



An optimization model for planning limited-stop transit operations

Mahmood Mahmoodi Nesheli¹ · Siva Srikukenthiran¹ · Amer Shalaby¹

Accepted: 3 February 2022 / Published online: 8 April 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Surface transit lines in North America commonly feature a basic service pattern consisting of a single branch of all-stop service, with stops usually tightly spaced. Such a configuration is inefficient for the operator and unattractive for the users, particularly if the prevailing passenger demand is unevenly distributed along the line. In such cases, it is more effective to tailor the scheduled services to passenger demand, both spatially and temporally. Public Transit agencies have increasingly adopted various stop and service pattern strategies in order to provide high-quality services while reducing operating costs. This study is focused on one such strategy, namely limited-stop operation. It proposes a new mathematical programming model to find the best candidate route stops for this strategy to minimize the total passenger travel time. The adopted approach consists of three steps: optimization, post-optimization, and simulation. An agent-based simulation platform, called Nexus, is used to represent real-life operating conditions, generate input data for the optimization model, enable post-optimization pattern recognition for grouping trips, and finally help assess the optimization results and present a best possible strategy. The developed approach is tested in a case study of a transit system in Hamilton, Ontario, Canada. Multiple analysis and algorithm test cases are demonstrated.

✉ Mahmood Mahmoodi Nesheli
m.nesheli@utoronto.ca

Siva Srikukenthiran
siva.srikukenthiran@mail.utoronto.ca

Amer Shalaby
amer.shalaby@utoronto.ca

¹ Department of Civil Engineering, University of Toronto, Transportation Research Institute (UTTRI), 35 St. George St., Toronto, ON, Canada

1 Introduction

Public transit (PT) systems rely increasingly on information extracted from operators' intelligent data-gathering equipment. Such data can enable more effective matching of supply and demand of the modern-day transit systems. With the help of newly available smart card data and recent advancements in communication technologies, it is possible to develop better-customized scheduling services that are spatially and temporally sensitive to passenger demand. This paper employs the emerging wealth of sensor data on both transit users and service vehicles for the design and optimization of stop and service patterns (SSP) operations.

The design of SSP operations is a critical task in transit planning and can have remarkable impacts on user travel times and system efficiency (Vuchic 2005). Stop and service patterns include several specialized service strategies designed to expedite operations and tailor service to unevenly distributed passenger demand along a public transit route; examples include stop consolidation, limited stop service, zonal service, and short turn routes. Agencies may consider one strategy at a time or a few in combination. For example, a line could be re-designed by consolidating some stops or introducing a limited-stop branch in conjunction with a basic all-stop branch. These decisions on SSP need to be supported by appropriate analytical tools built using detailed data on both demand and performance. The transit industry is enjoying increased availability of such data due to the wide implementation of automated data collection systems, such as Automatic Fare Collection (AFC), Automatic Passenger Counting (APC), and Automatic Vehicle Location (AVL). In this study, one specialized service strategy, the limited-stop strategy, is investigated to examine its potential to better align the provided service with demand. It is to note that a limited-stop service in this study refers to a bus service that operates between two terminals and skipping (i.e., non-serving) some stops in a given bus route. We introduce a new mathematical model to select the best candidate route and stops for this strategy so as to minimize the total passenger travel time. The optimization model is formulated in a way that minimizes the changes to the level of service (LOS) (in terms of walk time to/from stops and waiting time).

2 Literature review

Several studies have discussed different strategies and approaches for addressing the inefficiency of the surface transit system. Early research into planning SSPs includes the zonal operation studies by Jordan and Turnquist (1979) and the bi-directional extension of zonal service optimization by Furth (1986). Eberlein (1995) addressed the limited-stop problem as an integer nonlinear programming model with a quadratic objective function and solved it using a heuristic method. Leiva et al. (2010) examined the optimization of express operations, and Ulusoy et al. (2010) showed that these strategies could yield significant time savings and

social benefits. In another study, Larrain and Muñoz (2016) presented a model that can be used to predict the performance of limited-stop bus services as a bus rapid transit management policy. Wirasinghe and Vandebona (2011) analyzed grid and non-grid patterns for coordinated express services. Social welfare benefits of limited-stop services, when coordinated with regular services making all stops, were investigated by Chiraphadhanakul and Barnhart (2013).

