Title

Using Multi-Modal Path-Specific Transit Trips in Transportation Social Sustainability Analysis: Case Study in Atlanta, GA

Authors

Fan, Huiying
Lu, Hongyu
Guin, Angshuman
et al.

Publication Date

2024-09-01

DOI

10.7922/G2HH6HD0

Data Availability

The data associated with this publication are available at:
https://doi.org/10.5281/zenodo.13917429

Using Multi-Modal Path-Specific Transit Trips in Transportation Social Sustainability Analysis: Case Study in Atlanta, GA

September 2024

A Research Report from the National Center for Sustainable Transportation

Huiying ("Fizzy") Fan, Ph.D., Georgia Institute of Technology

Hongyu Lu, Georgia Institute of Technology

Angshuman Guin, Ph.D., Georgia Institute of Technology

Randall Guensler, Ph.D., Georgia Institute of Technology

National Center for Sustainable Transportation
A USDOT University Transportation Center

Georgia Tech School of Civil and Environmental Engineering
College of Engineering

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.
NCST-GT-RR-24-30

2. Government Accession No.
N/A

3. Recipient's Catalog No.
N/A

4. Title and Subtitle
Using Multi-Modal Path-Specific Transit Trips in Transportation Social Sustainability Analysis: Case Study in Atlanta, GA

5. Report Date
September 2024

6. Performing Organization Code
N/A

7. Author(s)
Huiying ("Fizzy") Fan, Ph.D., https://orcid.org/0000-0002-0351-386X
Hongyu Lu, https://orcid.org/0000-0002-0170-7169
Angshuman Guin, Ph.D., https://orcid.org/0000-0001-6949-5126
Randall Guensler, Ph.D., https://orcid.org/0000-0003-2204-7427

8. Performing Organization Report No.
N/A

9. Performing Organization Name and Address
Georgia Institute of Technology
School of Civil and Environmental Engineering
790 Atlantic Drive, Atlanta, GA 30332-0355

10. Work Unit No.
N/A

11. Contract or Grant No.
USDOT Grant 69A3551747114

12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
1200 New Jersey Avenue, SE, Washington, DC 20590

13. Type of Report and Period Covered
Research Report (February 2021 – July 2022)

14. Sponsoring Agency Code
USDOT OST-R

15. Supplementary Notes
DOI: https://doi.org/10.7922/G2HH6HD0
Dataset DOI: https://doi.org/10.5281/zenodo.13917429

16. Abstract
A previous National Center for Sustainable Transportation (NCST) study examined pandemic-related changes in MARTA transit system service and ridership in Atlanta, GA, and the combined effects on energy use and per-passenger energy use (Fan, et al., 2022). For that previous study, General Transit Feed Specification (GTFS) and the Automated Passenger Counter (APC) datasets were used to develop the transit network and derive distance and passenger load information within the TransitSim analytical framework. The research coupled ridership data with energy use and emission rates from MOVES-Matrix to assess how the changes in transit service and ridership affected energy use and emissions on a per passenger-mile basis. Research performed in this supplemental NCST study improved model algorithms to increase analytical efficiency and to integrate ridership demographics, so that energy use impacts could be assessed across demographic groups for use in social sustainability analysis. This report summarizes improvements that generated TransitSim 3.0 and provides a social sustainability modeling demonstration.

17. Key Words
Transit service, transit energy use, transit ridership, transit demographics, social sustainability analysis

18. Distribution Statement
No restrictions

19. Security Classification (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	76	N/A

Form DOT F 1700.7 (8-72)

NCST

About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: University of California, Davis; California State University, Long Beach; Georgia Institute of Technology; Texas Southern University; the University of California, Riverside; the University of Southern California; and the University of Vermont. More information can be found at: ncst.ucdavis.edu.

Disclaimer

Acknowledgments

This study was funded, partially or entirely, by a grant from the National Center for Sustainable Transportation (NCST), supported by the U.S. Department of Transportation (USDOT) through the University Transportation Centers program. The authors would like to thank the NCST and the USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project. The authors would also like to thank staff from the City of Atlanta and Atlanta Regional commission for providing network data and assistance.

Using Multi-Modal Path-Specific Transit Trips in Transportation Social Sustainability Analysis: Case Study in Atlanta, GA

A National Center for Sustainable Transportation Research Report

September 2024

Huiying "Fizzy" Fan, Research Engineer I, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

Hongyu Lu, Graduate Research Assistant, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

Angshuman Guin, Senior Research Engineer, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

Randall Guensler, Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta

[page intentionally left blank]

TABLE OF CONTENTS

List of Tables

List of Figures

Using Multi-Modal Path-Specific Transit Trips in Transportation Social Sustainability Analysis: Case Study in Atlanta, GA

EXECUTIVE SUMMARY

A previous National Center for Sustainable Transportation (NCST) study examined pandemic-related changes in MARTA transit system service and ridership in Atlanta, GA, and the combined effects on energy use and per-passenger energy use (Fan, et al., 2022). In that previous NCST report, General Transit Feed Specification (GTFS) and the Automated Passenger Counter (APC) datasets were used to develop the transit network and derive distance and passenger load information using the TransitSim 2.0 shortest path analytical framework (see also, Fan et al., 2023a). Ridership data were coupled with energy use and emission rates from MOVES-Matrix to assess how the changes in transit service and ridership affected energy use and emissions on a per passenger-mile basis. This report summarizes the improvements that generated TransitSim 3.0 to increase analytical efficiency and to integrates ridership demographics from the Atlanta Regional Commission on-board transit survey (a 10% sample of riders) to provide a social sustainability modeling demonstration. With the enhanced modeling tool, energy use impacts can now be assessed across demographic groups for each route in the system for use in social sustainability analysis. The methods outlined in this report can also be applied to passenger counts and demographics for individual routes when data are available.

The report first summarizes the model improvements and provides a modeling demonstration for social sustainability analysis. Trip-level energy use for MARTA trips from April 19, 2019, to June 1, 2019) was estimated using route data and MOVES-Matrix energy use rates. The analyses show that the average trip-level energy used by MARTA riders during the study period was 20,773 Btu, and the per-passenger-mile energy use was 2,798 Btu. The energy estimates were then allocated across demographic groups for comparison. Energy use across the race/ethnicity groups largely reflected the differences in average trip miles traveled by each group. There was no significant difference in per-passenger-mile energy use across different race/ethnicity groups or genders. Transit riders with a driver's license are associated with higher trip-level energy use compared to those without a license, but these riders also tend to have slightly lower energy use per passenger-mile. The analyses demonstrated how TransitSim 3.0 energy use analyses can be coupled with data from on-board ridership surveys to allocate energy use across demographic groups for use in social sustainability analysis in assessing potential impacts in transit investment and changes in operations.

Introduction

Properly designed and operated transit modes achieving reasonable ridership goals can significantly improve transportation sustainability by reducing energy use and emissions (Li, 2019; Li, et al., 2016; Pradono, Kusumantoro, and Retapradana, 2015; Litman, 2012), and by discouraging urban sprawl (Belzer and Autler, 2002; Freilich, 1998). Alternative transportation modes also have the potential to support social sustainability efforts in terms of enhancing transportation equity (Carleton and Porter, 2018; Griffin and Sener, 2016; Welch and Mishra, 2013). While disadvantaged groups are recognized as relying more heavily on transit and other alternative modes for mobility (Garrett and Taylor, 1999), studies have found that higher-income groups also tend to use transit for short-distance trips (Zhu et al., 2017). It is intuitive that transit service naturally benefits those in need. However, a regional planning emphasis on economic development may incentivize the provision of less sustainable modes (Linovski, Baker, and Manaugh, 2018). It is important to differentiate between the presence of transit and the ability of these systems to support mobility and accessibility. It is also important to disentangle mandatory trips from leisure travel and the activity of transit-dependent riders (those without automobile access) from choice-riders (those with automobile access who choose to take transit) in developing a more comprehensive understanding of transportation equity. The research presented in this report develops path-specific transit routes for more than one million transit trips in the Atlanta Regional Commission's travel demand model outputs and investigates the differences in transit use across demographic groups.

The social sustainability analysis analyses reported herein are supported by a path retention algorithm developed by the Georgia Tech research team (Zhao, 2021). Traditional travel demand models do not retain specific trip route information, even though the models calculate these routes internally calculated during iterative model loops for congestion estimation. Hence, every final trip between each origin-destination pair us reported with single values for travel time and distance. The path retention approach retains the origin-destination paths, allowing researchers to compare travel across synthetic household demographic groups (Zhao, 2020). The Activity-Based Model (ABM) developed by the Atlanta Regional Commission (ARC) is capable of outputting detailed trip characteristics, including travel time and mode for every trip conducted by individual trip-takers (WSP | Parsons Brinckerhoff, Atkins, 2017). The path retention algorithm then retains the detailed path information of each trip from the model. This approach has been successfully applied in analyses of demographic differences in travel demand and energy use (Zhao et al., 2019) and pollutant exposure analysis (Guensler, et al., 2022; Guensler, et al., 2022). The availability of demographic and path data is crucial for the analysis of transportation equity. However, due to the lack of a corresponding Cube Voyage API (Zhao et al., 2019), the ABM path retention algorithm is so far only available for the automobile travel mode. Although the ARC travel demand model does retain transit trip travel times, actual path information is not retained (Zhao, 2020). Many analyses model transit at the aggregated system-level (Gbologah, et al., 2014; Li, et al., 2016; Xu et al., 2015). Li, et al. (2016) produce spatially-explicit simulation and calculate emission reduction from re-assigning dead-head trips. Yet their approach is still constrained at a relatively coarse resolution. TransitSim allows users to generate the shortest origin-destination paths that are generated internally by the travel demand model and retain these trips with high spatial- and temporal-resolution. TransitSim is a part of the TransportSim family of models (RoadwaySim, TransitSim, SidewalkSim, BikewaySim, and CarpoolSim) developed by the research team at Georgia Tech. The TransportSim modules are capable of performing trip assignments using input network and trip origin-destination data (latitude and longitude), and output detailed path information for each trip modeled (Li et al., 2018).

While studies have investigated transit social sustainability through various lenses, including mobility and accessibility (Bennett and Shirgaokar, 2016; Aman and Smith-Colin, 2020), crime rates and incidents (Turner, 2013), service availability (Carleton and Porter, 2018; Griffin and Sener, 2016), connectivity (Welch and Mishra, 2013), operational efficiency (Wei et al., 2017), and inclusion in planning practices (Linovski et al., 2018; Karner and Levine, 2021), most of this research has been conducted at an aggregate level, such as traffic analysis zones (TAZs) or cities. However, it is difficult to assess differences in behavioral factors. The analyses performed for this report are conducted at the trip-level and at the household-level, such that demographic parameters can be attached to every trip for use in assessing how transit use, benefits, and impacts play out across a variety of demographic clusters.

By exploring transit social sustainability at the individual trip level, results can more-reliably assess similarities and differences in mobility and accessibility provided by transit systems across different demographic groups. This research focuses on the service experienced by individuals throughout their entire trip, analyzed at a high resolution using TransitSim 3.0. This report serves three audiences; 1) academics and researchers can use the framework and methods presented in this research to integrate trip-level system thinking into their analyses; 2) TransitSim is open-source and can be readily applied by transit agencies, metropolitan planning organizations, and departments of transportation in their planning practices; and 3) the report aims to inform the general audience about their travel conditions and offer insights into social sustainability.

The following chapters first introduce TransitSim 3.0, a model designed to simulate chosen routes, generate detailed trip trajectories at a second-by-second level, and perform modeling across large spatial and temporal scopes for multiple transit agency services and over various time periods as routes and schedules change. The subsequent chapter summarizes the case study methods and results, demonstrating the potential use of TransitSim 3.0 in transit social sustainability analysis. The demonstration also shows the integration of TransitSim 3.0 with the MOVES Matrix model for Energy Use Modeling. Finally, the conclusion and future work chapter summarizes the report and discuss implications for future research.

TransitSim 3.0 - Updated Modeling Routines

To integrate demographics from transit survey data for complex transit networks across an entire region, such as the metropolitan area of Atlanta, a model that is compatible with multiple providers is needed. Interagency analysis is complicated and often involves running multiple scenarios. For example, a case study is performed for transit networks from five providers that change over time, yields 12 different date ranges and 20 different time ranges, totaling 240 separate network scenarios. This poses high requirements on TransitSim, including: 1) computational efficiency; 2) file system complexity; and 3) algorithm, compatibility across various data conditions. To support the comparison across agencies and time periods, TransitSim 3.0 was updated to support interagency analyses and to support the integration of individual passenger trip data (second-by-second travel from origin to destination via transit with access and egress).

The TransitSim modeling scheme includes five sections: 1) distance calculation, 2) network construction, 3) routing analysis, 4) interagency modeling, and 5) visualization and extensions. Each section has various modules to conduct major types of analysis. Within each module, functions and tools support the program. Sections 1, 2, and 3 exist in the original TransitSim version. Despite the similarity in structure and some logic of the analyses, a big portion of the model is re-written to reduce the computational burden and complexity. Section 5 is also present in the old TransitSim model. These sections are recoded as standalone optional extensions so that when the project does not need a module, the main program can run smoothly and efficiently. Section 4 is newly added features to support the interagency modeling of the transit network. Table 1 summarizes the functions of the major modules in each section.

The report sections that follow the table describe the improvements to the TransitSim model in more detail. Section 1, 1. TransitSim Model Modifications for Enhanced Efficiency) presents the various strategies taken to recompile the TransitSim model for enhanced efficiency and reduced complexity. Section nteragency and Multi-Period Transit Routing Analysis) introduces the steps involved in downloading, preparing, and processing interagency transit network data and the strategy to use it in interagency transit analysis. Section xtension and Visualizations discusses some of the most important extensions to connect with other models and commonly used visualization options and proposes potential extensions that can be developed in future research. Finally, Section 4summarizes the differences between the previous TransitSim 2.0 and the updated TransitSim 3.0.

Table 1. Summary of TransitSim Sections and Modules

Module	Description	Improvements in TransitSim 3.0
	1. Distance Calculation	
1-1. Load data	Import formatted-GTFS files	One less file needed
1-2. Sub-sampling	When the sample size is too large (over 1 million stops time record)	New feature to enhance computational efficiency
1-3. Spacing	Add details to transit GPS trajectory to enhance analysis accuracy	1) A laddered system of approaches with two functions, spacing, and back-space; 2) Individual trip-based conditions to decide the number of iterations; 3) Thresholds are created based on the results of sensitivity analysis
1-4. Register	Register (match) GPS trajectory to segments corresponding to transit links (stop-to-stop pairs) to calculate accurate travel distance for each link	1) New algorithm based on spatial analysis (buffering and distance calculation), in place of the original time-consuming greedy algorithm; 2) Network analysis element removed
1-5. Distance Calculation	Calculate the length of previously generated GPS segments, and match them to corresponding links	Combination the previous make links and register back modules. Inclusion of add-ins for more optional data filtering and processing steps to ensure smooth running of the step
	2. Network Construction	
2-1. Prepare Transit Links	Convert stop-level information to trip segment level. Make transit links based on calculated distance and time.	1) Multiple networks avoided; 2) Similar links combined
2-2. Prepare Transfer Links	Based on the proximity of time and location, find stop pairs in different routes that can serve as transfer	Interagency transit routings
	3. Routing Analysis	
3-1. import Samples	Import multiple formatted sample files (O-D trips)	N/A
3-2. Samples Matching	Match the origin and destination in each sample to nearby walkable stops	Remove the process to first match to ABM nodes before matching to nodes
3-3. Shortest Path Finder	Build a network from links and run O-D pairs through the network to find the shortest path	One master network instead of running separate networks for each provider
	4. Interagency Modeling	
4-1. Download GTFS files	From a list of providers, automatically download GTFS files for the desired period to the specified location	Newly added features

Module	Description	Improvements in TransitSim 3.0
4-2. Generate Provider Information	Identify dates each version of GTFS is valid for each provider, and combine them to form a list of date ranges ("period") throughout which the same version of GTFS is used for each provider	
4-3. Add Sample Period	Based on the date and time samples are collected, assign them a date range ("period") and time range	
4-4. Combine Networks	Process the GTFS networks from multiple service providers into one master network, save one version for each date and time range	
5. Visualization and Extensions		
5-1. Park-and-Ride Information	Process the GTFS networks from multiple service providers into one master network, save one version for each date and time range	Remove the process to first match to ABM nodes before matching to nodes
5-2. Export Shapefiles	For each network, export the transit stops and routes as shapefiles for visualization and postprocessing purposes	Separated from the main sections as optional extensions
5-3. Export Links	Export shapefiles for network links	
5-4. Visualize Trajectories	Create the .png format travel trajectories for every single O-D trip samples	
5-5. Export link-by-link Trajectories	Export travel trajectories as link-by-link data frames	
5-6. Parallel Computing	Extension to use parallel computing to run the model	New feature

1. TransitSim Model Modifications for Enhanced Efficiency

This chapter details the strategies taken to modify TransitSim to improve efficiency. Seven major feature improvements are highlighted in the TransitSim module. Each improvement is discussed, as well as the advantages of each set of changes. To facilitate understanding of the difference between the original TransitSim and TransitSim, process flowcharts in Appendix A illustrate the modules presented in the first three for the two models respectively.

Modified Features

Various strategies were implemented to enhance TransitSim efficiency. For example, redundant or similar analyses were combined to reduce processing time and new algorithms were introduced. To avoid excess program run time in some complicated cases, sensitivity analyses are conducted to decide on cut-off points to switch the internal strategy. In many cases, a laddered system is introduced, where the higher rungs on the ladder having a high computational load and provide high accuracy, while the lower part of the ladder provides a lower computational load. Data are first processed higher up the ladder. When cut-off point is reached, the simpler approach is used to save processing time. In all of these applications, a back examination is incorporated to ensure that the efficiency enhancement does not take place at the price of reduced efficiency. The latter part of this section will reveal how these modifications enhance the accuracy at the same time of reducing complexity.

Sub-sampling

Analyses with a single service provider do not generally run into problems with computer memory; however, interagency analysis can sometimes be problematic due to the large dataset. This is especially true for big cities that rely heavily on transit. The strategy to solve this problem is to remove some of the repetitive trips to avoid the network being too large. First, trips are grouped by their features (e.g., route, time, etc.). Then, trips with the same features are grouped to be randomly sampled to a maximum size of 10. Because the link features are calculated based on the average travel time and distance between pairs of stops in such trips, the down-sampling only reduces the variance of the data, but with a relatively big sample size (10), the analysis results should not be significantly impacted. This strategy has been proven effective and reliable in test analysis with major cities in the United States, including New York and Los Angeles.

Sensitivity Analysis

For the time-consuming processing tools, particularly those that are involved in loops with varied sample sizes, a sensitivity analysis is conducted to assess how sample size impacts processing time. Based on the results of the sensitivity analysis, a cut-off point is identified for the program to adopt a different algorithm. This is intensively used in the spacing algorithm, to achieve a trade-off between program efficiency and results accuracy. As an example, two repetitively used tools are assessed for their efficiency:

1. Spacing Module: To prepare for the register modules, where GPS trajectories are matched to stops, an adequate density of GPS trajectory links is needed. Lower-resolution GPS data may cause a stop to be matched to an incorrect location and subsequently reduce accuracy, but higher-resolution GPS data increases processing time. Therefore, a tradeoff is needed. The time to add density (“space”) for different sizes of trajectory samples is compared in Table 2. To constrain the spacing time for each iteration to below 30 seconds, a cut-off is created to switch to a different algorithm when the sample size is larger than 200,000.
2. Point Distance Calculation: Point distance calculation is often used in matching transfer stops, matching stops with trajectory, matching the origin and destination (with stops), etc. While it is often the more direct solution to certain problems, point distance calculations can take a long time to run for a large sample size. In each of such use cases, TransitSim 3.0 provides a backup solution when the processing time is larger than

20 seconds. A sensitivity analysis is conducted to find out the corresponding sample size (100,000), as shown in Table 2.

Table 2. Example of Sensitivity Analysis for Spacing and Distance Calculation

Sample Size	Processing Time for Spacing	Processing Time for Distance Calculation
10,000	0.8s	1.6s
50,000	4.2s	7.7s
100,000	10.0s	15.4s
200,000	29.2s	31.4s
500,000	140.5s	77.2s

Laddered System of Approaches

A laddered system of approaches is incorporated in many modules. The idea is to use a more time-consuming and rigorous approach first. If the data become too large to process (decided by the cut-off points identified from the sensitivity analysis), the program moves on and tries a less rigorous but faster approach. For example, in the spacing module, instead of aiming for a fixed number of GPS points, TransitSim 3.0 decides the number of iterations based on individual trip conditions. This strategy improves running efficiency and at the same time reduces the chance of any inaccurate matches to GPS points that are too far away.

New Algorithm for Distance Calculation

The distance calculation section is simplified to allow faster processing. The prior TransitSim version uses a network-based greedy algorithm for the register module, which takes a long time to complete and puts a high demand on computing power. TransitSim 3.0 adopted a spatial-focused approach, reducing the computational burden for trips with fewer stop points.

New Algorithm for Network Construction

The main change to the network construction section is the adoption of interagency transit routing. The prior TransitSim version created a different network for each service provider. While this approach also provides users the possibility to adopt any available service providers in their travel, it does not allow transfer from one provider to another during a single trip. This brought a huge computing burden to the model. When there are multiple service providers, for example, the Atlanta Metro Area has five available provider datasets; hence, the model has to be run five times to create five networks. In larger cities such as New York (where there are ten transit service providers in Manhattan alone), this computational load will be too heavy to resolve. More importantly, it is non-realistic and less efficient for users to stay on one single provider for their entire trip, especially for those who travel from the city center to a suburb and often need to transfer from the metro to the local bus, or the cases of larger cities where the same provider are coded into multiple datasets for data accessibility. Another important feature TransitSim 3.0 is the combination of similar links. In the same date range and time range (for example, between 3:00 P.M. and 4:00 P.M. on weekdays in late August and September 2019), travel times between two adjacent stops are typically not different at the minute level. The links with the same feature can thus be combined into one link to reduce the computational load for handling multiple links in a network.

New Algorithm for Routing Analysis

The sample matching process in TransitSim 3.0 is greatly simplified. The previous version of TransitSim matched all samples to ABM nodes and created a connection between ABM nodes and transit stops. While this process makes it easier to link TransitSim to other models (e.g., SidewalkSim and RoadwaySim), it introduces a huge computational burden and makes the model incompatible for analyses in cities where a roadway network is not readily available. In comparison, the problem of incompatibility with other travel modes is resolved in TransitSim 3.0. The park-and-ride mode can be established with PNR information and an extension to RoadwaySim. The walking option can also be established through an extension to SidewalkSim. In fact, given the simple conditions surrounding the walk mode (no vehicle or bike required, hence continuity is not an issue), SidewalkSim networks can be integrated into TransitSim as another network provider.

Simplification of the File System

Because the processing time is significantly reduced, there is less of a need to save temporary files to prevent data loss from sudden program shutdowns. TransitSim 3.0 also reduces the complicated file structure to a more intuitive file system while allowing room to save multiple scenario outputs for interagency applications.

Advantages of the Modified Features

The modified features in TransitSim 3.0 have contributed two main categories of improvements to the model, the significantly reduced computation load and enhanced compatibility with different data and user needs. Because the computation load is greatly reduced, the model has higher running efficiency and lower demand for hardware features. Moreover, the updated TransitSim model structure simplifies the file system and lowers the learning bar for new users. While the program is simplified, reliability is not compromised. Instead, the model has been modified to be more compatible with different data and user needs. These improvements create the potential for an interagency model and multi-modal features incorporation. These advantages of the modified features have been proven empirically.

