finally, section 3.6 concludes this chapter with a summary.

3.2 Overview

In this project, the indicator of transit service (supply) will be denoted by a transit service score, which consists of the following: 1. transit service coverage (by all stops/stations within one blockgroup); and 2. Per capita maximum daily available seats (for each person within specific blockgroup). As mentioned in Section 3.1, the availability of transit service features stored in the GTFS data largely determines why and how these two components are developed. Features associated with the required files and required fields in GTFS data as shown in Figure 2.2 can be utilized under most circumstances. For transit demand, transit dependent (TD) populations will be calculated based on the 2010 US Census and the 2012-2016 American Community Survey (2016 ACS) data at the blockgroup level. The spatial estimation of residential units will be computed based on the parcel data of the city under study and will also be employed in the determination of both indicators of public transit supply. Finally, the transit service gap index will be obtained by comparing the differences between supply and demand in a standardized manner.

3.3 Transit Supply

3.3.1 Transit Service Coverage

In this study, the transit service coverage is defined as a ratio, which can be shown as follows:

$$TSCR_j = \frac{RUC_j}{RUT_j}$$

where $TSCR_j$ is the transit service coverage ratio of blockgroup *j*, RUC_j is the number of residential units covered by all stop within 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 "accepted walking distance" (O'Sullivan et al., 1996; Jiang et al., 2012; Daniels et al., 2013; Zhao et al., 2013; El-Geneidy et al., 2014; 8-12). 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 "rule of thumb" guidelines indicate (¼ miles, or 400 meters). This research further exhibited that with other factors being the same, at two miles from a transit stop there is a 50% chance that people will walk to a stop, and this probability will increase to 80% for one mile. Thus, in order not to underestimate the transit service coverage, a 0.5-mile walking distance has been applied in this study.

As has been discussed in Section 2.6, the 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. This can be simply demonstrated as follows:

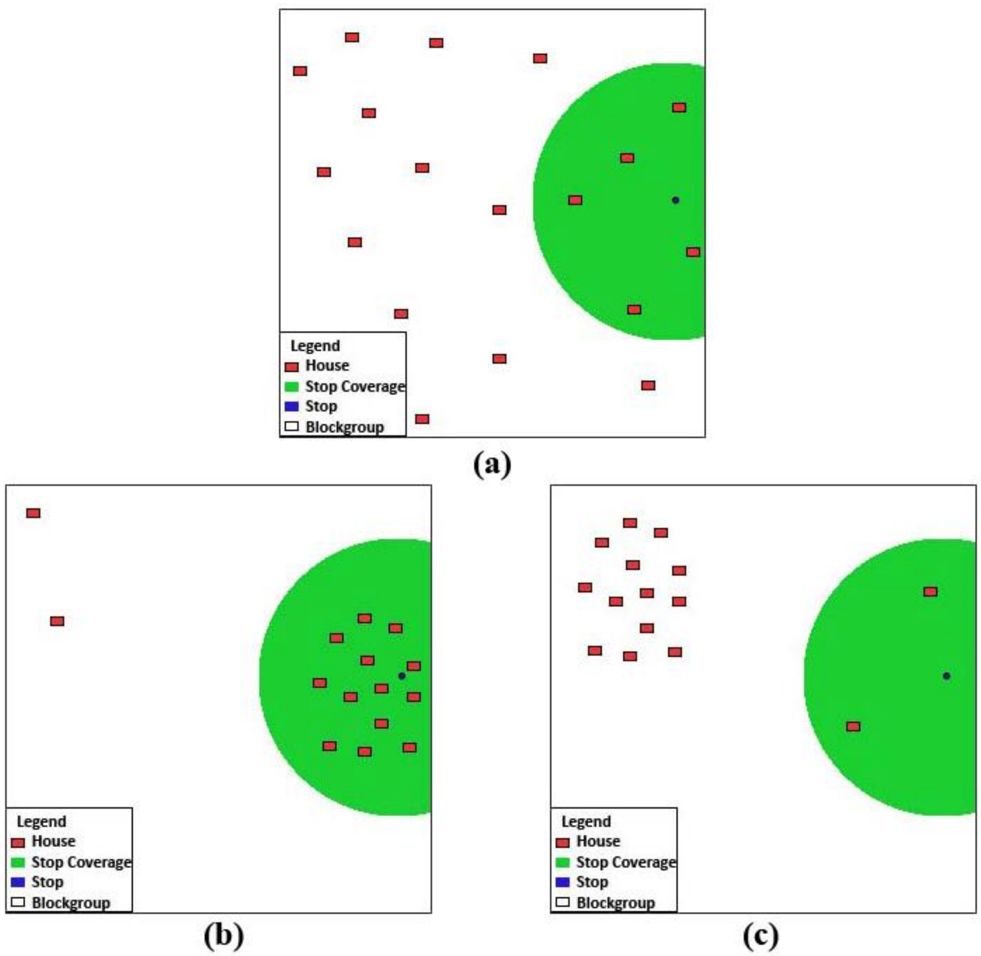


Figure 3.1 Spatial Relationship between Residential Units and Transit Stop Coverage

It is easy to see that the distribution of residential units largely determines the accuracy of the coverage ratio for those using the served area in the calculation, particularly when extreme distributions of the residential units occur within one blockgroup as shown in Figure 3.1 (b) and (c). In such cases, it will result in the non-mapping between the actual and calculated coverage, producing an underestimation in some cases or overestimation under others. On the contrary, the "transit service coverage ratio" used here will be more realistic for use to reflect the actual coverage by a transit stop/station.

3.3.2 Per Capita Maximum Daily Available Seats

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

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

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.

3.3.3 Transit Service Score

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

$$TSS_j = TSC_j \times D_j$$

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).

3.4 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, or 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 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)

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

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

$$TDS_j = \frac{TD_j}{TTP_j}$$

where TTP_j is the total population of blockgroup *j*.

3.5 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 supply and demand will be standardized in a scale of 0 to 1 based on the equation as shown below:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

And then the transit gap index can be calculated by subtracting TDS'_j from TSS'_j for blockgroup *j*:

$$TGI_j = TSS'_j - TDS_j{}'$$

3.6 Summary

The general procedures for developing the transit gap index are presented in this chapter. Detail information about each component included in the index is provided, in which criteria used for choosing variables and parameters are also discussed. The formulations provide a solid basis for the future developments in the overall GIS-based solution approaches, which will be discussed in detail in the following chapters.

Chapter 4. Solution Framework

4.1 Introduction

As discussed in the previous sections, the "gap analysis" is categorized as spatial analysis where ArcGIS is usually suitable to be deployed to undertake the task. Chapter 3 introduces the detail information about how to construct the transit gap index. This chapter will present the ArcGIS-based solution framework in detail.

The remainder of this chapter is organized as follows. Section 4.2 provides a description of the general solution framework. Section 4.3 illustrates how to integrate GTFS data with ArcGIS and how to use GTFS data to find unique stop-route pairs. Section 4.4 describes the implementation of calculating "transit service score" in ArcGIS that contains two subsections: 1) section 4.4.1 discusses the implementation of calculating "transit service coverage ratio" in ArcGIS, and 2) section 4.4.2 discusses the implementation of calculating "per capita maximum daily available seats" in ArcGIS. Section 4.5 provides simple steps to follow in order to obtain the "transit dependent score" and "transit gap index" in ArcGIS. Finally, section 4.6 concludes this chapter with a summary.

4.2 Solution Framework

Figure 4.1 provides a flow chart of the proposed solution framework for the "gap analysis" in this research. The major procedure within the general framework can be presented as follows: 1) Integration of GTFS data with ArcGIS; 2) Implementation of calculating "transit service coverage ratio" in ArcGIS; 3) Implementation of calculating "per capita maximum daily available seats" in ArcGIS; and 4) obtaining "transit dependent score" and "transit gap index" in ArcGIS.

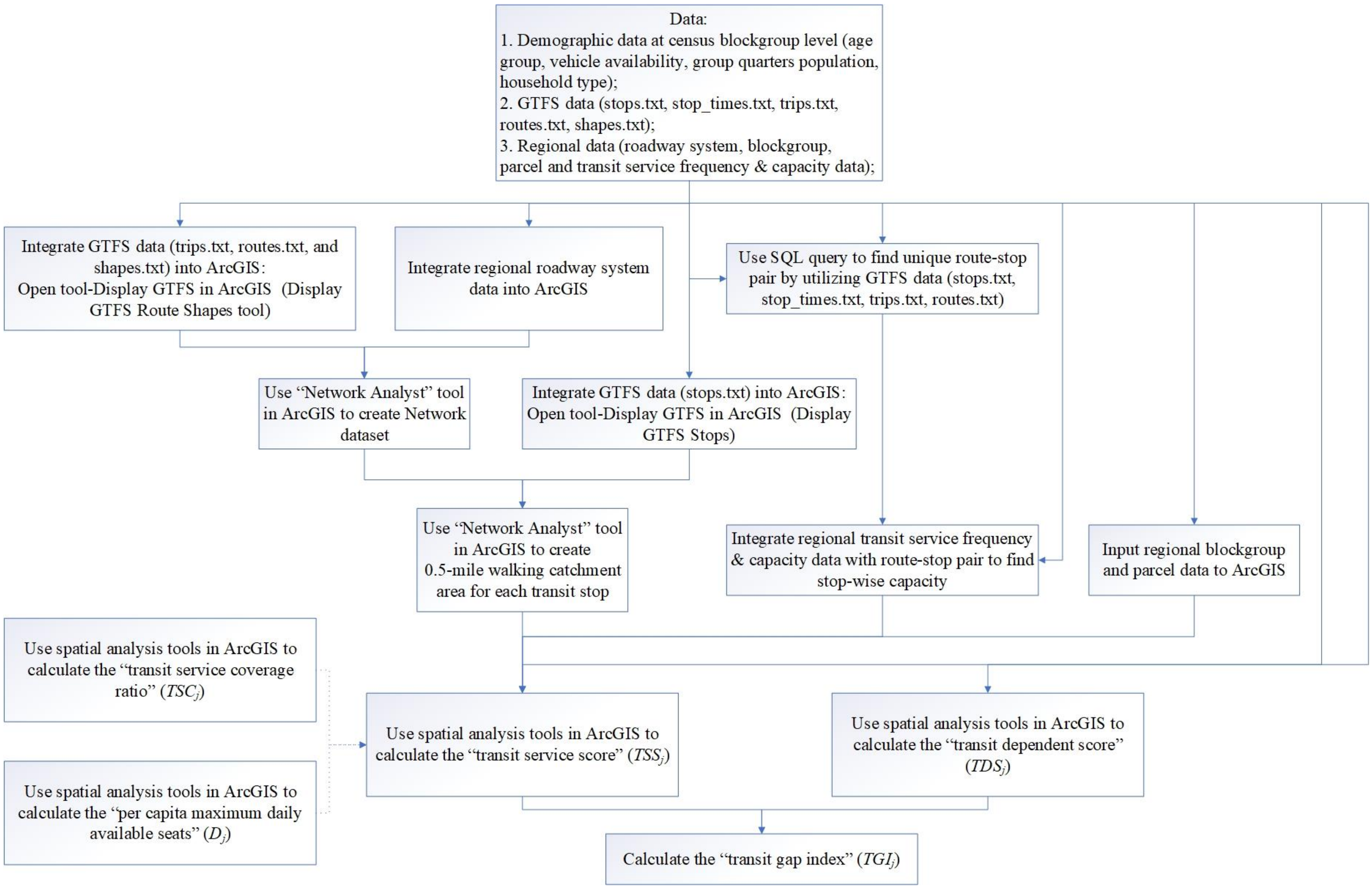


Figure 4.1 Flow Chart of the Developed "Transit Gap Index" Solution Methodology

4.3 GTFS Data in ArcGIS

As mentioned in section 4.2, there are two processes associated with the utilization of GTFS data in this study: 1) integrating GTFS data with ArcGIS; and 2) using GTFS data to find unique stop-route pairs. This section will give an introduction to both contents.

For integration of GTFS data with ArcGIS, as shown in Figure 4.1, two tools within the category of "Display GTFS" in ArcGIS toolbox have been applied to create shapefiles of public transit routes and stops in ArcGIS, respectively: 1) The "Display GTFS Stops" tool; and 2) The "Display GTFS Route Shapes" tool. Both tools are freely available tools, which can be found and downloaded on the website (https://github.com/Esri/public-transit-tools/tree/master/display-GTFS-in-ArcGIS). Figure 4.2 shows the dialog of the "Display GTFS Stops" tool and the input file for this tool is the stops.txt file in the GTFS data. The stops.txt file in GTFS data contains a series of fields that provide detailed information about stop locations, names, geocoding and id. Thus, the primary purpose of this tool is to convert such data to an ArcGIS feature class and to create a shapefile containing all the necessary information on public transit stops for further analyses.

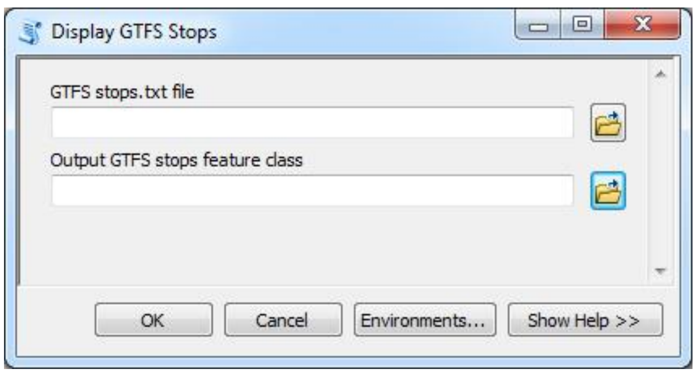


Figure 4.2 "Display GTFS Stops" Dialog

Figure 4.3 presents the dialog of the "Display GTFS Route Shapes" tool. It has the functionality that is similar to the "Display GTFS Stops" tool. Instead of generating the ArcGIS feature class and file for transit stops, a file containing public transit route will be created. The required files are trips.txt, routes.txt, and shapes.txt in GTFS data. The output will contain one-line feature for each unique shape in the GTFS data. The attributes for each line contain all the information about the routes represented by the shape.