Tétreault and El-Geneidy (2010) used AVL and APC data to select stops and estimate running times for the limited-stop service. Their finding shows that limited-stop service usage would result in substantial travel time savings for both limited-stop and existing local services. Chen et al. (2015) proposed a mathematical model for the optimal stopping design of a limited-stop bus service. The study considered the vehicle capacity and stochastic travel time, and it restricted skipping stops by two consecutive vehicles. Soto et al. (2017) developed a bi-level optimization method to design limited-stop services by considering bus capacity, transfers, and different types of behavioral models for the passengers. The recent study by Wang et al. (2018) also proposed a mathematical model to design limited-stop service operation strategies. The problem was formulated as a mixed-integer nonlinear program with equilibrium constraints. The model presented a global optimal solution for small to medium-size transit corridors. In a practical transit guide, Hart (2016) developed a methodology to assess the potential of limited-stop service along with the existing local service where the net benefits of in-vehicle travel time savings would outweigh the negative impact of increased wait, access, and egress times when limited-stop service is implemented. Concerning operational control strategies, Eberlein et al. (1999) categorized control strategies into three main categories: stop control, inter-stop control, and others. The first category includes two main classes of strategies known as holding and stop skipping/limited-stop. The second includes speed control and traffic signal priority. The third consists of strategies like adding vehicles, splitting trains, and more. Although the operational stage (i.e., real-time) problem is different from the planning stage problem considered in this study, it is possible to learn from both problems, especially in the development of the optimization formulation.

Sun and Hickman (2005) studied the possibility of implementing a limited-stop strategy for operations control in a real-time manner. A non-linear integer programming problem for two different stop-skipping policies was formulated to examine how the performance of the two policies could change with the variability of effective parameters, such as passenger distribution patterns, and vehicle travel time variability on the route. Nesheli and Ceder (2015) investigated how hybrid special operations, including skip-stop strategies, could be used for transfer synchronization. In another study, a systematic PT control procedure to increase service reliability was proposed (Nesheli and Ceder 2017). The main task was to develop a library of various strategies for PT operations to attain optimal strategies for reducing passenger travel times and improving PT systems' serviceability. A recent study by Cao et al. (2020) showed the application of a skip-stop strategy to implement the optimal/best-adjusted urban rail timetable at a tactical level.

Contribution and objectives

The contribution of our work is briefly presented as follows:

1. The previous efforts mainly focused either on developing complicated optimization methods, which were difficult to implement in practice or used some simplified analytical models to address the effects of any given SSP configuration. In contrast, this study utilizes a new mathematical model in combination with a route simulation model to provide more realistic metrics for the optimization model.
2. The study proposes a new mathematical programming model that explicitly considers the heterogeneity in travel patterns and the associated passengers' travel experiences between their stop origins and destinations.
3. A set of penalty functions in the model is introduced to consider the extent of drawbacks of the developed strategy on the impacted passengers.
4. The study introduces a practical procedure for finding feasible SSPs. Specifically, the developed framework is undertaken in two stages. First, an optimization model generates the optimal (or close-to-optimal) SSPs at the trip level, followed by a pattern recognition algorithm employed to define the best possible branches (distinct groups of trips) for the transit service route. Finally, the simulation platform tests the quality of solutions on passengers' travel experiences.

The objective of this work is to develop a framework combining an optimization model with a pattern recognition approach within a simulation platform for determining the best limited-stop strategy at the planning stage.

3 Model development

The limited-stop (LS) strategy is beneficial for considering high-frequency routes, where high-demand stops are served by all route vehicles while stops with low demand are skipped by some vehicles. The LS strategy benefits mainly accrue to "through" passengers aboard a vehicle skipping a given stop and those who will board at a downstream stop served by the vehicle. However, this strategy has an adverse effect on passengers who want to board at the skipped stop, since they have to wait for the next available vehicle or walk to the adjacent stop. In this section, we introduce an optimization model to develop an optimum bus service operation for a given bus route. The proposed model seeks to modify the bus timetable by optimally reassigning local trips to operate as LS service. Therefore, the LS service operates in parallel to the local service with the overall objective of minimizing the total passenger travel time.