First of all, TransitSim 3.0 improves processing speed by 20-50 times, depending on the size of the sample inputs. As an example, when processing the trips between 3:00-4:00 P.M. in late August and September 2019, Section 1 (distance calculation) took 180 minutes to finish on a supercomputer with 128 RAM with the previous version of TransitSim, TransitSim 3.0 only used 45 minutes on a personal computer with 8GB RAM. Section 2 (network construction) used to take 90 minutes to finish on a supercomputer with 128GB RAM, TransitSim 3.0 only used 15 minutes on a personal computer with 8GB RAM. Section 3 (samples routing) used to take about 1-3 minutes for an O-D pair, but now takes less than 1 minute to finish all (544) sample runs on a personal computer with 8GB RAM. While the prior TransitSim version is highly sensitive to processing large samples, TransitSim 3.0 can finish the model run relatively quickly with most sample sizes.

Second, TransitSim 3.0 has reduced the demand for computer hardware. Depending on the complexity of the network, the prior TransitSim version had high hardware requirements. Most of the 240 total scenarios (from combinations of various time and date ranges) require a random-access memory (RAM) of 64GB, 18 of them require 128GB to run, and another 5 need 256 GB. These requirements are hard to meet even in high-performance supercomputer systems, let alone individual users on personal computers. TransitSim 3.0 has been tested for multiple of these scenarios, including one (between 3:00-4:00 P.M. in late August and September 2019) that cannot be successfully run on the Georgia Tech PACE supercomputing cluster with 128GB allocated RAM. Using TransitSim 3.0, the analysis for this scenario is completed in about 70 minutes on a personal laptop with 8GB RAM. The computation load in the previous version of TransitSim was not a significant burden when applied to smaller networks from one single provider (e.g., MARTA). However, when five providers were combined into a large master network in this analysis, the drawback of redundant programming and complicated algorithms becomes noticeable. The Atlanta metro area transit network is complicated. Major cities like New York and Los Angeles have networks of multiple magnitudes more complicated than that of Atlanta. TransitSim 3.0 is better for regional applications.

TransitSim 3.0 is now compatible with a wider range of trip characteristics and therefore yields more accurate and comprehensive results. In the same model run mentioned earlier, out of 359 unique trip trajectories for 3:00-4:00 P.M. in late August and September 2019, TransitSim 3.0 found 4 errors. All four of these errors were related to mistakes in the dataset, two were related to too few stop records, and two were related to the miscoding of the stop location (over 1 mile far away from its actual trajectory). In comparison, the previous TransitSim version found 26 errors, most of which resulted from memory errors (the dataset was too large for the computer to process). Note that the above-mentioned model run is conducted on 128GB allocated RAM on the Georgia Tech PACE supercomputing cluster. Although the memory errors are not actual systematic errors that are embedded in the model itself, the fact that they cannot be solved in such a system means that in reality, they will be actual errors for the vast majority, if not all, of users. Furthermore, the occurrence of memory error is negatively associated with the hardware conditions of the computer. This suggests that when running complicated scenarios, average users on personal computers might experience an elevated error rate, sometimes too high to be reliable.

Fourth, the laddered system in TransitSim 3.0 improves the model’s flexibility in handling different types of trips and user needs. The carefully designed notification system also makes the model more intuitive for average users. For example, in the spacing module, the previous version of TransitSim used a fixed rate of spacing decided from the initial number of GPS trajectory points supplied in the GTFS data. However, a high number of points does not necessarily indicate higher network density. Urban-suburban transit lines tend to be longer and thus have a higher number of network nodes. When stops are not properly spaced, the subsequent functions of the system can easily get into trouble with either excessive run time or memory error. TransitSim 3.0 improves on this feature by using stop-specific filtering criteria to find out the GPS point density around each stop. The laddered system also allows the program to run more rigorous functions until a certain processing time is reached. Then, a less rigorous approach is used to find an initial solution, which is back-propagated to find similar-quality results as the direct solution. This feature is applied in multiple modules throughout TransitSim 3.0 and has been demonstrated to yield improved results (screening fewer trips as potential errors) with increased efficiency.

Last, but not least, the simplified algorithm and enhanced compatibility have allowed TransitSim 3.0 to be applied in more use cases. Thanks to a combination of a simplified algorithm and a laddered approach, TransitSim 3.0 reduces trip errors noted in previous versions the model by providing more flexibility when deviations from expectation arise. The model is thus more flexible and adaptable to other use cases. In a separate modeling effort, TransitSim 3.0 was successfully applied in 19 major cities in the United States, including New York, Los Angeles, San Diego, and other cities that employ regional transit services with complex operating conditions (Fan et al., 2023b). The new algorithms have been tested in multiple use cases and recalibrated to be compatible with situations that arose in those analyses. While there may be issues that will arise that have not been identified, for a general use case, TransitSim 3.0 delivers significantly improved performance and supports interagency route modeling for multi-modal applications.

This section discusses the modified features in TransitSim 3.0, and presents an empirical verification of the advantage over the previous version of TransitSim. TransitSim 3.0 significantly reduces model run time, requires less computing power, creates more reliable outputs, has higher flexibility to represent various trip conditions, and is compatible with more use cases.

2. Interagency and Multi-Period Transit Routing Analysis

This chapter introduces the newly added interagency and multi-period transit routing analysis. The four modules in the interagency modeling section are introduced one by one, followed by a summary of corresponding changes in other sections. Then, the potential for future deployment and extensions to other applications are discussed.

Download GTFS Files

Module 4-1 is comprised of four functions, GTFS data download, file filtering and unzipping, file cleaning, and ID field recoding. For a city with multiple transit service providers, the GTFS data is downloaded separately for each provider into an individual folder location. Module 4-2 will introduce the strategy to combine multiple service providers into one network corresponding to each time period.

In the first function, GTFS static data are downloaded from open-source data sources (e.g., https://transitfeeds.com/) using manually-adjusted agency and time period codes. Each transit agency updates its transit routes and service schedules from time to time. Once an update is made, a new GTFS file is uploaded. The GTFS files are downloaded as zipped folders containing several text files in GTFS format. TransitSim 3.0 uses the stops, stop times, trips, and shapes files.

The second function scans through all downloaded zipped folders, and deletes those that are empty (this is determined by examining the size of the zipped folder). To allow automation of the download process, the user is assumed to not know the exact dates such updates happen. An iteration is designed to loop through all possible dates within a given date range and download the corresponding zipped folders. Dates with no valid data will return a download of empty zipped files. This function identifies those files and removes them. The remaining zipped folders are extracted and the original zipped folders are deleted.

The third and last function first scans through all downloaded files and spots any apparent errors, for example, miscoded column names, and cleans them up. Then, to prepare for interagency network combination, the ID fields of all files are re-coded to be unique to each service provider. Depending on the preference of the provider, some IDs are initially coded as numbers, some others as text. For consistency, all IDs are converted to text, and a text label indicating the service provider (e.g., "MARTA_") is concatenated at the beginning of each ID. Once all processing is complete, the files are saved to a folder created specifically for the provider and version of GTFS.

Generate Provider Information

Module 4-2 Identifies dates that each version of GTFS is valid for each provider and combines them to form a list of date ranges ("period") throughout which the same version of GTFS is used by each provider. That is, anytime a GTFS feed changes for one agency, a new period is established for all agencies. The module includes three parts, creating a list of all versions of GTFS reliable for each service provider, identifying unique periods corresponding to the various versions, and finalizing the start and end dates as filtering criteria for sample processing.

The first function scans through all GTFS versions available in the local folder (downloaded from the previous module) for each provider. For each available version, the date information is extracted from the folder name and the calendar dates file. A dictionary is created to temporarily store the GTFS version information, with keys indicating service providers and values being a list of all possible versions indicated by start dates. For example, the MARTA 190317 version indicates the GTFS version that starts in effect on March 17, 2019.

In the second function, the dictionary of GTFS versions is first converted to a data frame. All possible combinations of GTFS versions (e.g., MARTA version 1 and GRTA version 2; MARTA version 2 and GRTA version 3; etc.) are generated using the product function in the itertools library. Unreasonable combinations are filtered out, based on the basic assumption that a new date range of the data is created when one of the providers updates its routes and all others are using their latest routes. Therefore, if the latest version change of a certain provider is combined with an older version of another provider, the combination is invalid for the specified date range. For example, MARTA version 191116 and Gwinnett version 190423 are a new combination compared to MARTA 190927 and Gwinnett 190423, but MARTA 191116 and

Gwinnett 190114 were not, because the Gwinnett routes had not been at the time MARTA changed its routes (November 16). In Python, the function is realized in four steps (an example shown in Table 3):

1. A table is created with each column being a service provider, and each row being a possible combination of GTFS data versions.
2. For convenience of analysis, all start dates are converted to numbers in “yymmdd” format. For example, the MARTA GTFS version updated on November 16, 2019, is coded as 191116.
3. The “updating” transit provider(s) is identified as the one with the latest (largest numerically) start date of its GTFS version. Every “updating” transit time point should create one and only one new date range.
4. Due to the nature of the product function, more than one date range will inevitably be generated for each “updating” transit time point. With only one being the actual update, all the others are reasonable as explained above. The filtering is done by summing up all dates in the row.
5. For each “updating” date, only the row with the largest sum of dates is kept.

An example of the final list of transit version combinations is shown later in the report when data processing is discussed in more detail (see Table 10 columns 4 to 8).

Table 3. Example Filtering Steps

MARTA	GRTA	Atlanta Streetcar	Gwinnett	Cobb	Updating Date	Sum of Dates	Keep
191116	190131	170630	190423	170923	191116	913223	True
191116	190131	170630	190114	170923	191116	912914	False

Once a unique list of transit service versions is identified, the start date of each identified period is the “updating date” identified in the previous step. The end date of each period is one day before the start date of the following period, which can be easily found using a shift function in Python. Table 10 shows an example of the creation of the 12 periods in the case study.

Add Sample Information

Based on the date and time samples are collected, modules 4-3 assign them a date range (“period”) corresponding to the GTFS periods. This is done by first converting the date information into a numeric format using the Python *date_to_integer* function. Once the date becomes an integer, the identification of periods is an easy task to compare the values of the date in samples and that of the GTFS information table. Any dates that fall between the start date and end date of a period (both sides included) are marked as a sample taken from that period.

The boarding times of the samples are re-coded into the time period format that can be matched to the time records in the GTFS stop times files. These samples are saved into a "sample in" file for each date and time range to be used in analyses in other sections. Table 12 is an example of sample processing, with each cell showing the sample size in the corresponding scenario.

Combine Networks

The network combination module recompiles the GTFS information and saves it for analysis in other sections. The original GTFS file is saved based on versions of data and providers. The interagency analysis needs a master network containing all sub-networks from service providers. Therefore, all ID re-coded GTFS files are first combined and then recompiled into several networks each with an individual date and time range.

Before combining the networks, some data processing steps are conducted. First of all, unrealistic coordinates are filtered out from the stops and shape files as data errors. Second, files with N/A values in major columns (ID columns and coordinates) are removed as invalid data. Finally, the boarding time in the stop times files is processed into time range bins directly corresponding to the time range in the samples. Boarding times are classified into their corresponding one-hour bin that is used in the sample files. Any times between 23:00 and 23:59 P.M. are classified as "11:00 P.M. and after" and times between 0:00 and 5:00 A.M. are classified as "Before 5:00 A.M."

The recompiled network files are saved to a predefined file structure. Treated as raw data files, the files are saved in the GTFS directory under the Data folder. Twelve folders are created to correspond to the 12 date ranges and under each folder. Each GTFS file is saved to its corresponding folder. When in use, other sections will query the period and time to pull data from their folder.

Modules 4-1 to 4-4 lay the foundation for interagency multi-period transit routing analysis, by collecting and processing necessary input data and recompiling and saving them in predetermined formats. This allows routing analysis using a master network of sub-networks from all agencies and supports more detailed analysis of different time and date periods with their specific input files.

Connection to Other Sections and Modules

While modules 4-1 to 4-4 prepare the data and set up the overall framework for interagency multi-period analysis, the conduct of interagency routing is handled by features in the other sections. These features include adjustment and recalibration for enhanced compatibility, construction of one master network, and synchronized sample processing procedure.

As mentioned earlier, the modified features include a sensitivity analysis and a laddered system of approaches that have greatly enhanced the compatibility of TransitSim 3.0 with various data types. These modified features also set the backbone for interagency routing modeling. Transit trips from different providers have different routes and schedules, and thus different characteristics for data processing. Taking Atlanta as an example, the GRTA/SRTA express bus network tends to extend for longer distances with fewer stops, while the MARTA network often has more stops and more detailed trip trajectories that could sometimes bring new issues in the data processing. Therefore, the new features in Section 1 need to ensure the capability to handle large datasets, span across a large area of service, and at the same time be able to capture enough details. TransitSim 3.0 has been calibrated to ensure the smooth processing of information from multiple providers at the same time. In terms of data quality, while smaller agencies sometimes are behind in updating their data and thus some samples cannot be properly matched to schedules, larger agencies have a higher chance of leaving errors in their data due to frequent updates. Therefore, the program is modified to cope with both types of errors. These modified features are important for the smooth conduct of interagency routing modeling.

The updates in Section 2, from running multiple separate networks to combining sub-networks into a large master network are also important to the concept of interagency transportation. While the sub-networks from multiple transit agencies are prepared and combined in Modules 4-1 to 4-4, the change in network construction strategies is incorporated in new features in Module 2-1. Features such as sub-sampling and taking the average of trip segments are incorporated to make the model adaptable to handle one large network instead of multiple smaller ones.

In Section 3, the prior TransitSim version takes different networks as different travel modes (e.g., MARTA walk-to-transit, GRTA walk-to-transit, MARTA park-and-ride, GRTA park-and-ride, and driving) and runs each sample through all possible travel modes and collect the outputs with the lowest travel time. This approach has a couple of drawbacks. First, it is very time-consuming. Intuitively, running the same data through five networks will take five times of processing time compared to one single network. Furthermore, the comparison function and post-analysis constrains the analysis to the sample-to-sample level. That is, there is no option to run multiple samples through the same network. Instead, each trip is processed through the network one-by-one. Given the fact that the processing time of network routing is not linear with sample size (processing 1,000 samples takes much less time than processing 10 samples 1,000 times), this approach further increases the computational load. Last and perhaps most importantly, this approach hinders the users' ability to transfer from one transit agency to another during a trip. This is unrealistic, especially in cities with large transit providers who tend to separate one single transit system into multiple GTFS feeds. The use of an interagency network in Section 3 not only enhances computational load but also allows the users more choices to mimic realistic use cases.

While the successful conduct of interagency modeling heavily relies upon the new features in the program, the interagency modeling approach itself further facilitates a more efficient model run. The interagency and trip-level integration features of TransitSim 3.0 represents a change in modeling philosophy, simplifying individual model runs and allowing potentials to combine with other use cases and modes.

3. Extension and Visualizations

In TransitSim 3.0, some visualization and file export modules are recompiled into optional extensions. The purpose of the change is to reduce the computation burden, simplify the file structure, and allow more user customizations. In addition, some features are added as extensions to further automate the program and enhance efficiency, including the park-and-ride information matching and parallel computing modules.

Park-and-Ride Information

In the previous TransitSim version, park-and-ride information needed to be manually entered in the middle of the program. This can be a burden for some users who are not familiar with searching for park-and-ride information or are not experienced in dealing with the translation from address to coordinates. TransitSim 3.0 automates this park-and-ride information matching process through the park-and-ride module (module 5-1) which contains three functions: geocoding, distance calculation, and park-and-ride information updates.

In the first function, park-and-ride information (as addresses) are geocoded into latitudes and longitudes. While there are several mature geocoding applications (e.g., Google geocoding API), many of them are not open source. To accommodate the needs of future users who may not have the necessary license to use fee-charging geocoding, a free Python library (Nominatim) is used as the geocoding engine. Compared to licensed geocoding engines, the current application has inevitable drawbacks in reduced reliability of processing. As of May 2022, manually-verified park-and-ride location data are provided as inputs to TransitSim 3.0, with address information extracted from MARTA (https://martaguide.com/parking/). To fully-automate the park-and-ride information module, future research could develop a web scraping algorithm to automatically collect updated park-and-ride information, and integrate improved geocoding approaches.

The second function decides if a park-and-ride station is accessible for transit passengers using certain transit stops. While most park-and-ride parking lots indicate their dedicated use station in their names, many transit stops are clustered together as a group of stations (e.g., Lindbergh MARTA train station and bus station). A distance calculation function is used to calculate the distance between previously identified parking lots and the transit stations. Any transit station that falls within a 1.0-mile distance of a parking lot is identified as park-and-ride viable. Figure 1 shows an example of all park-and-ride parking lot locations of the transit stations in Atlanta (May 2019).

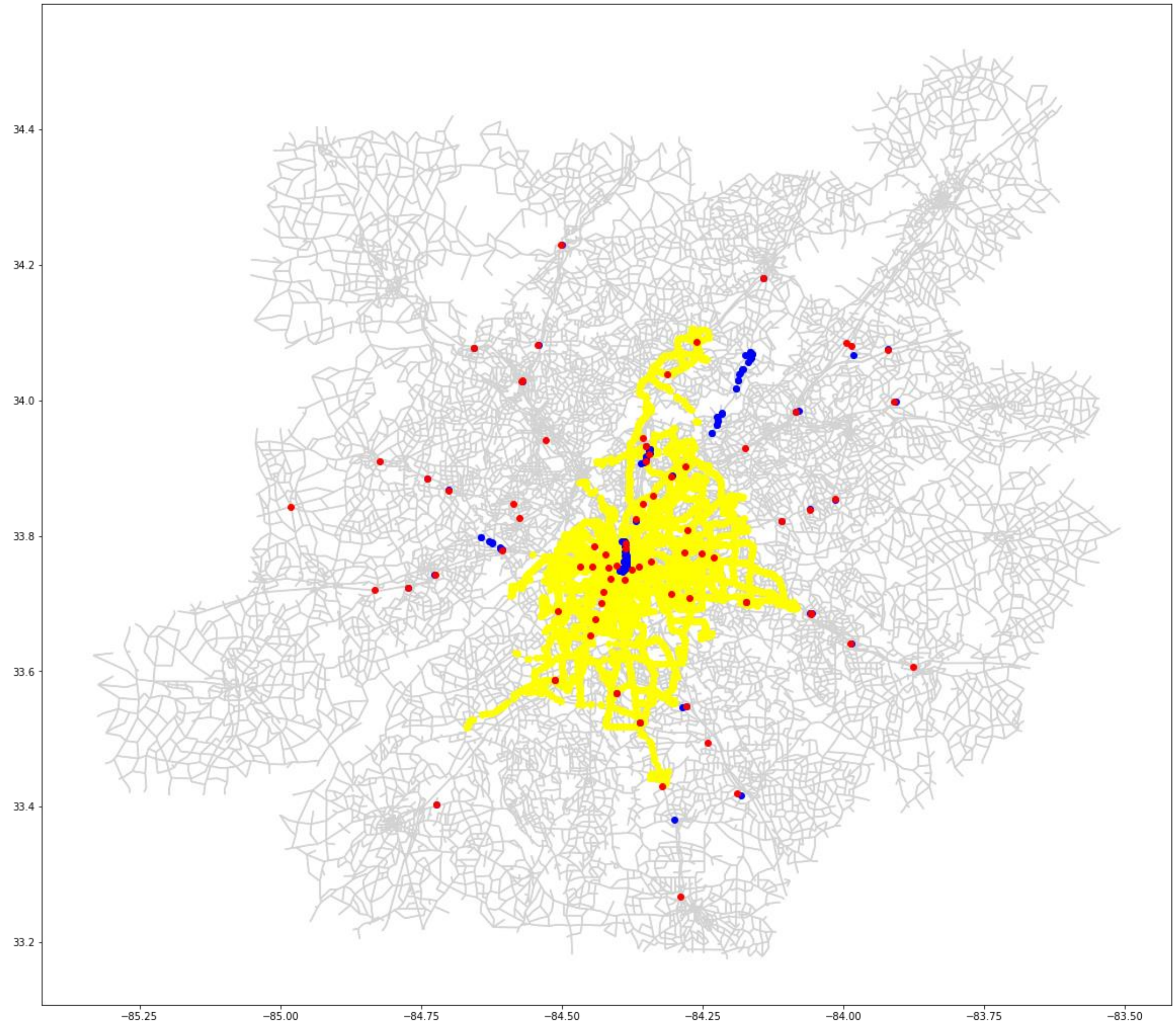


Figure 1. MARTA Stations (yellow), GRTA Stations (blue), and Park-and-Ride Parking Lots (red), as of May 2019

The last function in the park-and-ride information module combines the park-and-ride information to the stops file in the GTFS data and saves the updated file. The park-and-ride module will be included in TransitSim 3.0 as an optional extension.

Visualization and Shapefiles

In the previous TransitSim version, several files and visualizations are created to facilitate user understanding and post-processing of the data. While some of them may be important for the user's needs, the inclusion of all such modules elevates the complexity of file structure and reduces model efficiency. In TransitSim 3.0, these modules are trimmed off from the main running sections (Sections 1-3) and included as optional extensions based on user-specified parameters.

Module 5-2 exports stops and routes shapefiles. The first module of both the old TransitSim and TransitSim 3.0 creates transit routes and stops from the GTFS format network data. In some use cases, it might be important to have these networks visualized, and thus saving these files as an interim output may be helpful for users. However, running this module as mandatory may create an unnecessary burden for users who do not need such information, and for those who are running multiple scenarios. Moreover, one ESRI shapefile is saved as multiple files and thus a separate folder is needed to contain each of the outputs. In TransitSim 3.0, this module is provided as an optional extension.

Module 5-3 saves the transit links (stop-by-stop trip segments) as shapefiles. Similar to Module 5-2, these files are removed from the main running Sections and provided as an extension. One other change to this module is the option to save the links trip by trip. The previous version of TransitSim not only saved all links in a single shapefile, but also saved links for each trip in a separate file. This application was seldom used by users (typically only for visualization routines) therefore the default is now set to not run in the Module 5-3 extension. However, considering some special use cases that might still need this information, the user can still turn on this option in the setup parameters.

Module 5-4 visualizes the link-by-link travel trajectories for individual transit users. This module is already an optional extension in the old TransitSim file, and thus not much change is made. In TransitSim 3.0, the option to produce link-by-link travel trajectories (Module 5-5) has also been made optional. Therefore, a prerequisite for this module to be run is that Module 5-5 has been turned on. This will be described in more detail later in this document.

Modules 5-2 to 5-4 are visualizations and file export modules that were originally embedded in the previous version of TransitSim and now turned into an optional extension. The significance of such a change lies not only in efficiency enhancement but also in reducing the complexity of file structure and the size of the output files. For example, the case study analysis that involved 240 scenarios generated main Sections alone that reached 6.4GB. This is not a convenient size to transfer, especially in cloud computing. Turning off the unnecessary visualization modules helps reduce the file size and ease the process of model run and file transfer.

Link-by-Link Travel Trajectories

One of the biggest advantages of TransitSim 2.0 was its ability to produce link-by-link travel trajectories, which is easily transformed into a second-by-second location of the traveler through their entire trip (walking and driving in a vehicle). Second-by-second trip trace data (from origin-to-destination) is very useful in a variety of analyses for energy use, emissions, pollutant impact assessment, mobility, accessibility, level of service, equity, etc. The output of Section 3 is path information for individual O-D trip samples. This information provides link-by-link travel, but requires further processing for use in second-by-second analyses.

Taking one of the case study examples, a survey respondent walks to and boards on transit at MARTA station number 073637, and alights on MARTA station 073821 before walking to the destination. TransitSim 3.0 Section 3 outputs record this trip by its total travel distance, total travel time, and the path of "MARTA_073637__MARTA 073700__MARTA_073821". Although the basic route information is included in this file structure, the file requires further post-processing to generate a second-by-second trace. In TransitSim 2.0, all link-by-link travel trajectories of the same trip were output in a tabular format (see Table 4).