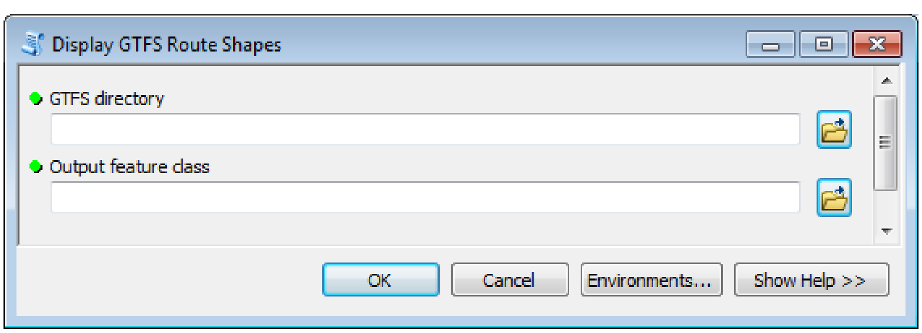


Figure 4.3 "Display GTFS Route Shapes" Dialog

Both tools are simple and straightforward to use, and the outputs will be utilized in the later analysis and computation. Figure 4.4 shows the example outputs of both "Display GTFS" tools in ArcGIS and different colors of polylines in ArcGIS represent different routes of the public transit services.

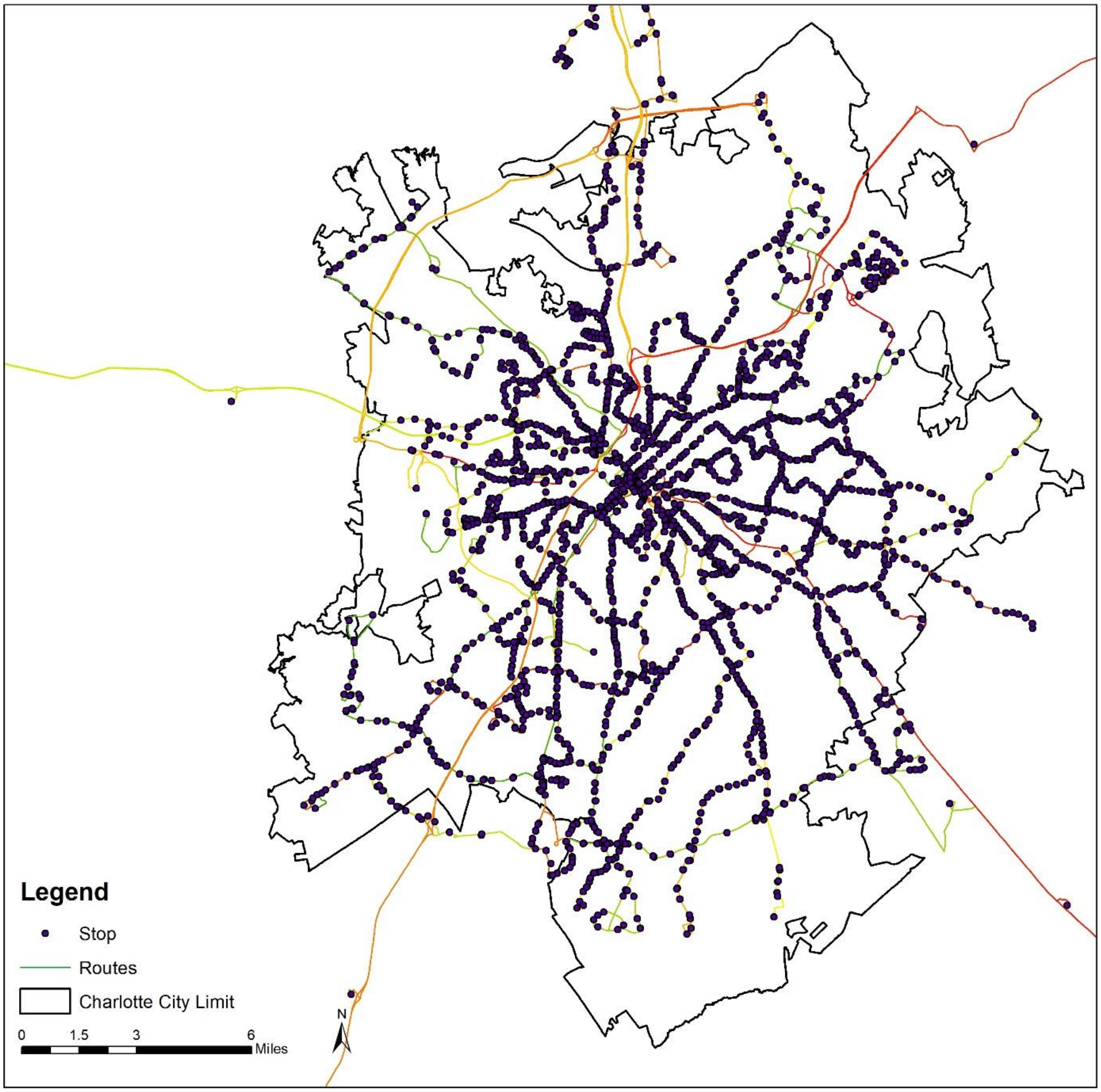


Figure 4.4 Outputs of "Display GTFS" Tool in ArcGIS

Another benefit of using GTFS data is that the data format provides an easy way to find the unique stop-route pairs and will be really useful for the following network analysis. Figure 4.5 illustrates the SQL query used in this research for finding the unique stop-route pairs. It can be seen that the relationship between stops.txt and routes.txt is not direct, and is connected by stop_timees.txt and trips.txt. However, such characteristics and interconnections among different files in GTFS data enable the data processing and make the processing easy to operate. The output of the unique stop-route pair dataset will be used to generate the stop-route matrix and will later be used to relate the capacity of each route to the capacity of each stop within the whole public transit service system. Figure 4.5 displays the SQL query process for finding unique stop-route pairs by utilizing the four files (stops.txt, stop_time.txt, trips.txt, and routes.txt) in GTFS data.

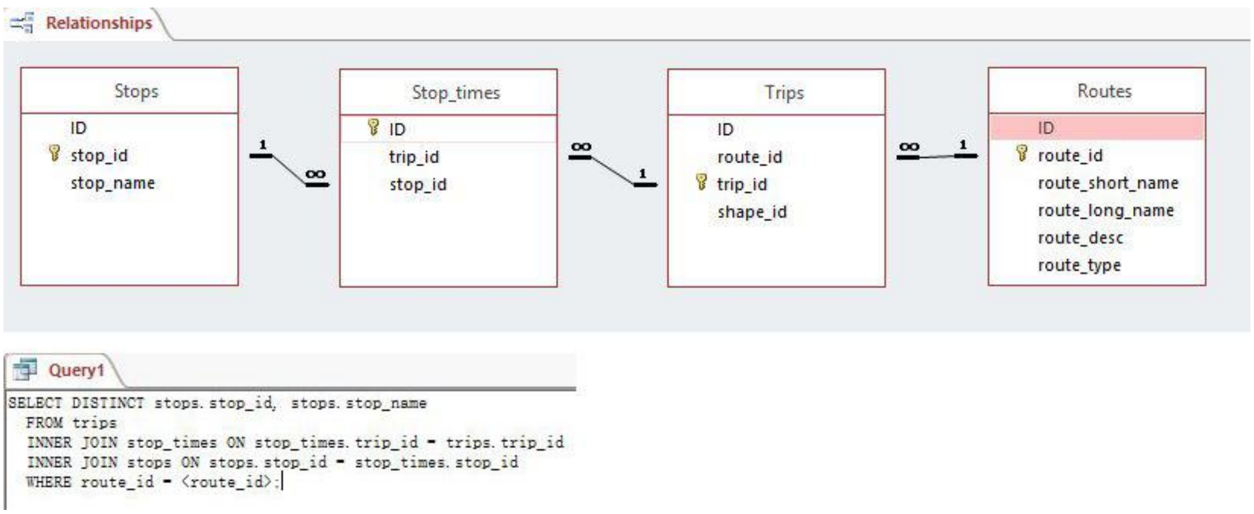


Figure 4.5 SQL Query for Finding Unique Stop-Route Pairs

4.4 "Transit Service Score" in ArcGIS

The process of obtaining the transit service score (TSS_j), as introduced in the previous section, contains two major subprocesses: 1) transit service coverage ratio ($TSCR_j$), and 2) per capita maximum daily available seats (D_j). Before getting into these two subprocesses, several preparations need to be made for further operations, including the preparation of "Network dataset" in ArcGIS, inputs of demographic data and regional data, and using the stop-route matrix to find stop-wise maximum potential capacities, which can be seen in Figure 4.1. Based on the network dataset and the public transit stop information, the 0.5-mile walking catchment area for each public transit stop/station can be created by employing the "Network Analyst" tool in ArcGIS. Once the 0.5-mile walking catchment area is obtained, these two major subprocesses can be conducted. The following subsections 4.4.1 and 4.4.2 will show the details of the two major subprocesses.

4.4.1 "Transit Service Coverage Ratio" in ArcGIS

Figure 4.6 displays the flow chart of calculating "transit service coverage ratio" ($TSCR_j$) in ArcGIS. As described in section 3.1, $TSCR_j$ is the ratio between the sum of the number of residential units within specific blockgroup covered by all related stops/stations (RUC_j) and the total number of residential units in the specific blockgroup (RUT_j). There are two main steps that are shown in Figure 4.6 and are associated with RUC_j and RUT_j , respectively:

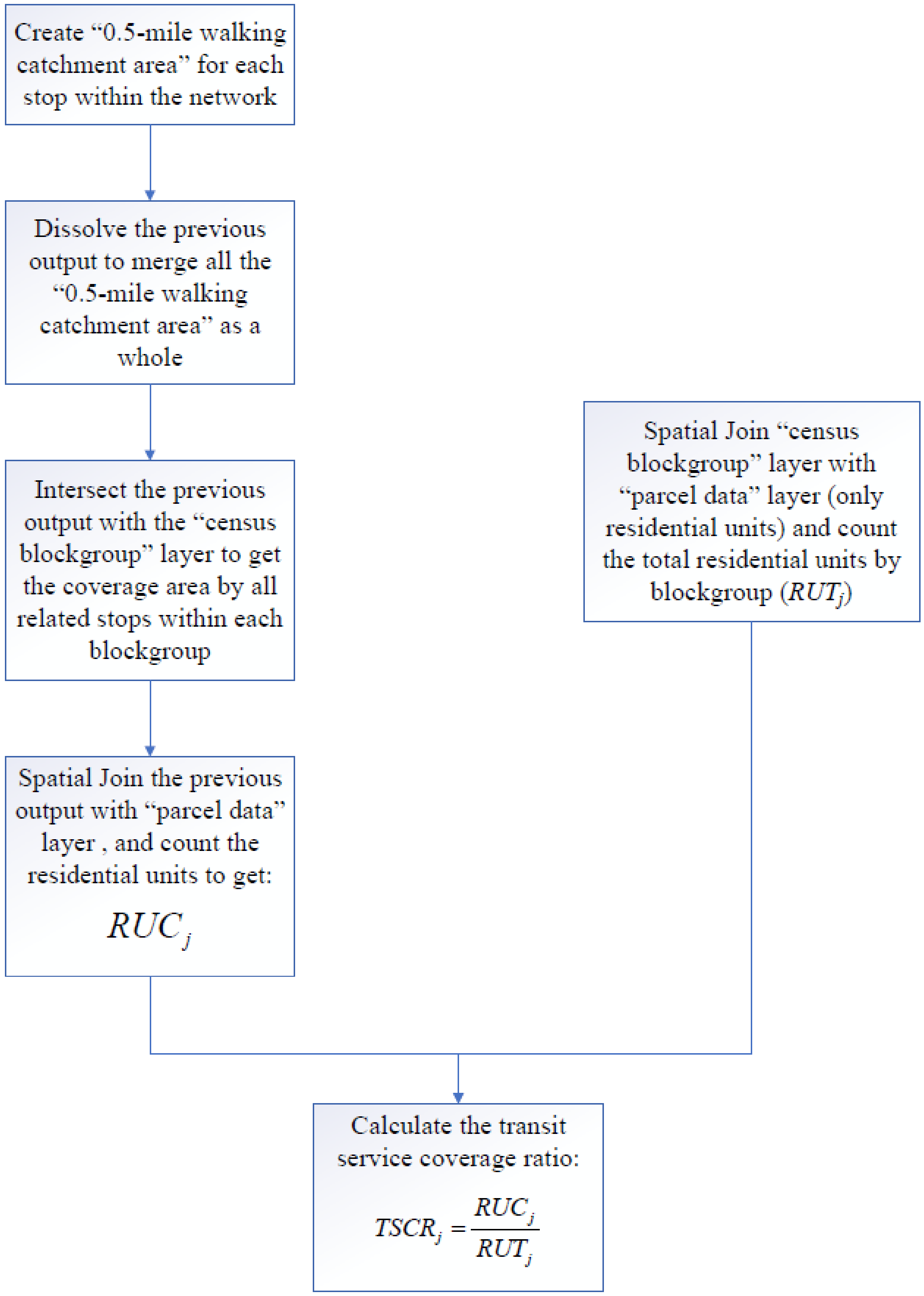


Figure 4.6 Flow Chart of Calculating "Transit Service Coverage Ratio" in ArcGIS

1. RUT_j : Using "spatial join" function in ArcGIS to join the "census blockgroup" layer with regional parcel data (residential units only), and sum up the total residential unit counts within each blockgroup area.