3.1 Assumptions

Usually, the LS strategy is constrained by the passengers who want to alight the vehicle at a candidate stop for skipping. It is assumed that passengers aboard

a vehicle are aware of a vehicle’s specific service pattern before boarding. It is also assumed that any LS introduction does not modify rider behavior in terms of mode choice and route choice. The model assumes the availability of origin-destination (O-D) demand matrices for routes under consideration. Such data can be derived from AFC systems with or without tap-off requirements. Because the LS strategy is treated here at the planning phase, the route information and any possible changes as a result of LS strategy deployment is assumed to be known to passengers and the service operator and fixed over the period concerned. It is assumed that passengers have knowledge about their choices and make rational decisions based on shortest travel time. It is also assumed that stops where passengers want to transfer cannot be skipped.

3.2 Model formulation

Consider a single corridor with $\mathbb{N} = \{1, 2, \dots, |\mathbb{N}|\}$ stops. A route is made up of a collection of “trips” in each direction. Each trip k represents a single vehicle run, with specific departure times at the stops along the route. The notations used in the model formulation are defined as follows:

- \mathbb{N} set of stops, with $n \in \mathbb{N}$;
- \mathbb{K} set of vehicle trips, with $k \in \mathbb{K}$;
- $b_{k,n}$ the number of boarding passengers at stop n for trip k ;
- $a_{k,n}$ the number of alighting passengers at stop n for trip k ;
- $l_{k,n}$ passenger load on trip k at stop n ;
- $c_{k,n}$ vehicle running time on trip k at stop n which is defined as the vehicle running time between stops $n - 1$ and n ;
- $d_{k,n}$ vehicle dwell time on trip k at stop n ;
- H_k vehicle planned headway time on trip k ;
- $S_{k,n}$ binary decision variable, 1 if trip k at stop n is skipped, and 0 otherwise.

3.2.1 Terms and parameters

While the benefit of any LS strategy is due to the in-vehicle time saved for through passengers, it increases the travel time for passengers who want to alight or board at the skipped stops. To consider the impact of an LS strategy on the waiting times of different passenger groups the following terms and parameters are defined.

Let $\Theta_{k,n} = \sum_{i=n+1}^{\mathbb{N}} b_{k,i}$, be the total number of passengers boarding trip k downstream of stop n . If an LS strategy requires trip k to skip-stop n , while keeping all other elements of the route schedule unchanged, trip k will have a shorter travel time to all downstream stops and the passengers who normally board this trip (i.e., k) from those stops can experience a shorter headway (between trips $k - 1$ and k), hence shorter wait times.

On the other hand, the passengers' travel times will change by applying the LS strategy at some previous stops (i.e., upstream of stop n). These time changes will also change the route headway and should be considered in the model formulation. Let $\Lambda_{k,n} = \sum_{i=1}^{n-1} (S_{k,i}d_{k,i})$ be the time change due to the introduction of an LS strategy at the upstream stops. That is, a bus will arrive at stop n (to be a skipped stop) in a shorter time. Therefore, $\Lambda_{k,n}$ is an additional wait time for the out-of-vehicle passengers of trip k at stop n due to the applied LS strategy at upstream stops, resulting in longer wait time. Consequently, $\Lambda_{k,n}$ will be added to the route headway, which increases the time interval for the next service to visit this stop.

In order to generalize the formulation and encompass different situations, it is also important to consider those passengers who are affected by the use of the LS strategy on the previous trip ($k - 1$) at stop n and could not board the vehicle and need to take the next available trip. Let $\Gamma_{k,n} = S_{k-1,n}b_{k-1,n}$ be the number of extra boarding passengers on trip k at stop n due to the applied strategy to the previous trip. These passengers will experience longer waiting times and will board the next available vehicle. Figure 1 illustrates the defined parameters.

One of the challenges in using any strategy is evaluating how these strategies affect the passengers' travel experiences between their origins and destinations. Recent advancements in data and information technology, such as smart card data, have made it possible to obtain highly disaggregate O-D passenger demand (stop to stop) for each route. Given passenger demand heterogeneity, using disaggregate O-D demand matrices offers more prudent modeling possibilities. This O-D information has been considered in the formulation as follows. First, a passenger-travel time O-D matrix ($PTOD$) for each trip is generated from the O-D matrix of passenger demand and the O-D matrix of travel time. A ratio based on those passengers who are affected by applying the LS strategy is computed with the numerator referring to the O-D matrix of affected passenger-travel time and the denominator referring to the O-D matrix of total passenger-travel time. Let $\Psi_{k,n}$ be the defined O-D ratio of trip k at stop n . As the following expression shows, the value of $\Psi_{k,n}^a$ depicts