Table 4. Output Trajectory of an Example Trip

Stop A	Stop B	Distance	Time	Mode
Origin	MARTA _073637	0.06	1.70	walk
MARTA _073637	MARTA 073700	0.24	0.38	Transit
MARTA _073700	MARTA_073821	0.15	1.35	Transit
MARTA _073821	Destination	0.07	1.98	Walk

While saving the entire trajectory as a table is easier to process and analyze in subsequent analyses, this file structure demands a significantly higher storage space and read-write processing time. Therefore, Module 5-5 was made an optional extension in TransitSim 3.0 that the user can select to post-process the output from Section 3 to full tabular format.

Parallel Computing

Module 5-6 is used to create a set up for parallel computing (parallel computing functionality was added to the previous version of TransitSim in earlier NCST model improvement efforts). As discussed in Part 1, 1. TransitSim Model Modifications for Enhanced Efficiency, the most time-consuming section of the prior TransitSim version is routing analysis for the samples. Therefore, the focus of the previous parallel computing module was on sub-dividing input O-D trip samples into multiple files to run synchronized on multiple computers. Considering the nature of TransitSim 3.0, where individual model runs typically takes less than 2 hours to complete, but users tend to include more scenarios, each with larger network data. In Module 5-6, the focus of parallel computing thus shifts to handling multiple scenarios. A dictionary object is added to the beginning of the main code bulk, to allow users to enter a list of periods and times that they desire to process in each model run. This supports running the model in a parallel computing set up, which is convenient feature for processing multiple model runs.

4. Model Improvement Summary

This chapter discusses the new features added to TransitSim 3.0:

- Main Run Sections 1-3 were rewritten to enhance computation efficiency while enhancing model compatibility with different types of data and user needs;
- An Interagency Multi-Period Routing Section was added to process and prepare network data from multiple transit service providers; and
- A couple of visualization and file export modules were made optional extensions, to further reduce model run time in the main running sections while allowing users the flexibility to output the files they need for post-processing.

TransitSim 3.0 provides several advantages over the previous version of TransitSim. First and furthermost, TransitSim 3.0 significantly reduces computation load, model complexity, and risk of memory errors. Second, TransitSim 3.0 is more compatible with a variety of trip characteristics and user needs. Third, TransitSim 3.0 can now be fully automated with minimal manual processing time for parameter set up. In the previous version of TransitSim, the sections were no directly connected and the file structure was more complicated. Therefore, the previous version required users to connect the sections by manually processing the outputs from one section and preparing it for use in the next section. This burden is eliminated in TransitSim 3.0 with the simplified file system. Finally, the flexibility and compatibility of TransitSim 3.0 allows for a wider variety of future variations, in terms of forms of application, combination with other travel modes, and use in other cities.

The prior TransitSim version had the advantage of a complete set-up framework for multi-modal analysis and user-customizable inputs such as desired arrival time, day, etc. However, TransitSim 3.0 is much more efficient in large-scale data analysis for research and analytics, the prior TransitSim version is better for multi-modal routing software deployment. A few improvements that were made between TransitSim 1.0 and 2.0 to support specific analysis types (such as an automated snap from TAZ origin and destination to a network node) were lost in TransitSim 3.0. However, considering the amount of time that TransitSim 3.0 saves in performing model runs was worth the tradeoff. The team is now exploring the integration of other “networks” (e.g., roads, bike paths, and sidewalks) that conform to the same structure (i.e., integration of SidewalkSim as a connected module to further support transit access analysis in the creation of second-by-second trajectories).

Case Study Data and Methods

1. Study Area and Data

Atlanta Metropolitan Area

The Atlanta Metropolitan Area ("Metro Atlanta") is the major urban cluster in Georgia. By 2020, metro Atlanta will be home to over six million population with over eight thousand square miles of land area. It encompasses nine counties, including Cherokee, Clayton, Cobb, DeKalb, Douglas, Fayette, Forsyth, Fulton, Gwinnett, Henry, and Rockdale (ARC, 2021). As home to 13 fortune 500 corporates, the city of Atlanta is the 10th largest economy, the 36th most populated city, and among the fastest-growing in the United States.

Despite a focus on freight transportation and a known reliance on automobile transportation (Crimmins and Preston, 1980; Henderson, 2002), metro Atlanta has established a mature transit system with various options for riders, including fixed-routes transit (rail, bus, express bus), circulators and shuttles, and on-demand services and paratransit (ARC, 2022). The public transportation in metro Atlanta is operated by different transit agencies (service providers) at the state, metropolitan, and county levels. This case study focuses on the fixed-route transit options in metro Atlanta, including:

- **Metropolitan Atlanta Rapid Transit Authority (MARTA)** is available to 1.7 million residents in metro Atlanta. MARTA has a wide coverage of services including heavy rail, buses, streetcar, and MARTA mobility paratransit (MARTA, 2021) (Table 5).
- **Xpress**, sometimes referred to as GRTA Xpress or SRTA Xpress is a regional express bus transportation system operated by GRTA (Georgia Regional Transportation Authority) – SRTA (State Road and Tollway Authority) – ATL (Atlanta Region Transit-Link Authority). Xpress is a state-level express bus transit option focused on metro Atlanta and its counties. It serves over 1.8 million riders each year. Different from MARTA's focus on short-distance to medium-distance trips, Xpress typically runs a longer distance with fewer stops. Most, if not all, Xpress trips use the highway (GRTA, 2022).
- **Gwinnett County Transit (GCT)** is operated by Gwinnett County to provide public transportation options for Gwinnett County residents. It includes three commuter express routes (co-operated with Xpress) and six local routes (Gwinnett County, 2022).
- **Cobb Link** is operated by Cobb County, including three commuter express routes (co-operated with Xpress) and twelve local routes (CobbLinc, 2022).
- **Hall Area Transit Services** is operated by Gainesville County. It operates a downtown Trolley to provide transit access to important locations around the county. The Hall Area Transit also operates the WeGo ridesharing service (Gainesville County, 2022).
- **The Cherokee Area Transportation System (CATS)** is operated by Cherokee County. It offers commuter express routes (co-operated with Xpress), local Canton fixed routes, and complementary on-demand paratransit (CATS, 2022).

Table 5. MARTA Service and Frequency (source: MARTA)

Summary of MARTA Service Tier Characteristics

Service Tier	Typical Frequency		Typical Distance Between Stops
	Peak Periods	Off-Peak Periods	
MARTA Mobility Paratransit	on-demand for eligible riders		origin-to-destination service
Community Circulator Bus	30 - 60 minutes all day		800-1200 feet (2-5 min. walk)
Supporting Local Bus	60 minutes	60 minutes	800-1200 feet (2-5 min. walk)
Frequent Local Bus	every 10 - 15 minutes	every 15 - 20 minutes	800-1200 feet (2-5 min. walk)
Peak Only Bus	60 minutes	– no service –	express service from park-ride lots
Limited Express	every 30 - 40 minutes	every 40 - 60 minutes	express service from park-ride lots
Streetcar	every 15 min. all day		900 feet (4 min. walk)
Heavy Rail	every 10 minutes	every 20 minutes	1 mile (20 min. walk)

Transit Onboard Survey

As an effort to update the travel demand model in metro Atlanta, the Atlanta Regional Commission (ARC) conducted a region on-board survey in the Spring and Fall of 2019. The purpose is to understand the origin-destination-based transit trip patterns on fixed-route transit in metro Atlanta, including MARTA, Xpress, Gwinnett County, Cobb Link, Hall Area Transit Services, and CATS. With a survey period of March to December 2019, this survey represents the latest demographic information obtained for transit riders before the pandemic. The survey collected a total of 43,398 survey responses, representing 10% of the total transit ridership based on 2018 estimates (ARC, 2020). The survey team took measures to ensure that at least 10% of the trips were collected for every transit agency (Table 6).

Table 6. Number of Surveys and Total Boarding Counts (source: ARC, 2020)

Agency Name	Number of OD Surveys	Average Daily Boardings	Sampling Percent
Cherokee	20	63	31.2%
Cobb Link	941	9,753	9.6%
Xpress	472	3,951	11.9%
Gwinnett	617	5,844	10.6%
CATS	94	552	17.0%
MARTA Bus	19,844	160,557	12.4%
MARTA Rail	21,410	136,891	15.6%

The survey collects a variety of information about the survey respondents, including the following categories:

- Transit services:
    - Last-mile transportation mode (from origin to the first transit station and from last transit station to destination)
    - Boarding and alighting stop
    - Previous trips information
    - Boarding and alighting time
    - Fare method
    - Whether the breeze card is used
    - Alternative mode of transit
    - Number of total transfers
- Demographics:
    - The transit riders' address
    - Household size
    - Number of vehicles in household
    - Number of 15+ y/o household members employed
    - Respondent's employment status
    - Respondent's student status
    - Whether the respondent has a driver's license
    - Age of respondent
    - Whether the respondent is of Hispanic, Latino, or Spanish origin
    - Race of the respondent
    - The language used in household
    - Respondent's English language ability
    - Annual household income
    - Respondent's gender
- Travel patterns of the riders:
    - Frequency of transit use
- Trip purposes
    - Origin and destination activity type
- Service coverage and quality
    - Expansion factors
    - Weights

GTFS data

The General Transit Feeds Specification (GTFS) data is a standard data format that was first developed by Google for app development purposes. Transit service providers upload their transit routes, station/stop data, schedules, fares, and other information in the standard GTFS format for developers. These data are open-access and can be downloaded from websites like Open Mobility Data (https://transitfeeds.com/).

There are two types of GTFS data, the static GTFS (or the “schedule component”) and the dynamic (or the “real-time component”). The static GTFS includes information about the routes, stops, schedule, fares, etc. The dynamic GTFS provides information on traffic conditions and predictions on transit arrival times and can be used with real-time tracking. This study uses static GTFS data. Most of the transit agencies in the study area have available static GTFS, except for Cherokee County and CATS (Table 7).

Table 7. Availability of Static GTFS Files for Transit Agencies in the Study Area

Agency Name	GTFS Available	GTFS Provider Name
Cherokee	**No**	
Cobb Link	**Yes**	Cobb
Xpress	**Yes**	GRTA
Gwinnett	**Yes**	Gwinnett
CATS	**No**	
MARTA Bus	**Yes**	MARTA
MARTA Rail	**Yes**	MARTA
MARTA Streetcar	**Yes**	Atlanta Streetcar

Static GTFS files typically contain the following text files:

- **Agency**: information about the transit agency
- **Calendar dates**: the effective date the GTFS file
- **Fare attributes**: fare amount
- **Routes**: name, ID, and other information about routes
- **Trips**: a beginning-to-end run of the route at a specific schedule is a trip, a route typically has multiple trips in a day
- **Stops**: stations information, containing the name, ID, and geographic locations
- **Stop times**: schedule of the transit trip in the format of arrival and departure time at each stop
- **Shapes**: GPS points comprising the trajectory of a trip

2. Data Pre-Processing

Data Quality

A few data filtering techniques were employed to remove data with missing values or mismatched information. The analyses consist of two main parts, the equity analysis, and the network routing analysis. As is shown previously in Table 6, ARC collected a total of 47,730 on-board trip surveys for the study period. For the equity analysis, samples with missing demographic information (1,355 samples), and some specific demographic categories (4,334 samples), are removed. This leaves 37,711 samples (86.9% of total trips) for the route analysis.

Transit network routing analysis employs the filtered samples; however, some additional samples are lost in this process, based on the network data. In TransitSim 3.0, when a trip does not meet certain minimum criteria for the program to run, the entire trip is removed. These criteria include:

1. The trip must have a valid schedule (time and date) information.
2. A trip must have at least two stops associated with it to form a link.
3. The minimum distance between any GPS trajectory points recorded in the GTFS file to a transit stop must be less than a user-defined threshold (the default is one-half mile). Because this study uses the GPS trajectory to calculate the distance between stops, if the GPS points are too far away from the stop, the calculated distance will be biased.
4. When the GPS trajectory is matched to the stops, the sequence by which the GPS trajectory is taken needs to match the order of stops in the trip. For example, if the first stop is matched to the 50th GPS point and the second stop is matched to the 30th, this adjacent stop pair (stop 1 – stop 2) is flagged as an error. The maximum number of continuous error-free stop pairs in a trip must be smaller than a user-defined threshold, the default being three, for a trip to be retained in the analyses.
5. A single model iteration should not take more than 30 seconds.

TransitSim 2.0 also used criteria 3 and criteria 4 above to identify erroneous trips. To handle trips that were not identified by these two criteria, a network was constructed every time a trip was processed. Hence, the previous version of TransitSim looped the data through a single approach, relaxing criteria until conditions were met. However, this method sometimes leads to system computations piling up and an exponential growth of run time. However, relaxing the criteria would also sometimes result in problematic runs, due to memory errors which disrupted some modeling processes. The latest version of the model now uses a laddered approach to reach closure more quickly. The case study samples are run through both TransitSim model versions (see Table 8). TransitSim 3.0 identified 1,624 (4.3%) stops as erroneous and removed them from the analysis. The final dataset contained 36,087 stops, 83.2% of total survey samples in the beginning. In comparison, the previous version of TransitSim removed 3,327 (8.8%) stops and left a final sample size of only 34,384 (79.2%).

Table 8. Data Information and Sample Sizes

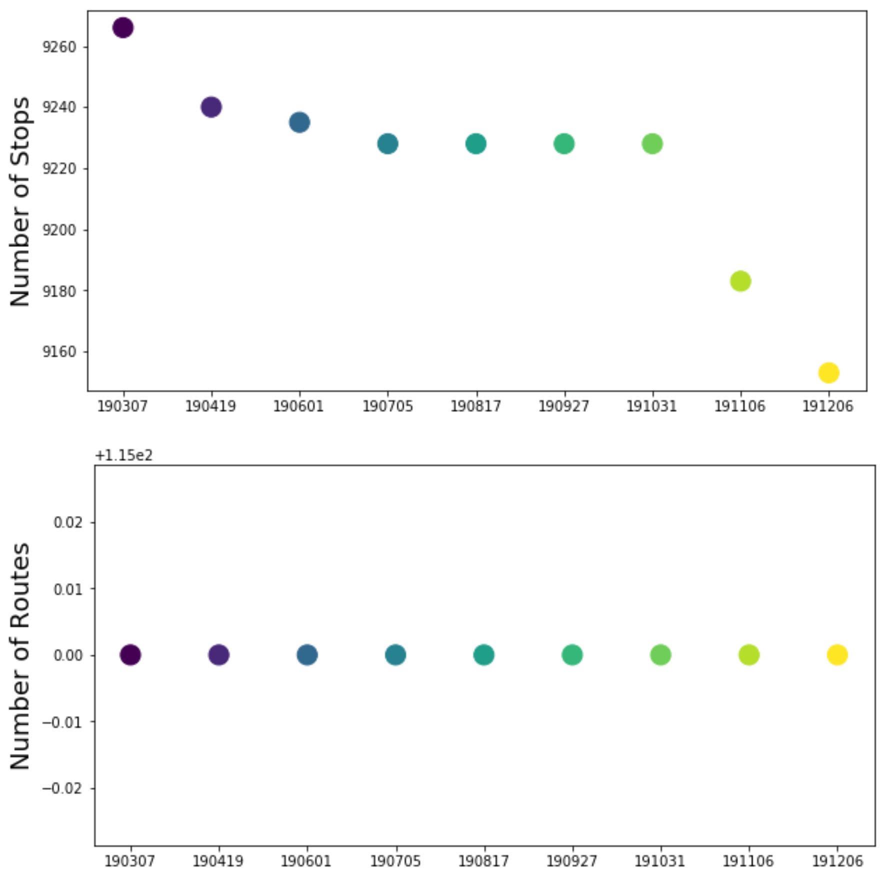
	TransitSim 2.0 (Previous Version)	TransitSim 3.0 (New Version)
Transit Onboard Survey Sample Size	43,398 (100.0%)	43,398 (100.0%)
Trips with missing demographic information	1,355	1,355
Trips with missing income information	4,334	4,334
Transit Onboard Survey with demographics	37,711 (86.9%)	37,711 (86.9%)
Stops not matched to the network	9,373	547
Screened Transit Onboard Survey Sample Size	28,338 (65.3%)	37,164 (85.6%)

Exploratory Data Analysis on GTFS changes over time

During the study period, some of the transit agencies updated their network (as reflected in GTFS data changes over time). As shown in Table 10, MARTA had eight updates and Gwinnett County had three updates to their network. Before conducting the analyses with TransitSim 3.0, it is important to first understand how these updates have changed the routes and trip schedules. When possible, periods can be combined to reduce the computation load. To understand this question, an exploratory data analysis is conducted. The analysis includes three parts, a general assessment of GTFS change in terms of transit composition, a detailed investigation of transit station relocations, and a detailed investigation of transit rerouting. The following paragraphs discuss these three aspects. Because the MARTA network experienced the most changes, the analyses use the MARTA network as an example.

General Assessment of GTFS Change in Terms of Transit Composition

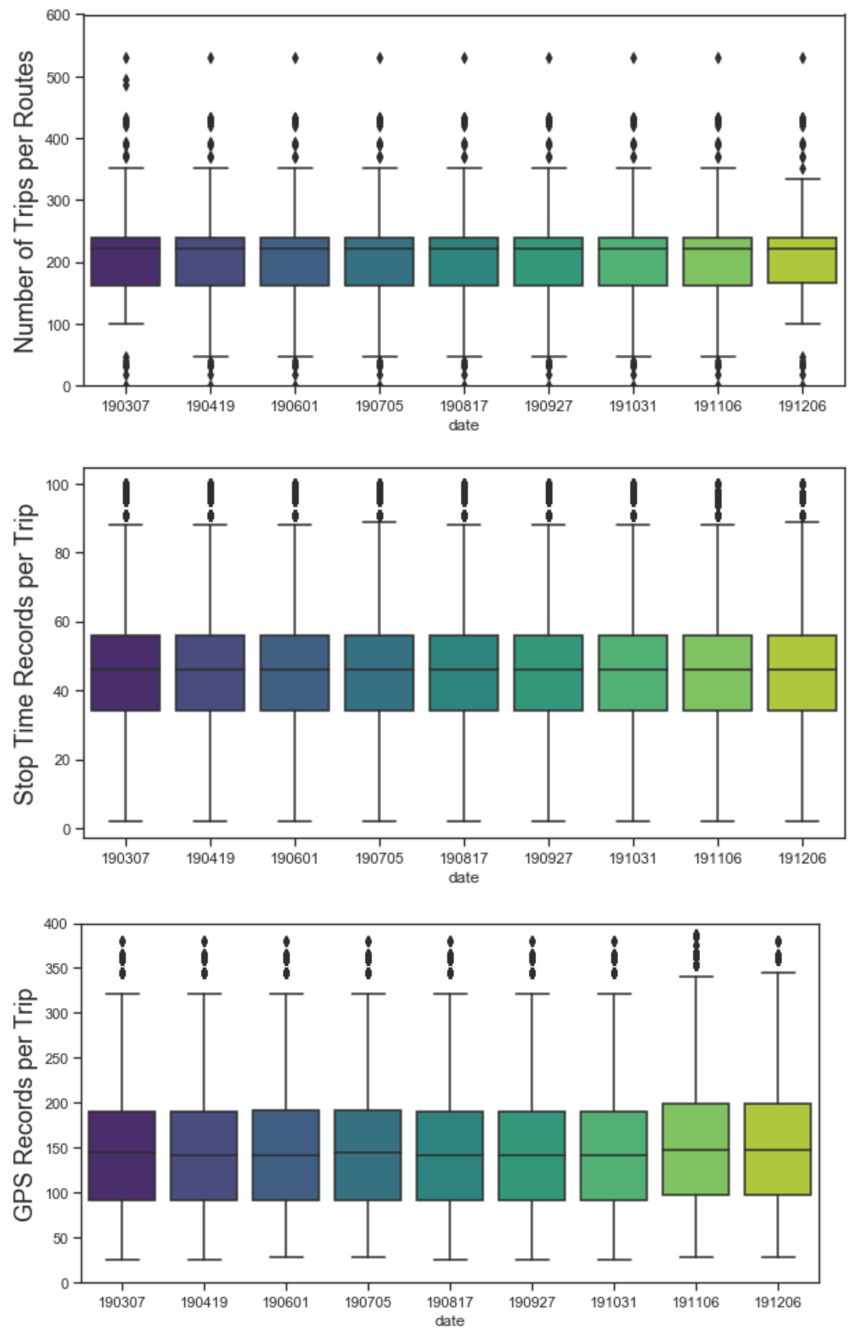
The general assessment of GTFS change looks at the compositional change in the network, such as changes in the number of routes and stops, and changes in the number of stops or trips per route. In the nine versions of networks involved, as shown in Figure 2, MARTA maintained the same number of routes throughout the study period. However, the number of stops decreased continuously over time. The GTFS analysis in this part is performed using Python.



Note: The scale on the y-axis has been reduced to highlight variations; x-axis date is coded as yymmdd

Figure 2. Changes in Number of Stops (top) and Routes (bottom) over the Study Period

In terms of the number of trips per route, over time MARTA has a similar distribution of the number of trips per route among its operating routes. However, there is a slight tendency of outlying samples (the routes with very high numbers of trips) to decrease in the number of trips. The number of scheduling records ("stop times") has also demonstrated a similar pattern. In terms of GPS points trajectory, MARTA seems to pick up more details in the GPS points recording based on an observed increase in the number of shape points per trip (Figure 3).



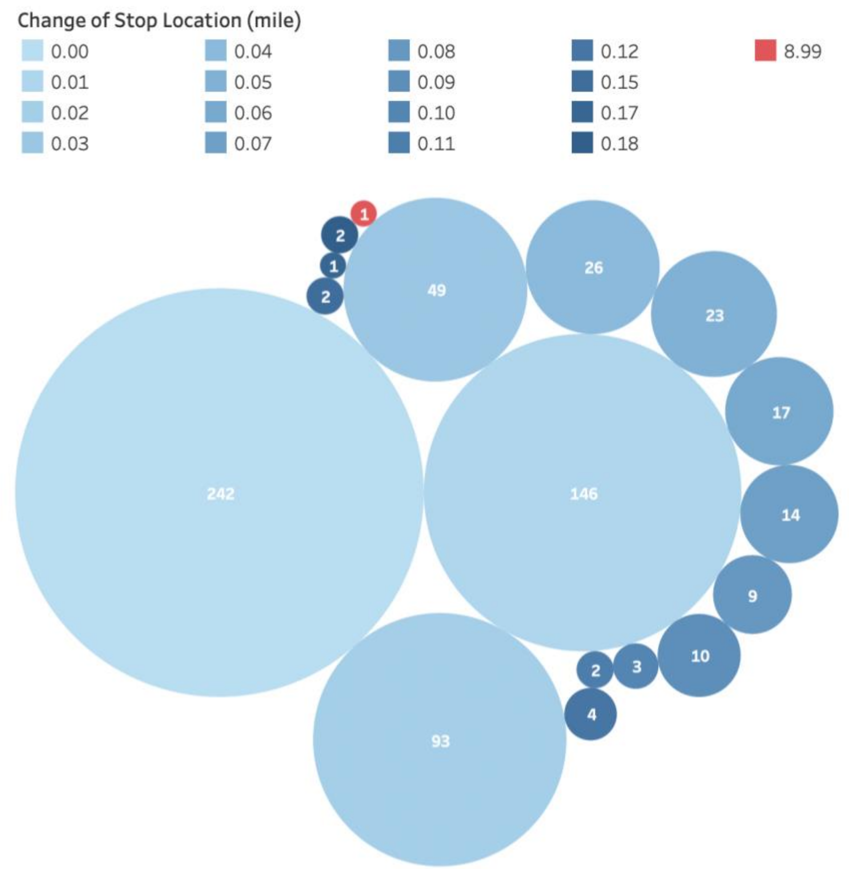
Note: The scale on the y-axis has been reduced to highlight variations; x-axis date is coded as yymmdd

Figure 3. Changes in Routes Arrangements and Scheduling over the Study Period

Detailed Investigation of Transit Station Relocations

The general investigation shows a general change in the reduced total number of stops from March to December of 2019. The next question is whether or not have the remaining stops changed. This is assessed through the distance by which later stops are separated from previous stops. The assessment was performed using R and visualized with Tableau.

