C2SMART
CONNECTED CITIES WITH SMART TRANSPORTATION
A USDOT University Transportation Center

New York University
Rutgers University
University of Washington
The University of Texas at El Paso
City College of New York

Evaluation of Bus Redesign Alternatives in Transit Deserts under Ride-Hail Presence

May 2020

Evaluation of Bus Redesign Alternatives in Transit Deserts under Ride-Hail Presence

Joseph Y.J. Chow
New York University
0000-0002-6471-3419

Ziyi Ma
New York University

Eric Goldwyn
New York University

Mina Lee
New York University
0000-0002-4945-4738

C2SMART Center is a USDOT Tier 1 University Transportation Center taking on some of today's most pressing urban mobility challenges. Some of the areas C2SMART focuses on include:

Urban Mobility and Connected Citizens

Urban Analytics for Smart Cities

Resilient, Smart, & Secure Infrastructure

Disruptive Technologies and their impacts on transportation systems. Our aim is to develop innovative solutions to accelerate technology transfer from the research phase to the real world.

Unconventional Big Data Applications from field tests and non-traditional sensing technologies for decision-makers to address a wide range of urban mobility problems with the best information available.

Impactful Engagement overcoming institutional barriers to innovation to hear and meet the needs of city and state stakeholders, including government agencies, policy makers, the private sector, non-profit organizations, and entrepreneurs.

Forward-thinking Training and Development dedicated to training the workforce of tomorrow to deal with new mobility problems in ways that are not covered in existing transportation curricula.

Led by New York University's Tandon School of Engineering, **C2SMART** is a consortium of leading research universities, including Rutgers University, University of Washington, the University of Texas at El Paso, and The City College of NY.

Visit c2smart.engineering.nyu.edu to learn more

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

This work built off the C2SMART's MATSim-NYC model to evaluate the bus network, so help from the lead modelers Brian Yueshuai He and Jinkai Zhou are much appreciated.

Executive Summary

Despite the emergence of many new mobility options in cities around the world, fixed route transit is still the most efficient means of mass transport. Transit efficiency, however, is well known to be interdependent with ridership demand and network design. As a result of this interdependency, bus operations are subject to vicious and virtuous cycles. Evidence of this can be seen in New York. Since 2007, travel speed reductions and increased congestion due to more mobility options competing for road space have led to a vicious cycle of ridership reduction and further increased congestion as former transit passengers take to other less efficient modes. In Brooklyn, bus ridership has declined by 21% during this period. The effects of this vicious cycle are particularly significant for captive riders and the vulnerable population that relies on efficient bus transport in areas where subway options are less available such as the transit desert areas of Brooklyn. Intervention is required to promote a virtuous cycle and make the bus more competitive, especially in the face of increased competition from ride-hail services. This can be done by redesigning the bus network in a way that does not increase operating cost while reducing user costs and increasing accessibility for more riders.

The research challenge lies in having an easy-to-use methodology to evaluate and compare two or more transit network designs in which one or more designs may be incomplete; i.e. only includes route alignments without either stop locations and/or service frequencies. This study addresses these gaps by presenting a systematic process that combines techniques from both analytical and simulation-based tools:

1. Given a route data set, use **analytical route-level modeling** to identify stop locations and/or frequency to minimize cost (both user and operator)
2. Create a **GTFS schedule** from the output of the analytical model
3. Use a **multi-agent simulation model** to derive the equilibrium for the network design

There is no commercial tool that combines all three of these methods together. The objective of the project was to put these steps together and show how it can provide insightful decision support to transit planners that these tools alone do not provide.

Drawing from lessons learned in the literature and the international community along with surveying 373 bus operators in Brooklyn, Dr. Goldwyn and Levy at the Marron Institute drafted a redesigned bus network for just this purpose. The route plan includes stop locations and frequencies. How does it compare to the existing system? Can those frequencies be improved upon? The study makes the following comparison between scenarios:

1. Existing Brooklyn bus network
2. Marron's proposed bus network redesign with their specified frequencies
3. Marron's proposed bus network redesign, with an analytical model used to optimize frequencies that includes demand feedback from a simulation model

For Scenarios 2 and 3, there is no GTFS data, so a GTFS schedule needs to be created for each. For Scenario 3, a simulation-based frequency optimization model is fitted to design frequencies, using the MATSim-NYC model developed by Chow et al. (2020a). These are then used to compare against Scenario 2. MATSim simulated the complete 24-hour day with departure times dynamically generated considering spillbacks and congestion on the road. A day-to-day adjustment component captures the demand response to the congestion. As a result, a simulation-based frequency setting approach accounts for demand response to the dynamic traffic propagation and selection of bus as a mode relative to a number of other modes: car, walking, bike, Citi Bike, taxi, and ride-hail services (see Chow et al., 2020a).

A methodology is designed for the pipeline of steps to take an initial route network shapefile with frequencies into a GTFS output, incorporation into MATSim-NYC, and output for comparison. The MATSim output with the existing model is compared in terms of route ridership to MTA data, indicating a ridership-weighted average difference in route ridership of 21%, which is an acceptable level.

The iterative algorithm updates the MATSim demand and the frequencies from the analytical model (see Tirachini, 2014). The numerical tests with the algorithm for Scenario 3 indicate that it can converge to a stable equilibrium, although a global optimum is not guaranteed. The model output can show bus boardings and alightings as well as load profiles for a route. Time of day distributions of the bus trips for the entire network can be compared.

Based on the methods in this study, we confirm that the proposed network redesign from Goldwyn and Levy (2020) should indeed increase ridership by their predicted 20% (we predict 23%), at a reduction of operating cost of 6%. However, by further adjusting frequencies throughout the day corresponding to the simulation, we are able to further reduce operating cost down to 25% reduction from existing condition while maintaining a ridership increase of 20%. The increased ridership draws primarily from passenger car use (nearly 75%), with a small 2.5% drawn from ride-hail services and another 5% from taxis. This suggests the redesigns should be effective in moving people away from less efficient transportation modes. The network redesign with our proposed frequency can essentially double the improvement in the farebox recovery ratio from Goldwyn and Levy's design.

Table of Contents

List of Figures

List of Tables

Section 1: Introduction

Subsection 1.1 Motivation for transit evaluation

Despite the emergence of many new mobility options in cities around the world, fixed route transit is still the most efficient means of mass transport (Walker, 2018). This is evident to the mobility providers as well as companies like Uber who experiment with "cheaper fares in exchange for more walking" (Hawkins, 2018). Transit efficiency, however, is well known to be interdependent with ridership demand and network design (Lampkin and Saalmans, 1967; Mohring, 1972; Newell, 1979; Ceder and Wilson, 1986; Desaulniers and Hickman, 2007; Daganzo, 2010; Tirachini et al., 2013).

As a result of this interdependency, bus operations are subject to vicious and virtuous cycles (Bar-Yosef et al., 2013). Evidence of this can be seen in New York. Since 2007, travel speed reductions and increased congestion due to more mobility options competing for road space have led to a vicious cycle of ridership reduction and further increased congestion as former transit passengers take to other less efficient modes. In Brooklyn, bus ridership has declined by 21% during this period. While the decrease in ridership has been steady throughout this period, there is an emerging concern that it will only get worse as for-hire vehicle services like Uber and Lyft add more trips to the road network (Sisson, 2018). According to a study from UC Davis, 49-61% of ride-hailing trips are substituting for walking, biking, or transit trips (Clewlow and Mishra, 2017).

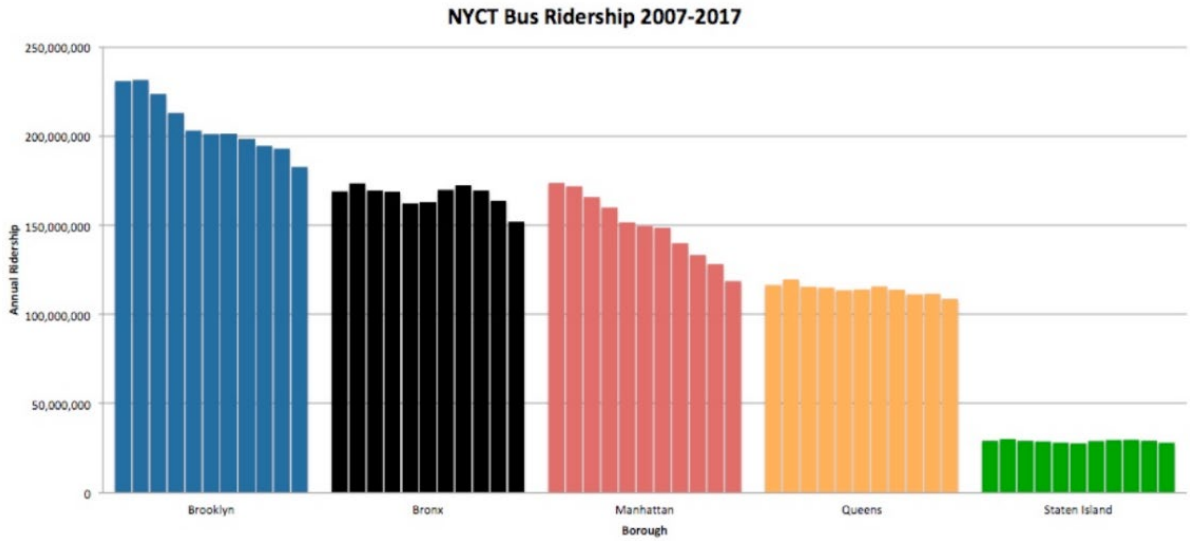


Figure 1: Reduction in bus ridership in NYC between 2007 and 2017 (source: MTA, 2020a).

The effects of this vicious cycle are particularly significant for captive riders and the vulnerable population that relies on efficient bus transport in areas where subway options are less available such as the transit desert (Allen, 2017; Jiao, 2017) areas of Brooklyn as highlighted in Figure 2.

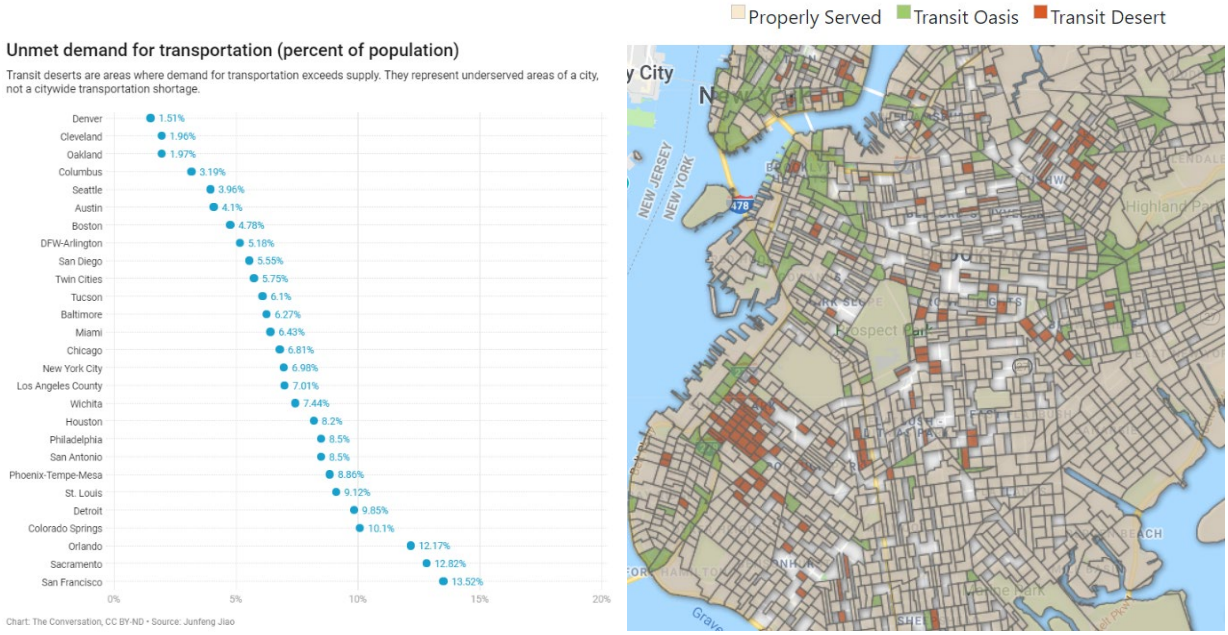


Figure 2: (left) Ranking of cities with transit deserts (source: Jiao and Bischak, 2018); (right) geographical distribution of transit deserts in NYC (source: http://www.transitdeserts.org/).

Planning a transit network requires taking all stakeholders into consideration. On one side, operators want to receive enough revenue to maintain the transit operation. However, like many public transit agencies, NY MTA relies heavily on dedicated taxes and state and local subsidies. According to MTA 2018 Adopted Budget February Financial Plan (MTA, 2018), it has an agency average farebox recovery ratio of 35.5%. Due to the increasing operation cost, MTA estimates in that plan to having an even lower farebox recovery ratio of 32.4% in 2021. Unfortunately, the agency also faces a higher operating deficit in the next few years due to growing pension and healthcare liabilities, as shown in Figure 3.

Figure 3: MTA Projected Financial Plan (MTA, 2020a).

On the other side, transit users want a highly available, reliable, comfortable, and safe transit service. According to a customer satisfaction survey done by MTA in 2010-2012 (NYCT, 2012), only 59% of the bus users were satisfied with the frequency of service, and only 62% of the bus users were satisfied with the overall value for the money of the local bus in 2010.

It is important to improve the existing bus service in order to address the issues mentioned above. Intervention is required to promote a virtuous cycle and make the bus more competitive, especially in the face of increased competition from ride-hail services (Warerkar, 2017). This can be done by redesigning the bus network in a way that does not increase operating costs while reducing user costs and increasing accessibility for more riders.

The research challenge lies in having an easy-to-use methodology to evaluate and compare two or more transit network designs in which one or more designs may be incomplete; i.e. only includes route alignments without either stop locations and/or service frequencies. Existing transportation planning tools (e.g. TransCAD, MATSim, EMME, Cube) require schedule information. This study addresses these gaps by presenting a systematic process that combines techniques from both analytical and simulation-based tools:

1. Given a route data set, use **analytical route-level modeling** to identify stop locations and/or frequency to minimize cost (both user and operator)
2. Create a **GTFS schedule** from the output of the analytical model
3. Use a **multi-agent simulation model** to derive the equilibrium for the network design

There is no commercial tool that combines all three of these methods together. The objective of the project was to put these steps together and show how it can provide insightful decision support to transit planners that these tools alone do not provide.

Subsection 1.2 Case study background: Brooklyn bus network redesign

Drawing from lessons learned in the literature and the international community along with surveying 373 bus operators in Brooklyn, Dr. Goldwyn and Levy at the Marron Institute drafted a redesigned bus network for just this purpose, as shown in Figure 4, and presented in CityLab and New York Magazine (Levy and Goldwyn, 2018; Goldwyn and Levy, 2018). The redesign features increased stop spacing along with other technological improvements like all-door boarding and transit signal priority.

The route plan includes stop locations and frequencies. How does it compare to the existing system? Can those frequencies be improved upon? The study makes the following comparison between scenarios:

1. Existing Brooklyn bus network (with volumes calibrated to average ridership levels provided by MTA)
2. Marron's proposed bus network redesign with their specified frequencies
3. Marron's proposed bus network redesign, with an analytical model used to optimize frequencies that includes demand feedback from a simulation model

For Scenarios 2 and 3, there is no GTFS data, so a GTFS schedule needs to be created for each. For Scenario 3, a state-of-the-art analytical model is fitted to design frequencies. These are then used to compare against Scenario 2.