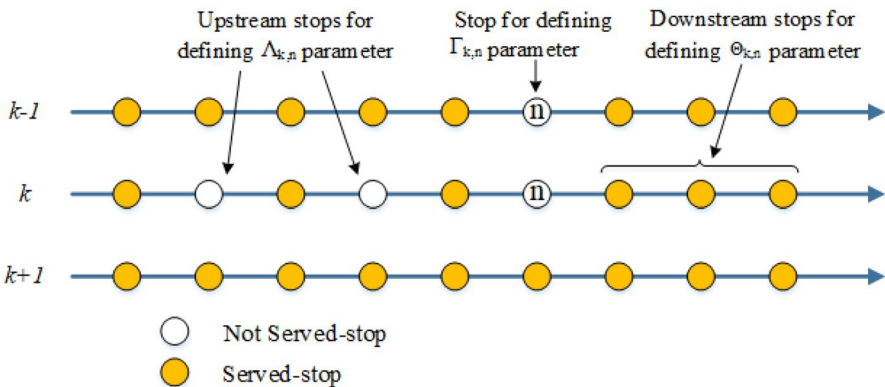


Fig. 1 The parameters of the LS strategy

the fraction of the travel times by those passengers who normally alight at the stop that is not served (i.e., skipped) and should instead use the alternative stops.

$$\Psi_{k,n}^a = \frac{\sum_{i=1}^n PTOD_{k,i,n}}{\sum_{i=1}^{|\mathcal{N}|-1} \sum_{j=i+1}^{|\mathcal{N}|} PTOD_{k,i,j}} \tag{1}$$

Similarly, for the alighting passengers, it is possible to consider the impact of O-D passenger travel time on those passengers who wanted to board at non-served stops using the following formula:

$$\Psi_{k,n}^b = \frac{\sum_{i=1}^n PTOD_{k,n,i}}{\sum_{i=1}^{|\mathcal{N}|-1} \sum_{j=i+1}^{|\mathcal{N}|} PTOD_{k,i,j}} \tag{2}$$

3.2.2 Penalty function

To consider the drawback of any LS strategy on the impacted passengers and to examine different alternatives for them, a set of penalty functions is introduced. The penalty function is a crucial element of the objective function that represents the passengers' inconvenience. If the LS strategy is found to cause more inconvenience than benefit, it will not be recommended. Figure 2 illustrates these penalties.

- (a) *Walking time penalty*: Let the “upstream served-stop” be the last stop served before the skipped-stop, and the “downstream served-stop” be the first stop

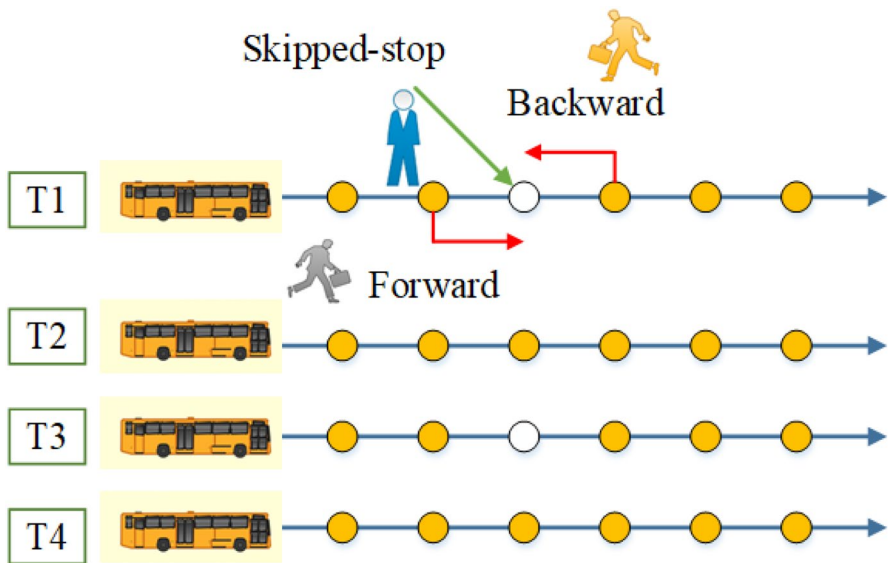


Fig. 2 The penalty time components

served after. It is possible to categorize two types of penalty functions to take into account the LS strategy's impact on disadvantaged passengers.