Figure 4 shows the distance that stops have relocated over the study period (out of the total of 9,100 to 9,300 stops). Each bubble represents a distance category (binned at 0.01 mile). The darker the color, the further away stops have been relocated. The size of the bubble represents the number of stops that are in this category. As is shown in the bubble plot, the majority of stops were relocated by less than 0.05 miles. However, six stops were relocated by more than 0.15 miles, and one moved 8.99 miles (indicative of a likely stop coding error).



Note: Colors indicate distance relocated (from light to dark) and bubble size indicates number of stations

Figure 4. Number of Stops that have Changed Location by Distance Changed

Figure 5 shows a histogram plot of the relocation distance of all the stops except for the outlier (8.99 miles). Taking the average lane width of 16 feet and assuming that a person will tolerate up to three lanes' width distance of difference, any relocation beyond 0.1 miles might be an issue. Out of 9,100 to 9,300 stops, the relocation of 402 stops was too large to ignore.

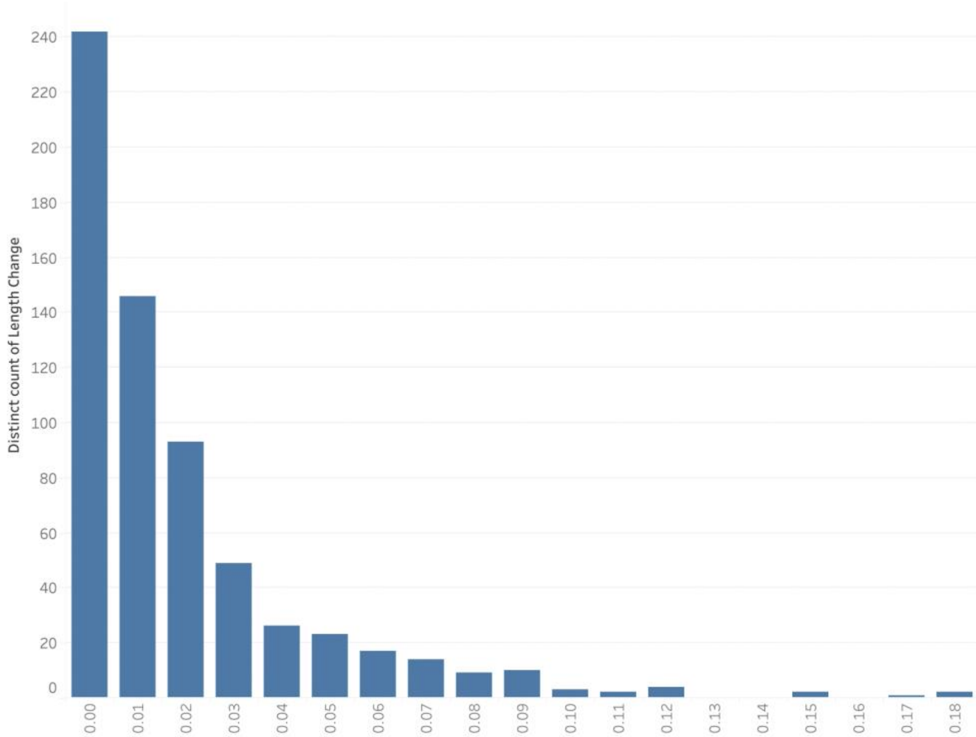


Figure 5. Distance Traveled (miles) Distribution (Excluding Extreme Values)

Detailed Investigation of Transit Rerouting

The final investigation was an assessment of transit routes that have been rerouted over the study period. Every transit route has a fixed sequence of stops (ordering). With a change in transit route, removal or addition of a stop, or change in sequence of a stop, the sequence of stops changes. Therefore, the investigation on rerouting is conducted based on the change in the stop sequence corresponding to each route. The stop sequence data in the stop_times.txt file is first aggregated by stops ID and route ID. By comparing the stop sequence for the same stop on the same route over different versions of GTFS, stops that have been changed sequence are identified. After that, the stops are aggregated again at the route level to find out, for each route, if any rerouting has occurred and how many stops have been relocated. This part of the analysis is conducted in R.

MARTA reported that 83 out of the 114 total routes had been rerouted (at least one stop had been changed). For the 114 MARTA routes, an average of 31.9% of the stops on each route had changed in some way. Figure 5 is a histogram plot of the percentage of stops that have a changed sequence in a route, for all 114 routes in the study period. Note that this analysis does not yet include the change in travel time. Some stops may simply be relocated, remaining in the same sequence on the route. However, even if all stops in a route have remained in the same sequence from one version of GTFS to another, that does not mean that the travel time in each trip segment will remain the same. For example, road construction over an extended period may mean that a trip segment now takes longer to traverse due to a new diversion.

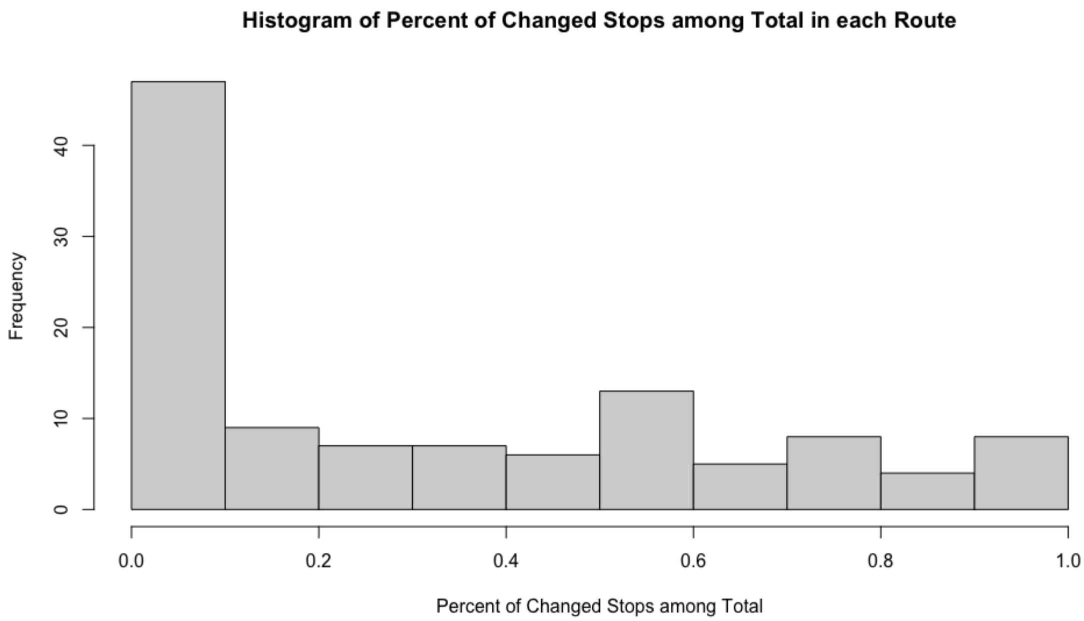


Figure 6. Percent of Changed Stops in Each Route

Summary of GTFS Changes over the Study Period

The network structure of MARTA, as represented in the GTFS data, changed eight times over the study period. The changes can occur in composition, distance, configuration, and travel time. This exploratory data analysis explored the changes to the composition, distance, and configuration of the network elements. Although the GTFS change at the compositional aspect is not significant, as represented by a small portion or none of stops and routes number change, the change at the distance and configuration are not negligible. A total of 402 stations were relocated by at least three lanes' width at some point during the study period, 83 out of 114 total routes were rerouted for at least one stop, and an average of 31.9% of stations in each route have been re-ordered. This analysis did not include an investigation of travel time change, but changes in the other three aspects are significant enough to address.

The fact that the difference across GTFS versions is too large to be ignored implies that the routing model has to be run for multiple scenarios. Because the Transit On-board survey includes 20 study times of day, each with different travel conditions (for example, travel conditions in 8:00-9:00 A.M. are very different from that in 10:00-11:00 A.M.), the number of scenarios that need to be run can easily exceed 200 (in fact, this study requires 240 model runs as will be discussed later). This analysis required the use of the routing model in TransitSim 3.0, so that scenarios could be processed much faster.

3. Approach to Transit Rider Equity Assessment

With the demographic information provided in the ARC Transit Onboard Survey, the equity analysis is conducted. In the first part of the analysis, some mutually exclusive household groups are identified based on their vehicle ownership, income, household size, income, and employment status. In the second part, personal demographics are identified (age, gender, race, English language ability) along with frequency of transit use and trip purpose for travel behavior assessment.

Demographic Groups

The demographic group categorization was based on a mutually exclusive categorization method developed by Zhao (2020) for an ARPA-E project. That original categorization created 16 mutually exclusive groups. However, because the Transit Onboard Survey 2019 does not include information about the number of children in the family, four of the groups could not be disentangled. Therefore, those four groups are lumped into two groups based on other demographic features. Hence, in this study, 14 demographic groups were identified (Table 9).

Table 9. Demographic Groups Assignment (14 Groups)

ID	Description	Own Vehicle	Low Income	HH Size	Annual Net Income	# of Workers	# of Vehicles	# of Responses	Percent
1	Non-Vehicle Owners	No	N/A	Any	Any	Any	0	16,672	38.4%
2	Low-Income Households with Vehicles	Yes	Yes	Any	< $25k	Any	1+	2,474	5.7%
3	Lower-Middle-Income Single-Unemployed-Person Households	Yes	No	1	$25 - $60k	0	1+	70	0.2%
4	Lower-Middle-Income Single-Employed-Person Households	Yes	No	1	$25 - $60k	1	1+	952	2.2%
5	Middle-Income Single-Person Households	Yes	No	1	$60k+	0 or 1	1+	553	1.3%
6	Lower-Middle-Income Multi-Unemployed-Person Households	Yes	No	2+	$25 - $60k	0	1+	138	0.3%
7	Lower-Middle-Income Two-Person Households with Workers	Yes	No	2	$25 - $60k	1+	1+	2,561	5.9%
8	Lower-Middle-Income Large Households with Workers	Yes	No	3+	$25 - $60k	1+	1+	7,782	17.9%
9	Middle-Income Multi-Person Households with One or Fewer Workers	Yes	No	2+	$60k - $120k	0 or 1	1+	691	1.6%
10	Middle-Income Dual-Worker Households	Yes	No	2	$60k - $120k	2+	1+	1,277	2.9%
11	Middle-Income Three-Person Households with Workers	Yes	No	3	$60k - $120k	2+	1+	1,140	2.6%
12	Middle-Income Large Households with Multiple Workers	Yes	No	4+	$60k - $120k	2+	1+	2,453	5.7%
13	High-Income Multi-Person Households with One or Fewer Workers	Yes	No	2+	$120k+	0 or 1	1+	151	0.3%
14	High-Income Multi-Person Households with Multiple Workers	Yes	No	2+	$120k+	2+	1+	797	1.8%
	Missing Data							5,687	13.1%

Among the 43,398 survey responses received, 42,045 (98.9%) have complete demographic information collected. However, among these responses, 7,354 respondents refused to put in their income, and 4,334 of the 7,354 own a vehicle. Because income is a mandatory field for group categorization across households that own a vehicle, there are only 37,711 (86.9%) survey responses that are categorized into groups. Table 9 shows the number of survey responses and their percent for each of the groups.

Other Socio-Demographic Characteristics

- Gender
- Age
- Race
- English
- Frequency
- Trip Purpose

4. TransitSim 3.0

As is shown in Exploratory Data Analysis on GTFS changes over time in Data Pre-Processing, the difference across networks over the entire study period is significant and cannot be ignored. Therefore, in this analysis, the different periods are analyzed separately (i.e., for the different transit scenarios generated by the Generate Provider Information module in Section 4). For Interagency Analysis using TransitSim 3.0, the analytical periods are shown in Table 10.

Table 10. Period Identification Process in the Case Study

Period	Start Date	End Date	Provider				
			MARTA	GRTA	Atlanta Streetcar	Gwinnett	Cobb
0	Mar. 07	April 18	190317	190131	170630	190114	170923
1	Apr. 19	April 22	190419	190131	170630	190114	170923
2	Apr. 23	May 31	190419	190131	170630	190423	170923
3	Jun. 1	July 4	190601	190131	170630	190423	170923
4	Jul. 5	Aug. 16	190705	190131	170630	190423	170923
5	Aug. 17	Sep. 26	190817	190131	170630	190423	170923
6	Sep. 27	Oct. 30	190927	190131	170630	190423	170923
7	Oct. 31	Nov. 5	191031	190131	170630	190423	170923
8	Nov. 6	Nov. 20	191116	190131	170630	190423	170923
9	Nov. 21	Dec. 5	191116	190131	170630	191121	170923
10	Dec. 6	Dec. 30	191206	190131	170630	191121	170923
11	Dec. 31	Dec. 31	191206	190131	170630	191231	170923

Studies have shown that people in different demographic groups tend to adopt different travel schedules (Lu and Pas, 1999). Consequently, they tend to experience different travel conditions and their travel time will also be different. Using one single network for all times of the day could introduce endogeneity from different times of the day. To avoid such bias, while still being able to take advantage of the simplified network construction procedure provided in TransitSim 3.0's new algorithm for network construction, this study lumps the links temporally within every one-hour time range. To avoid excessive computation load for low travel volume periods (11 PM to 5 AM the next day), and to match the time range format in the transit on-board survey, this study uses twenty time ranges as shown in Table 11.

Table 11. Time Range Definition in the Case Study

	Start Time	End Time
12:00 AM - 5:00 AM	00:00	04:49
5:00 - 6:00 AM	05:00	05:59
6:00 - 7:00 AM	06:00	06:59
7:00 - 8:00 AM	07:00	07:59
8:00 - 9:00 AM	08:00	08:59
9:00 - 10:00 AM	09:00	09:59
10:00 - 11:00 AM	10:00	10:59
11:00 AM - 12:00 PM	11:00	11:59
12:00 P.M. - 1:00 PM	12:00	12:59
1:00 - 2:00 PM	13:00	13:59
2:00 - 3:00 PM	14:00	14:59
3:00 - 4:00 PM	15:00	15:59
4:00 - 5:00 PM	16:00	16:59
5:00 - 6:00 PM	17:00	17:59
6:00 - 7:00 PM	18:00	18:59
7:00 - 8:00 PM	19:00	19:59
8:00 - 9:00 PM	20:00	20:59
9:00 - 10:00 PM	21:00	21:59
10:00 - 11:00 PM	22:00	22:59
11:00 PM to 12:00 AM	23:00	23:59

The different periods (date ranges) and time ranges combine into the 240 running scenarios. Table 12 summarizes the number of Transit Onboard Survey samples that belong to each scenario. The sample sizes that are filtered for GTFS network routing mismatches are summarized in Table 13. These 240 run scenarios were compiled into a run package and sent to the Georgia Tech PACE Supercomputing Cluster. One machine with 128 GB was used, and 10 nodes were run each time for an average of 24 iterations. This model run took about 20 hours to complete.

Table 12. Sample Sizes in Each Period and Time for the Case Study (demographics NA values, total: 42,045)

	0	1	2	3	4	5	6	7	8	9	10	11
12:00 AM - 5:00 AM	20	4	46	1	18	55	25	6	17	7	1	0
5:00 - 6:00 AM	124	9	212	0	60	422	266	40	117	183	69	0
6:00 - 7:00 AM	225	23	550	0	38	616	349	48	149	119	125	0
7:00 - 8:00 AM	333	42	688	0	65	705	421	51	258	169	152	0
8:00 - 9:00 AM	203	34	698	0	80	673	435	74	262	196	170	0
9:00 - 10:00 AM	302	26	665	0	50	580	517	84	273	228	172	0
10:00 - 11:00 AM	284	24	659	0	56	584	464	66	156	66	119	0
11:00 AM - 12:00 PM	284	25	589	1	54	539	386	70	127	67	89	0
12:00 P.M. - 1:00 PM	225	23	553	0	44	404	382	79	121	35	54	0
1:00 - 2:00 PM	204	18	527	0	24	362	422	68	85	16	44	0
2:00 - 3:00 PM	336	20	585	0	59	544	552	76	180	63	64	0
3:00 - 4:00 PM	379	28	755	0	76	900	596	77	334	177	80	0
4:00 - 5:00 PM	404	33	823	0	62	872	555	89	245	201	80	0
5:00 - 6:00 PM	335	31	777	0	61	819	542	86	268	189	76	0
6:00 - 7:00 PM	305	38	700	0	42	711	475	67	204	171	62	0
7:00 - 8:00 PM	296	31	640	0	44	576	391	60	191	87	48	0
8:00 - 9:00 PM	195	26	554	0	34	490	319	51	168	78	41	0
9:00 - 10:00 PM	88	17	334	0	27	277	207	38	123	54	41	0
10:00 - 11:00 PM	23	2	67	0	9	79	88	10	8	6	7	0
11:00 PM to 12:00 AM	0	0	21	0	1	12	18	5	6	1	3	0

Table 13. Number of Samples Filtered by GTFS network errors

	0	1	2	3	4	5	6	7	8	9	10	11
12:00 AM - 5:00 AM	0	1	0	0	0	0	0	0	0	0	0	0
5:00 - 6:00 AM	0	0	2	0	0	2	3	0	1	0	0	0
6:00 - 7:00 AM	0	0	0	0	8	14	0	0	0	0	0	0
7:00 - 8:00 AM	0	0	0	0	12	8	0	0	0	0	0	0
8:00 - 9:00 AM	3	0	0	0	14	7	0	0	0	0	0	0
9:00 - 10:00 AM	2	0	0	0	13	10	0	0	0	0	0	0
10:00 - 11:00 AM	0	0	0	0	17	14	0	0	2	0	0	0
11:00 AM - 12:00 PM	0	0	0	0	11	5	0	0	1	0	0	0
12:00 P.M. - 1:00 PM	0	0	0	0	10	7	3	0	1	0	0	0
1:00 - 2:00 PM	0	0	0	0	10	18	2	0	1	0	0	0
2:00 - 3:00 PM	1	0	0	0	9	13	1	0	0	0	0	0
3:00 - 4:00 PM	1	0	0	0	13	63	2	5	4	0	0	0
4:00 - 5:00 PM	1	0	0	0	13	39	2	1	4	0	0	0
5:00 - 6:00 PM	1	0	1	0	10	26	6	0	14	0	0	0
6:00 - 7:00 PM	0	0	0	0	0	8	1	0	0	0	0	0
7:00 - 8:00 PM	0	0	0	0	0	2	0	0	0	0	0	0
8:00 - 9:00 PM	0	0	1	0	0	0	1	0	0	0	0	0
9:00 - 10:00 PM	0	0	0	0	0	1	0	0	0	0	0	0
10:00 - 11:00 PM	0	0	0	0	0	0	0	0	0	0	0	0
11:00 PM to 12:00 AM	0	0	0	0	0	0	0	0	0	0	0	0

The data analysis was conducted in R and the analytical results were visualized in Tableau.

Transit Activity Results and Discussion

This section reports the findings on three aspects: the difference in travel time, the difference in trip purpose, and the difference in last-mile travel mode. The demographics of the transit users are analyzed. Then, the general trends of transit use in the Atlanta Metropolitan Area are discussed, followed by a demonstration of the results of the equity analysis, which is revealed by comparing the above-mentioned metrics across demographic groups, and different age, race, gender, and English language capability.

1. General Trends in Transit Use

On average, a transit user in Atlanta spends 19.6 minutes per a single trip, with a standard deviation of 12.1 minutes. The travel distance for transit users is an average of 7.3 miles, with a standard deviation of 13.5 miles. According to the American Public Transportation Association's 2021 Public Transportation Fact Book (APTA, 2021), the operating speed of rail in the United States averages 19.9 miles per hour (mph) in 2021, and that for buses is 12 mph. Based on our analysis, the average travel speed of transit trips in Atlanta is 22.4 mph. Considering that it includes regional express bus transportation (GRTA, etc.), the number is reasonable.

In metro Atlanta, 40% of transit riders take transit to travel between 0 and 5 miles. Another around 30% take transit between 5-10 miles, and 20% between 10-15 miles. Most people spend between 5 to 25 minutes on their transit trips, accounting for over 60% of all survey takers (Figure 7). Atlanta streetcar is commonly used for short-distance trips. All survey takers who used Atlanta streetcar report a travel distance lower than 5 miles, and a time between 5 to 10 minutes. The MARTA buses, which also serve mostly the City of Atlanta, are also used primarily as short-distance travel options, with some use cases for medium-distance. About 59% of reported MARTA bus trips travel a distance below 5 miles. Around 70% of all MARTA bus trips took a time between 5 to 25 minutes. The MARTA rails ("Atlanta Heavy Rail") are used mainly for short- and medium-distance trips, though also used in long-distance trips in some cases. Trip distances of 0-5, 5-10, and 10-15 miles predominate (each take up about 25-30% of total MARTA rail travel), with the remaining 20% of trips being 15-30 miles. Most people (more than 70%) travel on MARTA rail for 5 to 30 minutes. About 20% of all MARTA rail trips take a total of 30 to 60 minutes. SRTA serves short-, medium-, and long-distance travelers. Despite over 40% of SRTA riders traveling a distance of less than 5 miles, about 35% of trips are between 5 to 20 miles, and 20% are 20 to 35 miles. The travel time of SRTA trips varies widely, depending on distance, but all SRTA trips are than 25 minutes. Gwinnett County Transit serves primarily medium-distance trips (10-20 miles, about 60%), with some cases of long-distance trips (30-40 miles, about 7%). The majority of trips taken on Gwinnett County Transit are 30 to 55 minutes. CobbLinc serves primarily short-distance trips (0-10 miles, about 65%), although medium-distance trips also take up a notable share (10-25 miles, about 30%). Most people traveling with CobbLinc spend a time between 5 to 55 minutes on a single trip.