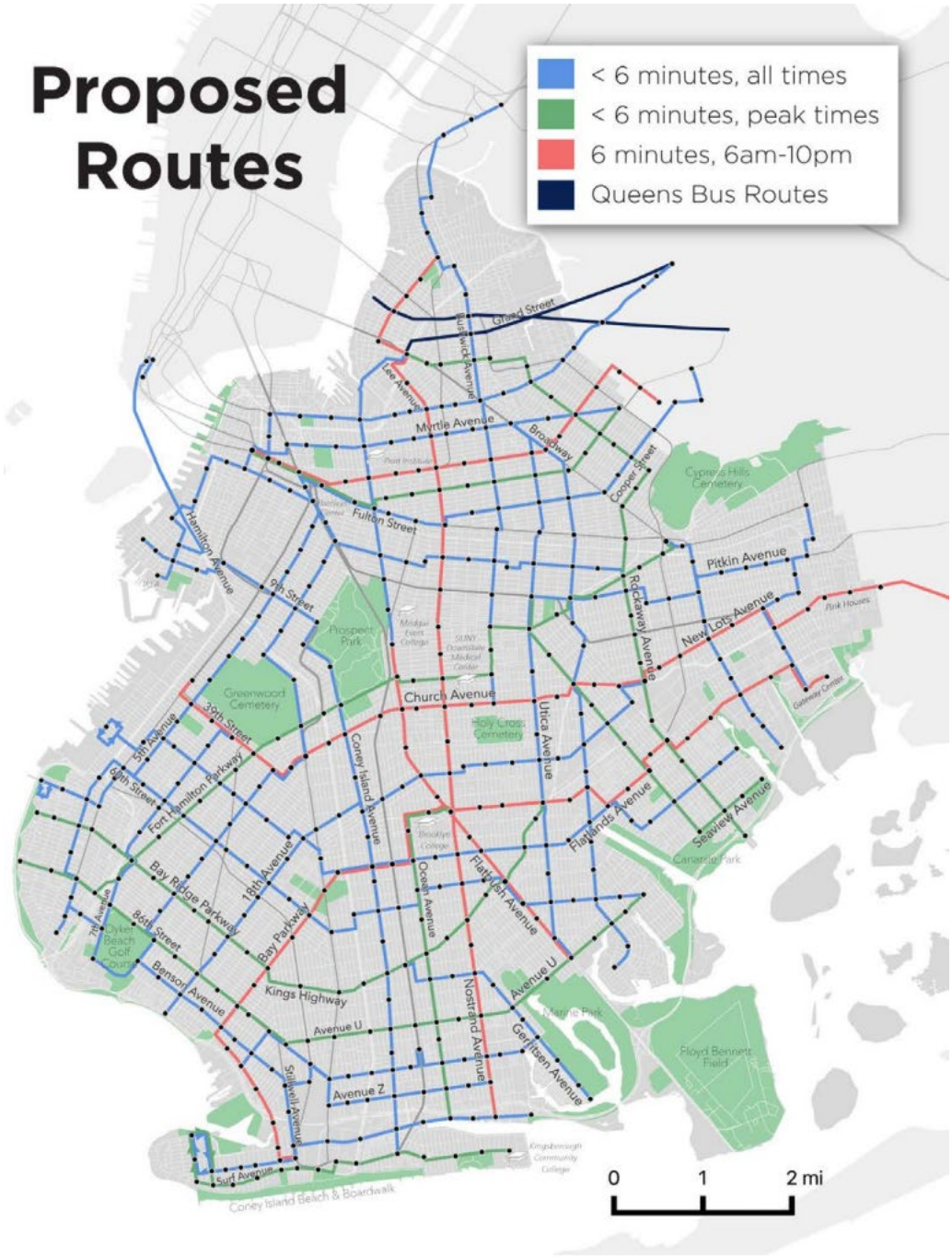


Figure 4: Brooklyn Bus Network Redesign Plan by Levy and Goldwyn (2020).

To evaluate how well any of the scenarios work, the performance measures of these designs are obtained using a multi-agent simulation, a citywide virtual test bed developed by C2SMART

to evaluate emerging transportation technologies and policies (He et al., 2020a,b). The simulation outputs the equilibrium demand for each scenario based on dynamic adjustments in a day-to-day process setting (see Djavadian and Chow, 2017a,b). In addition, the calibrated model also incorporates ride-hailing modes so that the scenario outputs take that into account when evaluating the performance. The output of this scenario is a recommended frequency of each route based on the simulated demand.

Furthermore, the analytical model in Scenario 3 assumes passenger demand is fixed. When designing the frequencies, the performance would inevitably impact ridership demand as that affects the level of service. The frequency setting problem is therefore a bilevel problem (Constantin and Florian, 1995; Szeto and Jiang, 2014; Verbas and Mahmassani, 2015). We propose an algorithm to find an equilibrium set of frequencies for the bus network that uses the MATSim model as the lower-level user equilibrium.

This report is structured as follows: Section 2 reviews previous research on bus network planning, and especially focuses on bus frequency setting (analytical modeling and bilevel modeling) and MATSim as a simulation platform; Section 3 introduces our NYC Virtual Testbed; Section 4 proposes the bilevel optimization algorithm to determine the optimal frequencies for Scenario 3; Section 5 introduces the overview of simulated scenarios will different configurations; Section 6 analyzes the result of different scenarios in terms of operation cost, users cost, total ridership, bus stop load, and vehicle load; Section 7 and 8 conclude and discuss the possible improvement of the project and potential future development of our transit simulation tool.

Section 2: Literature Review

Subsection 2.1 Transit Planning and Frequency setting

There are four main activities in public transportation planning: network route design, timetable development, vehicle scheduling, and crew scheduling. The activity diagram demonstrating these activities is shown in Figure 5. This project connects activity 1 and activity 2. An overview of public transit operations planning is provided in Chow et al. (2020b).

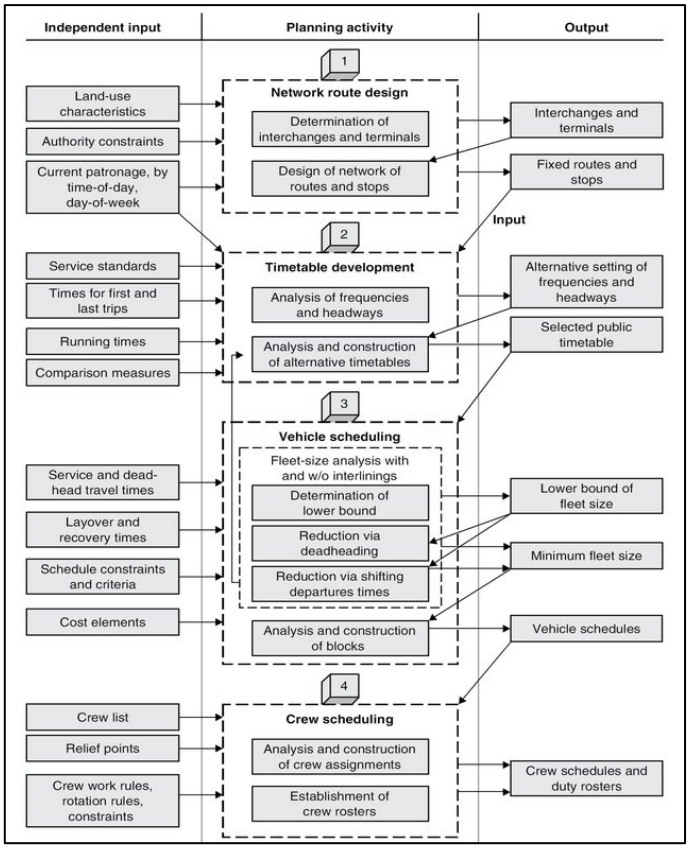


Figure 5: Typical Transit Planning Activity (Source: Ceder, 2016).

Strategic planning involves planning routes and frequencies, often called the line planning problem. Hasselström (1982) and van Nes et al. (1988) proposed early line planning optimization models for setting routes and frequencies jointly. Reviews of transit network design models and algorithms can be found in Desaulniers and Hickman (2007), Guihaire and Hao (2008), and more broadly in Farahani et al. (2013). Line planning has been shown to be NP-Hard in complexity (see Schöbel and Scholl, 2006) leading to the use of route construction heuristics like Ceder and Wilson (1986).

As such, line planning in practice may involve using permutations of simple structures. Fielbaum et al. (2017) explicitly tackle the problem of defining any city transit network using a parameterized network design structure. Fielbaum et al. (2016) used their network description to evaluate four different line structures: direct lines, exclusive lines, hub-and-spoke, and feeder-trunk. These are shown in Figure 6.

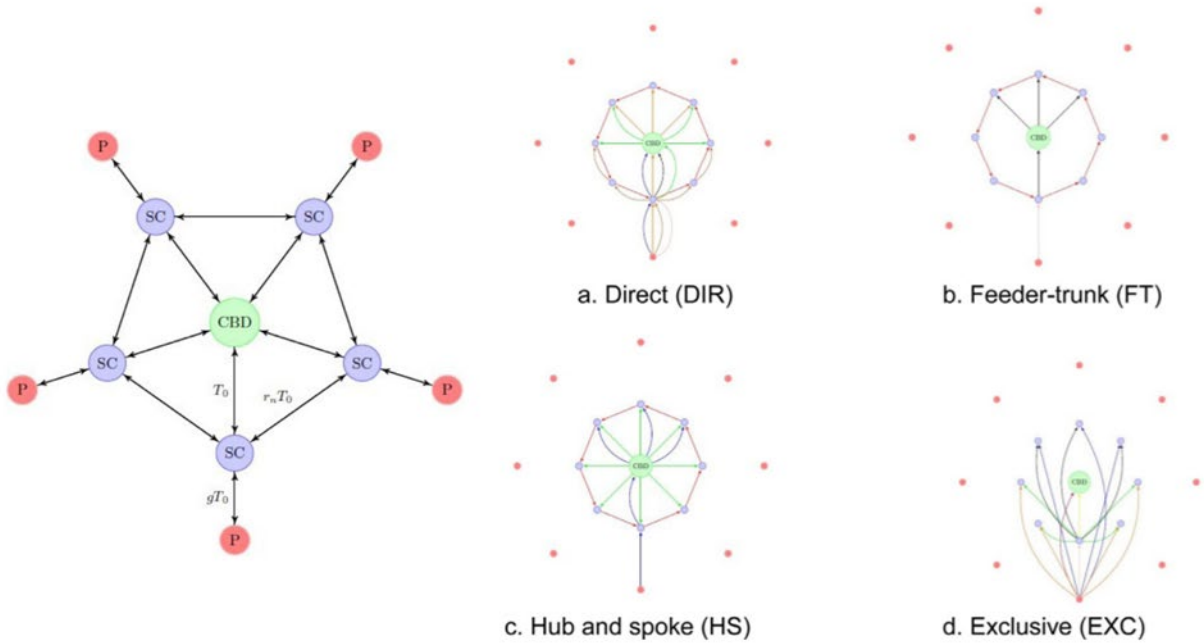


Figure 6. General network design that can be parameterized into different structures: (a) direct, (b) feeder-trunk, (c) hub and spoke, and (d) exclusive (source: Fielbaum et al., 2016).

For those interested in optimizing more custom designs, Byrne (1975) developed a continuous approximation model to optimize transportation line locations and headways for a region with uniform population density and demand. Newell (1979) also developed a model of this type to compare two bus network designs over a square street grid. Newell (1971) developed a model to set the service rate on a single route with a time-varying level of demand by minimizing the sum of user cost of delay and operator cost. His finding that the optimal frequency is proportional to the square root of the arrival rate of passenger is sometimes referred to as the square root rule. Mohring (1972) independently showed a similar square root rule using a simpler construct of the cost function. The objective in that study was to show that with user cost

considered, transit frequency for a given route exhibits economies of density, which serves as evidence for public subsidies of transit service.

Analytical models of this square-root form are still of use to contemporary researchers. Tirachini (2014) provides a simple analytical approach by considering operational cost, user waiting cost, user access and egress cost, and in-vehicle time cost. Though he mainly considers the optimal stop spacing in this paper, if the number of stops has been pre-determined, then the optimal frequency of a route can be calculated. The total cost function is shown in Eq. (1) and Eq. (2).

$$C_t = c_{operation} + c_{user\_access} + c_{user\_egress} + c_{user\_wait} + c_{user\_in-vehicle} \tag{1}$$

$$C_t = c \cdot f \cdot \left(\frac{L}{v_0} + \beta\frac{N}{f} + St_s\right) + P_a\frac{L}{2v_wS}N + P_w\frac{1}{2f}N + P_v\frac{l}{L}\left(\frac{L}{v_0} + \beta\frac{N}{f} + St_s\right)N \tag{2}$$

where c (\$/bus-h) is a unit bus operating cost, v_0 (mph) is bus operating speed, f (bus/h) is bus frequency, $t_c = \frac{L}{v_0} + \frac{\beta N}{f} + St_s$ (h) is the bus cycle time, β (sec/pax) is average boarding and alighting time per passenger, N (pax/h) is passenger demand, S is number of stops, t_s (h) is stopping delay, P_a (\$/h) is the value of access time, v_w (mph) is the walking speed, P_w is value of waiting time, P_v is value of in-vehicle time, and l is average travel distance (mi) per passenger. By jointly solving for (S^*, f^*) to minimize total cost, one can find the stop spacing and frequency to serve a route. Since solving them jointly is nonlinear, we discretized values of f to the nearest 0.1 increments, and optimized S^*. The optimum number of stops for a given frequency is obtained by taking the derivative of the cost with respect to number of stops. The value of f^* is found from the lowest total cost across all values of f. This relationship is visualized in Figure 7, where it shows the relationship between total cost, frequency, and number of fixed stops.

When demand is fixed, the optimal frequency is obtained by taking the derivative of Eq. (2) with respect to the frequency, resulting in Eq. (3). Since the bus stop delay is not modelled in in MATSim, t_s is dropped out in this study to yield Eq. (4).