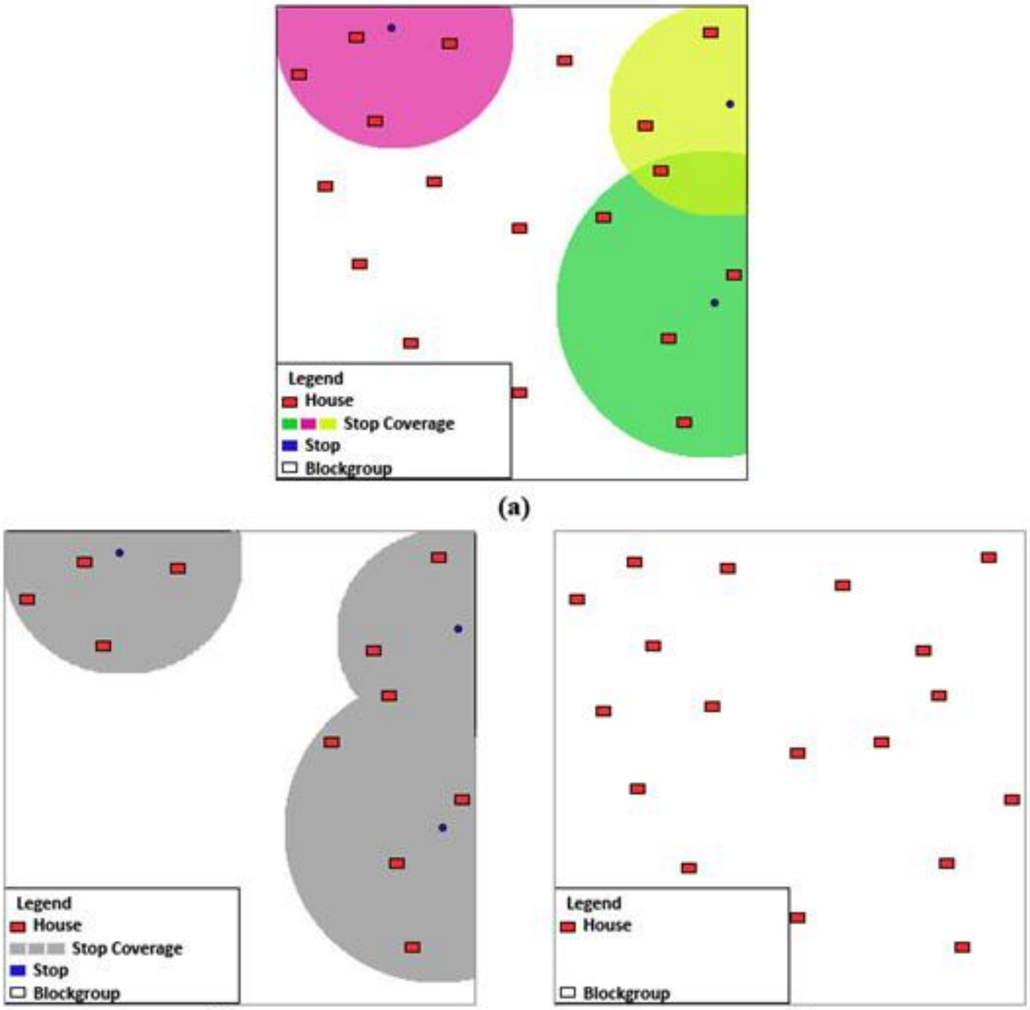
2. RUC_j : (A). Based on the "0.5-mile walking catchment area" for each stop/station, using "dissolve" function in ArcGIS to merge all the areas as a whole;

(B). Using the "intersect" function in ArcGIS to intersect the output from step (A) with the "census blockgroup" layer to determine the total coverage area by all related stops within each blockgroup;

(C). Using "spatial join" function in ArcGIS to join the output from step (B) with the regional parcel (residential units only), and sum up the total residential unit counts within the coverage area by all related stops within each blockgroup.

Figure 4.7 gives a clear and simple illustration of RUC_j and RUT_j. Once both RUT_j and RUC_j are obtained for each blockgroup, the transit service coverage ratio (*$TSCR_j$*) for each blockgroup can be calculated by using the following equation:

$$TSCR_j = \frac{RUC_j}{RUT_j}$$



(b) (c)

(a) Spatial Relationships Between Residential Units and Stop Coverage/Blockgroup Areas

(b) RUC_j

(c) RUT_j

Figure 4.7 A Simple Illustration

4.4.2 "Per Capita Maximum Daily Available Seats" in ArcGIS

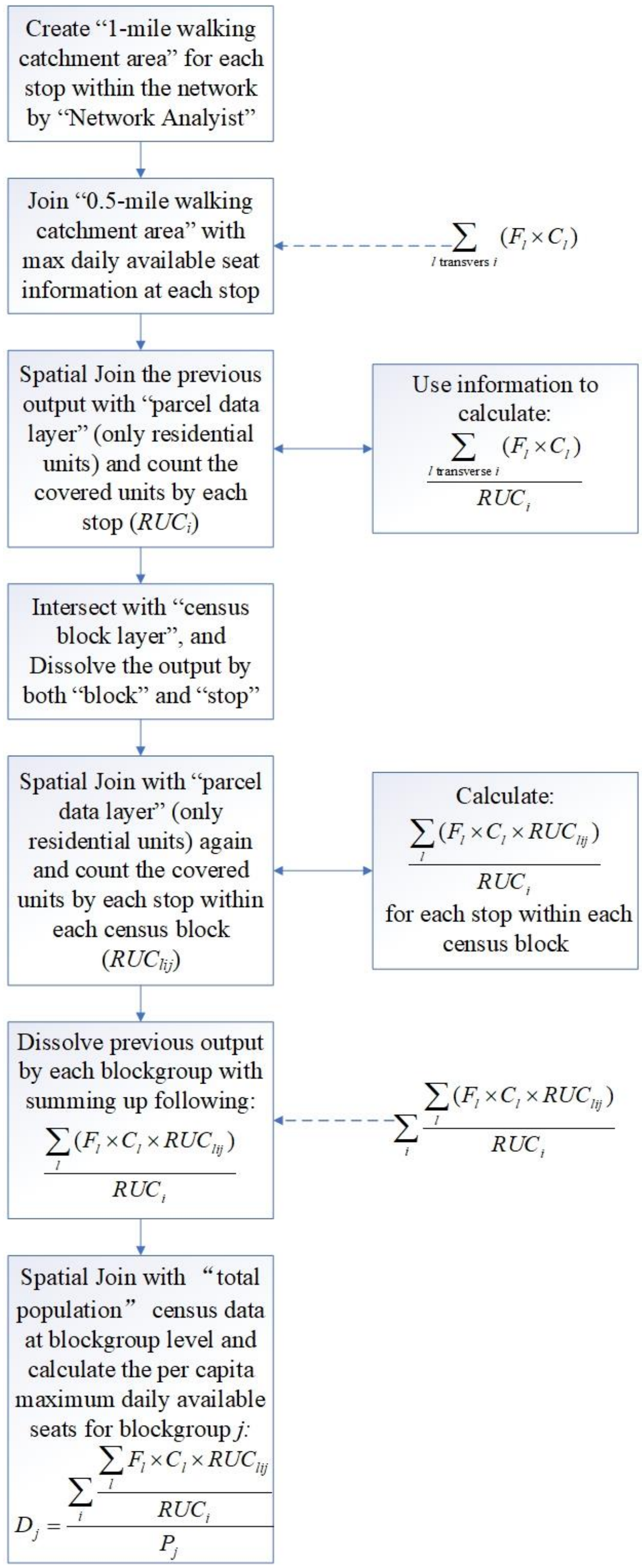


Figure 4.8 Flow Chart of Computing "Per Capita Maximum Daily Available Seats" in ArcGIS

Another important component of transit service score (TSS_j) is the per capita maximum daily available seats (D_j). Compared to $TSCR_j$, D_j is a little more complicated and cannot be directly achieved in ArcGIS. Figure 4.8 shows the flow chart of computing "per capita maximum daily available seats" (D_j) in ArcGIS. The steps can be described as follows:

(1) Built upon the "0.5-mile walking catchment area" for each public transit stop/station, and using "join" function in ArcGIS to obtain the information on maximum daily available seat at each stop (generated from the stop-route matrix), the total capacity of all transit routes transvers the stop can be represented as ($\sum_{l \text{ transvers } i} (F_l \times C_l)$);

(2) Using "spatial join" function in ArcGIS to join the output from step (1) with regional parcel data layer, and summing up the residential unit counts by each stop (RUC_i), and assembling the information with $\sum_{l \text{ transvers } i} (F_l \times C_l)$ to calculate $\frac{\sum_{l \text{ transverse } i} (F_l \times C_l)}{RUC_i}$, as well as binding it with each stop;

(3) Employing "intersect" function in ArcGIS to intersect the output from step (2) with census blockgroup layer and then merging the output via conducting "dissolve" function in ArcGIS by both "blockgroup (unique id)" and "stop (unique id)". Such operation will split the coverage area of each stop while binding the spatial information with the blockgroup that each split part falls into. Moreover, each split part will only belong to one stop and one blockgroup and will also bind with the information about $\frac{\sum_{l \text{ transverse } i} (F_l \times C_l)}{RUC_i}$ by stop (id). This can be simply illustrated using Figure 4.9 and the transformation from Figure 4.9 (a) and (b) to Figure 4.9 (c) and (d) sequentially shows the operation in this step;

(4) Using the "spatial join" function in ArcGIS to join the output from step (3) with regional parcel data layer again, and summing up the residential unit counts by each stop within each blockgroup (i.e., sum the counts within each split part obtained from (3)), and then assembling this information ($\underline{RUC_{lij}}$, or actually RUC_{ij}) with $\frac{\sum_{l \text{ transverse } i} (F_l \times C_l)}{RUC_i}$ to calculate $\frac{\sum_l (F_l \times C_l \times RUC_{lij})}{RUC_i}$ for each stop within each blockgroup;

(5) Applying the "dissolve" function in ArcGIS to merge the output from step (4) by each blockgroup and summing up $\frac{\sum_l (F_l \times C_l \times RUC_{lij})}{RUC_i}$ to get $\sum_i \frac{\sum_l (F_l \times C_l \times RUC_{lij})}{RUC_i}$ for each blockgroup;

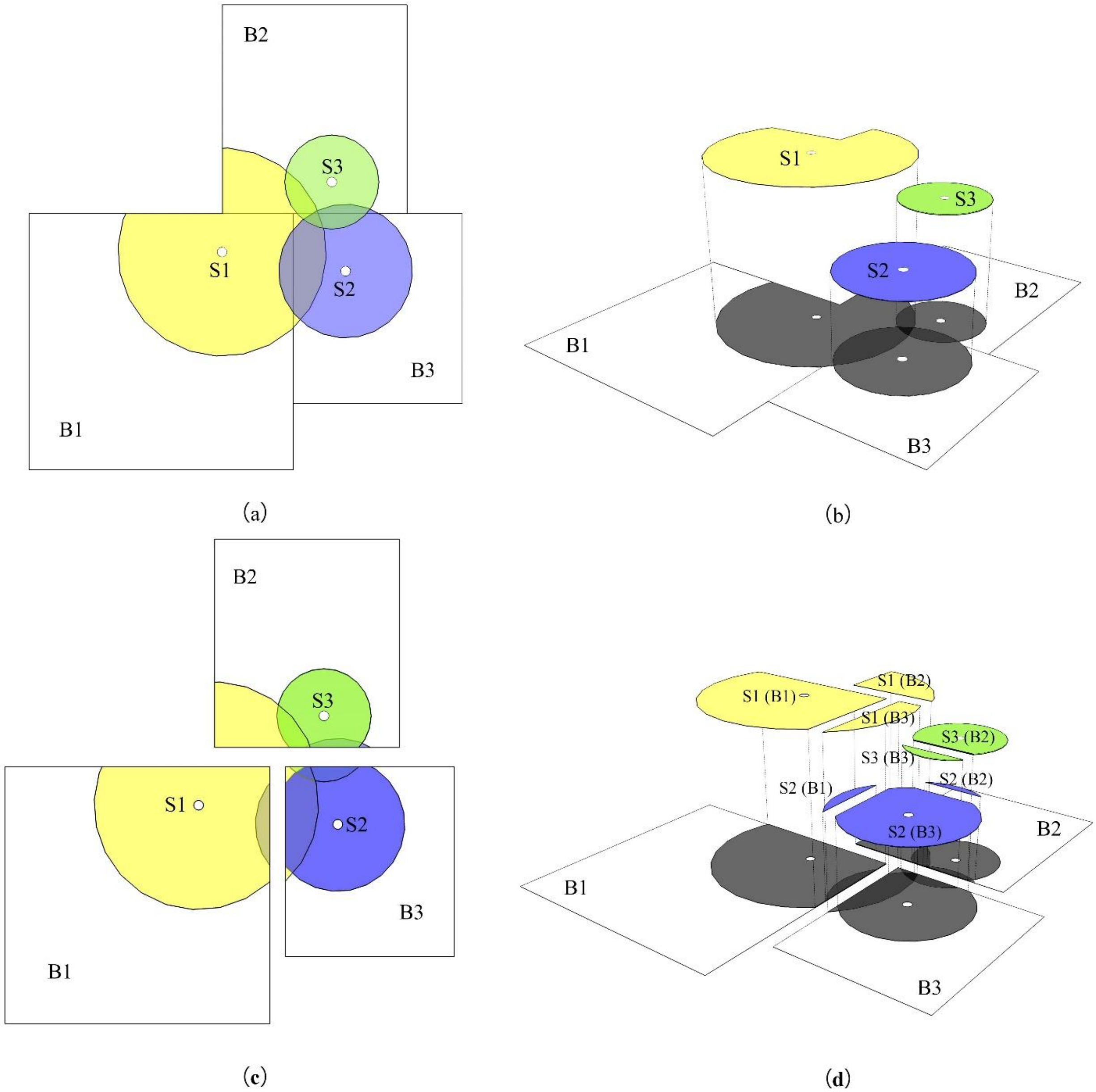


Figure 4.9 Illustration of the Step (3) of Calculating "Per Capita Maximum Daily Available Seats" in ArcGIS

(6) Using the "spatial join" function to obtain total population data at blockgroup level for each blockgroup, and then using the following formula to calculate the per capita maximum daily available seats (D_j) for each blockgroup:

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

Once both TSS_j and D_j are obtained, the transit service score (TSS_j) can be computed using the formula as follows for each blockgroup:

$$TSS_j = TSC_j \times D_j$$

4.5 "Transit Dependent Score" and "Transit Gap Index"

Transit dependent score (TDS_j) simply depends on the calculation of transit dependent population defined in this study, and can be easily done by importing all required demographic data into ArcGIS and then using field calculator to exclude the non-transit dependent population. The process is the same as discussed in section 3.4, and is shown as follows:

Household drivers = (population age 16 and over) – (persons 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)

As mentioned in Section 3.4, negative values will be adjusted to zero.

After obtaining the TD_j, a transit dependent score (TDS_j) is assigned for each blockgroup:

$$TDS_j = \frac{TD_j}{TTP_j}$$

The transit gap index can then be calculated by subtracting TDS'_j from TSS'_j for each blockgroup by using field calculator:

$$TGI_j = TSS'_j - TDS_j'$$

The purpose of the gap analysis is to identify both deficiencies and redundancies between the public transit supply and demand. By using the Jenks natural breaks classification method, *TSS'* and *TDS'* are classified into seven categories (i.e., Very Low, Low, Medium-Low, Medium, Medium-High, High and Very High). Then as a supplement to *TGI*, comparisons between *TSS'* and *TDS'* can be conducted to further determine the deficiencies (high demand, low supply) and redundancies (low demand, high supply) based on the public transit supply and demand.

4.6 Summary

The objective of this chapter is to present the basic framework for the "gap analysis" (transit gap index) and its major procedures. Detailed information about the procedures and outputs is also presented in this chapter.

Chapter 5. Case Study

5.1 Introduction

As discussed in both Chapters 3 and 4, the data required to implement the methodology that has been developed in this study include the following major components: 1) GTFS data; 2) demographic data; 3) transportation network data; and 4) other regional data. This method is developed and applied to the City of Charlotte, Mecklenburg County, North Carolina. Note that the city has established the transit-oriented development planning and includes this requirement in the city code of ordinances (City of Charlotte, North Carolina, Code of Ordinances, 2018).