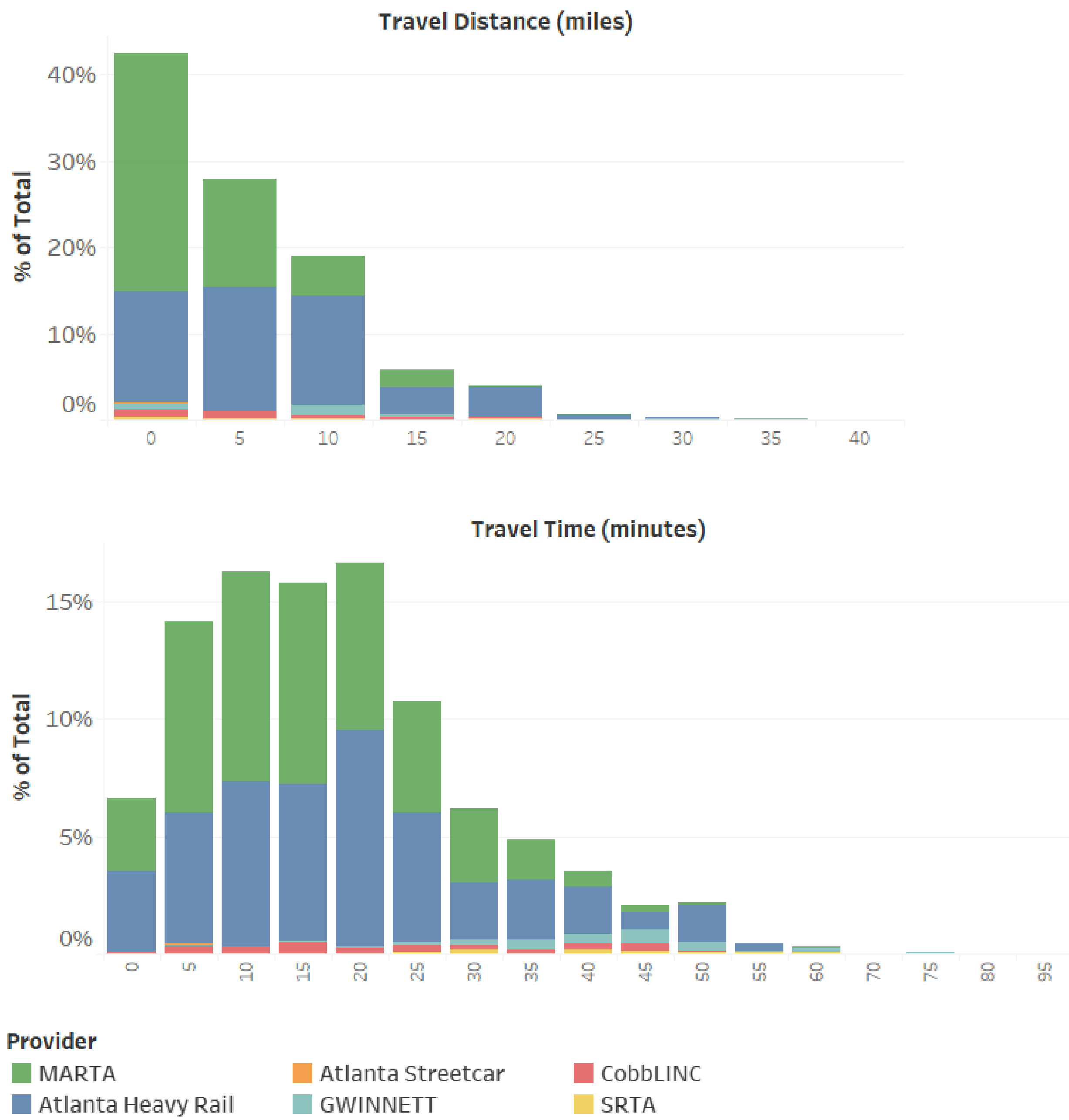


Figure 7. Distribution of distance (top) and travel time (bottom) by providers

As reflected in this survey, most people who use transit in metro Atlanta are frequent transit users (using transit 5 or more days a week), accounting for 68.6% of all survey respondents. The second most common transit user type uses transit 2-4 days a week. Together, over 97% of transit users surveyed use transit at least 2 days a week (Figure 8). Interestingly, Figure 9 and Figure 10 indicate that people who take transit weekly or more (at least once a week) tend toward shorter travel times (19.6 to 21.8 minutes) and distances (7.8 to 8.5 miles) per trip than those that use transit less frequently (26.3 to 30.5 minutes and 11.1 to 13.7 miles).

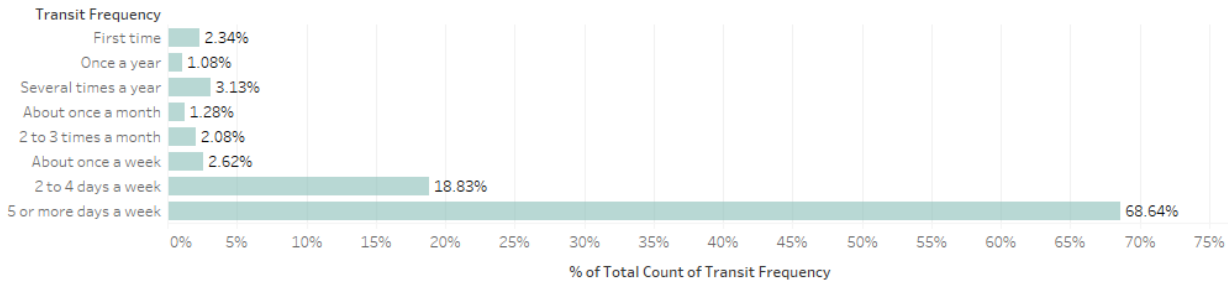


Figure 8. Frequency of Using Transit

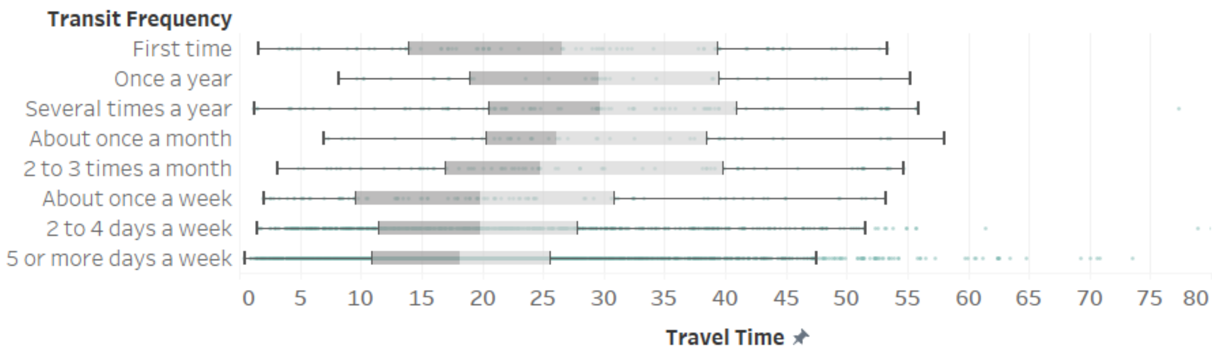


Figure 9. Travel Time Distribution of Transit Riders with Different Frequency of Transit Use

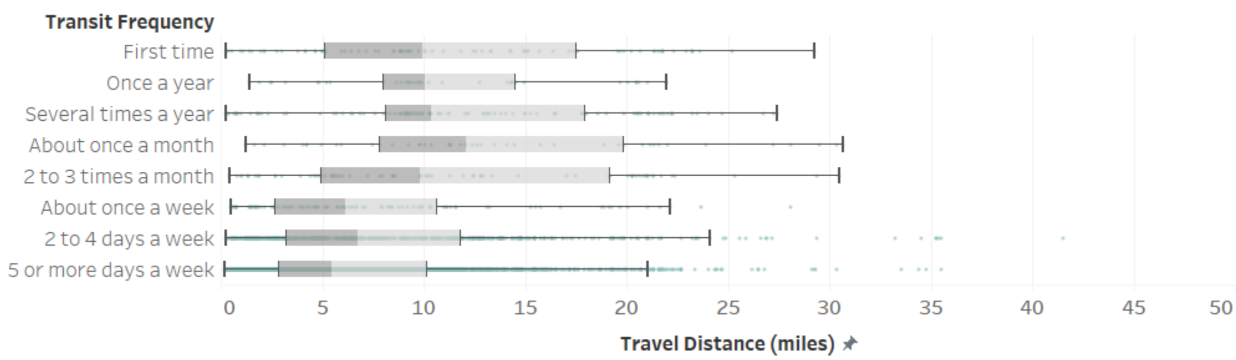


Figure 10. Travel Distance Distribution across Transit Riders by Transit Use Frequency

The trip purpose of the transit trips is represented by destination type. This study investigates six main types of destinations, home, work, college and school, recreational and shopping, hospital, and personal business. Frequent transit users generally use transit most frequently for their home-based trips (accounting for 39.0% to 47.2% of transit trips for users of different frequencies), as well as a higher share of their transit trips for work trips (37.2% for those who use transit more 5 or more days a week, 25.1% for 2 to 4 days a week, less than 10% for those using transit on monthly basis, and less than 5% for those using transit on yearly basis). Less frequent transit users, on the other hand, tend to use transit more frequently for airport trips (over 20% for transit users who use transit monthly, and more than 25% for transit users who use transit yearly), see Figure 11.

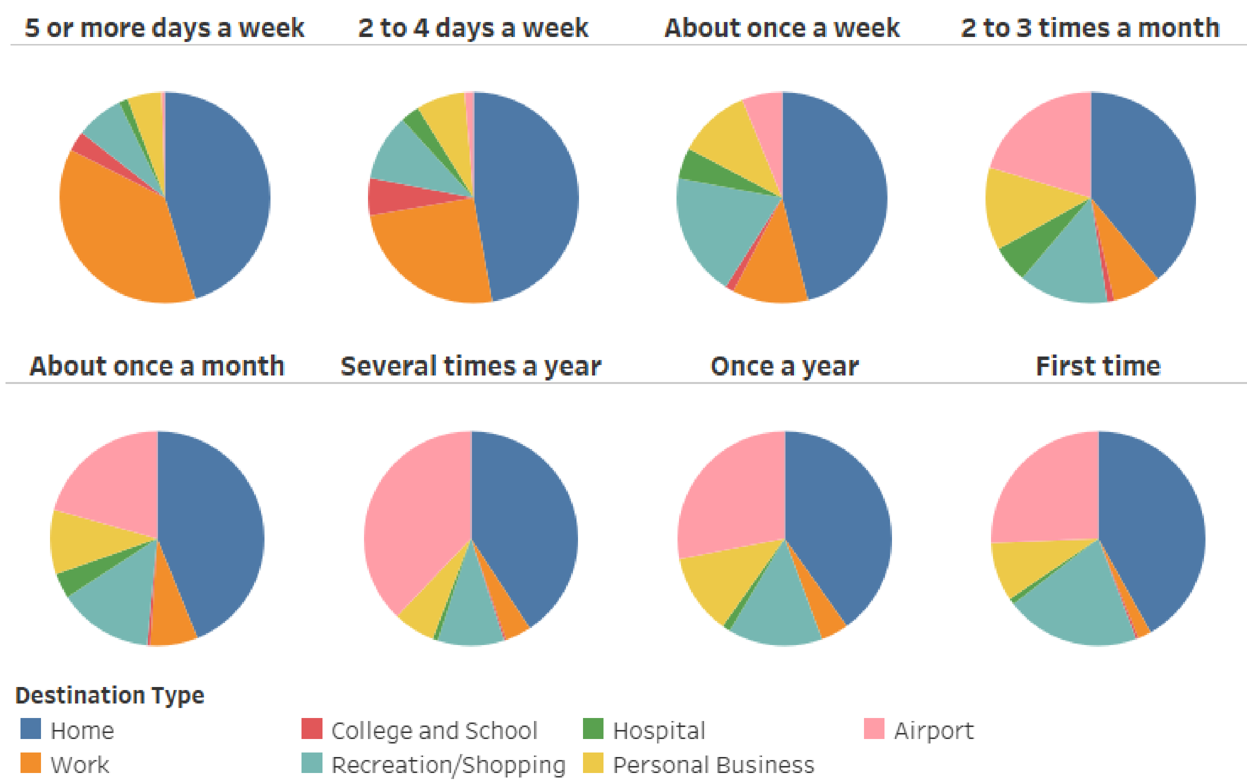


Figure 11. Trip Purpose (as Destination Type) by Transit Use Frequency

An investigation of the travel time by different trip purposes (destination type) might reveal the reason for the differed travel time by transit use frequency. As is shown in Figure 12 and Figure 13, airport trips have significantly higher travel time (a median of 39.6 minutes versus 20.3 minutes for the highest of other categories) and significantly higher travel distance (a median of 15.5 miles versus 6.4 miles for the highest of other categories).

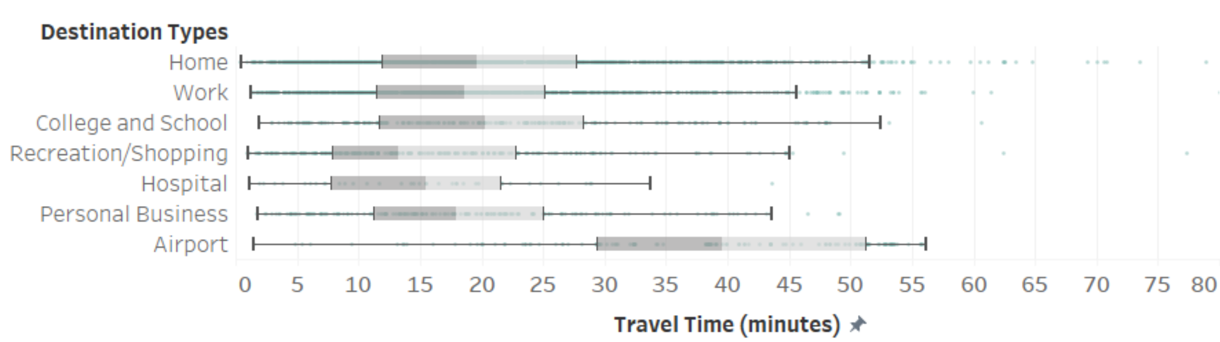


Figure 12. Travel Time Distribution by Trip Purpose

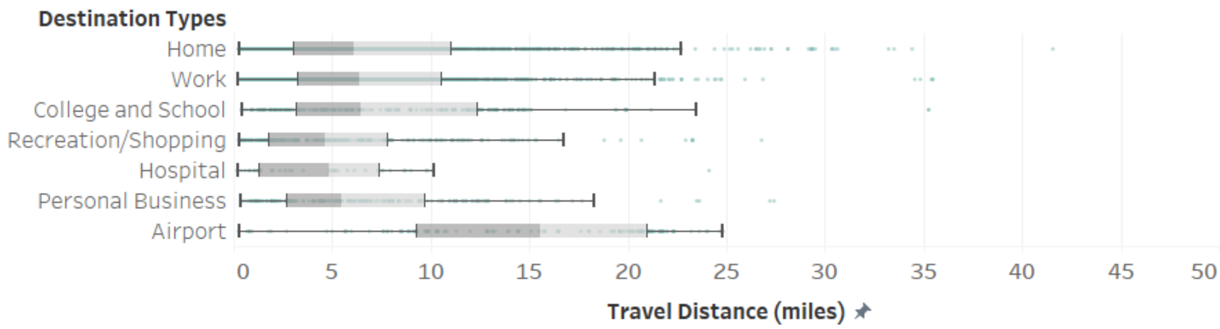


Figure 13. Travel Distance Distribution by Trip Purpose

2. Transit Riders across Demographic Groups

Within the same income and car ownership category, smaller households and households with fewer workers tend to have longer average travel times (Groups 5, 6, 9, and 13), indicating longer trips used for these groups. An examination of trip purpose also reveals that certain demographic groups (groups 3, 5, 9, 13, and 14) tend to use transit more for airport trips than other purposes, compared to other demographic groups, which may contribute significantly to the longer travel times.

Table 14 presents a summary of travel data across the different demographic groups, including mean travel times. Figure 14 illustrates the distribution of these data around the median with box plots. The average travel times vary widely across demographic groups, from 17.9 minutes in Group 1 to 27.7 minutes in Group 13. Demographic groups with higher income tend to have a higher average travel time as well as a larger standard deviation. Groups 13-14 belong to the highest income category (Annual Net Income > $120k), and they also have the highest average travel time (26.3 – 27.7 minutes) and standard deviation (15.4 – 15.5 minutes). This is followed by the next income category (Group 5, Annual Net Income > $60k, and then by Groups 9-12, Annual Net Income $60k - $120k), with average travel times of 21.7 and 25.7 minutes (standard deviations of 13.0 and 14.0 minutes). Zero-car households have the lowest average travel time (17.9 minutes) and lowest standard deviation (11.0 minutes). While the low-income group (Annual Net Income < $25k) with cars has an average travel time of 19.0 minutes and a standard deviation of 11.4 minutes. These results imply that households with higher income and more vehicles tend to use transit for longer trips (higher average travel time) with greater variability (higher standard deviation).

Table 14. Summary of Results Across the 14 Demographic Groups

Group	Average Travel Time (minutes)	Standard Deviation of Travel Time (minutes)	Average Distance (miles)	Standard Deviation of Distance (miles)	Sample Size
1	17.88	11.02	6.01	6.57	16,414
2	18.96	11.41	6.70	14.93	2,436
3	19.96	11.86	7.44	5.22	70
4	20.57	12.38	7.79	6.22	943
5	23.72	13.88	10.80	22.91	539
6	20.97	12.16	8.07	6.47	138
7	19.55	11.90	7.08	6.36	2,533
8	19.62	11.79	7.63	18.35	7,704
9	25.72	14.03	11.10	21.88	676
10	21.70	13.04	8.59	6.32	1,262
11	21.83	13.07	9.45	22.51	1,124
12	22.60	13.82	9.73	22.26	2,416
13	27.73	15.46	13.52	21.97	148
14	26.32	15.35	12.59	31.74	761

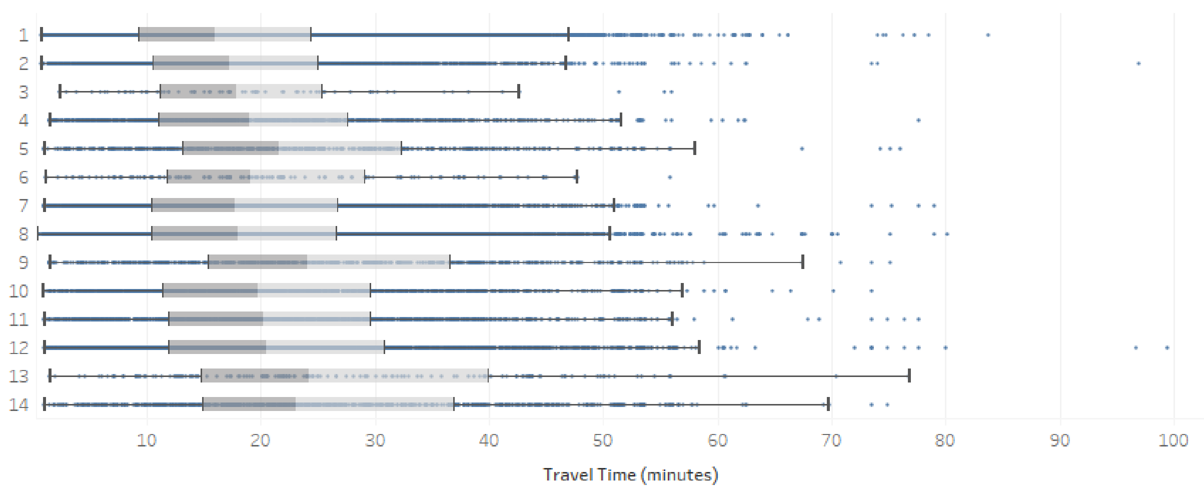


Figure 14. Travel Time Distribution across the 14 Demographic Groups

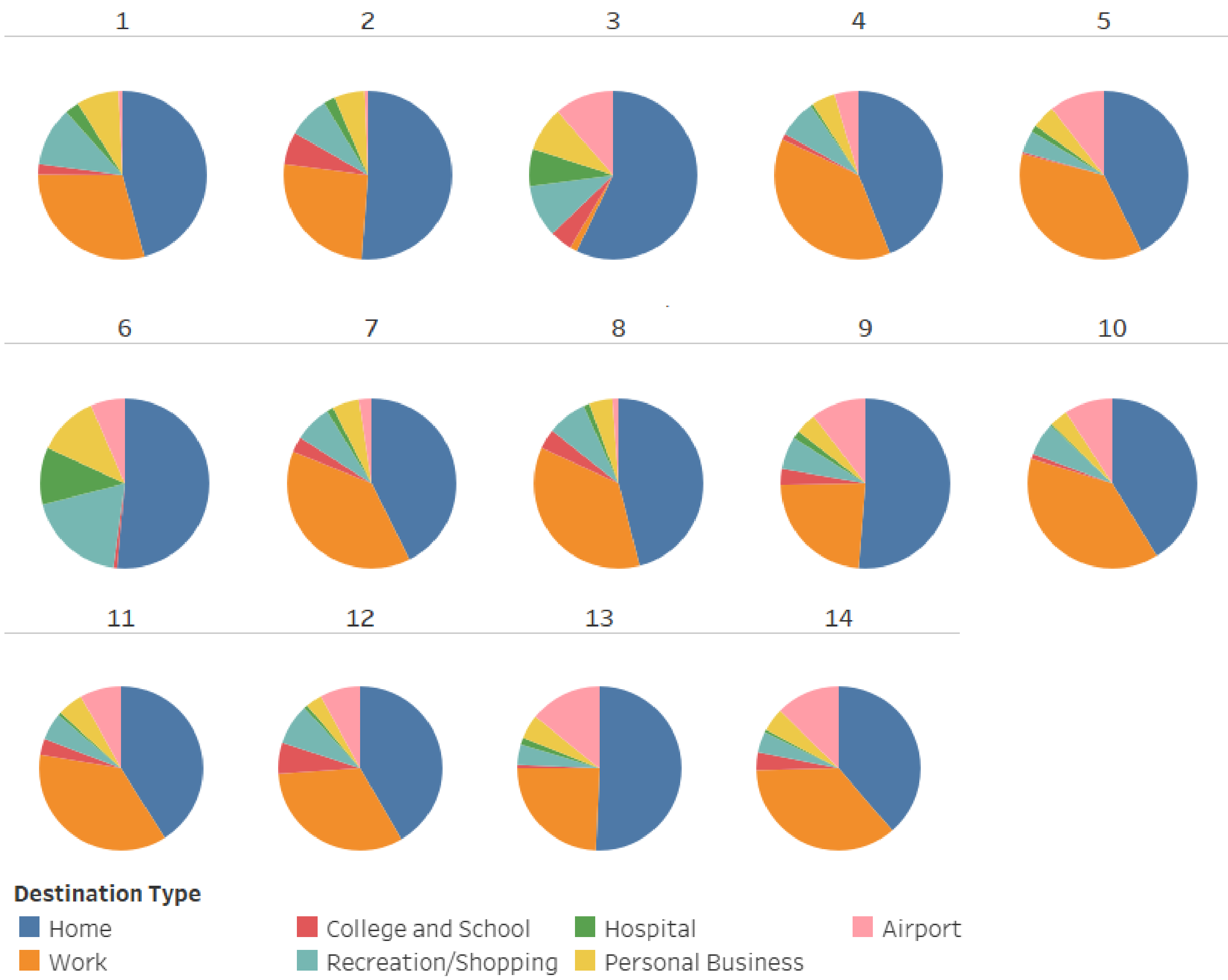


Figure 15. Trip Purpose across the 14 Demographic Groups

Looking at the last-mile travel mode of the different demographic groups, transit users most frequently use walking as their last-mile travel mode. However, Groups 5 and 9-14 also have a high percentage of park-and-ride trips. Because these groups belong to the higher income groups (Table 9), this finding may indicate that the higher income groups have more resources and flexibility in their transit trips.