In the first type, passengers who want to alight at the non-served (i.e., skipped) stop but will instead alight at the downstream served stop will need extra time to reach their destination; this is termed forward time penalty (P^F) and formulated as follows:

$$P_{k,n}^F = (f - 1) \sum_{i=1}^n c_{k,i} \prod_{q=i}^n S_{k,q} \quad \forall q, n \in \mathbb{N}, \{1 \leq q < n\} \quad (3)$$

where f is the ratio of an average bus speed over the average walking speed of a pedestrian. The impacted passengers are going from stop q (the stop that passengers decide to alight) to stop n (the stop that is not served by trip k). These passengers will not experience a prolonged bus running time along with the segments upstream of the non-served stop.

The second type of penalty, pertaining to passengers who alight at the “downstream served stop”, will need to return to their desired stop n , with this additional time given the term backward time penalty (P^B), and formulated as follows:

$$P_{k,n}^B = (f + 1) \sum_{i=n+1}^{\mathbb{N}} c_{k,i} \prod_{q=n}^i S_{k,q} \quad \forall q, n \in \mathbb{N}, \{n < q < |\mathbb{N}|\} \quad (4)$$

Note that these passengers will experience additional bus running time beyond the non-served stop because of staying in the bus and alighting at the next stop. We thus obtain the walking time penalty as follows:

$$P_{k,n}^{walk} = \min(P_{k,n}^F, P_{k,n}^B) \quad (5)$$

- (b) *Waiting time penalty*: The other possibility for passengers, which occurs when the walking distance is relatively long, is to wait for the next bus to reach their desired stops. This wait time could be associated with the effect of the LS strategy used on the previous trip. Thus,

$$P_{k,n}^{wait} = H_k + \sum_{i=1}^{|\mathbb{N}|} S_{k-1,i} d_{k-1,i} \quad (6)$$

Using passengers' waiting and walking times to construct the penalty function of those who wanted to use non-served stops, the total alighting penalty time (P^{alight}) takes the form:

$$P_{k,n}^{alight} = \min\{P_{k,n}^{walk}, P_{k,n}^{wait}\} \quad (7)$$

- (c) *Stop-waiting penalty*: Similarly, those who wanted to board at the stops that are not served (skipped) should wait for the next vehicle. Therefore, the formulation of penalty time of this group of passengers is associated with the upstream stops, where the LS strategy is used and can be written as follows:

$$P_{k,n}^{Swait} = H_k + \Lambda_{k,n} \tag{8}$$

Based on the above equations, the effect of the LS strategy on the change in the total travel time with respect to trip k and stop n can be expressed in Eq. (9). Note that the passengers’ wait time components are weighted by δ_1 for out-of-vehicle wait time and δ_2 for the in-vehicle wait, which can be set by the empirical data.

$$LS_{k,n} = S_{k,n} \left[(1 + \Psi_{k,n}^a) a_{k,n} P_{k,n}^{alight} + \delta_1 (1 + \Psi_{k,n}^b) b_{k,n} P_{k,n}^{Swait} + \delta_1 \Gamma_{k,n} (2H_k + \Lambda_{k,n}) - \delta_2 d_{k,n} (l_{k,n} + \Theta_{k,n}) \right] \tag{9}$$

The objective of the model is to minimize the total passenger travel time, which is computed as the total possible increase in out-of-vehicle passengers’ wait time and decrease in in-vehicle passengers’ travel time by the introduction of the LS strategy. That is, the objective function will be defined based on changes in the total passenger travel time. Note that if the optimization returns no passenger saving time, it means the LS strategy will not occur, resulting in no changes to the existing local services.

Consequently, the objective function for the proposed model to find the optimal LS strategy can be written as:

$$\min \sum_{k \in \mathbb{K}} \sum_{n \in \mathbb{N}} LS_{k,n} \tag{10}$$

3.2.3 Constraints

A set of constraints needs to be considered in the model to ensure the LS strategy meets the service requirements under different conditions. The following constraints ensure that the values of LS constitute valid service.

$$S_{k,1} = 0 \tag{11}$$

$$S_{k,n} Z_{k,n} = 0 \quad \forall Z_{k,n} \in \{0, 1\} \tag{12}$$

where $Z_{k,n}$ is 1 when n is a special stop that cannot be skipped, otherwise zero.