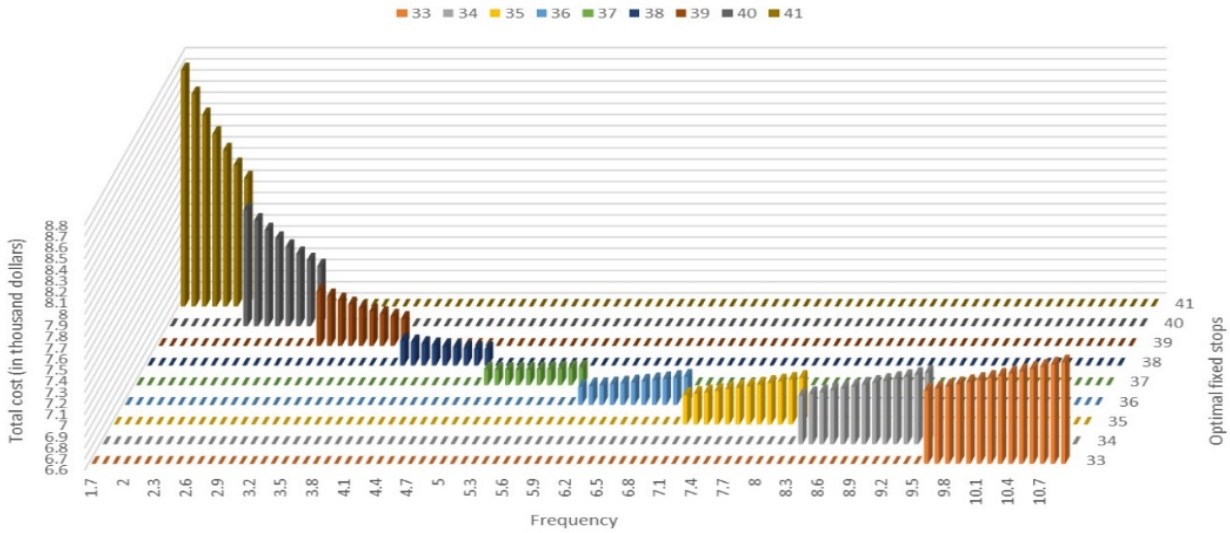


Figure 7: Relationship between cost, frequency, and number of fix stops (source: Chow et al., 2020b).

$$f^* = \sqrt{\frac{N(P_w L + 2P_v \beta l N)}{2cL(\frac{L}{v_0} + St_s)}} \tag{3}$$

$$f^* = \sqrt{\frac{N(P_w L + 2P_v \beta l N)}{2cL(\frac{L}{v_0})}} \tag{4}$$

When demand is elastic and subject to congestion, then the frequency setting problem in Eq. (3) – (4) need to include an equilibrium constraint. The resulting model is called a bilevel problem.

Subsection 2.2 Bilevel transit frequency setting algorithms

Bilevel network design problems refer to an optimization problem divided into two levels (see Chow, 2018). At the upper level a decision-maker determines design variables for a system in anticipation of the reaction of the users. At the lower level users have an objective that is

separate from the upper level objective which decides the decision variables used for the upper level problem. An optimal solution to this problem is considered a Stackelberg equilibrium (see Marcotte, 1986; Yang and Bell, 1998).

The model is known to be NP-hard (Bard, 1991) and non-convex (Bard and Moore, 1990). As a result, the problem requires heuristics to obtain satisficing solutions for in practice.

Bilevel problems have been applied to the frequency setting problem where users adjust their route choices as a user equilibrium constraint. In the case of static user equilibrium, examples include Constantin and Florian (1995), Gallo et al. (2011), Szeto and Jiang (2014), and Canca et al. (2016). In the case where the equilibrium constraint is based on dynamic assignment, then simulation-based methods are needed (e.g. Zhang et al., 2013, Verbas and Mahmassani, 2015).

In our study, we will use the cost function shown in Tirachini (2014) as an upper level objective where the user demand is determined from lower level dynamic equilibrium constraints (across multiple modes under a time-of-day activity scheduling process) based on MATSim.

Subsection 2.3 Multi-agent simulation

Agent-Based Modeling and Simulation (ABMS) (von Neumann, 1966; Bonabeau, 2002) can be used to model complex heterogeneous agents with interaction rules and agent learning. There are several well-known ABMS platforms designed to support decision-making, including but not limited to Transportation Analysis and Simulation System (TRANSIMS) (Nagel et al., 1999, Multi-Agent Transport Simulation Toolkit (MATSim) (Balmer et al., 2009), Sacramento Activity-Based Travel Demand Simulation Model (SACSIM) (Bradley et al., 2012) Simulator of Activities, Greenhouse Emissions, Networks, and Travel (SimAGENT) (Goulias et al., 2011), Polaris (Auld et al., 2016), SimMobility (e.g. Nahmias-Biran et al., 2019), etc. TRANSIMS was a first-generation tool developed by the Federal Highway Administration (FHWA), after which the creators took the lessons learned from it to produce the next generation tool MATSim.

MATSim is an open-source simulation toolkit implemented in Java. It has three desirable features that make it unique among other agent-based simulations. The first is the use of a synthetic population that includes activity schedules so that simulation incorporates activity

scheduling behavior. The role of MATSim as a simulation of activity scheduling is discussed at great length in Chapter 4 of Chow (2018). MATSim provides a feedback loop by using a day-to-day adjustment process, although the adjustment process is simplified with a heuristic (a genetic algorithm) and the use of only a single population.

The second desirable feature is that MATSim can simulate large-scale scenarios using a spatial queue model (Cetin et al., 2003) to simulate the traffic dynamics instead of car-following and lane-changing models (Zheng et al., 2013). To shorten the computation time, MATSim also adopts parallel computation for the spatial queue model.

Another advantage of MATSim is its numerous extensions as an open-source platform, which makes it easier for users to simulate and evaluate different scenarios. There are many applications of MATSim around the world, including Berlin (Neumann, 2016; Ziemke, 2016), Zurich (Rieser-Schüssler et al., 2016), Singapore (Erath and Chakirov, 2016) among others. These applications prove that MATSim is suitable for analyzing the complex urban transportation system in large cities. MATSim has also been used to evaluate several emerging technologies, including the following examples:

- Autonomous vehicle fleet (Hörl et al., 2019)
- Carshare (Ciari et al., 2016)
- Urban air mobility (Rothfield et al., 2018)
- Demand-responsive transit (Cich et al., 2017)
- MaaS (Becker et al., 2020)

As an agent-based simulation, MATSim can capture the behavior of each agent and the interaction between agents and transportation system. Each agent refers to an individual traveler. Traveler behavior is represented by a series of activities, travel modes and routes. MATSim uses an iterative framework for simulation, as shown in Figure 8. The goal of the iterative framework is to find the equilibrated state of the system. At each iteration of the MATSim loop, agents' activities of a typical day are simulated at each iteration. The activities are scored based on vary aspects based on the user-defined scoring function. A simulation will reach equilibrium as each agent optimize his score with the co-evolutionary algorithm. The overall simulation procedures are:

- Put the agents with the initial travel plans into MATSim and simulate their mobility in the physical system.
- Calculate the score (utility) of each agent's executed plan.

- Randomly select a certain proportion of agents and mutate their plans. Go back and re-run the simulation until the agents' scores converge.

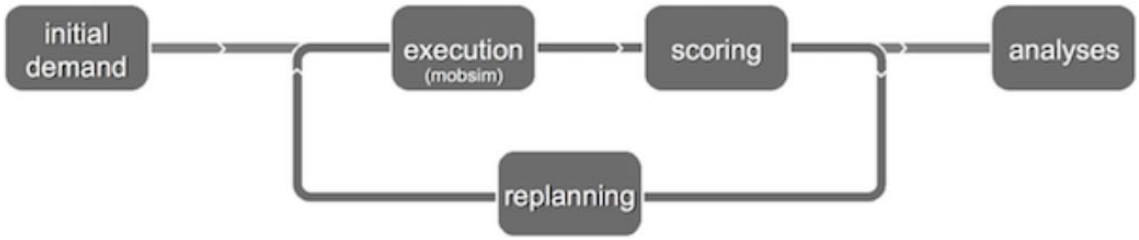


Figure 8: A MATSim Loop.

MATSim can address the following modeling needs:

- It needs to recognize dynamic traffic propagation to capture traffic technologies and policies like congestion pricing;
- It needs to recognize activity scheduling behavior of travelers (see Kang et al., 2013);
- It needs to recognize different segments of travelers in the population, e.g. low and high income, age groups, residents of different socioeconomic backgrounds;
- It needs to be flexible enough to adapt to new emerging technologies.

Based on these requirements, we chose to develop the initial test bed for NYC using MATSim, a Multi-Agent Transportation Simulation. MATSim models transportation networks using a mesoscopic simulation based on cellular automata. It is open source and many extensions have been quickly developed for it to handle a wide assortment of policy needs: autonomous vehicles, emissions modeling, parking, freight, electric vehicles, bikeshare, etc. MATSim makes use of a synthetic population which is useful for modeling heterogeneous population segments. It incorporates a day-to-day adjustment process that can reflect learning from the population (see Djavadian and Chow, 2017a,b) to achieve a social equilibrium under the technology scenario.

Section 3: Simulartion Framework

The C2SMART NYC Virtual Testbed is based on MATSim, called MATSim-NYC (Chow eta l., 2020). It was designed to help evaluate new policies and emerging technologies at a city level. It features a synthetic population representing eight-million NYC residents and a calibrated time-

variant multi-model network, both in a base year of 2016. The simulation can capture the interaction between each agent and the transportation system in both aggregated and disaggregated levels. There are three inputs: population, road network, and transit schedule.

Subsection 3.1 Road Network

Complex road networks like the one in New York City are impossible to be built on MATSim by hand. Our road network is generated from OpenStreetMap and converted it to MATSim XML format with JSOM. The network is visualized in Figure 9 on Simunto Via. To better reflect the real traffic work, link free speed and capacity are calibrated using 2016 bridges/tunnels volume and INRIX speed data.

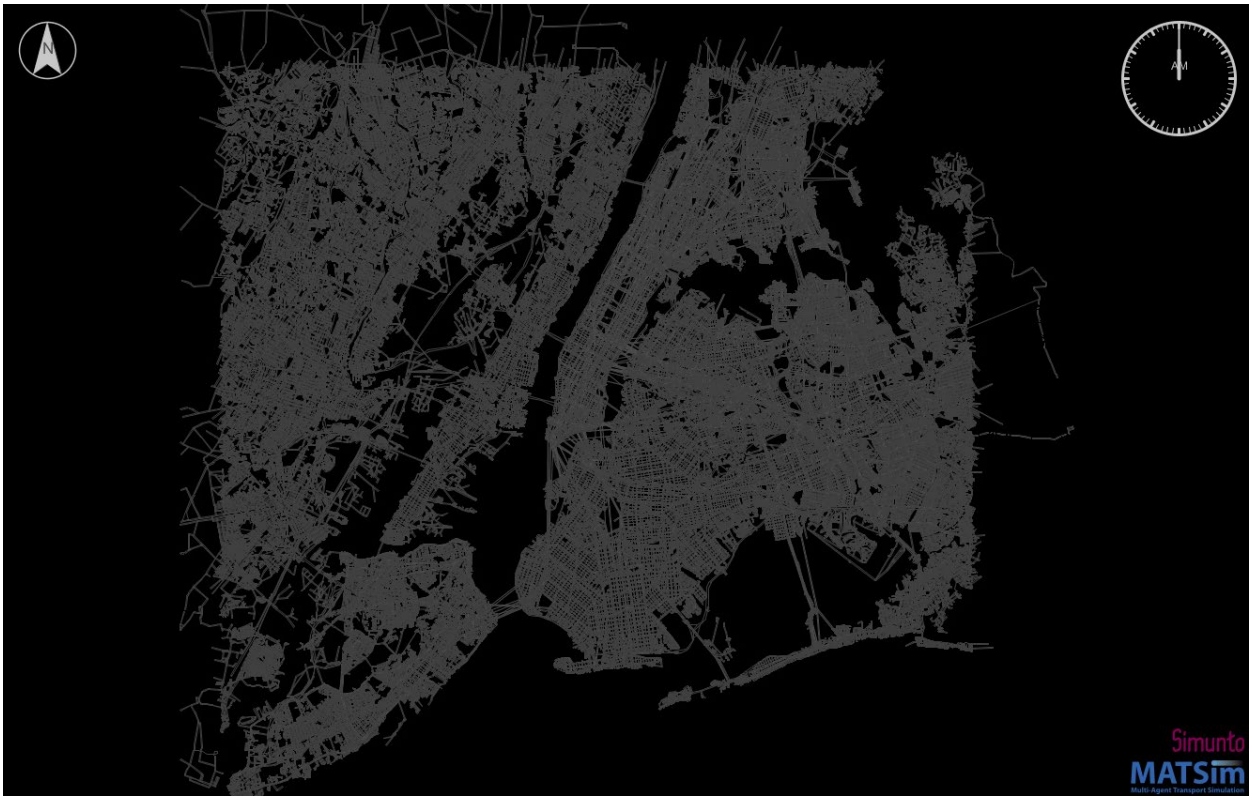


Figure 9: MATSim Road Network.

Subsection 3.2 Transit Schedule

The existing transit schedule is converted from GTFS data using pt2matsim, a MATSim's extension. In order to keep transit running based on historical record, subways and buses run on dedicated links, and they are not subject to road congestion. Figure 10 shows the existing location of the transit stops.

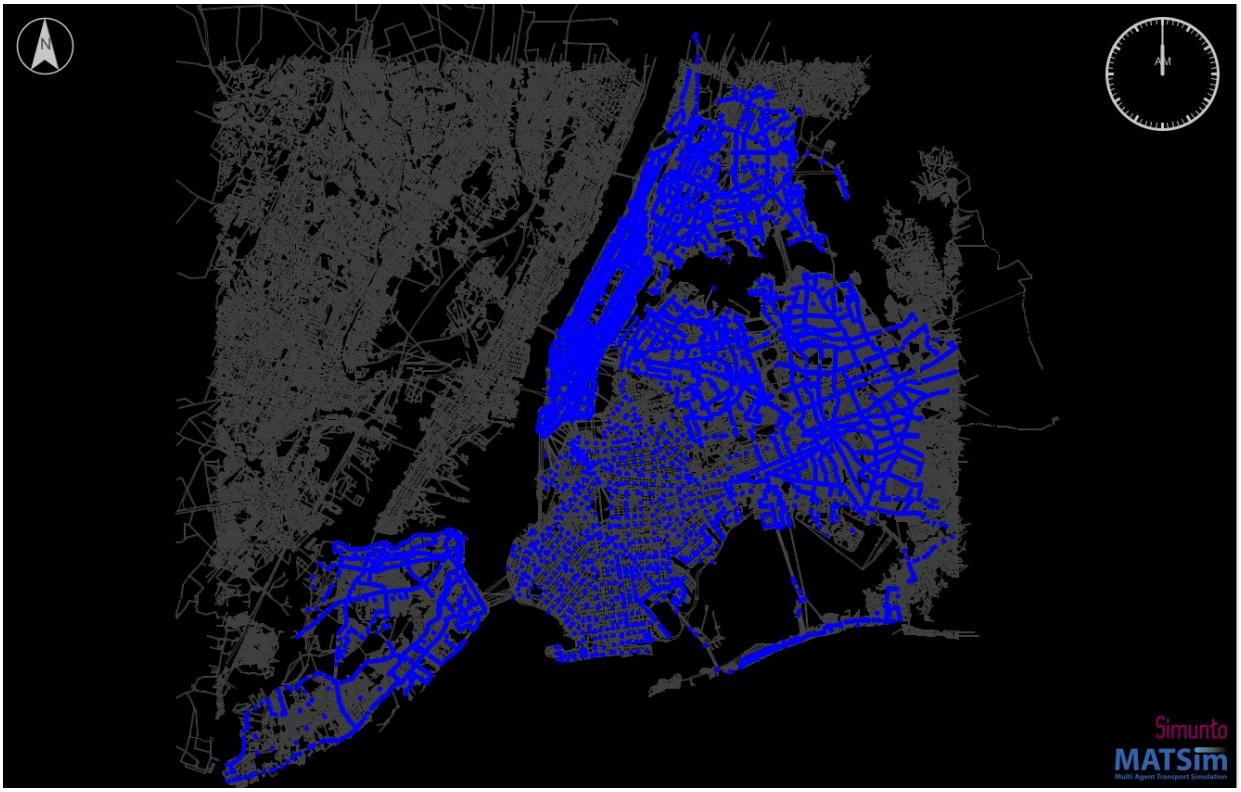


Figure 10: Existing Transit Stops in MATSim.