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. Finally, section 5.6 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

"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 section 4.3 in Figure 4.4, which is the output of "Display GTFS" tool in ArcGIS for the CATS.

"stops.txt", "stop_times.txt", "trips.txt" and "routes.txt" 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

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 of 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.

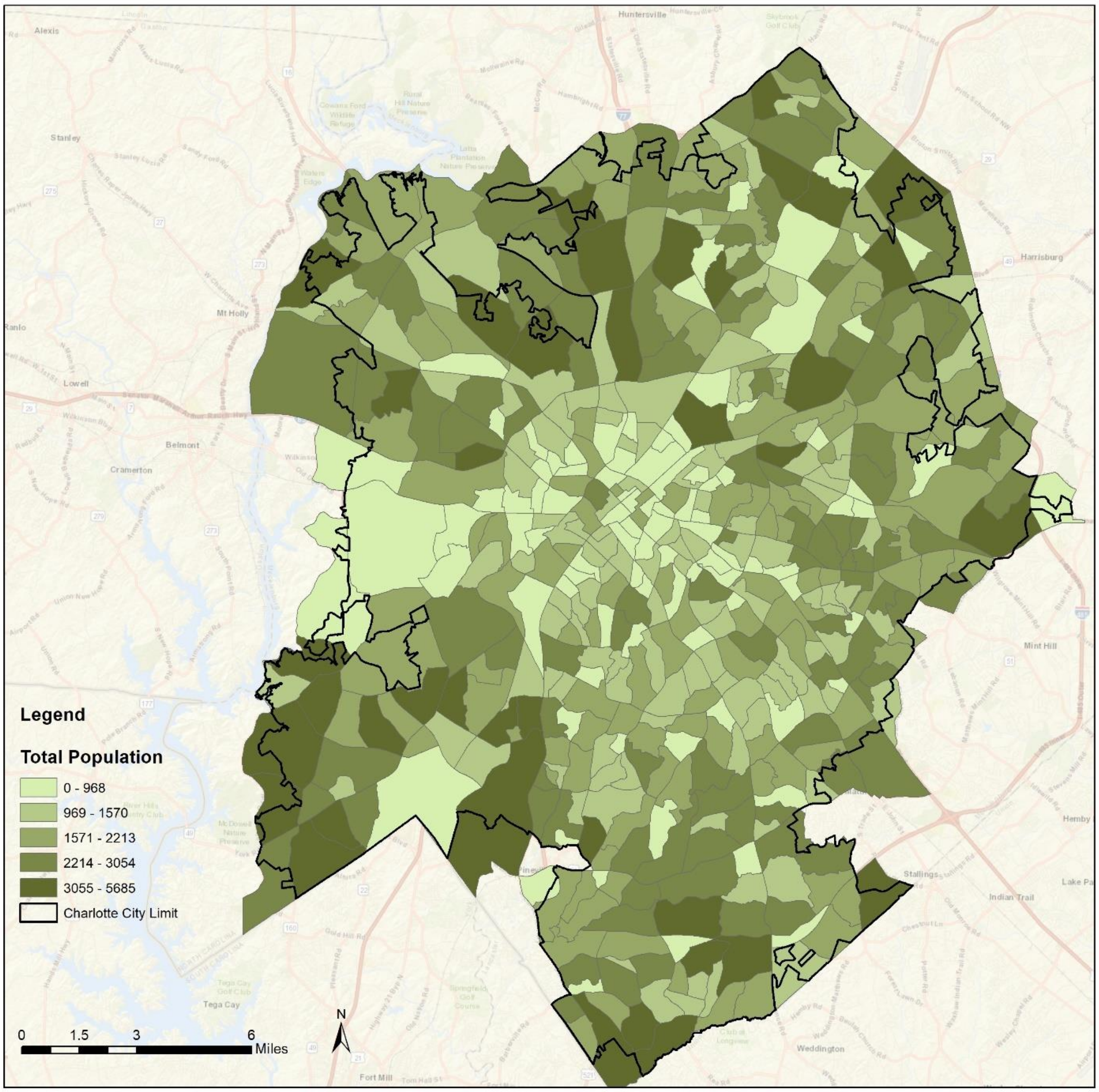


Figure 5.2 Distribution of Total Population in the City of Charlotte

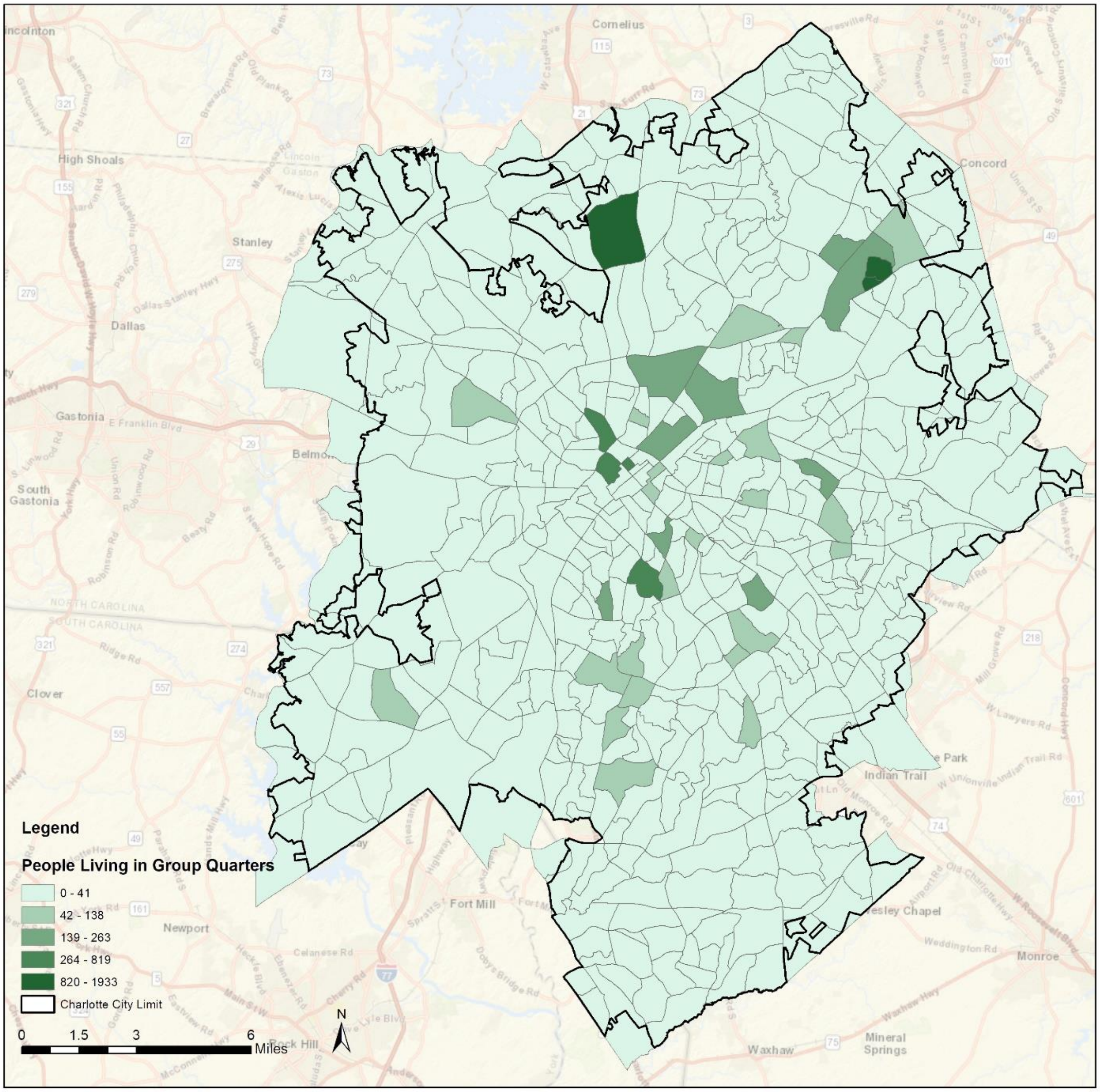


Figure 5.3 Distribution of People Living in Group Quarters in the City of Charlotte

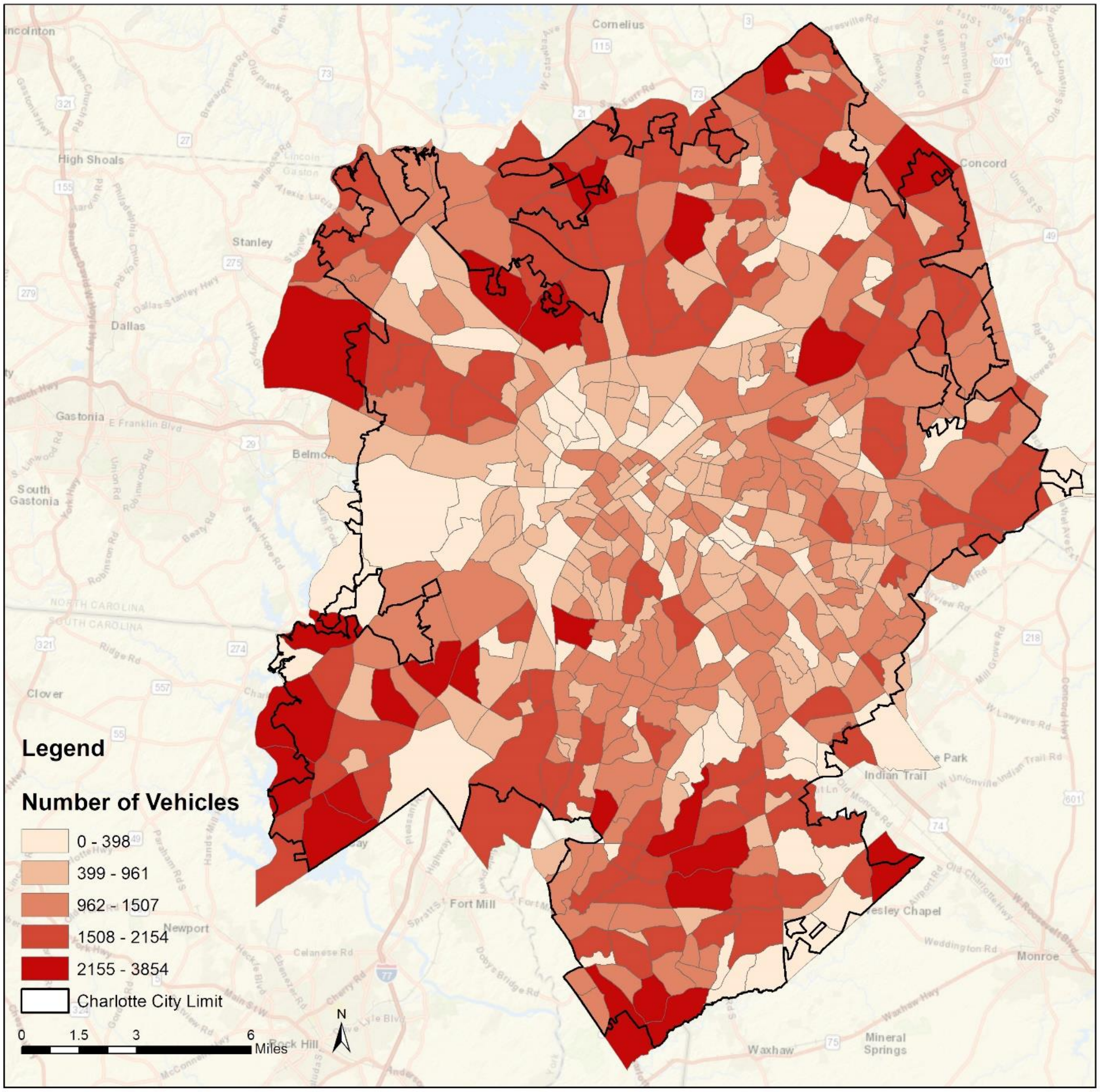


Figure 5.4 Distribution of Vehicles in the City of Charlotte

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 5.3 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.

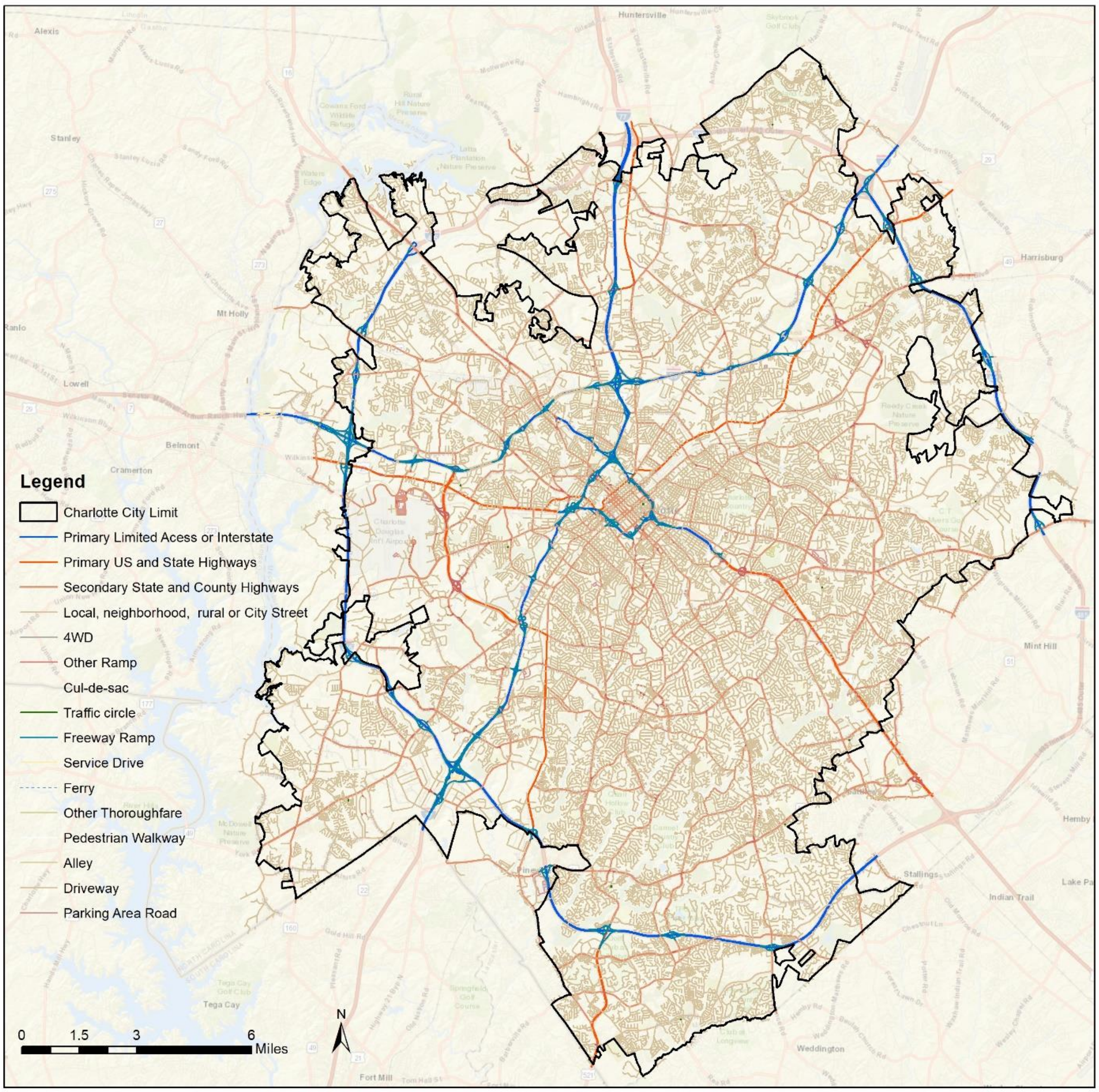


Figure 5.5 North Carolina Statewide System and Non-System Road System in the City of Charlotte

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).