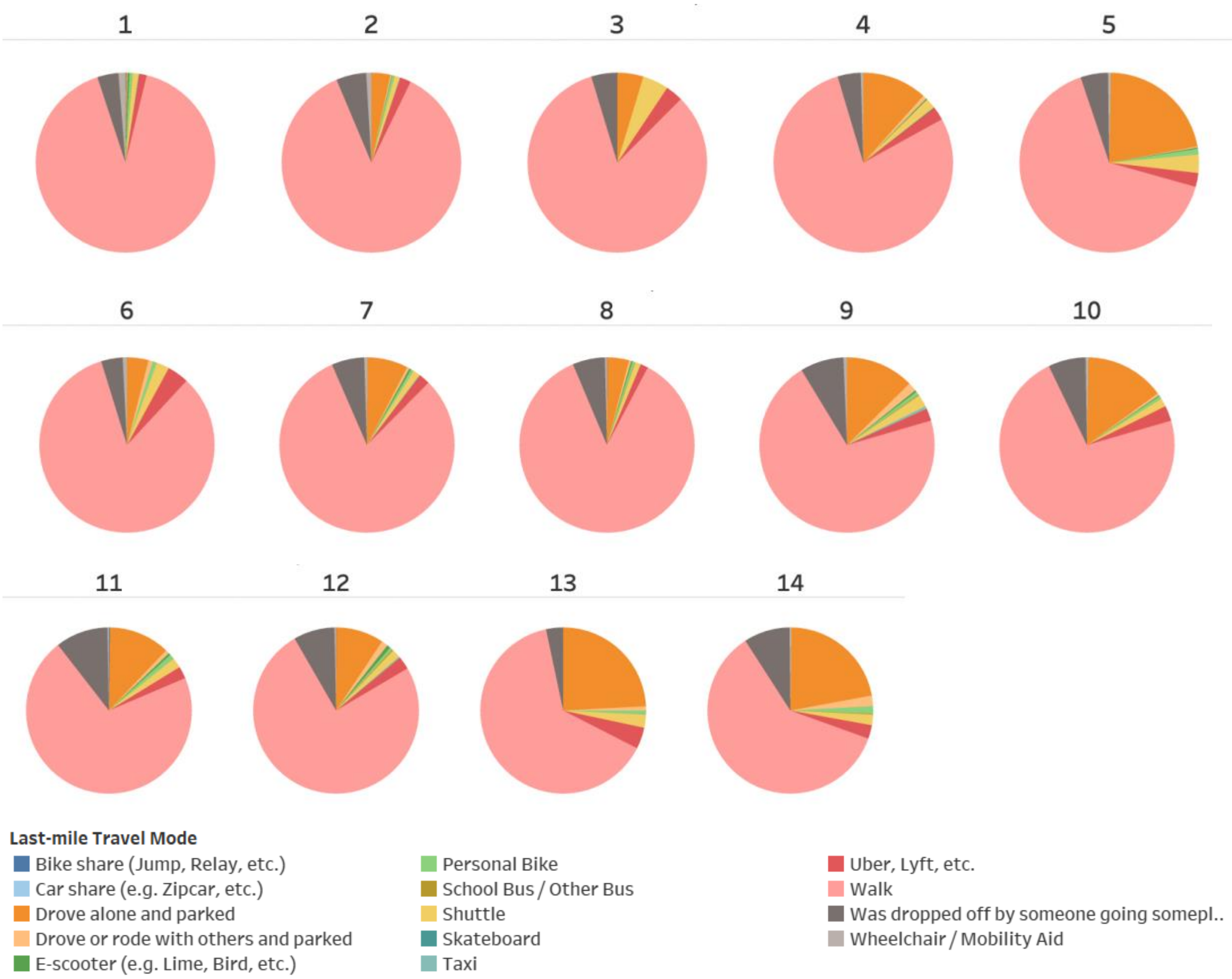


Figure 16. Last-mile Travel Mode across the 14 Demographic Groups

3. Transit Riders by Demographic Group

Transit Riders by Age Group

There is a general trend in transit travel conditions across different age groups. Travel time and distance increase as age increases for young and middle-aged groups, and then time and distance both decrease for ages over 55.

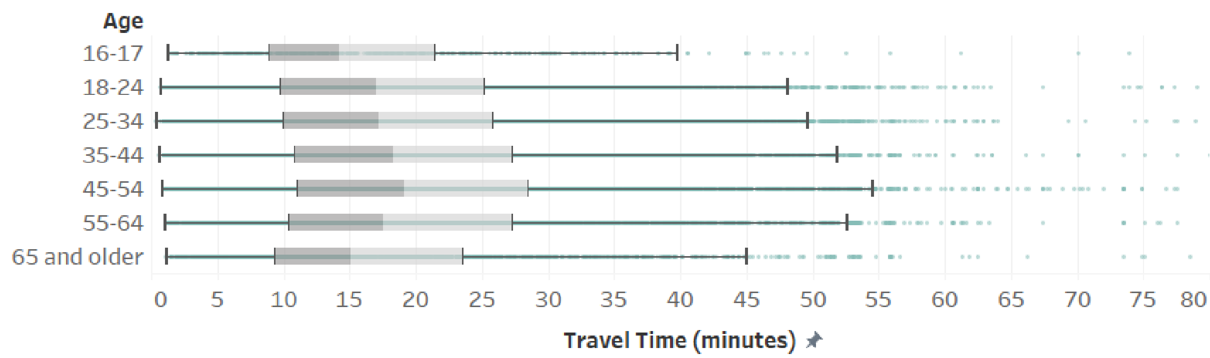


Figure 17. Travel Time Distribution by Age Category

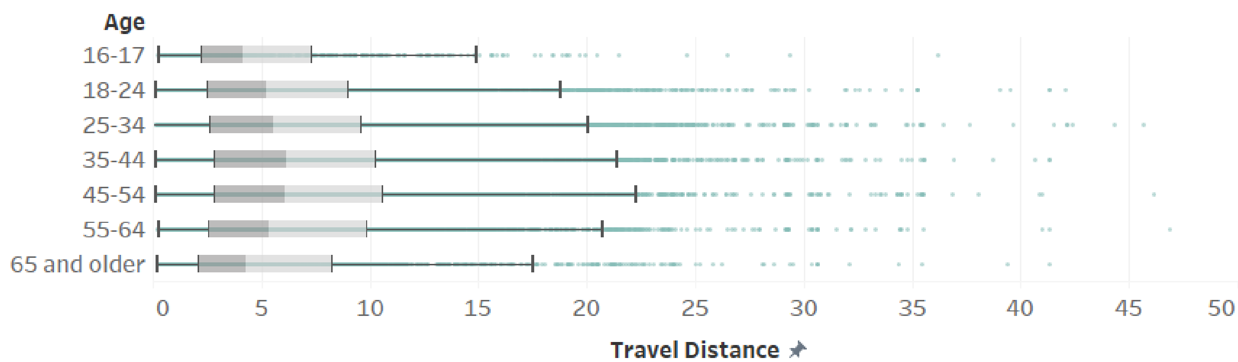


Figure 18. Travel Distance Distribution by Age Category

In terms of trip purpose, yunger riders use transit more frequently for school and college trips. Then as age increases, the middle-aged population uses transit most frequently for work. As the age continues to increase, the elder groups take more transit trips for recreation, hospital, and personal business.

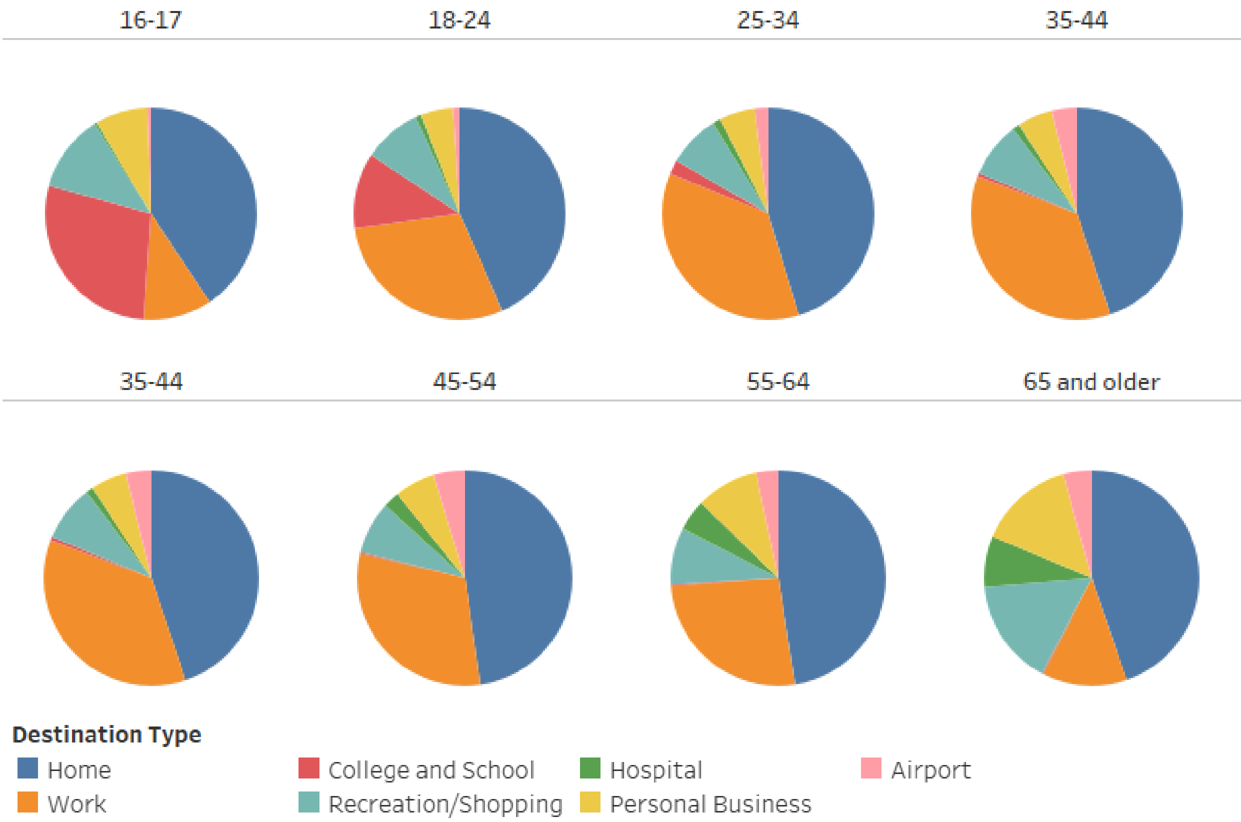


Figure 19. Trip Purpose Distribution by Age Category

Last-mile travel mode does not appear to show significant variation across different demographic groups. All age groups use walking as their primary last-mile travel method (taking up about 85% for all age categories). In terms of other modes, the young population (age 16-17) seldom drive alone for park-and-ride. As age increases, the older population has a notably higher dependence on wheelchairs and other mobility aids (less than 1% for all age groups younger than 55, 2.3% for ages 55-64, and 4.0% for age 65 and older).

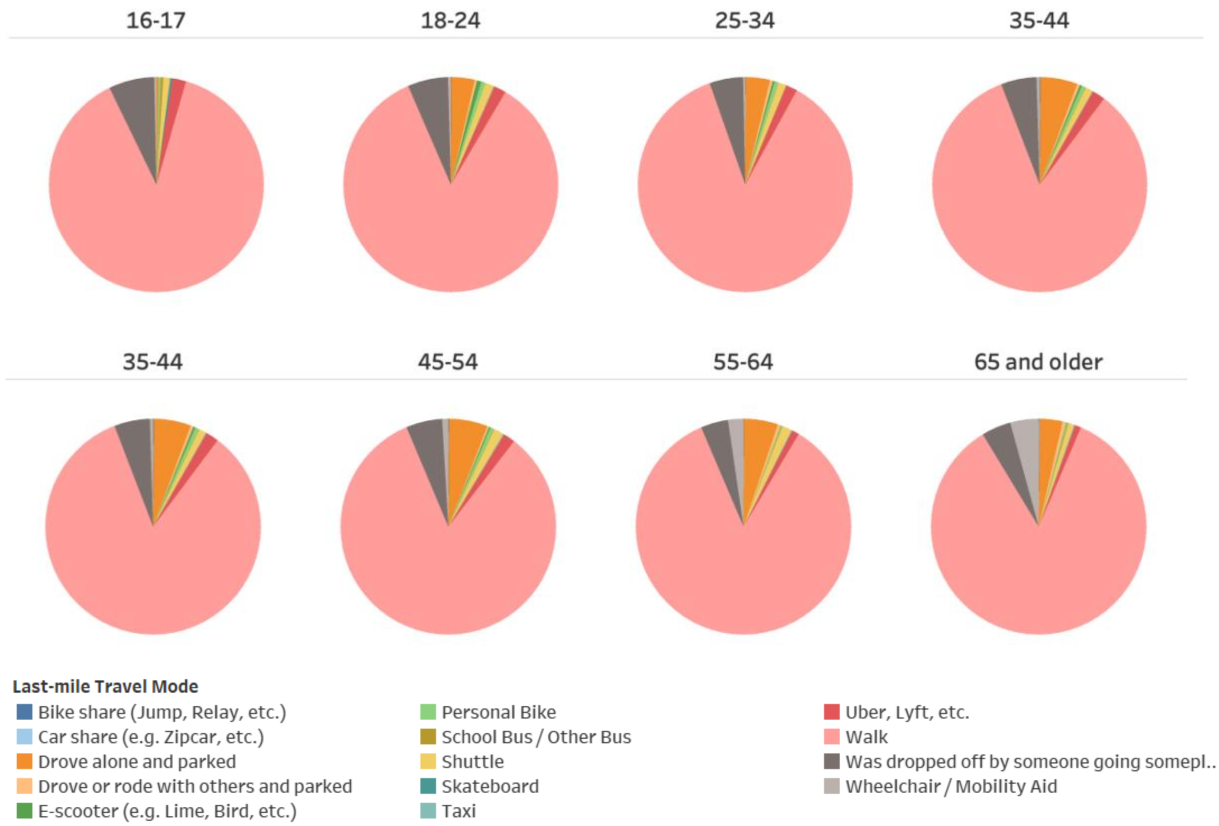


Figure 20. Last-mile Travel Mode by Age Category

Transit Riders by Race/Ethnicity Group

When comparing across different race groups, Asian, Native Hawaiian/Pacific Islander, and White/Caucasian respondents tend to travel longer distances and time than American Indian/Alaska Native and Black/African American respondents.

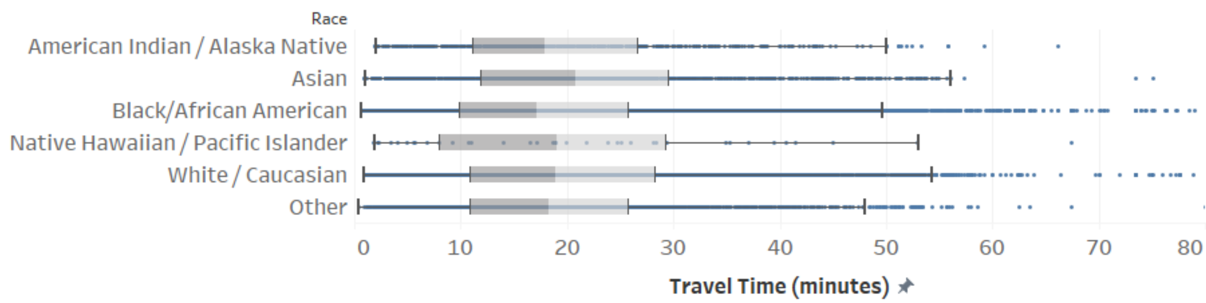


Figure 21. Travel Time Distribution by Race/Ethnicity Group

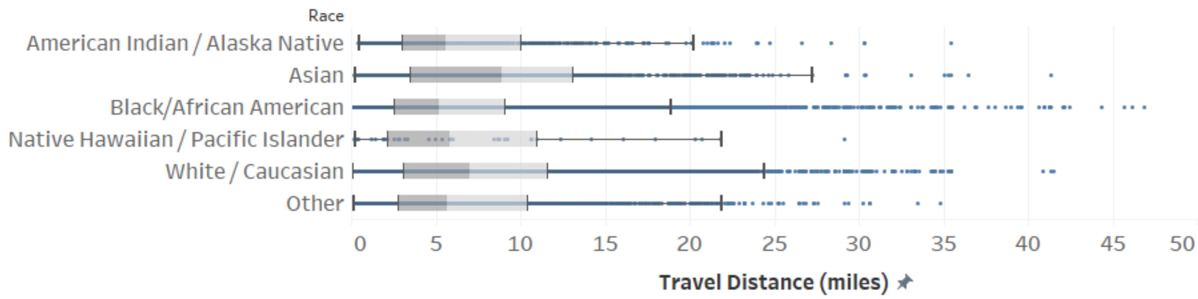


Figure 22. Travel Distance Distribution by Race/Ethnicity Group

When it comes to transit use frequency, over 75.4% of the Black/African American and 72.0% of the American Indian/Alaska Native respondents use transit five or more days a week. In comparison, only 61.1%, 54.8%, and 55.4% of Asian, Native Hawaiian/Pacific Islander, and White/Caucasian respondents use transit at such a high frequency. Only 1.9% of the Black/African American respondents are non-frequent transit users (once per year); whereas American Indian/Alaska Native, Asian, Native Hawaiian/Pacific Islander, and White/Caucasian respondents are infrequent users 5.1%, 8.8%, 6.5%, and 12.4% respectively.

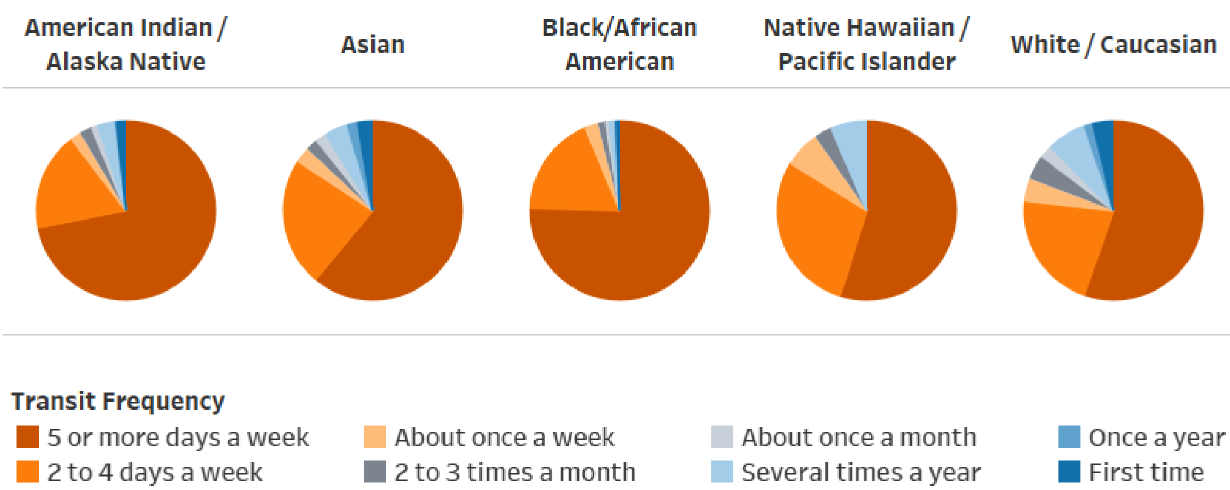


Figure 23. Transit Use Frequency by Race/Ethnicity Group

In terms of trip purpose, Asian and Native Hawaiian/Pacific Islanders spend a higher percentage of trips on school and college (12.6% and 10.8% respectively, compared to 2.9% to 4.2% for other groups). American Indian/Alaska Native and Black/African American respondents use transit more frequently for commuting (34.7% and 32.7% respectively, compared to 24.3% to 28.8% in other groups). Asian, Native Hawaiian/Pacific Islander and White/Caucasian respondents use transit more frequently for airport trips (6.2%, 5.4%, and 0.0% respectively, compared to 3.3% and 1.1% for American Indian/Alaska Native and Black/African American respondents).

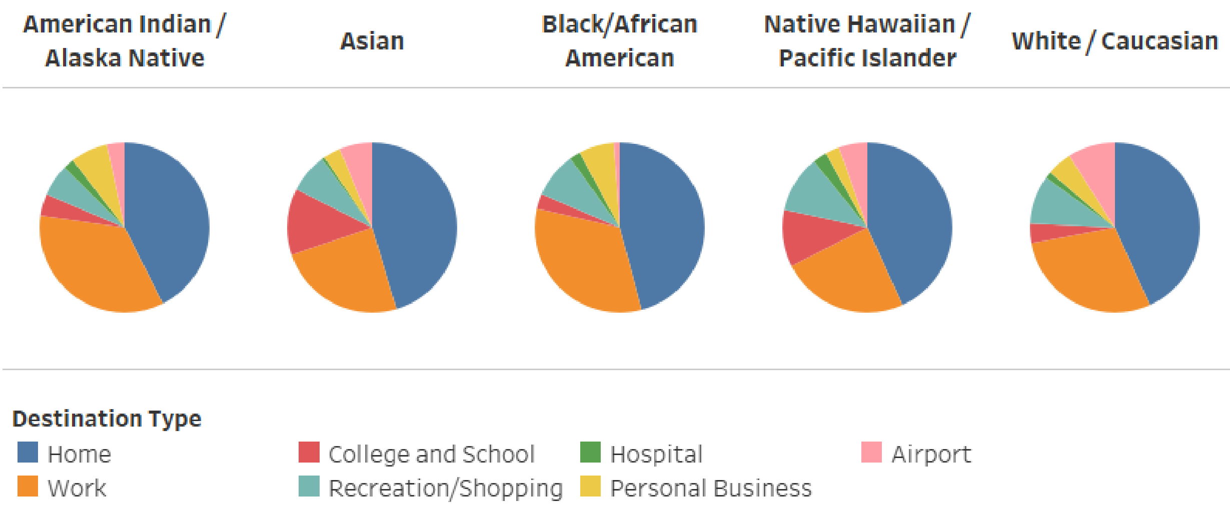


Figure 24. Trip Purpose by Race/Ethnicity Group

Last-mile travel mode also varies by race. Though all race groups use walking as their primary last-mile travel mode, American Indian/Alaska Native (85.1%) and Black/African American (87.6%) are frequent walkers (compared to 76.8% to 77.4% of other race groups). In comparison, Asians (10.8%), Native Hawaiian/Pacific Islanders (9.7%), and White/Caucasian (9.2%) use SOV-based park-and-ride more frequently than American Indian/Native American 4.4% and Black/African American (3.0%).

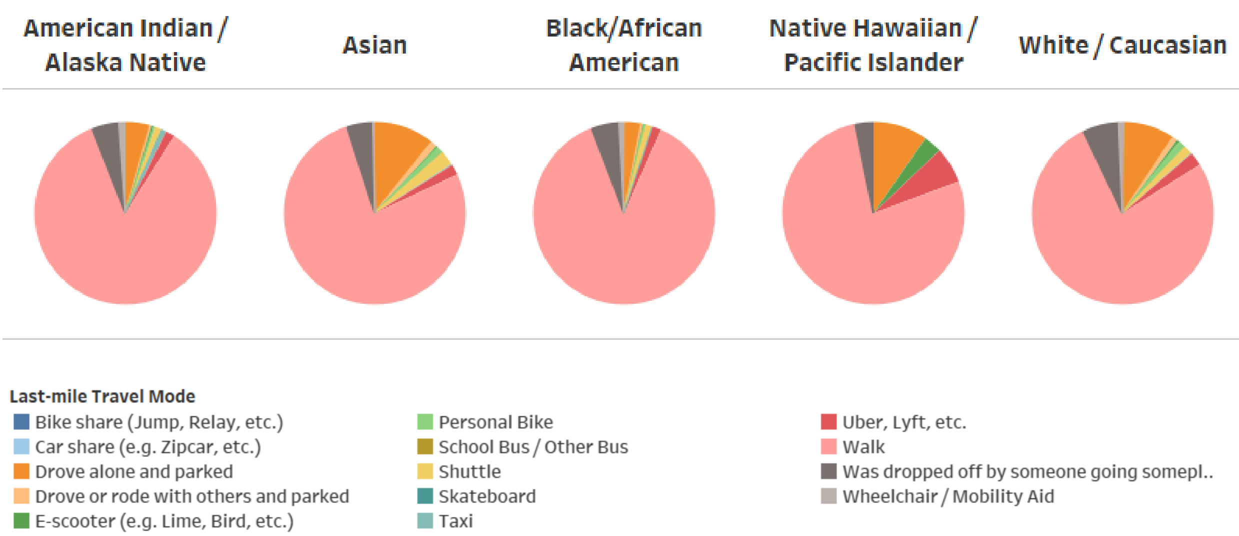


Figure 25. Last-mile Travel Mode by Race/Ethnicity Group

Transit Riders with Different Gender

Females and males in metro Atlanta have very similar travel patterns in terms of travel time (Figure 26), distance (Figure 27), trip purpose (Figure 28), last-mile mode (Figure 29), and trip frequency (Figure 30).