Subsection 3.3 Population

The population is synthesized using PopGen 2.0 (MARG, 2016), which is based on an enhanced iterative proportional updating algorithm. In total, a population of 8.38 million agents are generated and validated with LEHD data. Each agent receives an agenda, which consists of a series of activities of a regular business day. Their agenda are based on the 2010/2011 Regional Household Travel Survey (RHTS). More than 30 million trips are assigned, and each agent receives 3.5 trips on average. The mode of each trip is also updated using a nested logit model, which

considers agent's income, age, work status, travel time, and travel cost. Each agent is represented as one small white dot in Figure 11.

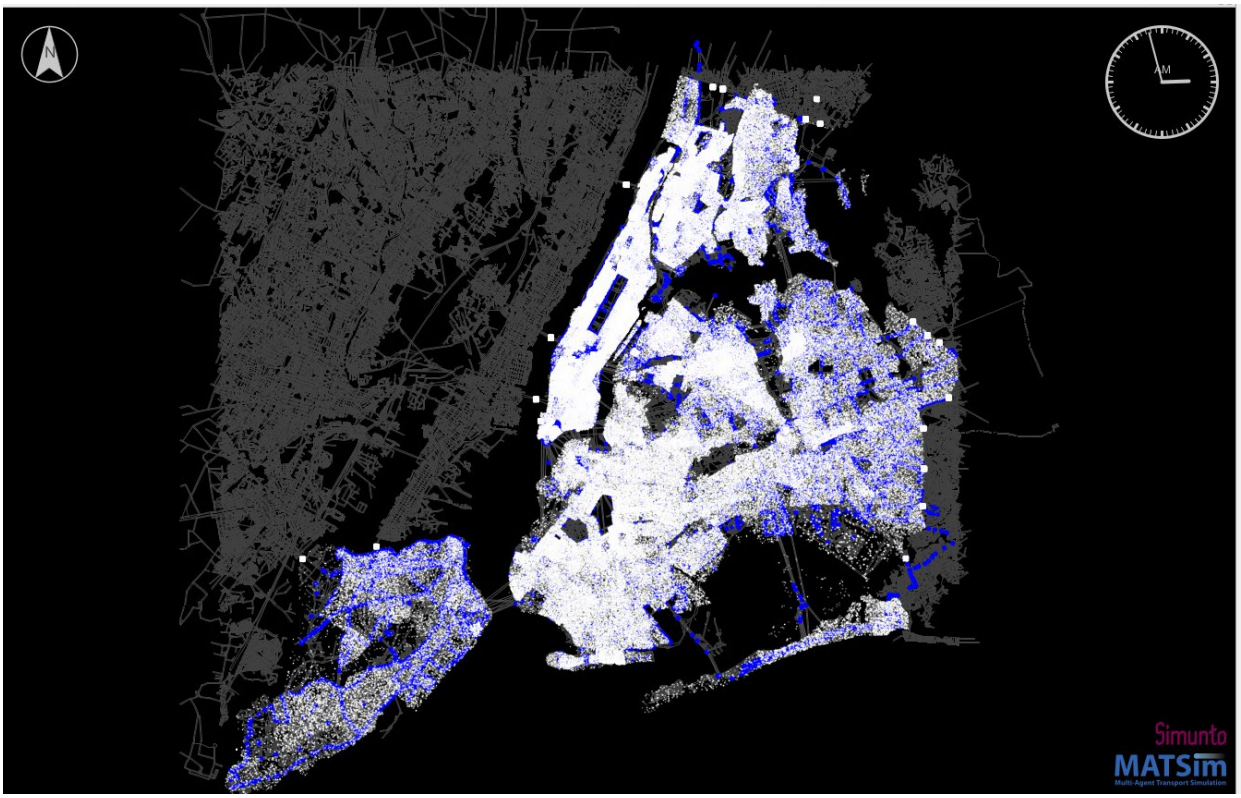


Figure 11: Agents in MATSim.

Subsection 3.4 Simulation summary

For the bus network scenarios, the evaluation is based on the whole city-wide simulation. The advantage of this approach is that trips for passengers going to different parts of NYC are all simulated instead of breaking the network into subnetworks. Details of the synthetic population and the whole network are provided in He et al. (2020a,b) and Chow et al. (2020a).

When evaluating the new scenarios, the Brooklyn bus network shown in Figure 12 is cut out and replaced with the proposed networks.

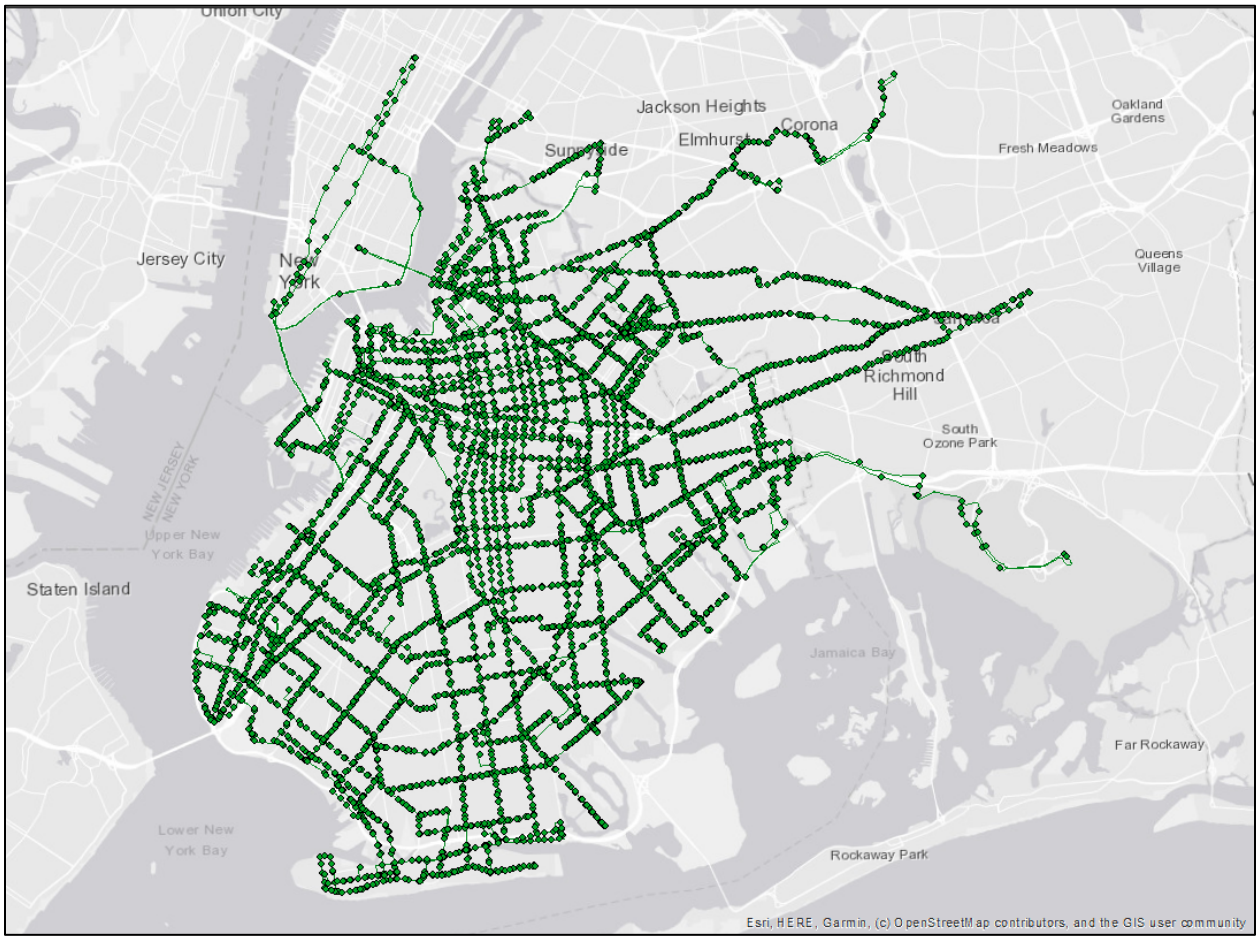


Figure 12: Existing Brooklyn bus network.

MATSim-NYC includes the following modes: driving, carpool, walking, biking, Citi Bike, public transit (subway and bus), taxi, and TNC (ride-hail). As a result, the simulation output considers the presence of ride-hail services as a competing mode in addition to other modes.

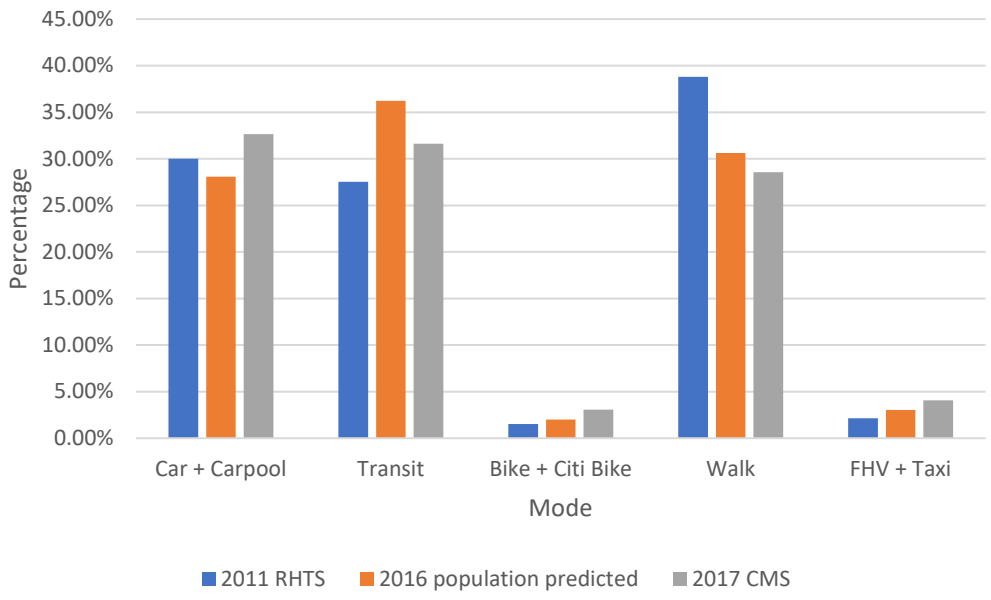


Figure 13: Predicted mode share from MATSim-NYC compared to NYC Citywide Mobility Survey (source: Chow et al., 2020a).

Section 4: Implementation and algorithm design

Subsection 4.1 Implementation Framework

Per our discussion in section 3, the NYC Virtual Test bed already contains the population and a calibrated road network. We only need to prepare the transit schedule, which is accomplished in 3 steps. First, convert the existing transit service from GTFS data obtainable on MTA Website to MATSim transit schedule XML format. MATSim's pt2matsim module is used to accomplished it, and this module also supports converting HAFAS and OpenStreetMap networks. Second, the existing Brooklyn bus network is replaced with the proposed network. Third, the frequency of the proposed network is inserted to transit schedule XML file.

The implementation framework is shown in Figure 14. On the first branch, the user can specify the frequency or headway of each route and run the simulation. If the user wants some optimization on the frequency, he or she can use Eq. (4) shown in section 2.1 to update the frequency based on the simulated demand. On the second branch, the operation frequency is not specified. A default initial frequency will be provided: 2 trips per hour on non-peak hours,

and 6 trips per hour on peak hours. The initial frequency is then simulated on MATSim and updated based on the simulation output. The frequency update loop stops when the demand changes is minimal. Minimal policy frequency is also considered to provide coverage to low demand hours. In NYC, the minimal frequency is set to be 2 trips per hours.

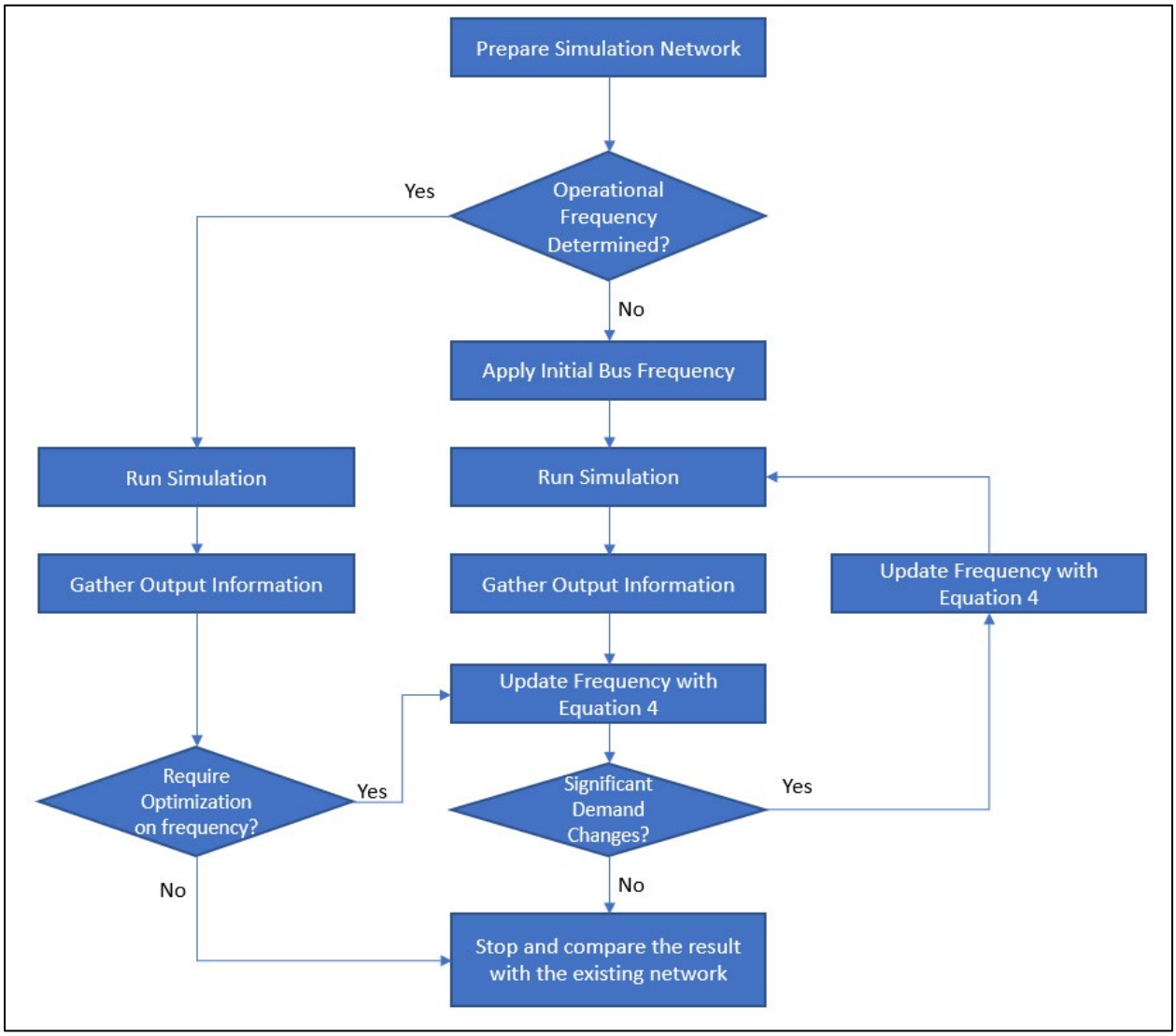


Figure 14: Simulation Flow Chart.