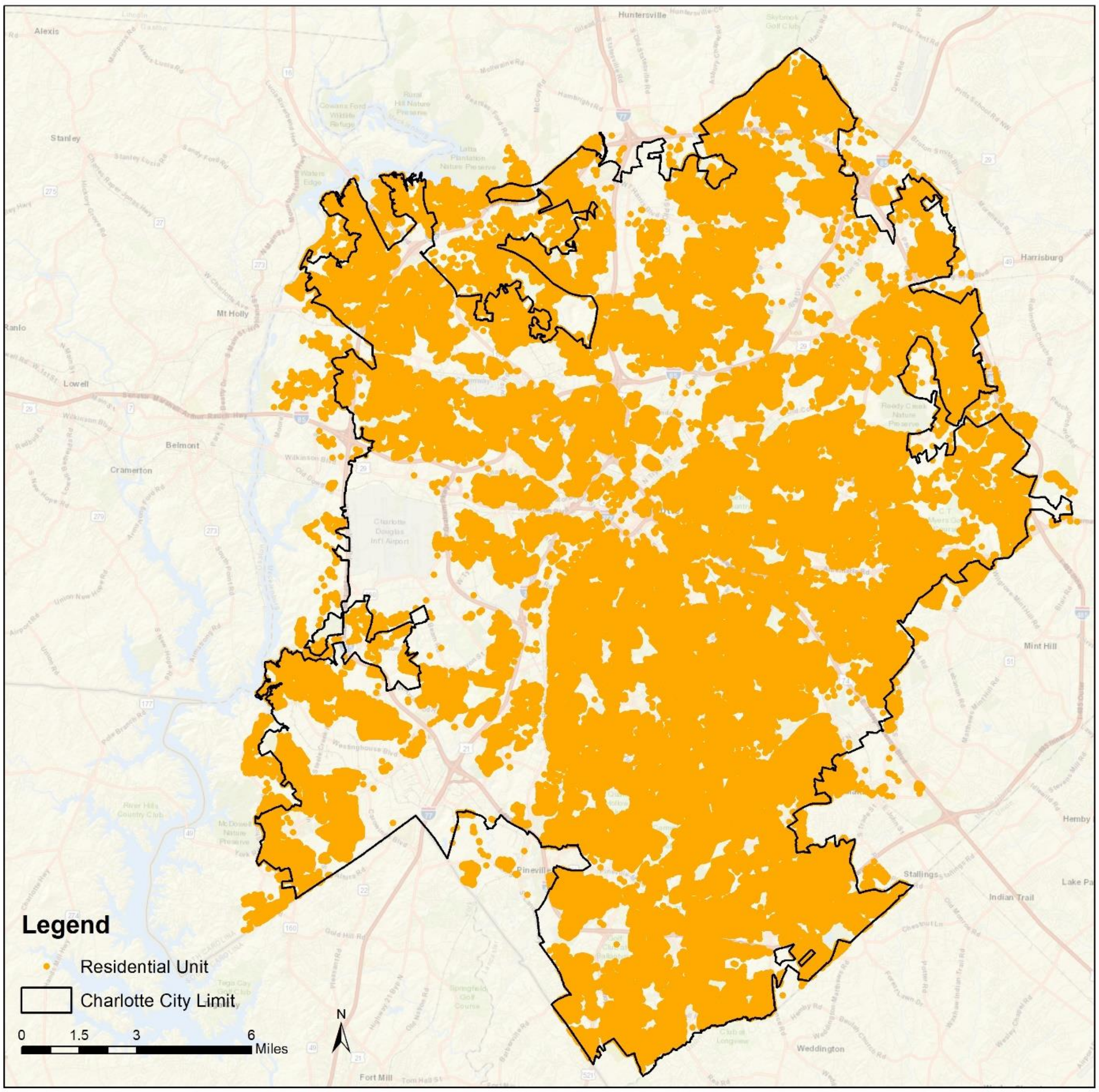


Figure 5.6 Distribution of Residential Units in the City of Charlotte

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 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 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 3, the transit gap index and its associated components are developed and presented to evaluate the public transit system to better understand the equity, accessibility and service gaps. Detailed solution framework and case study have been discussed in previous chapters. This chapter focuses on the numerical results of the developed methodology. Numerical results of the assessment of public transit equity and accessibility 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 results of using the 0.5-mile walking catchment area that is covered by public transit stops/stations. Section 6.3 gives detail results and analysis of the public transit supplies. Section 6.4 discusses the numerical results of the transit demand. Section 6.5 presents the gap analysis of the public transit system and provides a comprehensive discussion. Finally, a summary concludes this chapter in Section 6.6.

6.2 0.5-mile Walking Catchment Area

The 0.5-mile walking catchment area covered by public transit stops/stations is the starting point of the gap analysis conducted in this study. As discussed in Chapter 4, the 0.5-mile walking catchment areas are obtained by utilizing the "Network Analyst" and creating the "Service Area" for each public transit stop/station. Figure 6.1 displays the 0.5-mile walking catchment areas for each public transit stop/station. It can also be seen that there are a lot of overlaps among areas since some of the stops/stations are very close to each other.

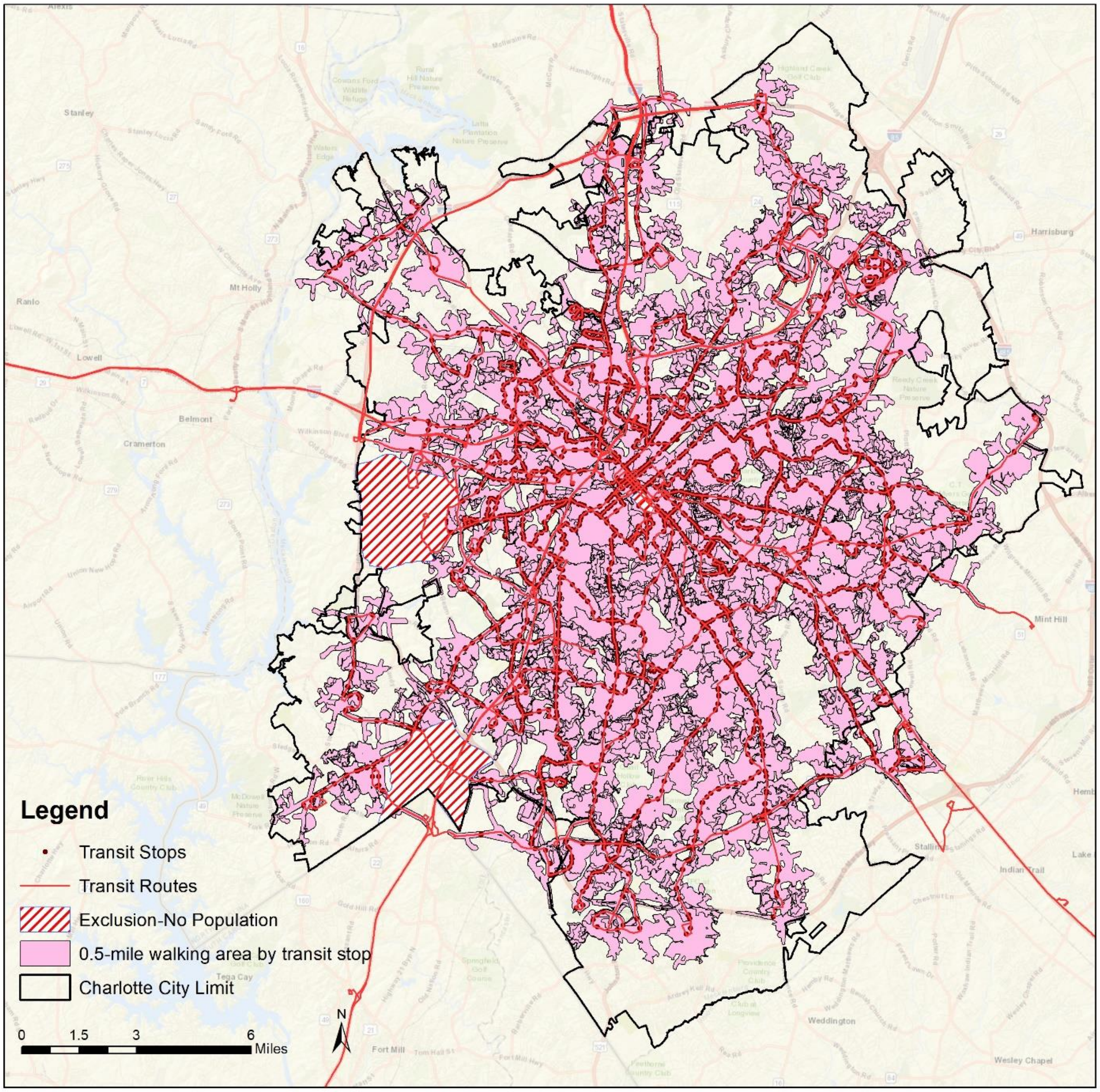


Figure 6.1 0.5-mile Walking Catchment Areas for Each Public Transit Stop/Station

6.3 Public Transit Supply

Transit service score and its associating two major components are obtained by following the steps that were designed in Section 4.4. As discussed in previous sections, the transit service score for each blockgroup is used to represent the public transit supply of CATS in Charlotte. After obtaining all the value of *TSS* for each blockgroup, the majority of the data fall into the range of 0 to 100 with only seven values greater than 100. In order not to underestimate the transit supplies, 100 is set as the TSS_{max} when further calculating the *TSS'*.

Table 6.1 shows different categories of transit service score (*TSS'*) with corresponding numbers of affected blockgroups and their transit dependent populations. Figure 6.1 presents the frequency distribution of TD population and number of blockgroups of each TSS' category. Figure 6.2 displays the spatial distribution of the transit supply (*TTS'*).

Table 6.1 Transit Service Score Categories with Corresponding Numbers of Blockgroups and TD Populations

Transit Service Score	Number of Blockgroups	Number of Transit Dependent Population	Total Population
Very Low	165	82,084	348,740
Low	82	41,177	161,424
Medium-Low	75	35,502	132,051
Medium	49	22,106	85,385
Medium-High	45	18,694	59,316
High	39	11,965	40,337
Very High	16	4,075	15,376
Total	471	215,603	842,629

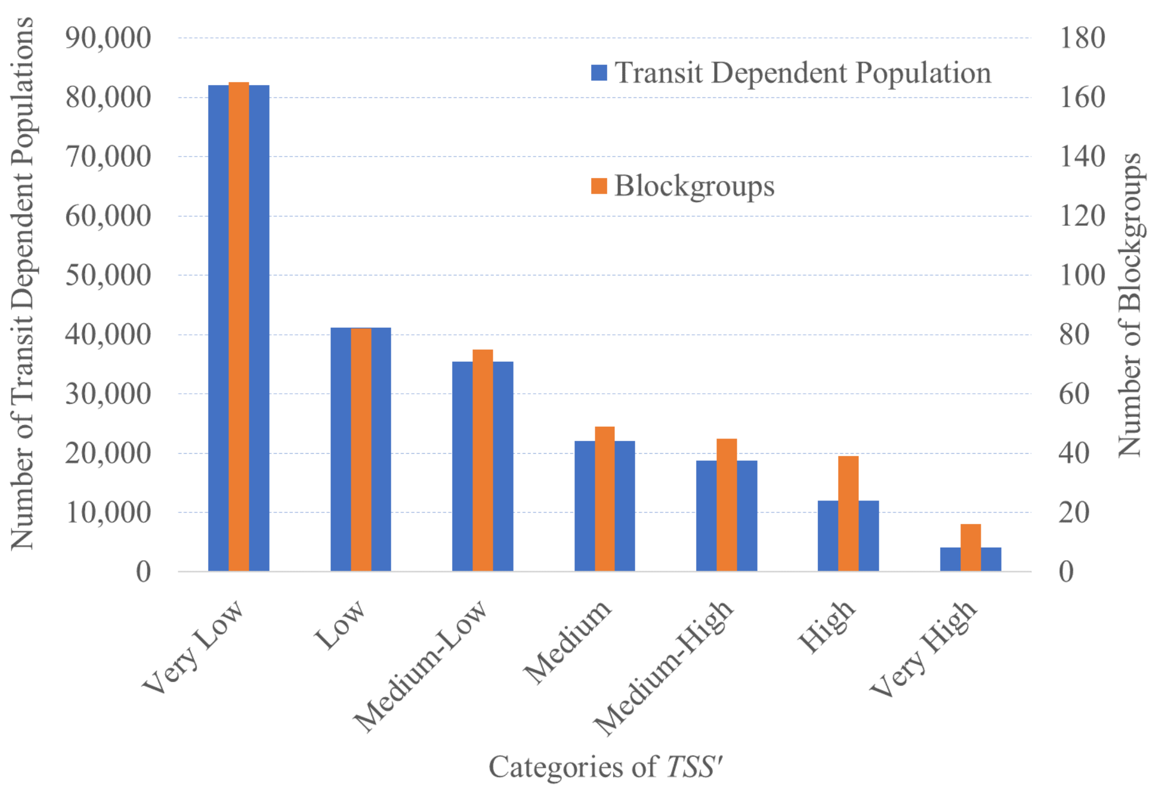


Figure 6.2 Frequency Distribution of TD Population and Number of Blockgroups of Each *TSS'* Category