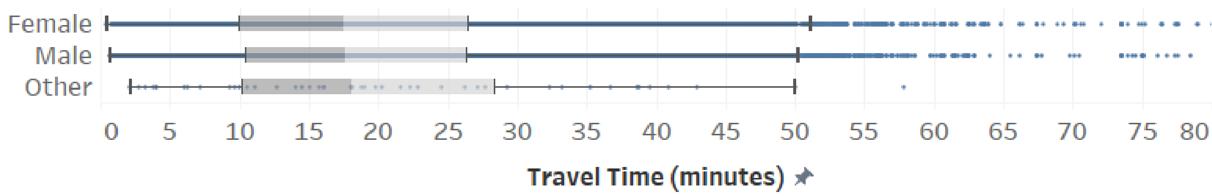


Figure 26. Travel Time Distribution by Gender

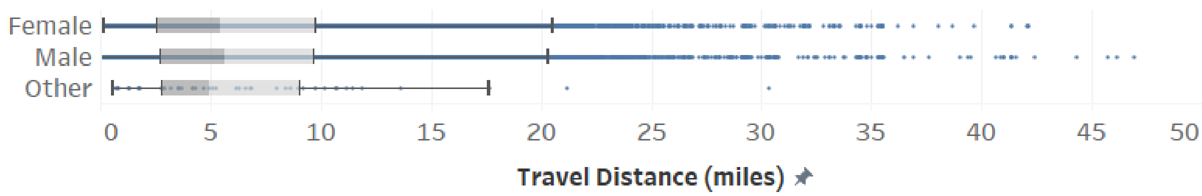


Figure 27. Travel Distance Distribution by Gender

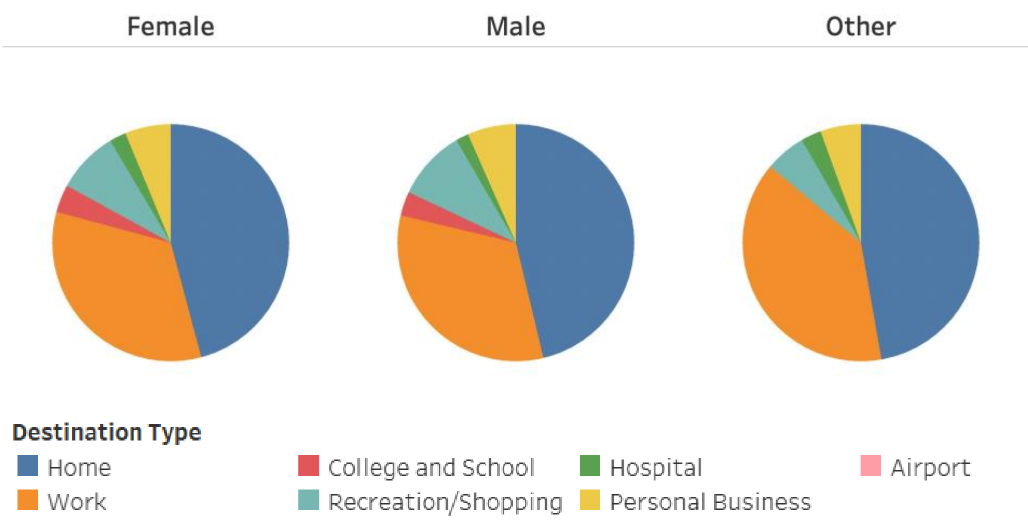


Figure 28. Trip Purpose by Gender

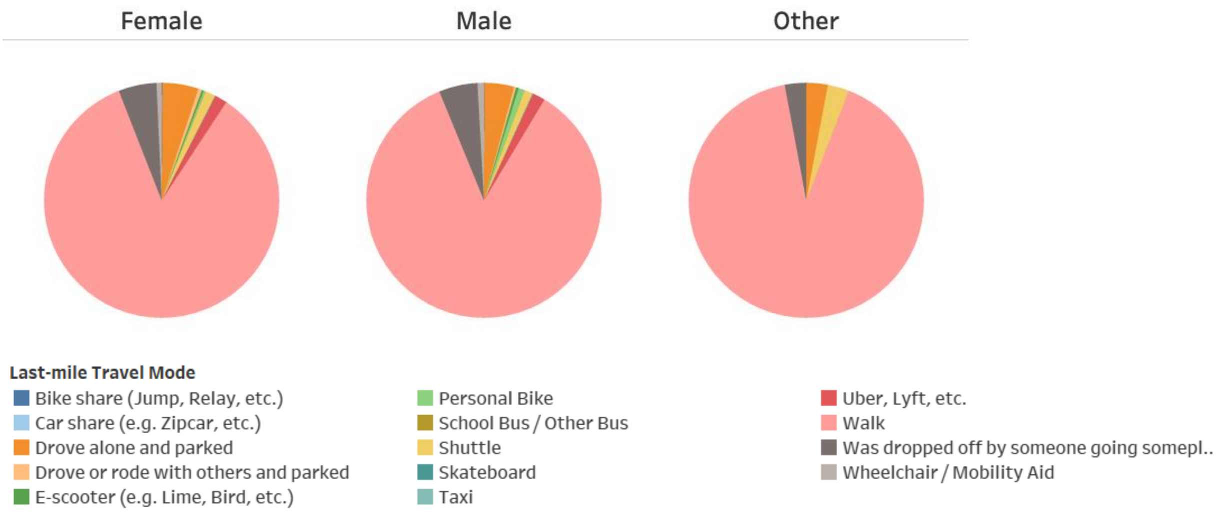


Figure 29. Last-mile Travel Mode by Gender

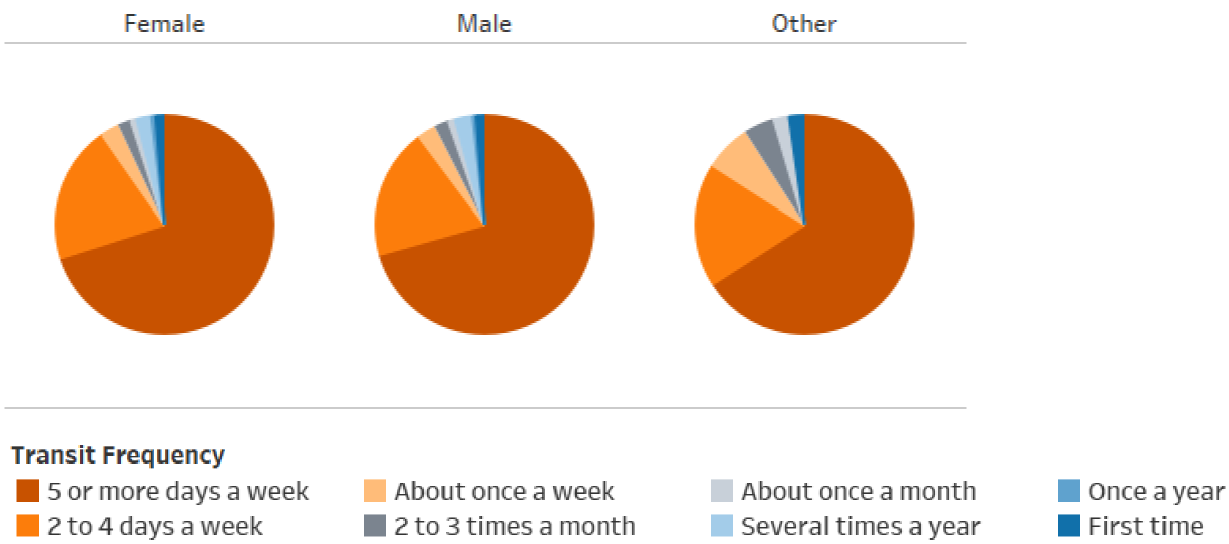


Figure 30. Transit Use Frequency by Gender

Transit Riders by English Language Ability

The Transit Onboard Survey has four categories for English language ability, Very Well, Well, Less than Well, and Not at All. Due to sample size constraints, the category Not at All (sample size of 4) is excluded from the analysis. As is shown in the results, the groups with lower English language capability tend to travel shorter distances (Figure 31) and time (Figure 32) and have fewer options (Figure 33) for last-mile travel mode.

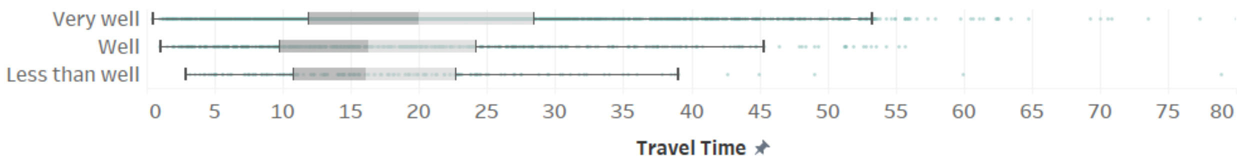


Figure 31. Travel Time Distribution for Riders with Different English Language Ability

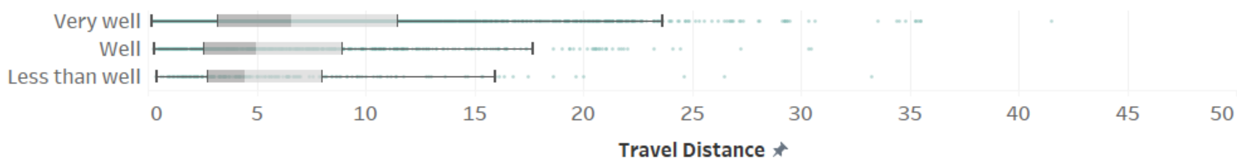


Figure 32. Travel Distance Distribution of Riders with Different English Language Ability

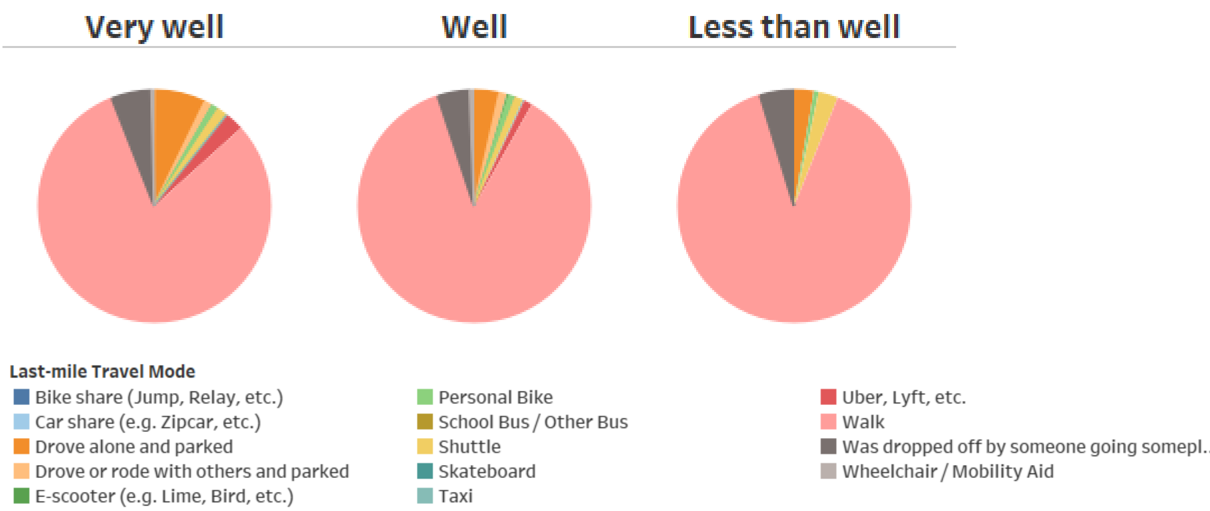


Figure 33. Last-mile Travel Mode for Riders with Different English Language Ability

Energy Use Modeling Across Demographic Groups

This section introduces the modeling method to estimate trip-level energy use associated with each trip, replicating the work conducted in the prior study (Fan, et al., 2022). The energy use modeling was conducted on a subset of the data, MARTA trips in Periods 2-3 (April 19, 2019, to June 1, 2019), to demonstrate the feasibility of such analysis in large-scale research. For this study, the energy analysis outputs are combined with demographic information for each route to assess energy use across demographic groups.

The analysis was performed by implementing MOVES-Matrix, developed by Georgia Tech to facilitate rapid applications of energy modeling using the same outputs as the MOVES regulatory model (Liu, et al., 2019; Xu, et al., 2018; Liu, et al., 2017; Guensler et al., 2016; Xu, et al., 2015). By running MOVES about thirty thousand times for a region (i.e., for each specific fuel specification and inspection and maintenance program), across all combinations of input variables that affect emission rates, a multi-dimensional matrix of 90 billion energy use and emission rates is generated. Users can query the emission rates directly from the matrix, significantly improving run-time efficiency (Liu, et al., 2019). Link-by-link average speed was derived from transit travel time between stops and link distance, and the source type distributions and transit vehicle age distributions were extracted from the fleet composition profiles provided by MARTA. Analyses also control for any other factors that affect energy use rates, such as ambient temperature and humidity. Instead of using daily meteorological conditions as individual model run inputs, meteorology information is estimated from the National Weather Service Climate Summary of May 2019 (National Oceanic and Atmospheric Administration, 2019). The average May temperature (70°F) and humidity (70%) in Atlanta are used as meteorology input for MOVES-Matrix (consistent meteorology settings for all periods).

To obtain an estimate of the per-passenger energy use, the average route-segment level passenger count in the first week of May 2019 (May 6 – 12, 2019), which was 5.8, was used as the denominator for all energy use factors. During the study period, MARTA riders used an average of 20,773 Btu per trip, and 2,798 Btu per passenger-mile (Figure 34, Figure 35). 100.0% of trips have over 2,000 Btu per passenger mile, 99.2% over 2,500, and 15.2% over 3,000. The post-analysis was conducted in R and the analytical results were visualized in Tableau. Supplemental analyses for a forthcoming journal manuscript of this case study will integrate automated passenger count data for each route to refine the analysis; however, the QA/QC process for these data will take more than a month to complete.

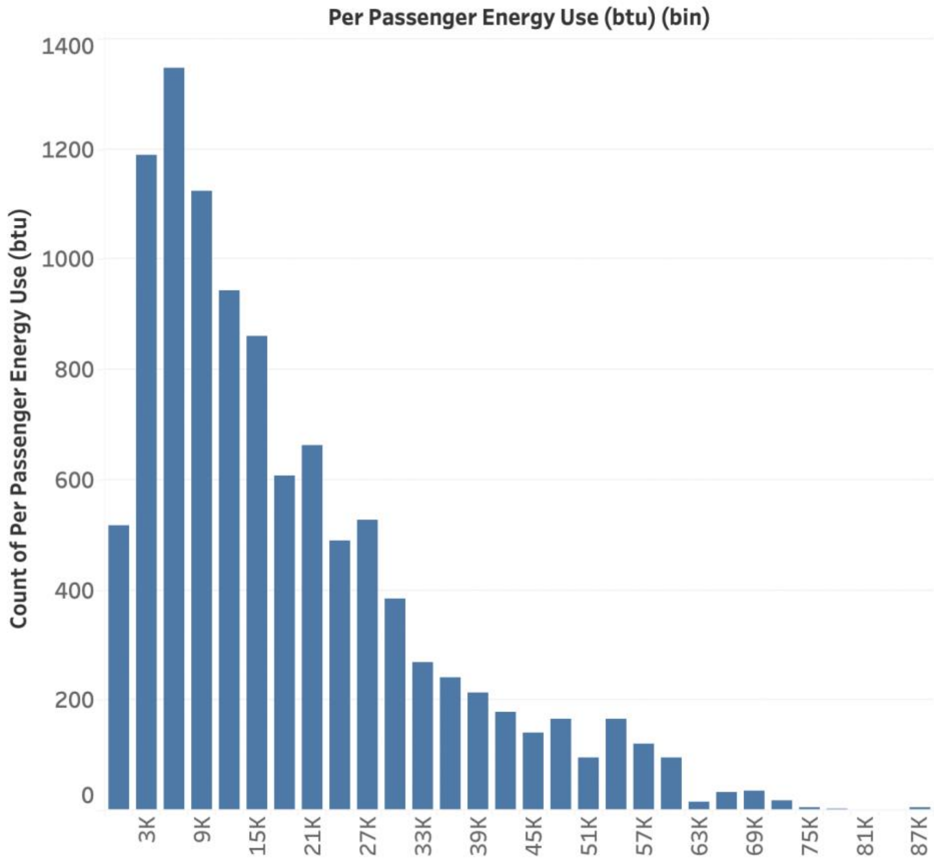


Figure 34. Energy Use Profile of All Passengers at the Trip Level

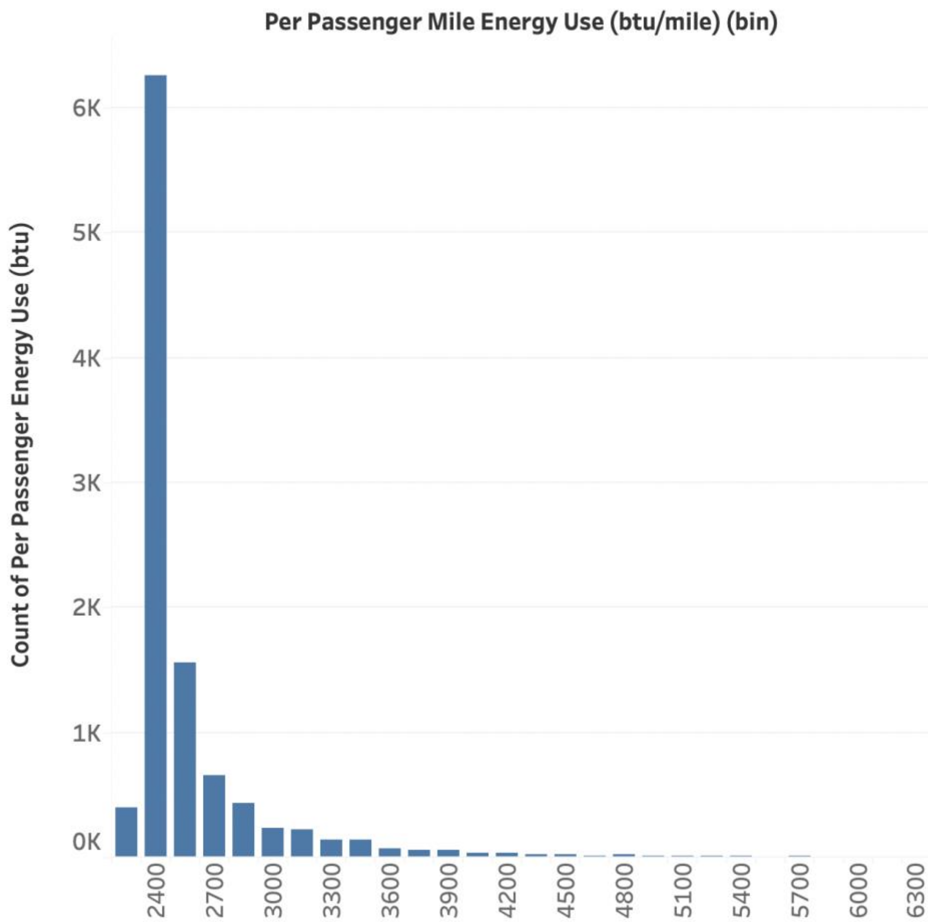


Figure 35. Per Passenger Mile Energy Use during the Study Period

The trip-level energy used by different racial groups largely reflects the average trip miles traveled by each group (Figure 36, Figure 37). Asian and White/Caucasians use the highest amount of energy at the trip level, and Black/African Americans the least. When it comes to per passenger mile energy use, however, there is no significant difference across the race/ethnicity groups (Figure 36). Note that the Native Hawaiian/Pacific Islander group is excluded from this analysis due to the small sample size (n = 5).

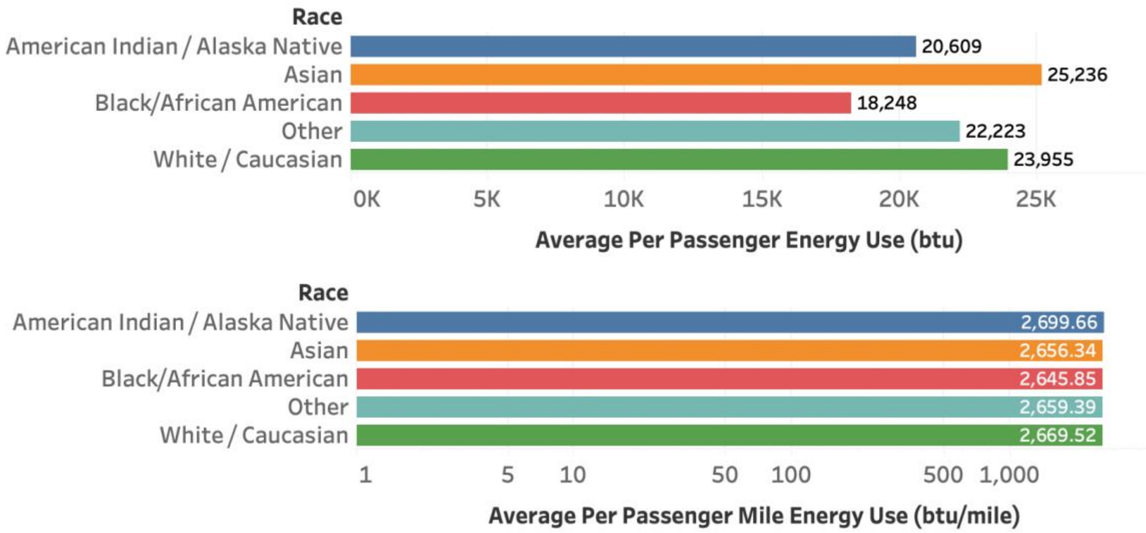


Figure 36. Energy Use by Race

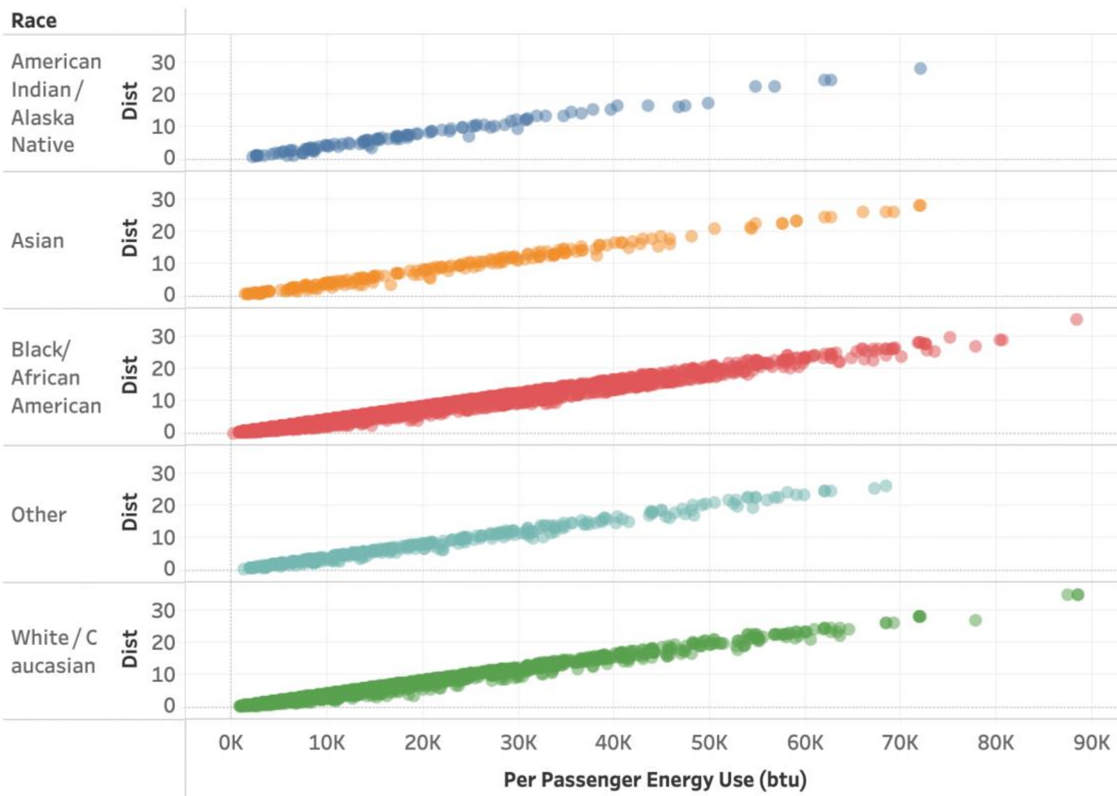


Figure 37. Relationship Between Energy Use and Distance by Race/Ethnicity

Energy use by gender shows a similar pattern to that of race (Figure 38). The trip level energy use differs by a small amount related to the average trip miles traveled, while energy use per passenger mile is identical between females and males. Note that the gender "Other" is excluded from this analysis due to the small sample size (n = 12).