Subsection 4.2: Algorithm design

The frequency updating algorithm is an iterative optimization algorithm.

First, an initial guess is generated to test the ridership. The initial guess consists of 2 trips per hour per direction per route from 10PM-4AM and 11AM-3PM, and 6 trips per hours for the remaining hours. Frequency information is stored in "TransitSchedule.xml". Function CallMATSim passes this file out to the Virtual Test Bed to start the simulation. By default, the transit network is simulated in MATSim for 100 iterations using 4% of the NYC population at each round.

Second, the simulated ridership is read from the MATSim output "ExperiencedPlan.xml" using the ReadMATSimOutput function. "ExperiencedPlan.xml" contains a detailed travel dairy of all agents throughout the day. For each leg of the travel dairy, it contains the selected mode, departure time, travel time, start link, end link, travel time, travel distance, vehicle ID, and traversed links. If a trip is made using public transit, it also contains the vehicle trip name, boarding and exiting stop links, for example:

<route type="experimentalPt1" start_link="pt_10000145" end_link="pt_10000384" trav_time="00:05:50" distance="1689.0534736915677">PT1===10000145.link:pt_10000145===NYU25===NYU_EV-Everyday_19200_NYU25_1===10000384.link:pt_10000384</route>

The ReadMATSimOutput function only targets transit trips. It extracts the agent ID, agent departure time, route chosen, traveled distance, departure stop ID, arrival stop ID, and travel direction. They are further summarized to average travel distance(l), and number of riders(N) for each route for each direction for 24 hours.

Third, based on the simulated ridership, the optimal frequency is calculated using the CalculateOptimalFrequency function. CalculateOptimalFrequency is an implementation of Eq. (4). Its inputs consist of the route usage metrics passed from ReadMATSimOutput and five constant parameters. The parameters are vehicle operation cost(c), vehicle travel speed(v_0), passenger boarding time(β), value of wait time(P_w), value of in vehicle travel time(P_v). The output of this function is an array of frequencies for each route for each direction for 24 hours.

Finally, "TransitSchedule.xml" is updated using function UpdateXML. This function first removes all vehicle trips of the targeted routes, then inserts new trips according to the new frequency. The output of this function is an updated "TransitSchedule.xml", which is ready for another iteration of simulation.

The flow of the iterative optimization algorithm is shown in Figure 15, and the pseudocode is shown in the Appendix. After the simulation using the initial guess, ReadMATSimOutput is used to summarize the simulated result. CalculateOptimalFrequency is used to calculate the new frequency. Finally, UpdateXML is used to update the "TransitSchedule.xml" file for another round of simulation. The algorithm can be set to a tolerance based on transit ridership change or simply to a maximum number of iterations.

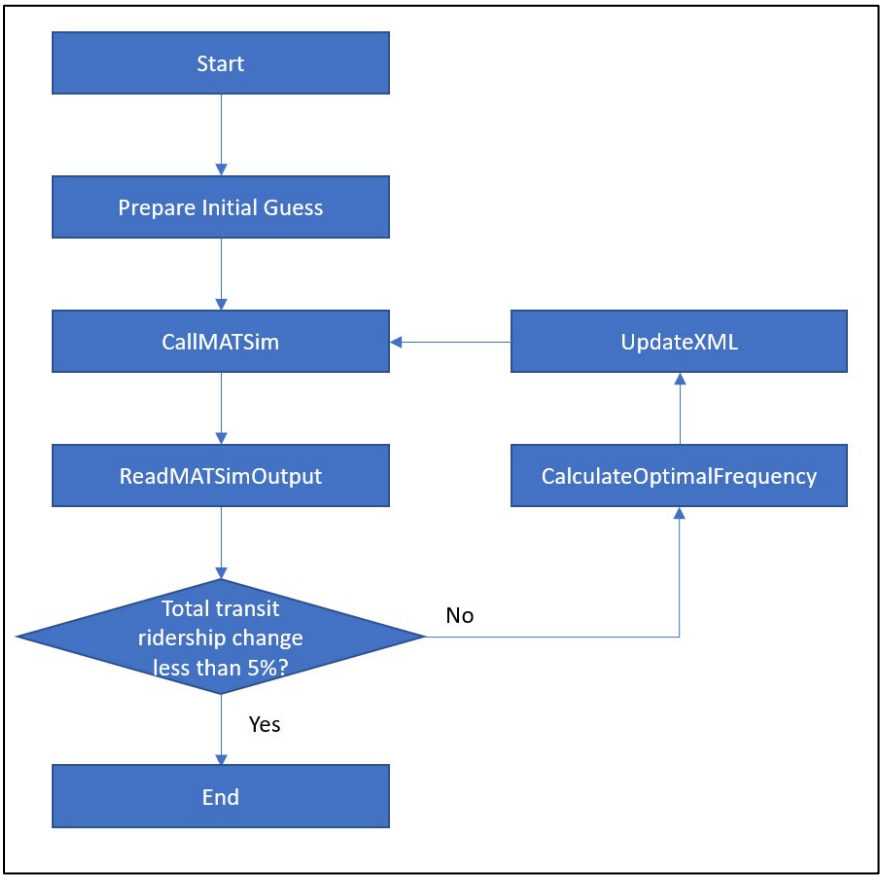


Figure 15: Iterative Optimization Algorithm Flow.

Section 5: Testing Scenarios

Three scenarios are tested in this project. The first scenario tests the existing bus operation. January 2020 MTA GTFS is imported to MATSim-NYC and tested to establish a baseline.

The output is analyzed in detail and compared with the published ridership. The simulated ridership is scaled to match the published ridership, and this scaling factor is kept the same for the rest of the scenarios.

The second scenario is to test the plan proposed in Goldwyn and Levy (2020). In this scenario, the frequency of each route is predetermined in their design.

The third scenario is to test the proposed network in which the frequency is set using the algorithm that accounts for the demand response from the MATSim-NYC model.

Subsection 5.1 Input Parameters

Lam and Small (2001) reports the value of time at Orange County, California is $22.87. Since the cost of living in Orange County is comparable to New York City, we decided to test scenario 3 with an in-vehicle travel time of $20/hr and a wait time cost of $35/hr (based on more than 50% wait time premium in Balcombe et al., 2004).

Table 1 lists the global parameters used in the simulation, where P_w is the cost of wait time, P_v is the cost of in-vehicle travel time, β is the average boarding and alighting time per passenger, c is the operation cost for buses, and v_0 is the bus travel speed.

MATSim-NYC is run to 100 days of iterations to adjust passengers demand to the scenario design. One run requires about 11 hours on a PC with an Intel Xeon E5-2637 CPU and 128 GB RAM. In practice, we recommend users to reserve at least 70GB of RAM to run our 4% MATSim-NYC model. The computation time is relatively high due to the large number of agents and NYC's complex network. It is highly recommended to use numerical optimization techniques such as the one shown in Figure 14 to reduce computation time.

Table 1: Testing Parameters for Scenario 3

Parameter	**Value**	**Unit**
P_w	35	$/hr
P_v	20	$/hr
β	7.2	second
c	215	veh/hr
v_0	15	km/hr

Subsection 5.2 Scenario 1 calibration results

The existing transit network is run in MATSim-NYC for one run and the results are compared to existing MTA ridership numbers. Because MATSim-NYC was calibrated overall at the citywide level considering both subway and bus for one transit mode, there are discrepancies to the total ridership values. Assuming the distribution of ridership is adequate, we add a further calibration for the local area study in Brooklyn by applying a scale factor to the ridership to make it match the total Brooklyn bus ridership. A factor of 3.69 was applied to the output MATSim-NYC bus ridership to scale it to the observed ridership. The resulting scaled MATSim-NYC ridership is compared at the line level to the observed ridership in Table 2. The average of the relative differences is shown to be 30% while the median difference is 17% (this means there are a few large outliers). The outliers greater than 50% are colored red. None of them have an observed ridership greater than 10,000 daily trips. When weighted by ridership, the ridership-average difference across the routes is 21%. This suggests the distribution of the ridership is within a reasonable range (see Flyvbjerg et al., 2005).

Table 2: Comparison of calibrated MATSim-NYC daily line ridership to MTA data

Route	Simulated Ridership (After Scaling)	Reference Ridership (MTA, 2020b)	% Difference
B1	17643	18180	3%
B11	12206	10377	18%
B12	10830	12124	11%
B13	6356	6084	4%
B14	3713	5900	37%
B15	17459	17977	3%
B16	11070	6184	79%
B17	7068	9382	25%
B2	2041	2088	2%
B20	8478	6315	34%
B24	3411	2449	39%
B25	9110	7874	16%
B26	7725	8234	6%
B3	8637	11309	24%
B31	2694	2619	3%
B32	1078	820	31%
B35	21626	27273	21%
B36	6434	12305	48%
B37	4573	2300	99%
B38	10424	18011	42%
B39	805	220	266%
B4	8379	6192	35%
B41	19537	22967	15%
B43	10055	9346	8%
B44	24896	32334	23%
B45	4894	4973	2%
B46	28580	38120	25%
B47	11605	9252	25%
B48	6607	3534	87%
B49	11276	10886	4%
B52	10150	9940	2%

B54	9516	9197	3%
B57	7825	6851	14%
B6	26705	35963	26%
B60	7663	8364	8%
B61	14056	8876	58%
B62	8183	6841	20%
B63	10712	11148	4%
B64	7246	5442	33%
B65	6301	2795	125%
B67	5160	4441	16%
B68	12535	12660	1%
B69	4031	4134	2%
B7	8146	5091	60%
B70	3946	6520	39%
B74	4056	3598	13%
B8	15348	18388	17%
B82	27868	25126	11%
B83	6090	7024	13%
B84	476	544	12%
B9	12265	14416	15%

Subsection 5.3 Proposed Transit Network Configuration

For Scenarios 2 and 3, we refer to Goldwyn and Levy's (2020) proposed bus network redesign. The network is visualized in Figure 16. Major attractions are marked with red triangles, and bus depots are marked with a green bus sign. Compared to the existing average operation speed of 7.2 mph for Brooklyn buses (NYCDOT, 2019), Levy and Goldwyn propose to speed up the buses to 9.32 mph (15km/hr) by implementing off-board fare collection, stop consolidation, dedicated lanes, and signal priority. In order to provide sufficient frequency, they also propose to consolidate the network from the current 550km to about 355km.

The network is then converted into XML, and it is visualized on Simunto Via (a MATSim visualization software), shown in Figure 17.

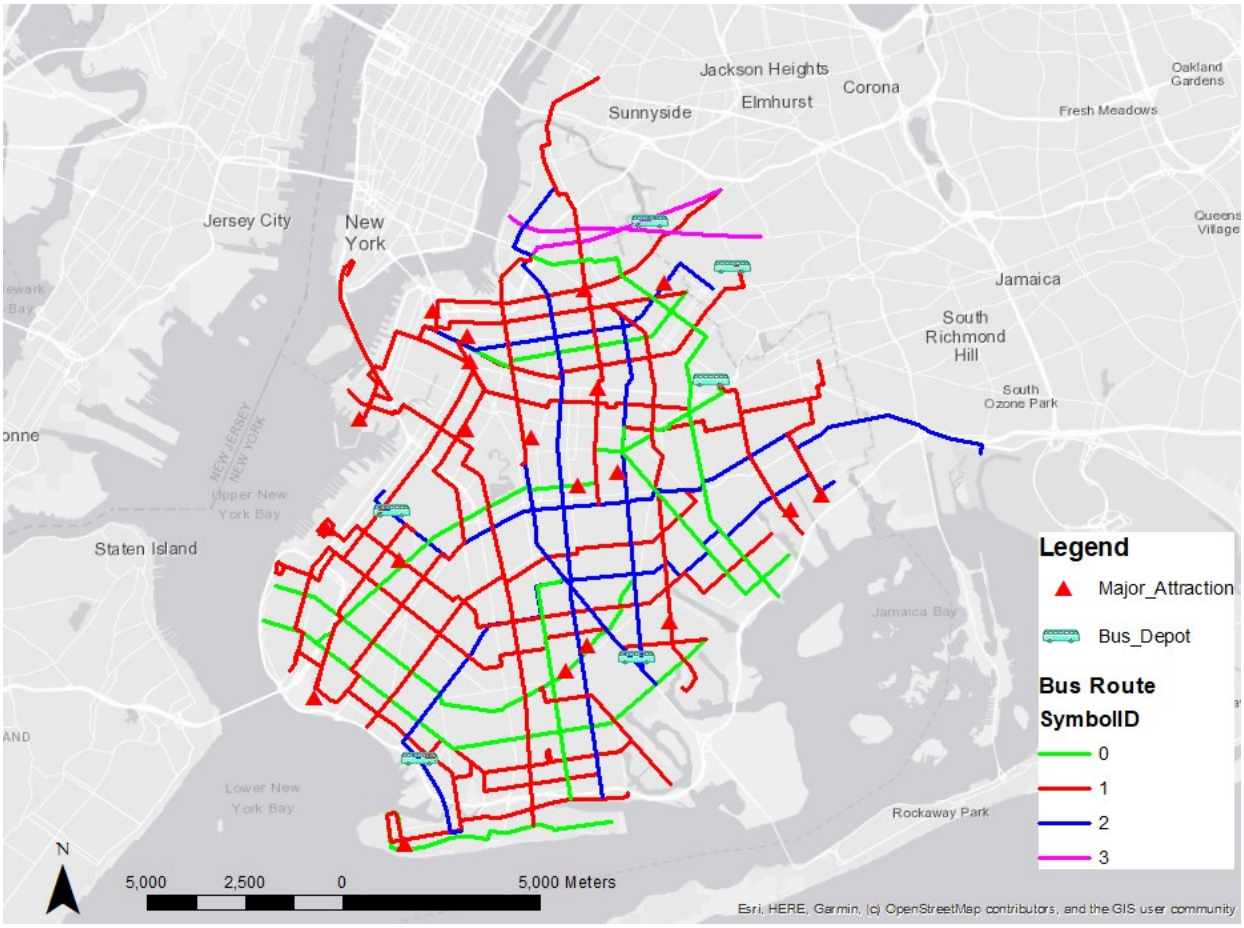


Figure 16: Goldwyn and Levy's (2020) Bus Redesigned Plan (on GIS).

Figure 17: Goldwyn and Levy's (2020) Bus Redesigned Plan (on MATSim).

Section 6: Results

6.1 Algorithm convergence for Scenario 3

For scenario 3, ten iterations of the proposed algorithm were run to find a stable solution. Figure 18 shows the trajectory of the algorithm for scenario 3 in terms of ridership and number of bus trips. The number of riders increases monotonically each iteration until iteration 5, after which the algorithm stabilizes. The ridership stabilizes at around 64,400 per day. The number of provided vehicle trips increases rapidly since iteration 0 and stabilizes at iteration 5 to around 12,000 vehicle trips per day. These are used as the final frequencies for the design in the comparison of scenarios.