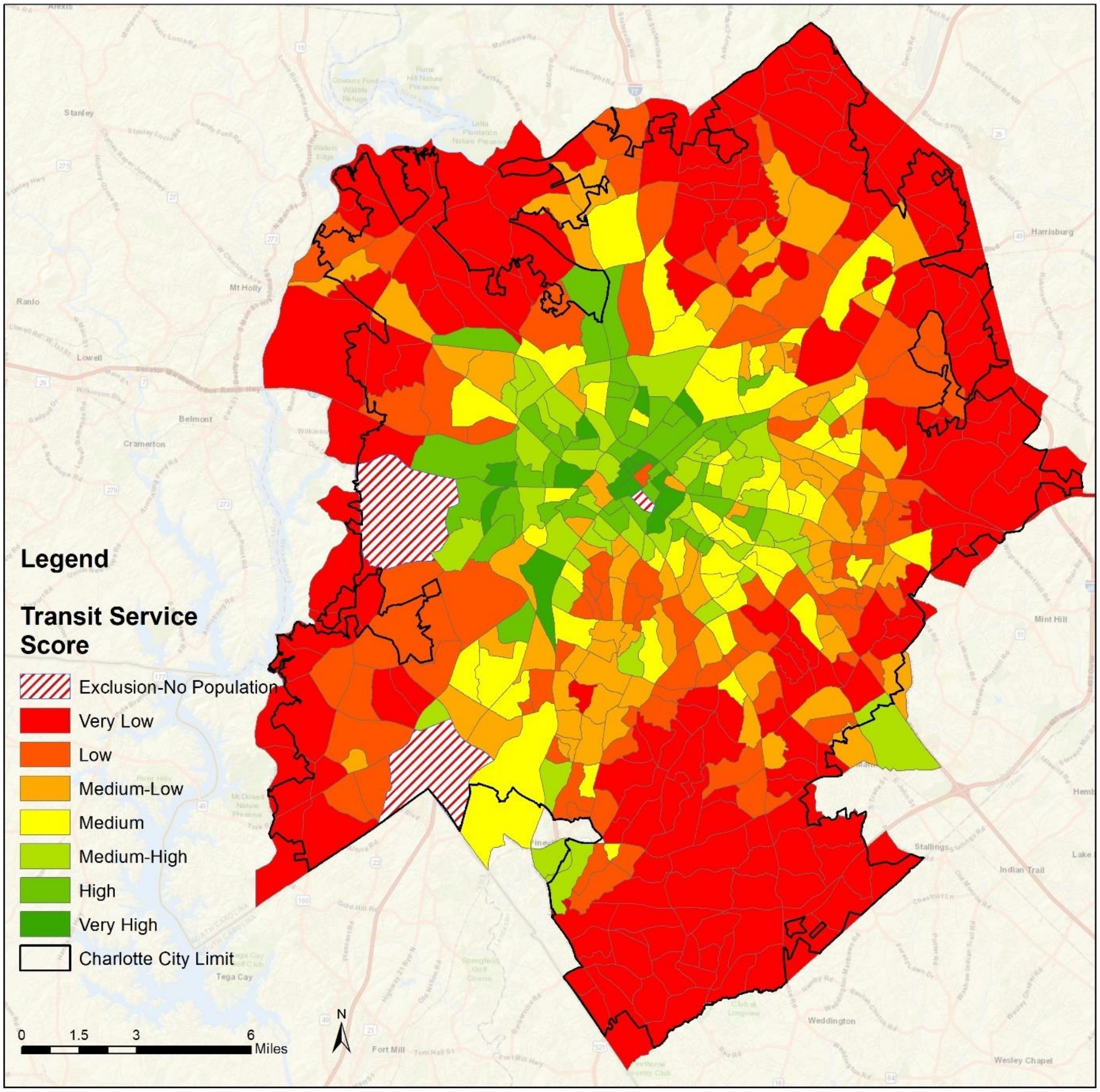


Figure 6.3 Spatial Distribution of Transit Supply (Transit Service Score, *TSS'*)

As can be seen in the Table 6.1, according to the *TTS'*, a total of 123,261 transit dependent people (which correspond to 57.17% of the total transit dependent population; 14.63% of the total population) within 247 blockgroups (i.e., 52.44% of the total blockgroups being analyzed) in the study area have poor public transit services (i.e., *TSS'* are either "Very Low" or "Low"). It is easy to see from the Figure 6.3 that the areas suffered poor transit services are distributed away or far away from the central business district of Charlotte. On the contrary, downtown Charlotte (actually called uptown Charlotte in a very unique way) areas (and adjacent areas) are served with much better public transit services.

6.4 Public Transit Demand

By excluding the specific type of populations from the total population, which is introduced in Chapters 3 and 4, transit dependent population are computed to further determine the transit demand score for each blockgroup representing the public transit demand. It is noted that the dataset, the "total population, household type (including living alone) by relationship, 2012-2016 American Community Survey (ACS) 5-Year Estimates" does not have the category of "non-institutionalized population living in group quarters". Therefore, "Population in group quarters, group quarters population by sex by age by group quarters type, 2010 Census Summary File 1 100% data" is used here to estimate the proportion of non-institutionalized population living in group quarters in the total population living in group quarters.

Table 6.2 presents different categories of transit dependent score (*TDS'*) with corresponding numbers of blockgroups and their transit dependent populations. Figure 6.4 shows the frequency distribution of TD population and number of blockgroups within each *TDS'* category. Figure 6.5 displays the spatial distribution of the transit demand (*TDS'*).

Table 6.2 Transit Dependent Score Categories with Corresponding Numbers of Blockgroups and TD Populations

Transit Dependent Score	Number of Blockgroups	Number of Transit Dependent Population	Total Population
Very Low	61	4,885	92,719
Low	85	18,784	158,809
Medium-Low	100	38,415	208,586
Medium	88	43,549	166,008
Medium-High	64	39,980	114,729
High	33	21,933	46,088
Very High	40	48,057	55,690
Total	471	215,603	842,629

Figure 6.4 Frequency Distribution of TD Population and Number of Blockgroups of Each *TDS'* Category

Category "Very High" of the *TDS'* has the largest portion of transit dependent population (i.e., with the greatest transit needs) compared to other categories (48,057, 22.29% of total transit dependent population and 86.29% of the total population within these blockgroups). Within this category, there are 40 blockgroups, which is less than 9% of the total number of blockgroups in the Charlotte area. However, many of them are distributed away or far away from the downtown Charlotte areas. Since transit service scores are higher within the downtown areas, such phenomenon would inevitably result in redundancies and deficiencies of public transit services. Despite the "Very High" category, most transit dependent populations are residing in areas with medium *TDS'* categories (i.e., "Medium-Low", "Medium" and "Medium-High"). As can be seen from Figure 6.4, 32.46% of the transit dependent population (69,990 people out of 215,603) in the Charlotte area are living in the area with high transit need (categories "Very High" and "High"). According to Figure 6.5, the spatial distribution of such areas with higher transit need disperse in the whole study area. Though not all of them locate in the fringe of the city, many of them are still in the suburban or rural portion of the city.

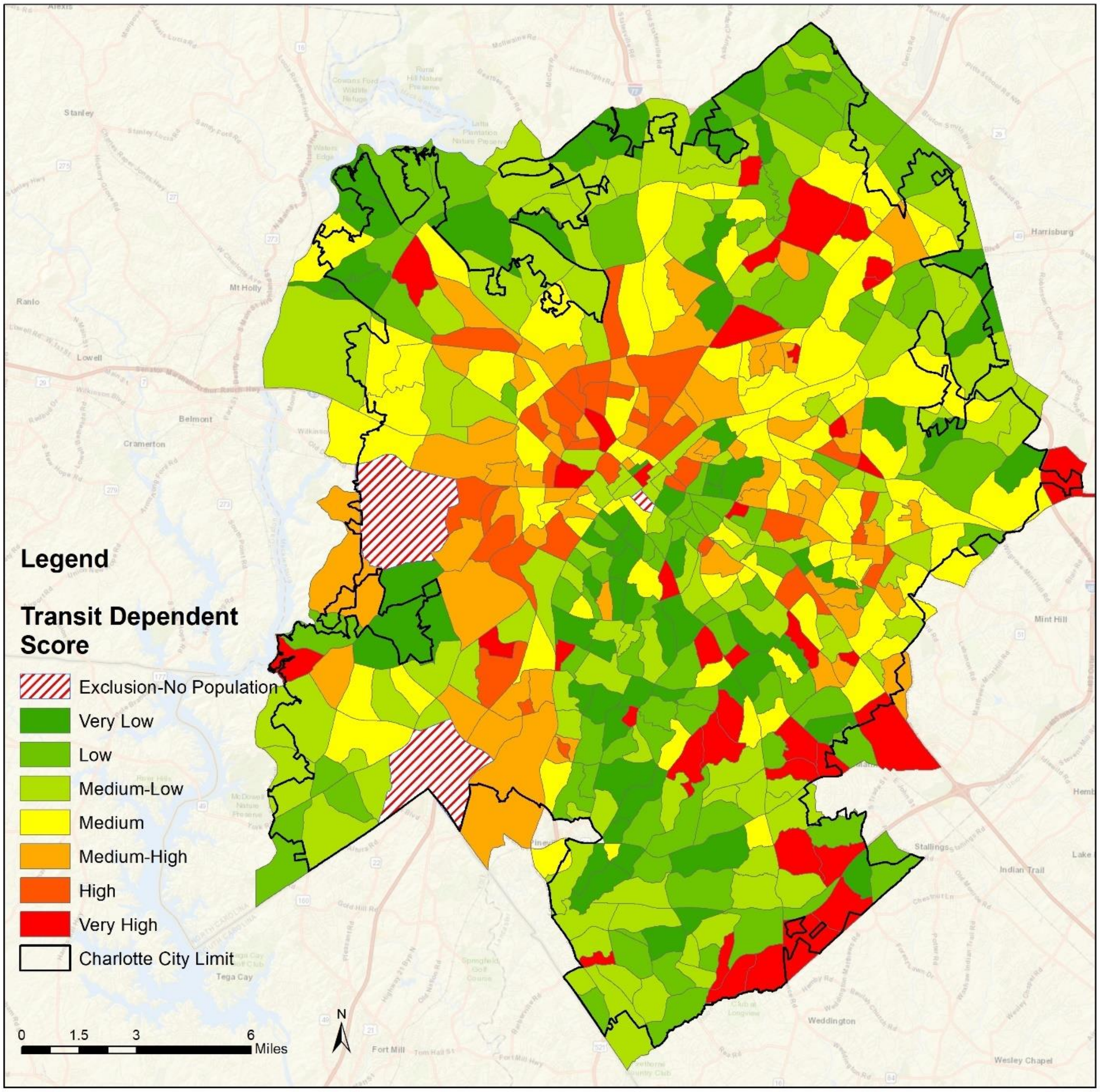


Figure 6.5 Spatial Distribution of Transit Demand (Transit Dependent Score, *TDS'*)

6.5 Public Transit Gap Analysis

Finally, according to Section 4.5, by subtracting *TDS'* from *TSS'*, the *TGI* for each blockgroup can be calculated. Figure 6.4 illustrates the spatial distribution of the *TGI*. Table 6.3 presents different categories of transit dependent score (*TGI*) with corresponding numbers of blockgroups and their transit dependent populations. Figure 6.7 shows the frequency distribution of TD population and number of blockgroups within each *TGI* category.

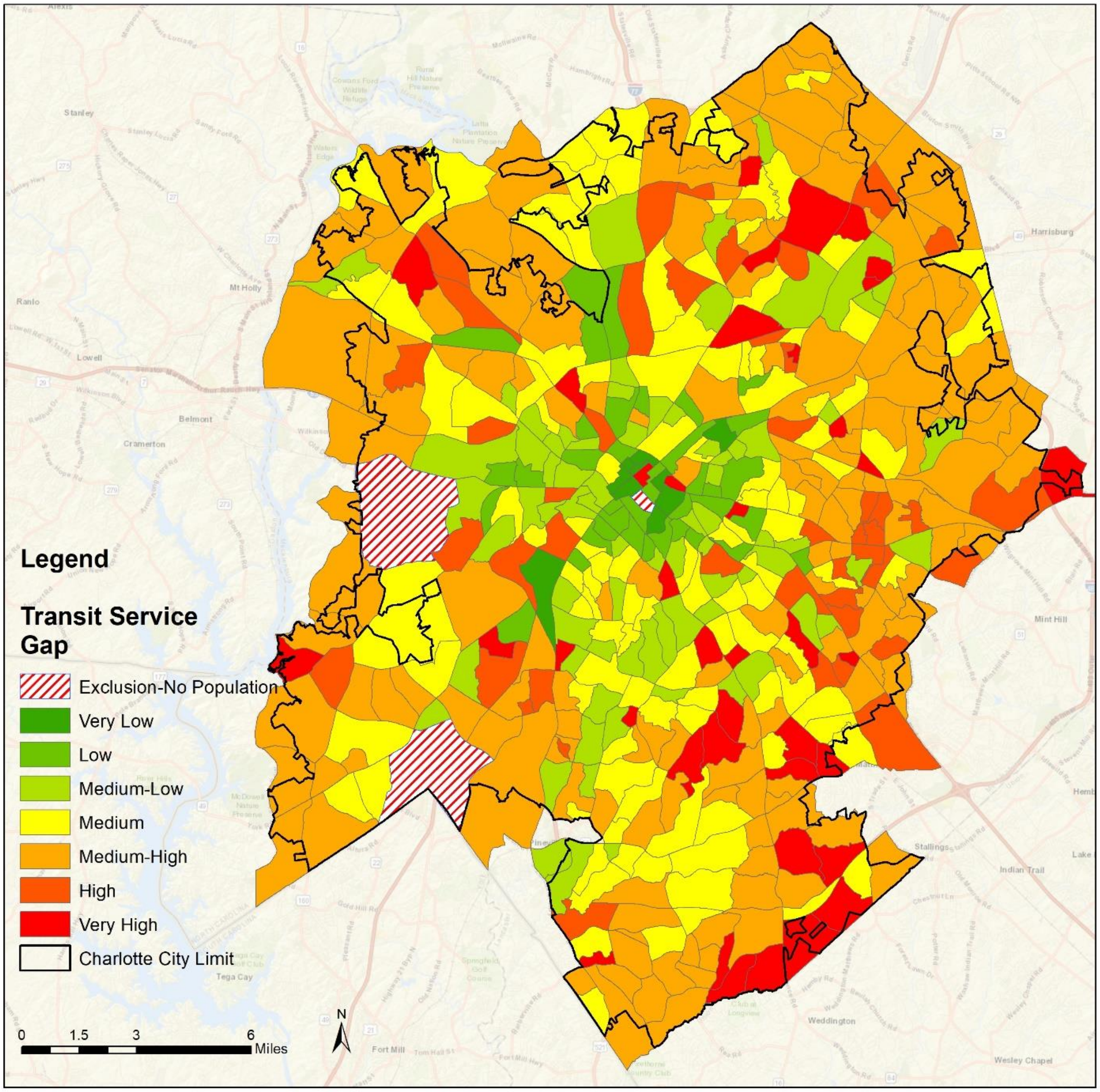


Figure 6.6 Spatial Distribution of Transit Gap (Transit Gap Index, *TGI*)