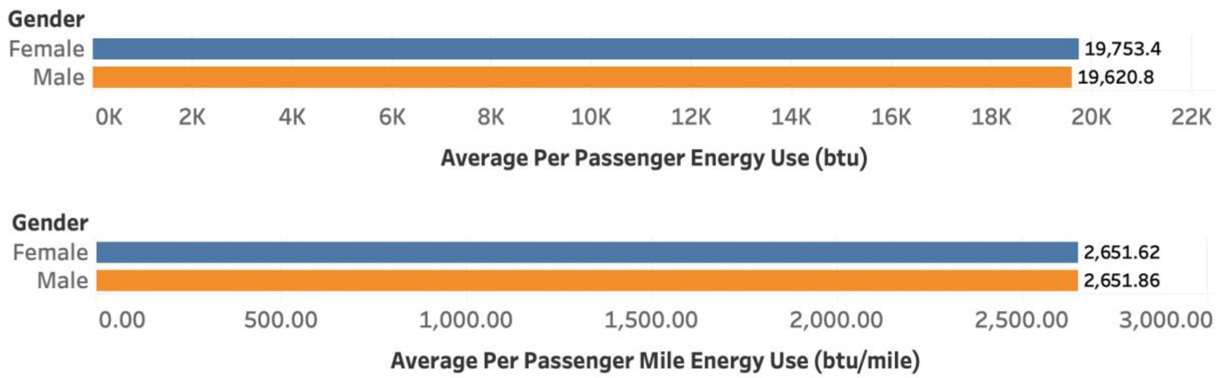


Figure 38. Energy Use by Gender

Transit riders with a driver's license are associated with higher trip-level energy use compared to those without a license (Figure 39). Several factors could help explain these results. First, people with a driver's license may be more closely associated with living in suburban areas and thus traveling longer distances. Second, there might be more airport-bound and longer-distance trips. Riders with a driver's license are also found to have a slightly lower per-passenger-mile energy use (Figure 39). This can also be possibly explained by the fact that people without a driver's license tend to live closer to Midtown and Downtown, where the transit travel speed is lower than 30 mph and slightly less energy efficient compared to travel at higher speeds (for example, 50 mph).

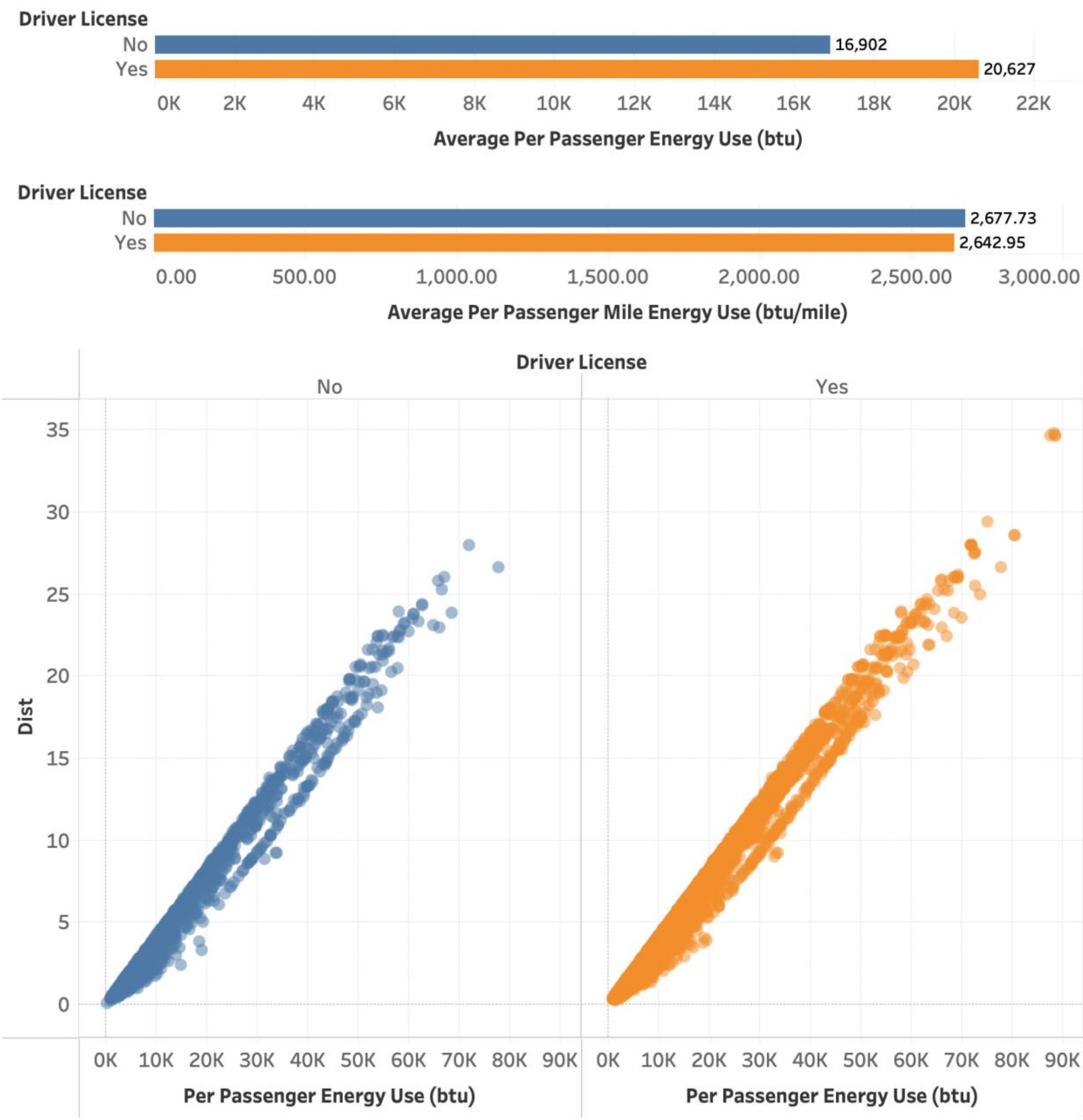


Figure 39. Energy Use vs. Possession of a Driver License

Conclusions and Future Work

Many transit providers changed their schedules and route configurations during the COVID-19 pandemic, providing more frequent bus service on major routes and curtailing other routes, to reduce the risk of COVID-19 exposure. The previous NCST study examined pandemic-related changes in MARTA transit system service and ridership in Atlanta, GA, and the combined effects on energy use and per-passenger energy use (Fan, et al., 2022). For that previous study, General Transit Feed Specification (GTFS) and the automated passenger count (APC) datasets were used to develop the transit network and derive distance and passenger load information within the TransitSim analytical framework. Ridership data were coupled with energy use and emission rates from MOVES-Matrix to assess how the changes in transit service and ridership affected energy use and emissions on a per passenger-mile basis. The prior results showed that the coupled effect of the shift in transit frequency and decrease in ridership from 2019 to 2020 increased route-level energy use and per-passenger-mile energy use for the vast majority of MARTA transit routes.

The research performed in this supplemental NCST study provided improvements to TransitSim to increase analytical efficiency and to integrate ridership demographics, so that energy use impacts could be assessed across demographic groups for use in social sustainability analysis. This report first summarized the model TransitSim improvements and demonstrated how the modeling could be used to support social sustainability analysis. Trip-level energy use for MARTA trips from April 19, 2019, to June 1, 2019) was estimated using route data and MOVES-Matrix energy use rates. The analyses show that the average trip-level energy used by MARTA riders during the study period was 20,773 Btu, and the per-passenger-mile energy use was 2,798 Btu. The energy estimates were then allocated across demographic groups for comparison. Energy use across the race/ethnicity groups largely reflected the differences in average trip miles traveled by each group. There was no significant difference in per-passenger-mile energy use across different race/ethnicity groups or genders. Transit riders with a driver's license are associated with higher trip-level energy use compared to those without a license, but these riders also tend to have slightly lower energy use per passenger-mile. The analyses demonstrated how TransitSim 3.0 energy use analyses can be coupled with data from on-board ridership surveys (or other trip-level data) to allocate energy use across demographic groups for use in social sustainability analysis and in assessing the potential impacts of transit investment and changes in operations.

This research proposed a novel framework and methods for trip-level system analysis that provides a number of benefits to different stakeholders: 1) the modeling tools provide academics and researchers the means to incorporate trip-level system thinking into their analyses; 2) the case study demonstrates to transit agencies, metropolitan planning organizations, and departments of transportation that an open-source, ready-to-apply modeling framework (TransitSim 3.0) can be used in planning processes, and 3) the general public and elected officials have some new insight into how social sustainability within transit systems varies at the transit-trip level.

References

Aman, J.J. C., and J. Smith-Colin (2020). Transit Deserts: Equity analysis of public transit accessibility. Journal of Transport Geography, 89, 102869.

APTA, American Public Transportation Association (2021). 2021 Public Transportation Fact Book.

ARC, Atlanta Regional Commission (2022). Transit Services in the Atlanta Region - ARC WWW Document. URL https://atlantaregional.org/transportation-mobility/transit/transit-in-the-atlanta-region/ (accessed 5.24.22).

ARC, Atlanta Regional Commission (2021). ARC 2021 Annual Report. Atlanta.

ARC, Atlanta Regional Commission (2020). Atlanta Regional Commission Regional On-Board Prepared for the Atlanta Regional Commission by ETC Institute. Atlanta.

Belzer, D., and G. Autler (2002). Transit oriented development: moving from rhetoric to reality (pp. 06-15). Washington, DC: Brookings Institution Center on Urban and Metropolitan Policy.

Bennett, K., and M. Shirgaokar (2016). Social Sustainability of Transit: An Overview of the Literature and Findings from Expert Interviews.

Carleton, P.R., and J.D. Porter (2018). A comparative analysis of the challenges in measuring transit equity: definitions, interpretations, and limitations. J Transp. Geography 72, 64–75. https://doi.org/10.1016/J.JTRANGEO.2018.08.012

Carleton, P. R., and J.D. Porter (2018). A comparative analysis of the challenges in measuring transit equity: definitions, interpretations, and limitations. Journal of transport geography, 72, 64-75.

CATS, Cherokee Area Transit Service (2022). CATS Transportation | Cherokee County, Georgia. WWW Document. URL https://www.cherokeecountyga.gov/Transportation/. Accessed 05/24/22.

CobbLinc (2022). Routes and Schedules | Cobb County Georgia. Online document. URL https://www.cobbcounty.org/transportation/transit/routes-and-schedules (accessed 5.24.22).

Crimmins, T.J. and H. L. Preston (1980). Automobile Age Atlanta: The Making of a Southern Metropolis, 1900-1935. Technol Cult 21, 122. https://doi.org/10.2307/3104013.

Fan, H., H. Lu, Z. Dai, T.R. Passmore, A. Guin, K. Watkins, and R. Guensler (2023a). Combined effect of changes in transit service and changes in occupancy on per-passenger energy consumption. Transportation Research Record, 2677(2), 1252-1265.

Fan, H., Z. Liu, and R. Guensler (2023b). Multi-layered transit network resilience to flooding: An analysis of 19 U.S. cities (TRBAM-22-05475). Paper presented at the Transportation Research Board 102nd Annual Meeting, Washington, D.C. January 2023.

Fan, H., H. Lu, Z. Dai, R. Passmore, A. Guin, K. Watkins, and R. Guensler (2022). Combined Effect of Changes in Transit Service and Changes in Occupancy on Energy Consumption per Passenger-Mile. Research Report from the National Center for Sustainable Transportation. December 2022.

Freilich, R. H. (1998). The land-use implications of transit-oriented development: Controlling the demand side of transportation congestion and urban sprawl. Urb. Law., 30, 547.

Gainesville County. (2022). Hall Area Transit Services | Gainesville, GA. Online document. URL https://www.gainesville.org/186/Hall-Area-Transit-Services (accessed 5.24.22).

Griffin, G.P., and I.N. Sener (2016). Public Transit Equity Analysis at Metropolitan and Local Scales: A Focus on Nine Large Cities in the US. J Public Trans 19, 126. https://doi.org/10.5038/2375-0901.19.4.8.

GRTA, Georgia Regional Transportation Authority (2022). Xpress | Commuting Made Easy! Online document. URL https://www.xpressga.com/ (accessed 5.24.22).

Gwinnett County. (2022). Bus Routes and Schedules. https://www.gwinnettcounty.com/web/gwinnett/departments/transportation/gwinnettcountytransit/routesandschedules. Accessed 05/24/22.

Garrett, M., and B. Taylor (1999). Reconsidering social equity in public transit. Berkeley Planning Journal, 13(1).

Gbologah, F., Y. Xu, M.O. Rodgers, and R. Guensler (2014). Demonstrating a Bottom-up Framework for Evaluating Energy and Emission Performance of Various Electric Rail Transit Options. Transportation Research Record. Number 2428. pp. 10-17. National Academy of Sciences. Washington, DC. 2014.

Griffin, G.P. and I.N. Sener (2016). Public transit equity analysis at metropolitan and local scales: A focus on nine large cities in the US. Journal of public transportation, 19(4), 126-143.

Guensler, R., H. Lu, and M.O. Rodgers (2022). High-Resolution Exposure Modeling with ABM, MOVES-Matrix, AERMOD, and Demographics. Atlanta Regional Commission (ARC) Model User's Group Webinar. September 30, 2022.

Guensler, R., H. Lu, W. Reichard, Z. Dai, T. Xia, A. Guin, and M.O. Rodgers (2021). AERMOD, RLINE, and RLINEXT Case Study Analyses in Atlanta, GA. Prepared for the Federal Highway Administration and Georgia Department of Transportation. Final Report (FHWA-GA-21-2024). https://trid.trb.org/view/1896759. November, 2021.

Guensler, R., H. Liu, X. Xu, Y. "Ann" Xu, M.O. Rodgers (2016). MOVES-Matrix: Setup, Implementation, and Application, in: Transportation Research Board 95th Annual Meeting. Washington, D.C. Gwinnett County, 2022.

Henderson, J.M., 2002. Contesting the Spaces of the Automobile: the Politics of Mobility and the Sprawl Debate in Atlanta, Georgia. University of Georgia, Athens.

Karner, A. and K. Levine (2021). Equity-advancing practices at public transit Agencies in the United States. Transportation Research Record, 2675(10), 1431-1441.

Li, H. “Ann,” Y. “Cody” Wang, X. Xu, H. Liu, A. Guin, M.O. Rodgers, M. Hunter, J.A. Laval, K. Abdelghany, and R. Guensler (2018). Assessing the Time, Monetary, and Energy Costs of Alternative Modes, in: Transportation Research Board 97th Annual Meeting. Washington, D.C.

Li, H., 2019. A framework for optimizing public transit fleet conversion to alternative fuels (Ph.D. Dissertation). Georgia Institute of Technology.

Linovski, O., D. M. Baker, and K. Manaugh (2018). Equity in practice? Evaluations of equity in planning for bus rapid transit. Transportation Research Part A: Policy and Practice, 113, 75-87.

Litman, T. (2012). Evaluating public transportation health benefits. Victoria, BC, Canada: Victoria Transport Policy Institute.

Liu, H., R. Guensler, X. Xu, Y. Xu, and M. Rodgers (2019). “MOVES-Matrix for High-Performance On-road Energy and Emission Rate Modeling Applications.” UAWM-2019-0014. DOI: 10.1080/10962247.2019.1640806. Journal of the Air and Waste Management Association. July 2019.

Liu, H., X. Xu, M. Rodgers, Y. Xu, and R. Guensler (2017). "MOVES-Matrix and Distributed Computing for Microscale Line Source Dispersion Analysis." Journal of the Air and Waste Management Association. DOI: 10.1080/10962247.2017.1287788. Volume 67, Issue 7, pp. 763-775. April 2017.

Liu, H., Xu, Y. “Ann,” Stockwell, N., Rodgers, M.O., Guensler, R., 2016. A Comparative Life-cycle Energy and Emissions Analysis for Intercity Passenger Transportation in the U.S. by Aviation, Intercity Bus, and Automobile. Transp Res D Transp Environ 48, 267–283. https://doi.org/10.1016/J.TRD.2016.08.027

Liu, L., Miller, H.J., Scheff, J., 2020. The Impacts of COVID-19 Pandemic on Public Transit Demand in the United States. PLoS One 15, e0242476. https://doi.org/10.1371/JOURNAL.PONE.0242476

Lu, X., Pas, E.I., 1999. Socio-demographics, activity participation and travel behavior. Transp Res Part A Policy Pract 33, 1–18. https://doi.org/10.1016/S0965-8564(98)00020-2

MARTA, 2021. MARTA Service Standards FY 2021. Atlanta.

Pradono, P., I.P. Kusumantoro, and A. Retapradana (2015). Towards carbon dioxide emission reduction through public transport revitalization: case of proposed Surabaya tram transit project. Journal of the Eastern Asia Society for Transportation Studies, 11, 967-978.

Turner, J. (2013). Urban mass transit and social sustainability in Jakarta, Indonesia. Global Report on Human Settlements.

Wei, R., X. Liu, Y. Mu, L. Wang, A. Golub, and S. Farber (2017). Evaluating public transit services for operational efficiency and access equity. Journal of transport geography, 65, 70-79.

Welch, T.F., and S. Mishra (2013). A measure of equity for public transit connectivity. Journal of Transport Geography, 33, 29-41.

WSP | Parsons Brinckerhoff and Atkins (2017). Activity-Based Model User Guide - Coordinated Travel – Regional Activity-Based Modeling Platform (CT-RAMP) for Atlanta Regional Commission.

Xu, X., H. Liu, Y. Xu, M. Hunter, and R. Guensler (2015). “Estimating Project-level Vehicle Emissions using Vissim and MOVES Matrix.” DOI 10.3141/2570-12. Transportation Research Record. Number 2570. pp. 107-117. National Academy of Sciences. Washington, DC. December 2015.

Xu, X., H. Li, H. Liu, M. Rodgers, and R. Guensler (2018). “Evaluation of Transit Ecodriving in Rural, Suburban, and Urban Environments.” DOI: 10.1177/0361198118797778 Transportation Research Record. Volume 2672, Issue 8, pp. 152-164. Washington, DC. September 2018.

Xu, X., Liu, H., Passmore, R., Patrick, T., Gbologah, F., Rodgers, M.O., Guensler, R., 2018b. Fuel and Emissions Calculator (FEC), Version 3.0, Summary Report. Atlanta, GA.

Yoon, S., Li, H., Jun, J., Ogle, J.H., Guensler, R., Rodgers, M.O., 2005. Methodology for Developing Transit Bus Speed–Acceleration Matrices for Load-Based Mobile Source Emissions Models. https://doi.org/10.1177/0361198105194100104 1941, 26–33. https://doi.org/10.1177/0361198105194100104

Zhao, Y. (2020). Distributional Impact Assessment of Transportation Policy Using Activity-Based Models With Path Retention (Doctoral dissertation, Ph. D Dissertation, Georgia Institute of Technology).

Zhao, Y. (2021). Distributive justice impact assessment using activity-based modeling with path retention (Doctoral dissertation, Georgia Institute of Technology).

Zhu, Z., Z. Li, Y. Liu, H. Chen, and J. Zeng (2017). The impact of urban characteristics and residents’ income on commuting in China. Transportation research part D: transport and environment, 57, 474-483.

Zhao, Y. (2020). Distributional Impact Assessment of Transportation Policy Using Activity-Based Models with Path Retention. Georgia Institute of Technology, Atlanta.

Data Summary

Products of Research

The research team collected no data for this study. The data employed include:

- General Transit Feed Specification (GTFS) Data - Open source and readily available online (link: https://transitfeeds.com/p/marta/65)
- Automated passenger count (APC) Data - Proprietary data procured from MARTA under a specific end-use agreement
- MOVES-Matrix Energy and Emission Rates - Open source data available through NCST at: https://tse.ce.gatech.edu/ncst/movesmatrix

Data Format and Content

- GTFS Data - Standard GTFS format
- APC Data - Proprietary
- MOVES-Matrix Energy and Emission Rates - Text arrays

Data Access and Sharing

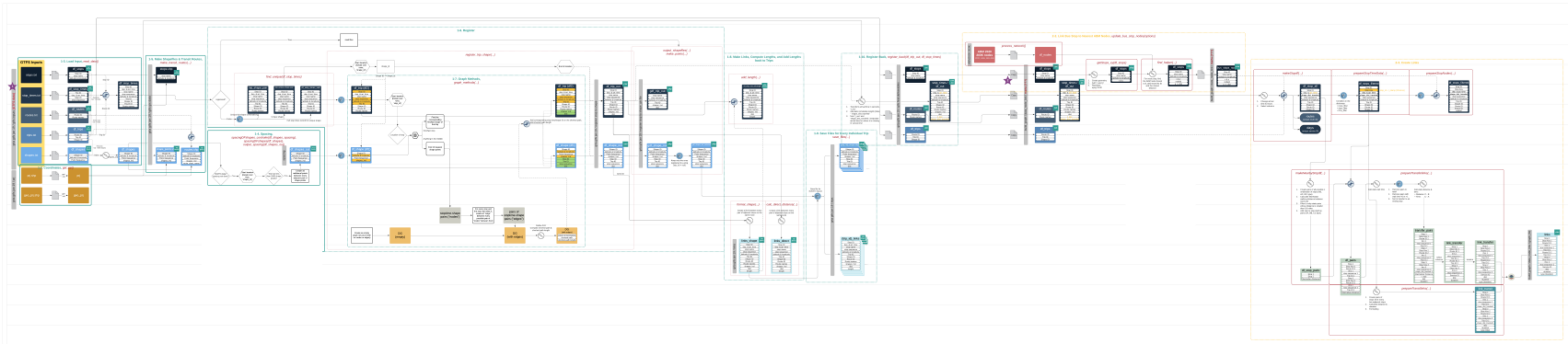
- GTFS Data - Open source available online
- APC Data - Proprietary
- MOVES-Matrix Energy and Emission Rates - Open source data available through NCST at: https://tse.ce.gatech.edu/ncst/movesmatrix

Reuse and Redistribution

There are no restrictions with respect to reuse and redistribution of the Python code, analytical results, and the dataset used to populate the analyses presented in this report. These materials for the Using Multi-Modal Path-Specific Transit Trips in Transportation Social Sustainability Analysis project are available through Zenodo at https://doi.org/10.5281/zenodo.13917429. The GTFS data are public domain. The proprietary MARTA data cannot be distributed by the research team. All MOVES-Matrix algorithms and data are in the public domain.

Appendix A – Process Flowcharts for Previous and New TransitSim Versions

TransitSim 2.0 Previous Version



TransitSim 3.0 New Version