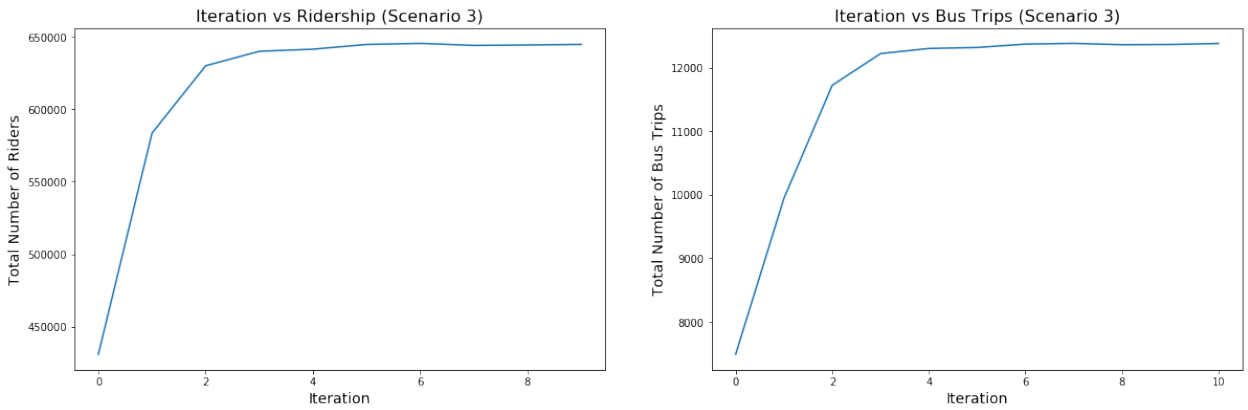


Figure 18: Ridership and Vehicle Trips Trajectories for proposed algorithm used in Scenario 3.

6.2 Scenario 3 output summary

Based on the experienced plans from each agent in the simulation, we can aggregate the number of riders who departure or arrive at each transit stop station per day. Figure 19 shows the stop boardings and alightings for scenario 3, where the red bars represent the aggregated number of agents who depart from a station, and the green bars represent the aggregated number of agents who arrive at a station.

It is important to know the rider demand at each station because it can provide guidance to engineers and planners on where to prioritize bus infrastructure investments. For example, NYC DOT should first deploy their Real-Time Passenger Information Signs at the stops with the most passengers to reduce waiting anxiety.

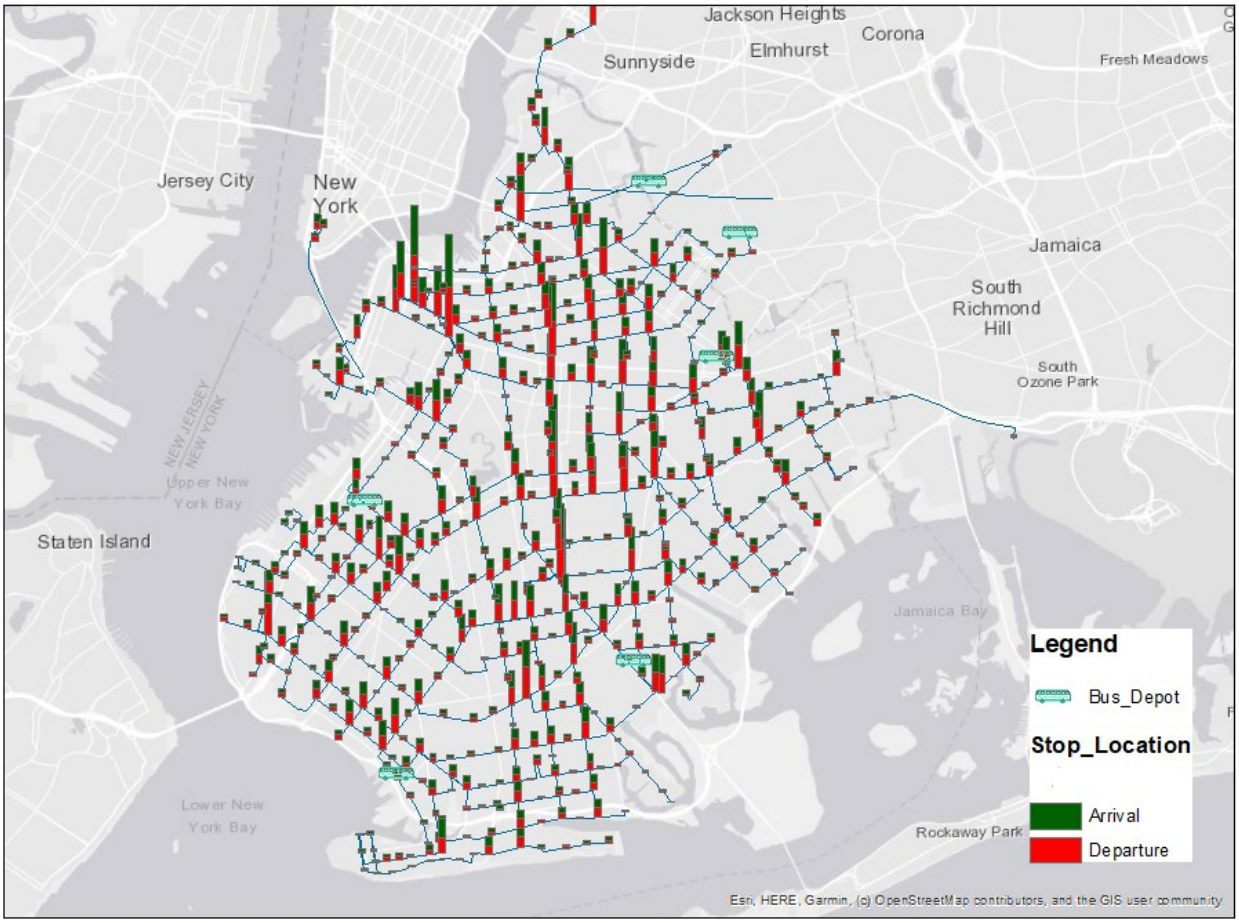
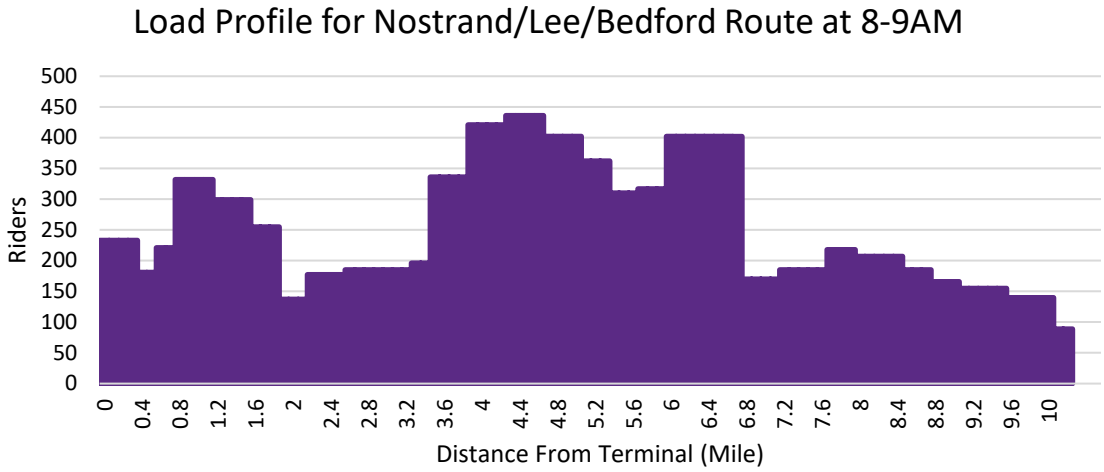


Figure 19: Bus stop boardings and alightings.

We can also aggregate the simulated demand on a route level. As an example, Figure 20 show the vehicle load for the proposed Nostrand/Lee/Bedford Ave route between 8 AM and 9 AM. This profile shows where the peak loads are which allows the bus operator to focus on those stops and segments.

(a)



(b)

Figure 20: (a) Nostrand route, and (b) simulated load profile for Nostrand Route at 8-9AM.

From our simulation, we can aggregate the number of vehicle trips provided each hour, which reflects the output frequencies for the routes in the network. Figure 21 shows the vehicle trips provided for each hour by stacking the frequency of each route at different hours of a day. The left part of the figure is based on Goldwyn and Levy's (2020) headway design, and the right part of the figure is the simulation output for Scenario 3. Scenario 3 provides a significantly lower number of vehicle trips at mid-day and evening. Engineers and planners can use these figures to guide their timetable, vehicle scheduling, and crew scheduling designs.

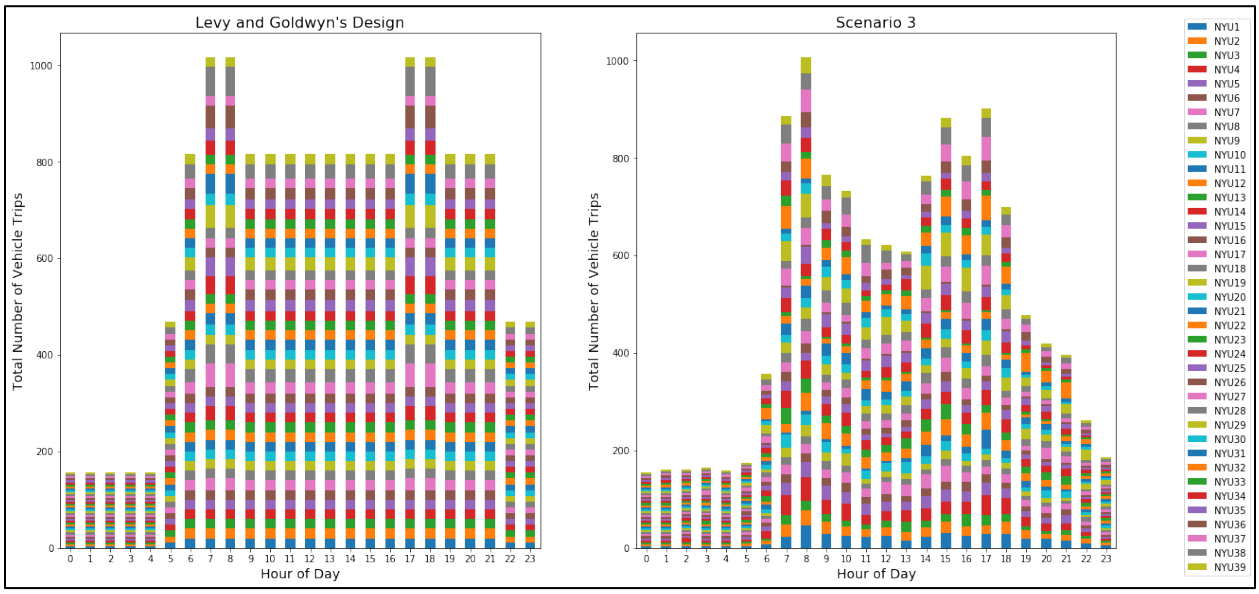


Figure 21: Trips Provided per Hour for the Proposed Design for Scenarios 2 and 3.

6.3 Comparison of Scenarios

The daily ridership across the three scenarios are summarized in Figures 22 – 24. They show that the Existing ridership is relatively low, but by redesigning the routes and frequencies, it is possible to attain higher ridership throughout.

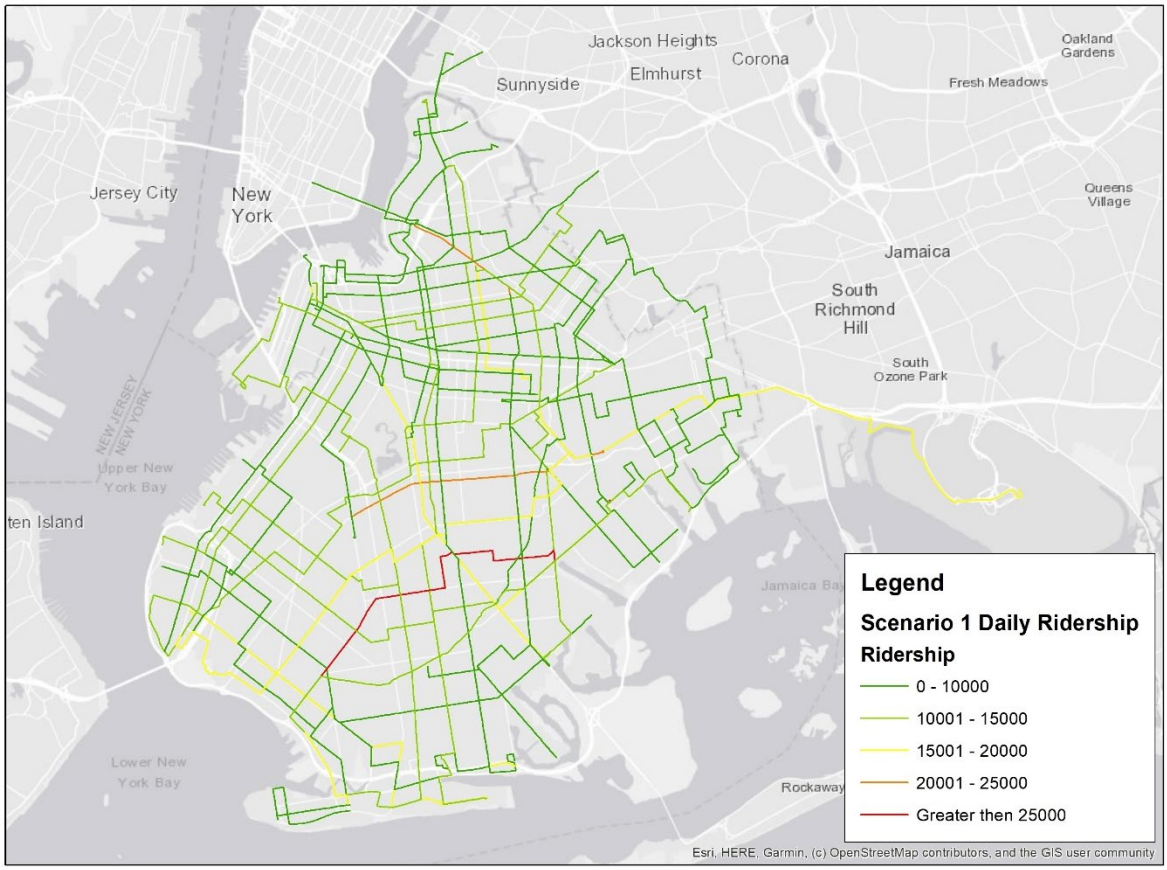


Figure 22: Scenario 1 (existing) Brooklyn bus network daily ridership.