Table 6.3 Transit Gap Index Categories with Corresponding Numbers of Blockgroups and TD Populations

Transit Service Score	Number of Blockgroups	Number of Transit Dependent Population	Total Population
Very Low	11	2,068	11,002
Low	35	7,157	36,779
Medium-Low	79	18,676	98,342
Medium	112	31,487	191,570
Medium-High	154	74,761	353,352
High	45	36,828	99,489
Very High	35	44,626	52,095
Total	471	215,603	842,629

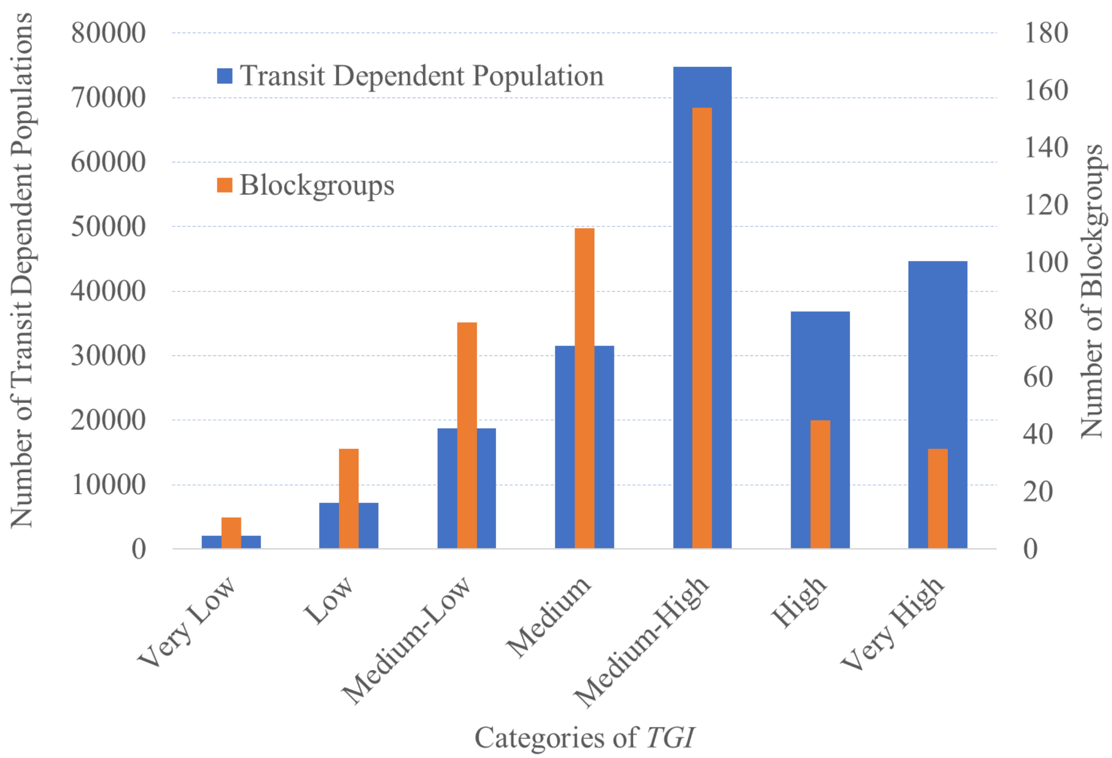


Figure 6.7 Frequency Distribution of TD Population and Number of Blockgroups of Each *TGI* Category

From Figure 6.6, downtown Charlotte areas have lower transit gaps. Even though the pattern is not obvious, for *TGI* below “High”, the gap increases as the distance between central business district of Charlotte and the blockgroup increases. For *TGI* categories above “Medium-High”, the gaps are dispersed over the study area. The category of “Medium-High” ranks the highest on both transit dependent population and number of blockgroups (74,761 people, 34.68% of the total transit dependent population; 154 blockgroups, 32.70% of the total number of blockgroups). The statistics show that 33.80% of the transit dependent population resides in high gap areas (81,454 people with “High” and “Very High” *TGI*, in 80 blockgroups).

In order to further explore the facts behind the *TGI*, other than the Jenks natural breaks, smaller even intervals are also applied to *TGI*. Figure 6.8 displays the frequency distribution of *TGI*. The majority of the public transit service equity and accessibility seem to be below average in the Charlotte area according to the analysis.

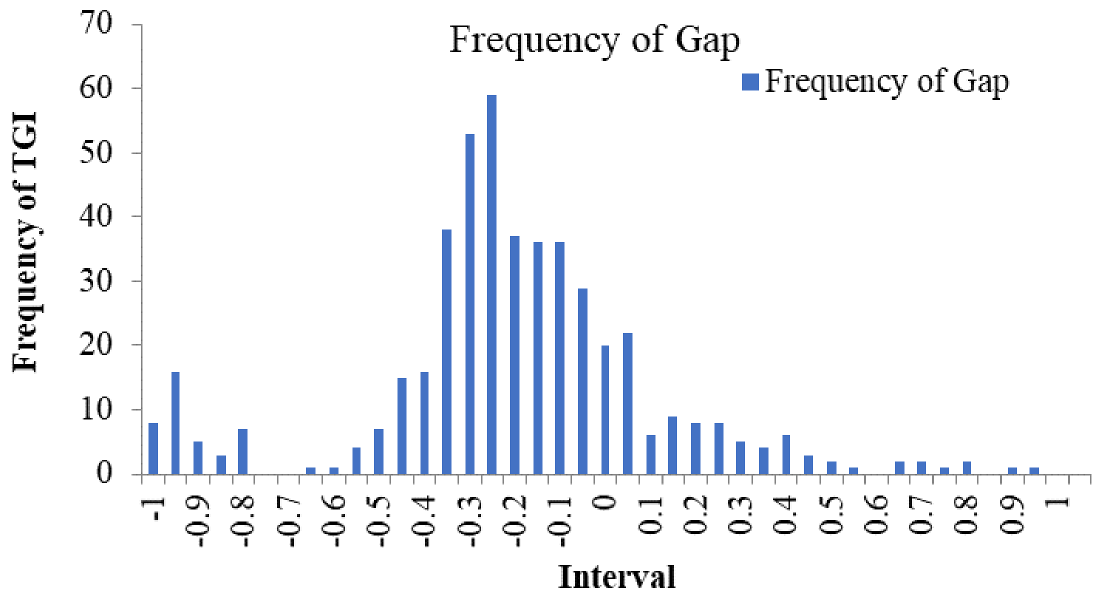


Figure 6.8 Frequency Distribution of *TGI*

Figure 6.9 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’*. No obvious relationship is found between the public transit supply and demand. Thus, this implies that the gaps are dispersed over the whole Charlotte area, which seems to be in line with other studies (Currie, 2010; Bejleri et al., 2018). Based on this scatter plot, areas 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’*) are identified. Figure 6.10 shows the areas with both public transit deficiency and redundancy.

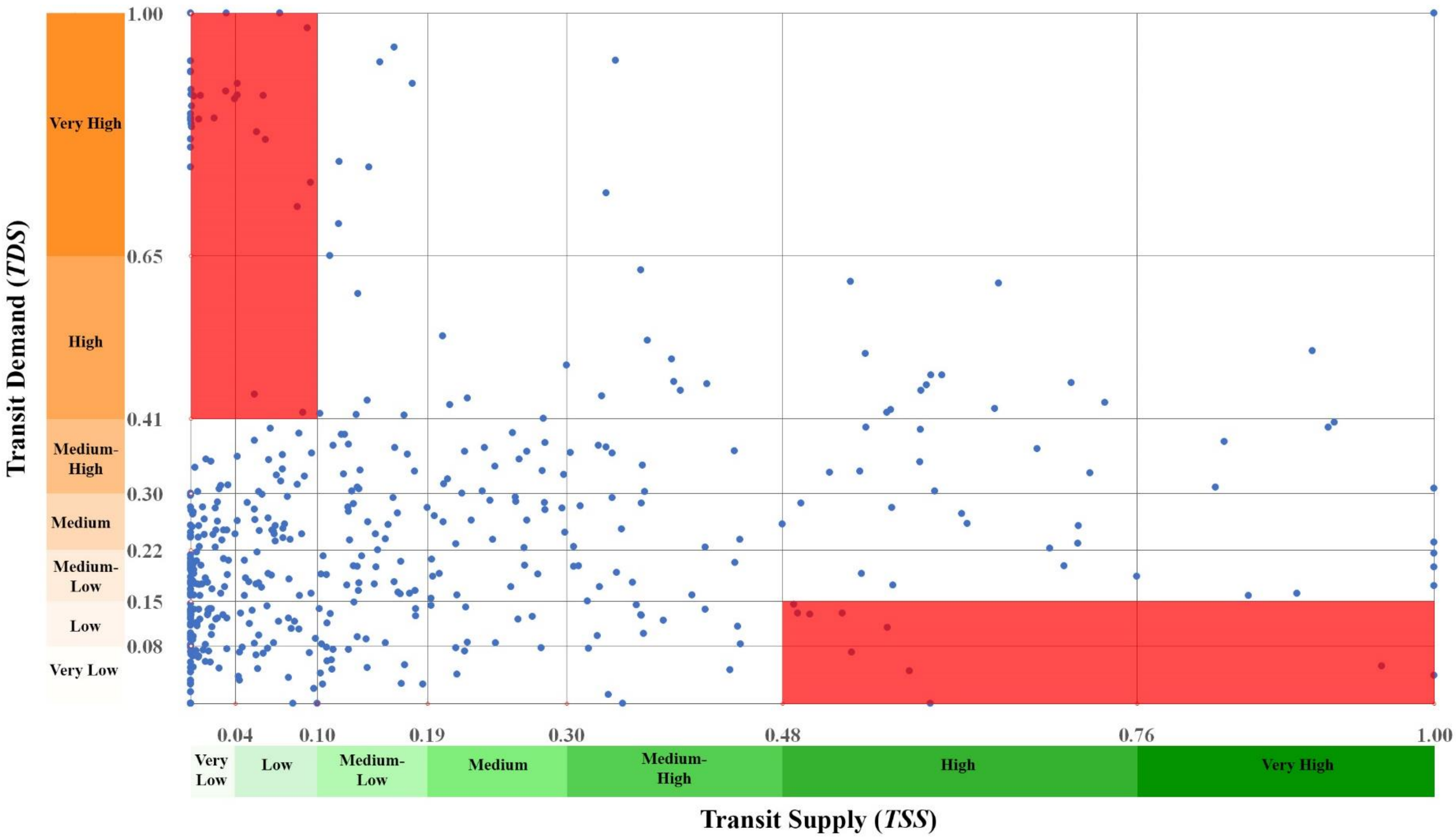


Figure 6.9 Scatter Plot of Transit Supply (*TSS*) and Transit Demand (*TDS*)

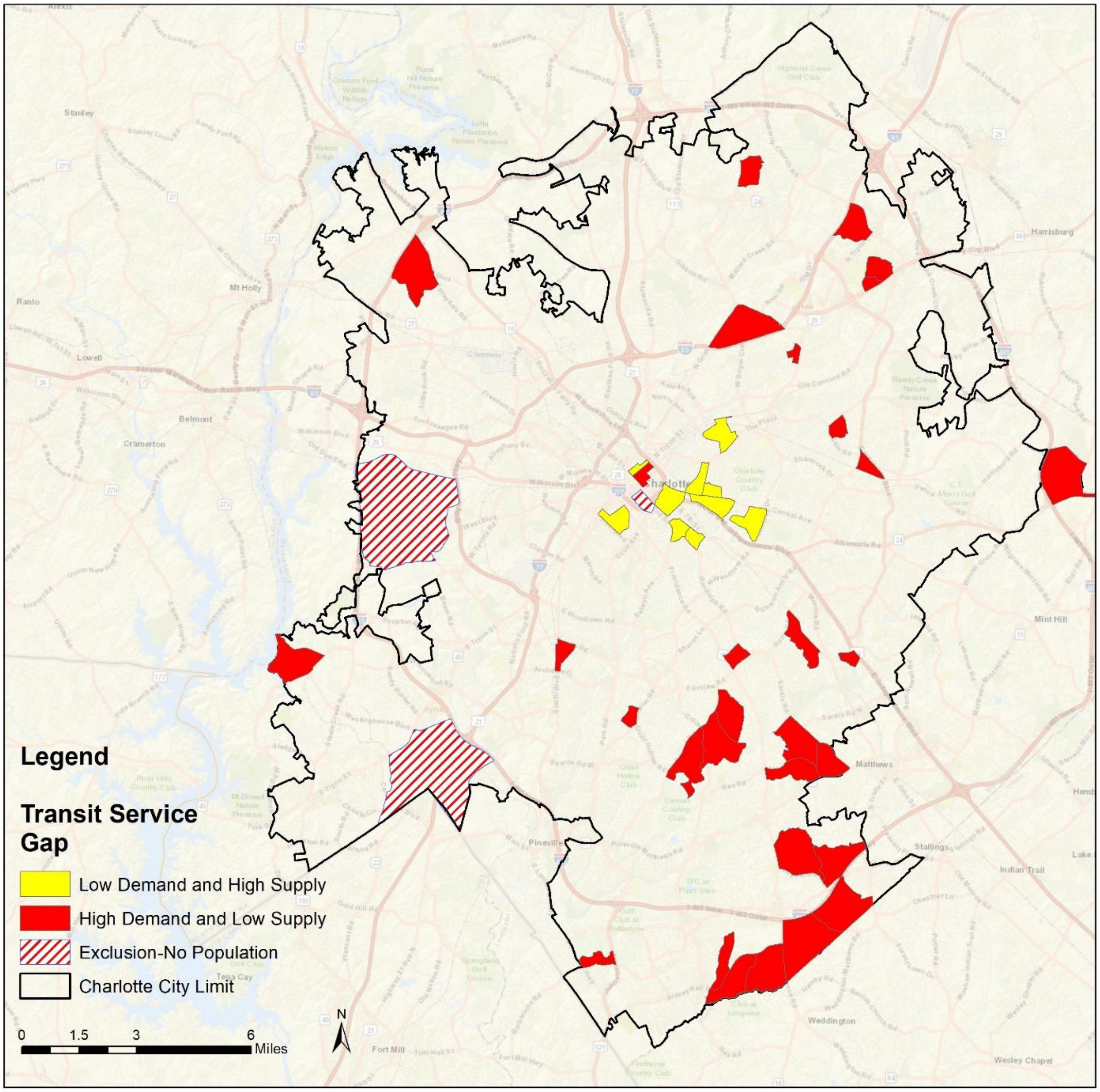


Figure 6.10 Spatial Distribution of Transit Supply Redundancies and Deficiencies

According to another transit assessment study that covers the City of Charlotte (Jiao, 2013), the results have things in common: 1) the central Charlotte areas are fairly well-served; 2) many of the most transit dependent population areas are residing in the fringe of the city, where most of the areas are suburban and rural portion of the city; and 3) there are still a few census blockgroups with greater demand than supply that are dispersed in the whole Charlotte area.