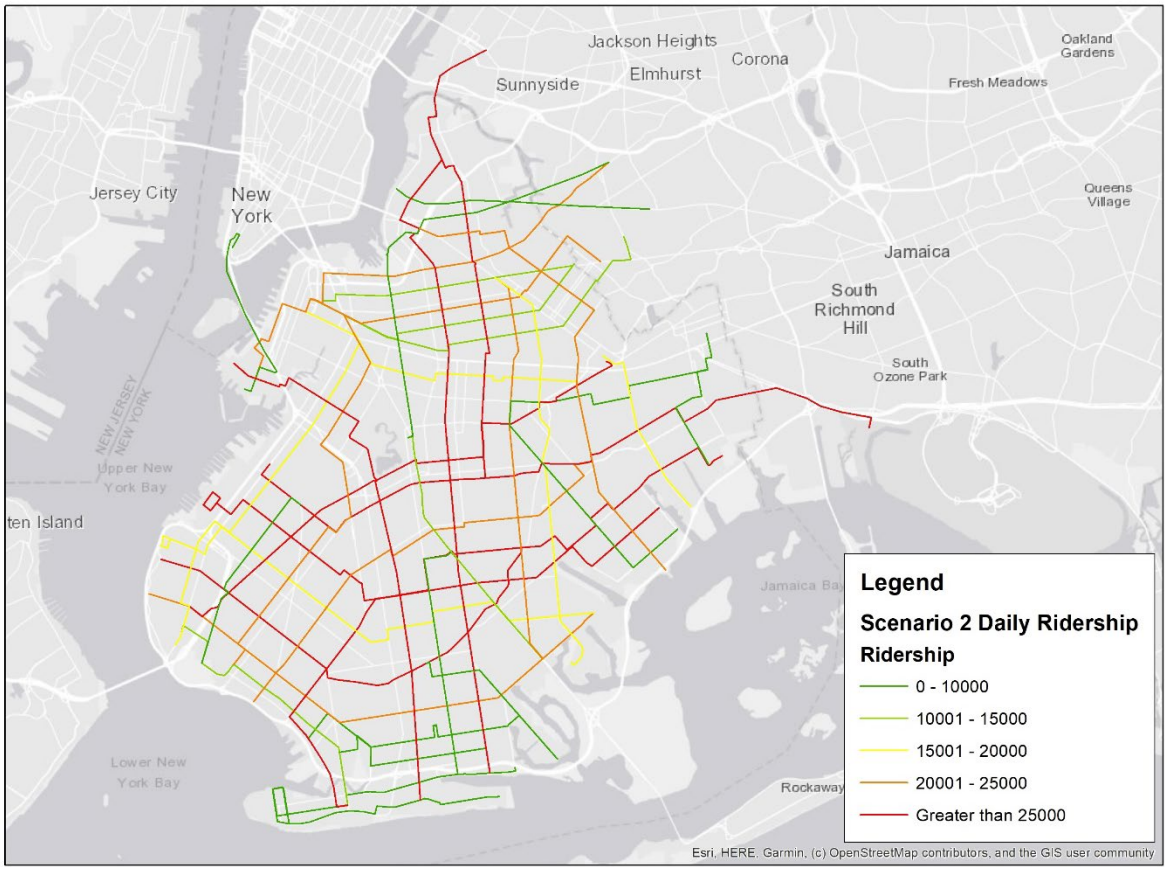


Figure 23: Scenario 2 Brooklyn bus network daily ridership.

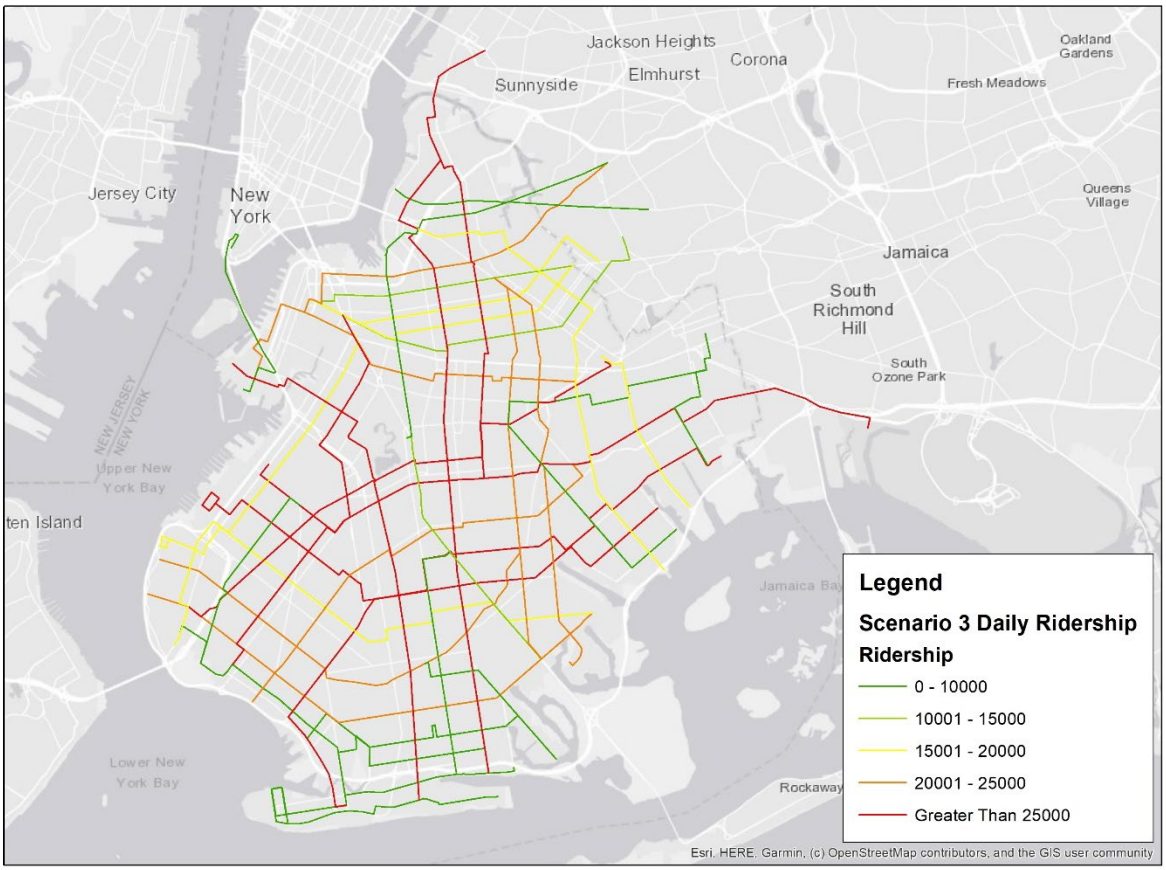


Figure 24: Scenario 3 Brooklyn bus network daily ridership.

The three scenarios are compared in terms of operating cost, ridership, fares collected, and farebox recovery ratio, as shown in Table 3. The operating costs are based on the same parameters used for determining operating cost shown in Table 1 used to compute the cost for each network of routes according to their frequencies. The revenue is based on the simulated ridership for all three scenarios in MATSim. This ridership reflects the costs of transit to the riders who use the system. Increased wait times (due to lower frequencies) or in-vehicle times would result in changes in ridership.

Table 3: Comparison of Simulated Daily Metrics Between Scenarios

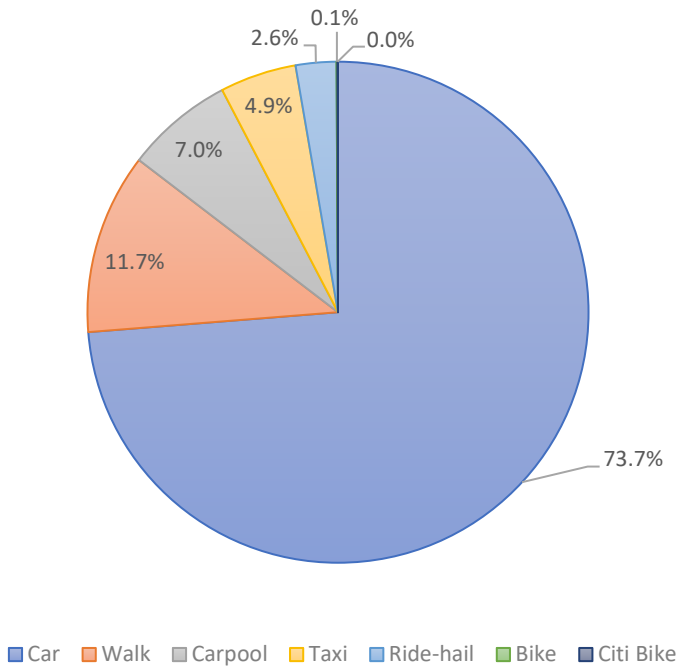
	Scenario 1	Scenario 2	Scenario 3
Vehicle Miles Traveled	93,447	113,480	90,252
Operation Cost	$2,790,451	$2,623,462	$2,081,994
[% change]		[-6%]	[-25%]
Ridership	537,217	661,755	644,788
[% change]		[+23%]	[+20%]
Equivalent Fare Revenue	$613,899	$756,214	$736,825
Farebox Recovery Ratio	0.22	0.29	0.35

The results suggest that the bus network redesign from Goldwyn and Levy (2020) would increase demand by 23% while reducing operating cost by 6%. This result is similar to their estimated ridership increase of 20% as well. Meanwhile, the propose frequency setting leads to a more balanced outcome: ridership improves over existing by 20% but also reduces operating cost by 25%. This is a small 3 percentage point drop from the Goldwyn and Levy (2020) design with an accompanying 19 percentage point reduction in operating cost.

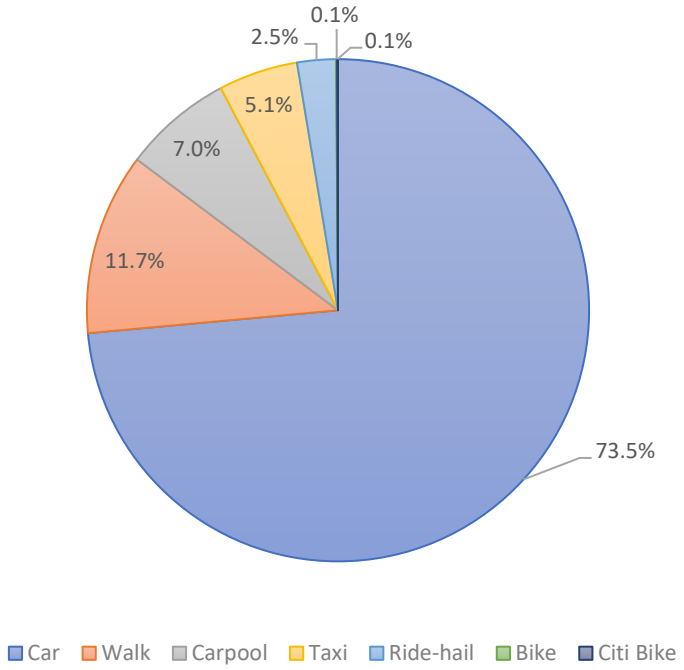
Because the simulation outputs individual agent choices, we can determine where all the new ridership is coming from in Scenario 2 and Scenario 3. This helps answer the question of how redesigning the bus network in the presence of ride-hail services as well as other modes would impact travelers' choices to choose bus. The results are shown in Figure 25. In both scenarios, the major mode drawn from is by car, which is a very encouraging result. Only 2.5 – 2.6% (this amounts to about 2400 daily trips by ride-hail) of the new trips in the bus redesigns would come from ride-hail, which suggests there is not so much competition between the two modes. Nearly

12% of new trips would be drawn from walking, which suggests that the redesign is able to provide a more convenient alternative to people who otherwise would have walked only.

Although the basic fare is $2.75, many riders have discounts or free rides through various programs for seniors, students, monthly passes, etc. To account for this, we determine an equivalent fare per passenger by setting it such that the farebox recovery ratio matches the observed value from the MTA (2018) of 0.22. The resulting equivalent fare is $1.14, which we apply to the other scenarios.



(a)



(b)

Figure 25: Modes shifted to new bus ridership in (a) Scenario 2 and (b) Scenario 3.

Goldwyn and Levy's (2020) design increases the farebox recovery ratio by 30% to 0.29 by providing faster and more frequent services. Scenario 3 provides a more balanced solution, whose vehicle mile traveled per day is similar to the existing MTA's system, but can attract more customers. It is more efficient compared to Levy and Goldwyn's design because many of the unnecessary trips during mid-day and evening are taken away based on the demand. The farebox recovery ratio for Scenario 3 is 0.35, which is a 60% improvement over Scenario 1.

Section 7: Conclusion

In sum, this project has two main contributions. First, we proposed a simulation-based optimization framework for bus frequency planning in a large-scale transportation network. Second, we implemented this framework with C2SMART's Open Source Multi-Agent Virtual

Simulation Test Bed to evaluate the existing Brooklyn bus network, the proposed network redesign by Marron Institute, and an alternative design based on our framework.

The MATSim-NYC model from Chow et al. (2020a) is generally able to simulate similar patterns to the existing bus network in Brooklyn with some calibration. A line-level comparison to observed ridership shows a ridership-weighted average of 21% difference between observed and simulated route ridership, with a few outliers for the smaller volume routes.

An iterative simulation-based frequency optimization method is proposed that uses an analytical model to set frequencies and a simulation model (MATSim-NYC) to update demand. Numerical tests show that the algorithm converged to an equilibrium outcome.

Comparisons of the Goldwyn and Levy (2020) design to the existing scenario confirms their claim that their design can increase ridership by 20% (our simulation result suggests an increase of 23%), at a reduction in operating cost of 6%. By using our simulation-based frequency setting approach, however, we can further improve operating cost (to 25% reduction from Scenario 1) while maintain a 20% increase in ridership from Scenario 1. As a result, our simulation-based optimization approach can improve upon Goldwyn and Levy's (2020) network redesign to increase farebox recovery ratio from an improvement of 30% up to 60% over the existing scenario.

The increased ridership draws primarily from passenger car use (nearly 75%), with a small 2.5% drawn from ride-hail services and another 5% from taxis. This suggests the redesigns should be effective in moving people away from less efficient transportation modes.

Section 8: Technology Transfer, Dissemination, and Broader Impacts

8.1 Technology Transfer

In this section we provide links to all the completed products we developed during this project as shown in Table 4. Each of the items in the table served on its own to transfer "new knowledge" or contributed in developing the product that did.

Table 4. Delivered products

Output	Description
Code to create GTFS from shapefile	This is the code to create GTFS from the network shapefile and frequency data.
GTFS for redesigned Brooklyn Bus network	This is the GTFS version of the network and frequencies proposed by Goldwyn and Levy (2020).
Code for simulation-based frequency setting	This is the iterative algorithm code that solves the optimal frequency for Eq. (4) given a demand, and calls MATSim-NYC to update the demand based on the new frequencies.
GTFS for redesigned Brooklyn Bus network v2	This is the GTFS version of the network proposed by Goldwyn and Levy (2020) with frequencies set by the proposed algorithm.
MATSim-NYC-Marron model	This is the MATSim-NYC-Marron model, which takes the baseline MATSim-NYC and replaces the Brooklyn bus network with the network and frequencies proposed by Goldwyn and Levy (2020).

MATSim-NYC-Marron-V2 model	This is the MATSim-NYC-Marron-V2 model, which takes the baseline MATSim-NYC and replaces the Brooklyn bus network with the network proposed by Goldwyn and Levy (2020) and the frequencies proposed in this project.
User guide	This is a user guide that provides instructions to users on going through the implementation process.

8.2 Dissemination

The work is being prepared in a paper to be submitted to Transportation Research Record and TRB Annual Meeting 2021. A Linkedin post will be prepared to share the findings with industry, particularly aiming to share the results with NYCT and Remix, who are involved in bus network redesign projects in Bronx and Brooklyn.

A webinar is being planned for the summer to showcase the work to C2SMART stakeholders.

8.3 Broader Impacts

In addition to the direct dissemination and technology transfer, this research has led to several broader impacts.

Student training and involvement: Two graduate students were trained and mentored through this project: Ziyi Ma and Mina Lee. Ziyi also received an Eisenhower Graduate Fellowship as a result of this work that he built upon in his application. He also led a team that won the Microtransit Hackathon hosted by IATR in 2019. The study output has been used in classroom setting, including in TR-GY 7133 Public Transport and in an undergraduate course CE-UY 3373 Transportation Systems Analytics.

In addition to the main research team, we participated in the ARISE program to expose K-12 STEM students to this research and other projects from our lab.

Public engagement: Through social media (Linkedin) we will present the findings of this work to publicize it.

Industry engagement: Through the webinar in the summer we will engage with industry and stakeholders that might be interested in this kind of tool.

References

Allen, D.J., 2017. *Lost in the Transit Desert: Race, Transit Access, and Suburban Form*. Routledge.

Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., & Zhang, K. (2016). POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. *Transportation Research Part C: Emerging Technologies*, *64*, 101-116.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., ... & White, P. (2004). The demand for public transport: a practical guide. Report TRL593.

Balmer, M., Rieser, M., Meister, K., Charypar, D., Lefebvre, N., & Nagel, K. (2009). MATSim-T: Architecture and simulation. In A. L. C. Bazzan & F. Klugl (Eds.), Multi-agent systems for traffic and transportation engineering (pp. 57–78). Hershey, PA: IGI Global.

Bar-Yosef, A., Martens, K. and Benenson, I., 2013. A model of the vicious cycle of a bus line. *Transportation Research Part B: Methodological*, *54*, pp.37-50.

Bard, J. F. (1991). Some properties of the bilevel programming problem. *Journal of optimization theory and applications*, *68*(2), 371-378.

Bard, J. F., & Moore, J. T. (1990). A branch and bound algorithm for the bilevel programming problem. *SIAM Journal on Scientific and Statistical Computing*, *11*(2), 281-292.

Becker, H., Balac, M., Ciari, F., & Axhausen, K. W. (2020). Assessing the welfare impacts of shared mobility and Mobility as a Service (MaaS). *Transportation Research Part A: Policy and Practice*, *131*, 228-243.

Bonabeau, E. (2002). Agent-Based Modeling: Methods and Techniques for Simulating Human Systems. PNAS, Vol. 99, No. 3.

Bradley, M., Bowman, J. L., & Griesenbeck, B. (2010). SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. Journal of Choice Modelling, 3(1), 5–31.

Brooklyn Bus Network Redesign. n.d. Retrieved from https://new.mta.info/system_modernization/brooklynbusredesign.

Byrne, B. F. (1975). Public transportation line positions and headways for minimum user and system cost in a radial case. *Transportation Research*, *9*(2-3), 97-102.

Canca, D., Barrena, E., De-Los-Santos, A., & Andrade-Pineda, J. L. (2016). Setting lines frequency and capacity in

dense railway rapid transit networks with simultaneous passenger assignment. *Transportation Research Part B: Methodological*, *93*, 251-267.

Ceder, A. (2016). *Public transit planning and operation: Modeling, practice and behavior*. CRC press.

Ceder, A. and Wilson, N.H., 1986. Bus network design. *Transportation Research Part B: Methodological*, *20*(4), pp.331-344.

Cetin, N., Burri, A., & Nagel, K. (2003). A parallel queue model approach to traffic microsimulations. Paper presented at the Transportation Research Board 82nd Annual Meeting, Washington, DC.

Chow, J. Y. J. (2018). *Informed Urban Transport Systems: Classic and Emerging Mobility Methods toward Smart Cities*. Elsevier.

Chow, J.Y.J., Ozbay, K., He, B.Y., Zhou, J., Ma, Z., Lee, M., Wang, D., Sha, D. (2020a). Multi-agent simulation-based virtual test bed ecosystem: MATSim-NYC. C2SMART Final Report.

Chow, J. Y. J., Rath, S., Yoon, G., Scalise, P., Alanis Saenz, S., 2020b. Spectrum of public transit operations: from fixed route to microtransit. FTA Final Report, No. NY-2019-069-01-00.

Ciari, F., Balac, M., & Axhausen, K. W. (2016). Modeling carsharing with the agent-based simulation MATSim: State of the art, applications, and future developments. *Transportation Research Record*, *2564*(1), 14-20.

Cich, G., Knapen, L., Maciejewski, M., Bellemans, T., & Janssens, D. (2017). Modeling demand responsive transport using SARL and MATSim. *Procedia Computer Science*, *109*, 1074-1079.

Clewlow, R.R. and Mishra, G.S., 2017. Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States. *UC Davis, Institute of Transportation Studies, UCD-ITS-RR-17-07*.

Constantin, I., & Florian, M. (1995). Optimizing frequencies in a transit network: a nonlinear bi-level programming approach. *International Transactions in Operational Research*, *2*(2), 149-164.

Daganzo, C.F., 2010. Structure of competitive transit networks. *Transportation Research Part B: Methodological*, *44*(4), pp.434-446.

Desaulniers, G. and Hickman, M.D., 2007. Public transit. *Handbooks in Operations Research and Management Science*, *14*, pp.69-127.

Djavadian, S., & Chow, J. Y. J. (2017a). Agent-based day-to-day adjustment process to evaluate dynamic flexible transport service policies. *Transportmetrica B: Transport Dynamics*, *5*(3), 281-306.

Djavadian, S., & Chow, J. Y. J. (2017b). An agent-based day-to-day adjustment process for modeling 'Mobility as a Service' with a two-sided flexible transport market. *Transportation research part B: methodological*, *104*, 36-57.

Erath, A and Chakirov, A. 2016. Singapore. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) The Multi-Agent Transport Simulation MATSim, Pp. 379–382. London: Ubiquity Press. DOI: http://dx.doi.org/ 10.5334/baw.57. License: CC-BY 4.0

Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., & Rashidi, H. (2013). A review of urban transportation network design problems. *European Journal of Operational Research*, *229*(2), 281-302.

Fielbaum, A., Jara-Diaz, S. and Gschwender, A., 2016. Optimal public transport networks for a general urban structure. *Transportation Research Part B: Methodological*, *94*, pp.298-313.

Fielbaum, A., Jara-Diaz, S. and Gschwender, A., 2017. A parametric description of cities for the normative analysis of transport systems. *Networks and Spatial Economics*, *17*(2), pp.343-365.

Flyvbjerg, B., Skamris Holm, M. K., & Buhl, S. L. (2005). How (in) accurate are demand forecasts in public works projects?: The case of transportation. *Journal of the American planning association*, *71*(2), 131-146.

Gallo, M., Montella, B., & D'Acierno, L. (2011). The transit network design problem with elastic demand and

internalisation of external costs: An application to rail frequency optimisation. *Transportation Research Part C: Emerging Technologies*, *19*(6), 1276-1305.

Goldwyn, E., Levy, A. (2020). Rebuilding bus ridership in America: A case study in Brooklyn, New York. Marron Institute Report, https://marroninstitute.nyu.edu/papers/rebuilding-bus-ridership-in-america-a-case-study-in-brooklyn-new-york.

Goulias, K. G., Bhat, C. R., Pendyala, R. M., Chen, Y., Paleti, R., Londuri, K. C., et al. (2011). Simulator of activities, greenhouse emissions, networks, and travel (SimAGENT) in southern California. Paper presented at the Transportation Research Board 91st Annual Meeting, Washington, DC.

Guihaire, V., & Hao, J. K. (2008). Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice*, *42*(10), 1251-1273.

Hasselström, D. (1982). Public transportation planning: A mathematical programming approach. PhD dissertation, University of Göteburg, Sweden.

Hawkins, A.J., 2018. Uber Express Pool offers the cheapest fares yet in exchange for a little walking. *The Verge*, Feb 21.

He, B. Y., Zhou, J., Ma, Z., Chow, J. Y. J., Ozbay, K., 2020a. Evaluation of city-scale built environment policies in New York City using an emerging mobility-accessible synthetic population, working paper.

He, B. Y., Zhou, J., Ma, Z., Wang, D., Sha, D., Lee, M., Chow, J. Y. J., Ozbay, K., 2020b. Calibration, validation, and application of multi-agent simulation test bed to New York City emerging transportation policies and technologies, working paper.

Hörl, S., Ruch, C., Becker, F., Frazzoli, E., & Axhausen, K. W. (2019). Fleet operational policies for automated mobility: A simulation assessment for Zurich. Transportation Research Part C: Emerging Technologies.

Horni, A., Nagel, K. and Axhausen, K. W. (eds), 2016. The Multi-Agent Transport Simulation MATSim. London: *Ubiquity Press*. DOI: https://doi.org/10.5334/baw

Jiao, J., 2017. Identifying transit deserts in major Texas cities where the supplies missed the demands. *Journal of Transport and Land Use*, *10*(1), pp.529-540.

Jiao, J., Bischak, C., 2018. People are stranded in 'transit deserts' in dozens of US cities. *The Conversation*, Mar 13.

Kang, J. E., Chow, J. Y., & Recker, W. W. (2013). On activity-based network design problems. *Transportation Research Part B: Methodological*, *57*(C), 398-418.

Lampkin, W. and Saalmans, P.D., 1967. The design of routes, service frequencies, and schedules for a municipal bus undertaking: A case study. *Journal of the Operational Research Society*, *18*(4), pp.375-397.

Lam, T. C., & Small, K. A., 2001. The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review*, *37*(2-3), 231–251. doi: 10.1016/s1366-5545(00)00016-8

Levy, A., Goldwyn, E., 2018. To build a better bus system, ask a driver. *Citylab*, Jun 18.

Levy, A., & Goldwyn, E. 2018. Get on the Bus: A Radical Plan for Brooklyn's Bus Network. *New York Magazine*.

Marcotte, P. (1986). Network design problem with congestion effects: A case of bilevel programming. *Mathematical programming*, *34*(2), 142-162.

Mohring, H., 1972. Optimization and scale economies in urban bus transportation. *The American Economic Review*, 62(4), pp.591-604.

MTA, 2018. Adopted Budget February Financial Plan. Retrieved from http://web.mta.info/mta/budget/pdf/MTA-

2018-AdoptedBudgetFebruaryFinancialPlan_2018-21.pdf.

MTA, 2020a. Operating Budget Basics. Retrieved from https://new.mta.info/budget/MTA-operating-budget-basics, last accessed May 25, 2020.

MTA, 2020b. Average weekday bus ridership. http://web.mta.info/nyct/facts/ridership/ridership_bus.htm, last accessed June 10, 2020.

Nagel, K., Beckman, R. L., & Barrett, C. L. (1999). TRANSIMS for transportation planning. Paper presented at the 6th International Conference on Computers in Urban Planning and Urban Management, Franco Angeli, Milano, Italy.

Nahmias-Biran, B.H., Oke, J.B., Kumar, N., Basak, K., Araldo, A., Seshadri, R., Akkinepally, A., Lima Azevedo, C. and Ben-Akiva, M., 2019. From Traditional to Automated Mobility on Demand: A Comprehensive Framework for Modeling On-Demand Services in SimMobility. *Transportation Research Record*, p.0361198119853553.

Neumann, A. and Nagel, K., 2013, January. Passenger Agent and Paratransit Operator Reaction to Changes of Service Frequency of a Fixed Train Line. In *ANT/SEIT* (pp. 803-808).

Neumann, A. 2016. Berlin I: BVG Scenario. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) The Multi-Agent Transport Simulation MATSim, Pp. 369–370. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/ baw.53. License: CC-BY 4.0

Newell, G. F. (1971). Dispatching policies for a transportation route. *Transportation Science*, *5*(1), 91-105.

Newell, G.F., 1979. Some issues relating to the optimal design of bus routes. *Transportation Science*, *13*(1), pp.20-35.

NYCT, 2012. Customer Satisfaction Survey. Retrieved from: http://web.mta.info/mta/news/books/docs/2012LocalBus_CSS.pdf

NYCDOT, 2019. New York City Mobility Report. Retrieved from: https://www1.nyc.gov/html/dot/downloads/pdf/mobility-report-singlepage-2019.pdf

Rieser-Schüssler, N, Bo¨sch, P M, Horni, A and Balmer, M. 2016. Zu¨rich. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) The Multi-Agent Transport Simulation MATSim, Pp. 375–378. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/baw.56. License: CC-BY 4.0

Rothfeld, R., Balac, M., Ploetner, K. O., & Antoniou, C. (2018). Agent-based simulation of urban air mobility. In *2018 Modeling and Simulation Technologies Conference* (p. 3891).

Schöbel, A., Scholl, S. (2006). Line Planning with Minimal Traveling Time. Presented at the 5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05), p.16.

Sisson, P., 2018. Are Uber and Lyft helping or hurting public transit? Curbed, May 8.

Szeto, W. Y., & Jiang, Y. (2014). Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B: Methodological*, *67*, 235-263.

Tirachini, A., 2014. The economics and engineering of bus stops: Spacing, design and congestion. *Transportation research part A: policy and practice*, *59*, pp.37-57.

Tirachini, A., Hensher, D.A. and Rose, J.M., 2013. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice*, *53*, pp.36-52.

van Nes, R., Hamerslag, R., & Immers, L. H. (1988). The design of public transport networks. *Transportation Research Record* 1202, 74-83.

Verbas, İ. Ö., & Mahmassani, H. S. (2015). Integrated frequency allocation and user assignment in multimodal transit networks: Methodology and application to large-scale urban systems. *Transportation Research*

Record, *2498*(1), 37-45.

Von Neumann, J. (1966). *Theory of Self-Reproducing Automata*. Edited by A. W. Burk. Urbana: University of Illinois Press.

Walker, J., 2018. The Bus is Still Best. *The Atlantic*, Oct 31.

Warerkar, T., 2017. Uber surpasses yellow cabs in average daily ridership in NYC. *Curbed.*

Yang, H., & H. Bell, M. G. (1998). Models and algorithms for road network design: a review and some new developments. *Transport Reviews*, *18*(3), 257-278.

Zhang, G., Zhang, H., Li, L., & Dai, C. (2013). Agent-based simulation and optimization of urban transit system. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 589-596.

Zheng, H., Son, Y. J., Chiu, Y. C., Head, L., Feng, Y., Xi, H., ... & Hickman, M. (2013). A primer for agent-based simulation and modeling in transportation applications (No. FHWA-HRT-13-054). United States. Federal Highway Administration.

Ziemke, D. 2016. Berlin II: CEMDAP-MATSim-Cadyts Scenario. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) The Multi-Agent Transport Simulation MATSim, Pp. 371–372. London: Ubiquity Press. DOI: http://dx.doi.org/10.5334/baw.54. License: CC-BY 4.0

Appendix

Pseudo Code for the Iterative Optimization Algorithm:

Function ReadMATSimOutput(ExperiencedPlan.xml, route attributes):

Generate a DataFrame called *SimulatedResult(i)* consisted of list of routes for 24 hours

Merge the total length of each route to the *SimulatedResult(i)*

Calculate and merge the *AverageTraveledDistance(i), Number-of-Riders(i)* from ExperiencedPlan.xml to the *SimulatedResult(i)*

return *SimulatedResult(i)*

Function CalculateOptimalFrequency(*OperationCost(c), VehicleTravelSpeed(c), PassengerBoardingTime(c), Value-of-WaitTime(c), value-of-in-VehicleTravelTime (c), SimulatedResult(i)*):

Calculate optimal frequency with Eq. (4)

Generate arrays of frequencies of all routes and hours for each direction of the transit traveled

return FrequencyArrays

Function UpdateXML(FrequencyArrays, TransitSchedule.xml):

Reformat FrequencyArrays

```
    Remove the existing vehicle trips
    Insert updated vehicle trips to TransitSchedule.xml with ElementTree
    Return UpdatedTransitSchedule.xml

Function CallMATSim(UpdatedTransitSchedule.xml):
    Use io command to start MATSim simulation using UpdatedTransitSchedule.xml as input
```