6.6 Summary

This chapter presents detailed results of the gap analysis methodology that is developed and applied in the case study of the City of Charlotte. Interpretations from both supply and demand sides are given, along with the transit service gap. Results show that the overall public transit equity and accessibility of CATS are below average. Moreover, areas of public transit service deficiency and redundancy are identified in this chapter.

Chapter 7. Summary and Conclusions

7.1 Introduction

Public transit as a component of our transportation system, plays a significant and necessary role in ensuring the mobility of people. The equity and accessibility of the public transit system are the determining factors to ensure such functionality. Many researchers have addressed the issues of the equity/accessibility assessment by considering various perspectives and impacts. However, efforts still need to be made to enhance and enrich the relevant set for better understanding the equity and accessibility of public transit service system.

Additionally, as a novel and standard format of transit feed data, GTFS receive unprecedented speed of development and attract more and more attention from many researchers. Due to the characteristics of convenience and efficiency, GTFS data could provide a greater chance for efficient and effective public transit performance evaluation and enable various analysis more than ever. Many studies have shown the potential benefits of utilizing GTFS data. By taking such advantages, public transit equity and accessibility assessment need to be put forward, since it is an important avenue of research that is in line with the FAST ACT priority area of "Improving Mobility of People and Goods".

The primary objective of this research is to develop a suitable metric/measurement to better evaluate and improve the equity and accessibility of public transit service system. The transit gap index (*TGI*) is formulated by taking demographic features, spatial and temporal transit service characteristics into considerations. A case study in the City of Charlotte and the associated comprehensive gap analysis based on the proposed methodology are also conducted and presented.

The rest of this chapter is organized as follows. In section 7.2, the principal features of the proposed methodology (i.e., *TGI* and its components) are reviewed and a summary of the conclusions made based on the numerical results derived from the case study is discussed. 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 employment of GTFS data as a basis in the public transit equity/accessibility assessment as it might help improve the quality of transit service and benefit the transit dependent populations. A comprehensive review of the state-of-the-art/practices on the transit equity/accessibility evaluation modeling has been conducted. Particularly, existing studies focusing on applying GTFS data to assess the public transit service have been explored. Based on the classification of horizontal/vertical equity, attentions have been paid to focus on transit gap analysis that is specifically associated with transit dependent population, which has been treated as vertical equity in this study. A transit gap index has been formulated along with the discussion of the solution framework and a case study in the City of Charlotte.

The developed transit gap index consists of two major components from both transit supply and demand points of view. In examining the transit supply, *TSS* is formulated concerning both the spatial and temporal characteristics of the transit service. 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 other than the covered area, which provides more accuracy. Based on several very recent research efforts, the distances that users are willing to walk to the transit stop/station might be larger compared to the results from previous studies. Thus, in order not to underestimate the transit service coverage, the 0.5-mile walking distance has been applied here. On the other hand, in order to eliminate and avoid the "double counts" by simply adding separate transit dependent populations together, a method that excludes non-overlapping non-transit-dependent populations from the whole has been used. As for data inputs, the proposed methodology requires less data and has easier access to all the datasets.

Furthermore, an ArcGIS-based solution framework has also been developed to conduct the gap analysis. The workflow is straightforward with simple operations that have been demonstrated in previous chapters. By taking advantage of the ArcGIS, a 0.5-mile walking catchment area for each transit stop/station, which is closer to the actual situation than previous method by simply applying circle buffer, can be obtained to enable more precise analysis. It has also been shown that the ArcGIS-based solution can handle the GTFS data very well in such gap analysis.

Additionally, a case study in which the proposed method and solution framework is applied has been conducted in the City of Charlotte by analyzing the CATS. The spatial distribution of the transit supply shows that the supply is centrally oriented with higher service coverages in the central business district of Charlotte and decreases as the distance to the downtown Charlotte area increases. However, even though not all of them locate in the fringe of the city, many of the most transit dependent areas are still in the suburban or rural portion of the city. A few areas with transit deficiency where high transit dependent populations are served by low transit supply have been identified, along with the areas with transit redundancy. As mentioned in Chapter 6, most of the results share things in common with several previous transit assessment related studies.

7.3 Directions for Future Research

In this section, some of the limitations of the developed gap analysis framework in this research are presented and directions for further research are also discussed.

According to the literature review in Chapter 2, public transit assessment related research efforts commonly utilize the travel survey data to generate OD to better understand and evaluate the transit service based on travel behaviors, directions, and purposes. Though this project tends to develop and use a simple method with fewer data inputs to accomplish the public transit equity/accessibility assessment, a potential future research direction could focus on integrating OD data to further improve the whole gap analysis.

It has been discussed at the very beginning that the public transit evaluation needs to include various perspectives and characteristics related to supply, demographics and other socioeconomic data. For example, one of the very important factors, transit fares, has not been considered when developing the transit gap index. Thus, considering more socioeconomic characteristics when building the assessment metrics could potentially enhance the accuracy by focusing more on transit dependent populations, since the classification of transit equity used here is vertical equity.

It should also be pointed out that as mentioned in Section 3.1, the walking distance of 0.25 or 0.5 mile might underestimate the capability of the public transit system. Conducting further sensitivity analysis of the walking distance could help. Meanwhile, transit related facilities (such as sidewalk, Americans with Disabilities (ADA) required infrastructures and equipment) sometimes limit the walking distance or the willingness of users to walk toward the stop/station. Therefore, applying various walking distances under different situations as mentioned above in the analysis could be another research direction.

References

1. 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.
2. 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.
3. 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.
4. Capital Area Transit Authority (CATA). (2011). *Michigan/Grand River Avenue Transportation Study*. Lansing, MI.
5. Catala, M., Dowling, S., and Hayward, D. (2011). *Expanding the Google Transit Feed Specification to Support Operations and Planning* (No. FDOT BDK85# 977-15).
6. Chen, Y., Ravulaparthy, S., Deutsch, K., Dalal, P., Yoon, S., Lei, T., and Hu, H. H. (2011). Development of Indicators of Opportunity-Based Accessibility. *Transportation Research Record: Journal of the Transportation Research Board*, *2255*, 58-68.
7. City of Charlotte, North Carolina, Code of Ordinances, Appendix A - Zoning, Chapter 9: - General Districts, Part 12: - Transit Oriented Development Districts. (2018)
8. Colopy, J.H., (1994). Road Less Traveled: Pursuing Environmental Justice Through Title VI of The Civil Rights Act of 1964. *The Stanford Environmental Law Journal* 13, 125–189.
9. Currie, G. (2010). Quantifying Spatial Gaps in Public Transport Supply Based on Social Needs. *Journal of Transport Geography*, *18*(1), 31-41.
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. Farber, S., Morang, M. Z., and Widener, M. J. (2014). Temporal Variability in Transit-Based Accessibility to Supermarkets. *Applied Geography*, *53*, 149-159.
14. Fayyaz, S. S.K., Liu X.C. and Zhang, G. (2017) An Efficient General Transit Feed Specification (GTFS) Enabled Algorithm for Dynamic Transit Accessibility Analysis. *PLoS ONE 12*(10): e0185333. Accessed December 18, 2017 from https://doi. org/10.1371/journal.pone.0185333.
15. Forester, J., and Krumholz, N. (1990). *Making Equity Planning Work: Leadership in the Public Sector*. Temple University Press, Philadelphia, PA.
16. Fortin, P., Morency, C., and Trépanier, M. (2016). Innovative GTFS Data Application for Transit Network Analysis Using a Graph-Oriented Method. *Journal of Public Transportation*, *19*(4), 2.

17. Fransen, K., Neutens, T., Farber, S., De Maeyer, P., Deruyter, G., and Witlox, F. (2015). Identifying Public Transport Gaps Using Time-Dependent Accessibility Levels. *Journal of Transport Geography*, *48*, 176-187.
18. Front Seat Management, LLC. "City-Go-Round." Accessed December 18, 2017 from http://www.citygoround.org/agencies/.
19. Gandavarapu, S. (2012). Using Google Transit Feed Specification in Travel Modeling. In *Submitted for presentation at 4th Transportation Research Board Conference on Innovations in Travel Modeling. http://onlinepubs.trb.org/onlinepubs/conferences/2012/4thITM/Papers-R/0117-000113.pdf* (pp. 0117-000113).
20. Garrett, M., and Taylor, B. (1999). Reconsidering Social Equity in Public Transit. *Berkeley Planning Journal*, *13*(1).
21. Grengs, J. (2001). Does Public Transit Counteract the Segregation of Carless Households? *Transportation Research Record* 1753: 3-10.
22. Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, *25*(2), 73-76.
23. Ingram, D. R. (1971). The Concept of Accessibility: A Search for An Operational Form. *Regional Studies*, *5*(2), 101-107.
24. Jaramillo, C., Lizárraga, C., and Grindlay, A. L. (2012). Spatial Disparity in Transport Social Needs and Public Transport Provision in Santiago de Cali (Colombia). *Journal of Transport Geography*, *24*, 340-357.
25. Jennifer Lee. (2009). "A Transit Voice Planner? 'Back Up.' Back Way Up." The New York Times. Accessed December 18, 2017 from http://cityroom.blogs.nytimes.com/2009/06/08/a-transit-voice-planner-back-up-back-way-up/.
26. 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.
27. Jiao, J., and Dillivan, M. (2013). Transit Deserts: The Gap between Demand and Supply. *Journal of Public Transportation*, *16*(3), 2.
28. 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*.
29. Rawls, J. (1971). *A Theory of Justice*, The Belknap Press of Harvard University Press (www.hup.harvard.edu). Accessed December 18, 2017 from http://en.wikipedia.org/wiki/A_Theory_of_Justice.
30. Kawabata, M. (2002). Job Access and Work among Autoless Adults in Welfare in Los Angeles. *The Ralph and Goldy Lewis Center for Regional Policy Studies*. Accessed December 18, 2017 from https://escholarship.org/uc/item/6bq3457v#page-3.
31. Litman, T. (2002). Evaluating Transportation Equity. *World Transport Policy & Practice*, *8*(2), 50-65.
32. Liu, Y., and Cirillo, C (2015) Measuring Transit Service Impacts on Vehicle Ownership and Use. *Public Transport*, *7*(2), 203-222.
33. Ma, T., and Jan-Knaap, G. (2014). Analyzing Employment Accessibility in A Multimodal Network Using GTFS: A Demonstration of the Purple Line, Maryland. In *the Association of Collegiate Schools of Planning (ACSP) Annual Conference, Philadelphia, PA.*

34. 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.
35. Morris, J. M., Dumble, P., and Wigan, M. R. (1979). Accessibility Indicators for Transport Planning. *Transportation Research Part A: General*, 13(2):91 – 109.
36. Nassir, N., Khani, A., Lee, S., Noh, H., and Hickman, M. (2011). Transit Stop-Level Origin-Destination Estimation Through Use of Transit Schedule and Automated Data Collection System. *Transportation Research Record: Journal of the Transportation Research Board*, *2263*, 140-150.
37. Nazem, M., Trépanier, M., and Morency, C. (2013). Integrated Intervening Opportunities Model for Public Transit Trip Generation-Distribution: A Supply-Dependent Approach. *Transportation Research Record: Journal of the Transportation Research Board*, *2350*, 47-57.
38. Neutens, T. (2015). Accessibility, Equity and Health Care: Review and Research Directions for Transport Geographers. *Journal of Transport Geography*, *43*, 14-27.
39. 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.
40. Owen, A., and Levinson, D. M. (2015). Modeling the Commute Mode Share of Transit Using Continuous Accessibility to Jobs. *Transportation Research Part A: Policy and Practice*, *74*, 110-122.
41. Owen, A., Murphy, B., and Levinson, D. M. (2016). Access Across America: Transit 2015. *University of Minnesota Digital Conservancy, Minneapolis, MN.* Accessed December 18, 2017 from www.cts.umn.edu/Publications/ResearchReports/pdfdownload.pl?id=2740.
42. Sanchez, T. W. (1998). *The Connection between Public Transit and Employment*. Presented at the Association of Collegiate Schools of Planning Annual Conference, Pasadena, CA. Accessed December 18, 2017 from http://www.upa.pdx.edu/CUS/publications/docs/DP98-7.pdf.
43. Sarker, A. A., Welch, T. F., Golias, M. M., and Kumar, A. Measuring Transit Connectivity using GTFS Data. Accessed December 18, 2017 from http://tfresource.org/images/e/ea/ITM16_Measuring_Transit_Connectivity_using_GTFS_Data.pdf.
44. 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.
45. Tomer, A., Kneebone, E., Puentes, R., and Berube, A. 2011. *Missed Opportunity: Transit and Jobs in Metropolitan America*. Washington, DC: The Brookings Institution.
46. TransitWiki. “General Transit Feed Specification” Accessed December 18, 2017 from https://www.transitwiki.org/TransitWiki/index.php/General_Transit_Feed_Specification.
47. Tribby, C. P., and Zandbergen, P. A. (2012). High-Resolution Spatio-Temporal Modeling of Public Transit Accessibility. *Applied Geography*, *34*, 345-355.
48. Tsou, K. W., Hung, Y. T., and Chang, Y. L. (2005). An Accessibility-Based Integrated Measure of Relative Spatial Equity in Urban Public Facilities. *Cities*, *22*(6), 424-435.
49. US Census Bureau. (2011). American Community Survey Puerto Rico Community Survey 2009 Subject Definitions.
50. Welch, T. F., and Mishra, S. (2013). A Measure of Equity for Public Transit Connectivity. *Journal of Transport Geography*, *33*, 29-41.

51. Wong, J. C. (2013). *Use of The General Transit Feed Specification (GTFS) In Transit Performance Measurement* (Doctoral dissertation, Georgia Institute of Technology). Accessed December 18, 2017 from https://smartech.gatech.edu/bitstream/handle/1853/50341/WONG-THESIS-2013.pdf?sequence=1&isAllowed=y
52. Wong, James (2013) Leveraging the General Transit Feed Specification for Efficient Transit Analysis. *Transportation Research Record*, *2338*, 11-19.
53. 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.