$$S_{k,n} Y_{k,n} = 0 \quad \forall Y_{k,n} \in \{0, 1\} \tag{13}$$

where $Y_{k,n}$ is 1 when n is a transfer stop that cannot be skipped, otherwise zero.

$$S_{k,n} + S_{k,n+1} \leq 1 \tag{14}$$

$$S_{k,n} + S_{k+1,n} \leq 1 \quad (15)$$

$$v = I_{k,n} + b_{k,n} - a_{k,n} \leq v^{max} \quad (16)$$

Constraint 11 ensures that the first stop is served. Constraints 12 and 13 guarantee that important stops, such as transferring stops and some special stops with amenities cannot be skipped. Constraint 14 ensures that two consecutive stops cannot be skipped. Constraint 15 ensures that the LS strategy cannot be applied to consecutive trips at the same stop. The reason behind designing constraints 14 and 15 is to reduce the negative impacts on people disadvantaged by skipping the stop. Skipping consecutive stops on a trip or skipping the same stop on consecutive trips will increase the chance of imposing extra delays on those passengers who want to use that specific stop or trip. Also, these constraints enable the flexibility for passengers to have the additional option of walking to the adjacent stops. Therefore, the constraints are designed to minimize the negative impact of walk time to/from stops and waiting time with the introduction of the LS strategy. This would also support the assumption of study that mode and route choices remain unchanged.

Constraint 16 is defined to check that the bus capacity meets the available demand when the bus departs the stop. The value of v indicates the capacity of a bus. When the vehicle capacity is reached ($v = v^{max}$), those denied passengers will have to wait for the next available vehicle. It is assumed that passengers will wait at their stop until a vehicle arrives (no one leaves the system without taking the first vehicle to arrive). Therefore, the number of passengers prevented from boarding due to capacity constraints corresponds to $\Gamma_{k,n}$, which is defined in the previous section.

4 Model implementation

4.1 Modelling process

As the limited-stop service design optimization process was developed to harness the emerging wealth of sensor data produced by modern-day transit systems, the method expects detailed input data on the movement of both transit users and service vehicles on a trip-by-trip basis. Emerging transit demand data sources are not yet at a level to provide the detailed input data required in this study, particularly for vehicle trip-level O-D volumes, but continuous advancements and adoption of AFC technologies can eventually provide such data. Another challenge stems from the need for an adequate representation of the complex dynamics of transit route actors (vehicles, stops, and passengers) in order to evaluate the impact of any proposed service changes properly. To address these challenges, an agent-based simulation model is firstly constructed using available sources of information to generate random input data for the proposed optimization model. Subsequently, the optimization model is used to find the optimal value of the decision variable for each trip and stop. Although these optimized solutions are theoretically sound, it would be difficult to implement a different LS pattern for every single bus trip in real-world

practice. Therefore, a post-optimization analysis using a pattern recognition algorithm is carried out in order to group trips into a reasonable number of distinct service patterns. Finally, more simulation runs are made, containing the new service pattern determined by the optimization program, and using a post-optimization procedure so as to finalize the results attained by the model. Figure 3 illustrates the general structure of the limited-stop service modelling framework.

4.2 Model data for simulation

The simulation model adopts inputs and parameters from real-life data so that the transit network can be considered under stochastic conditions. This study utilizes a mesoscopic simulation, namely the Nexus simulation platform previously developed at the University of Toronto (Srikukenthiran and Shalaby 2017). Nexus is an agent-based simulator with an integrated transit assignment, allowing for detailed metrics on both service and user experience. Initially developed to enable a highly detailed simulation of large-scale regional multi-modal networks, the platform facilitates the dynamic integration of multiple simulation software, each modelling different aspects of a transit system (rail, surface transit, and stations). Transit users, as agents, are transferred between the various software components as they travel through the network, with their overall route determined separately by a path choice decision-making component.

Since this study is focused on urban transit network service optimization, the utilization of Nexus was limited to its surface transit simulator in Nexus, which is described in Srikukenthiran and Shalaby (2017). The network is constructed dynamically using information provided in the Google Transit Feed Specification (GTFS), which, amongst others, describes stops, routes, schedules, and geographic paths taken by vehicles. Using census and GTFS data provides a feasible approach for measuring the gaps between PT supply and demand (Kaeoruean et al. 2020). Vehicles make a set of trips, with information on vehicle blocks also provided in GTFS. While for the initial Nexus simulation model, vehicle speeds were set to maintain schedule, this feature was adjusted in this paper,

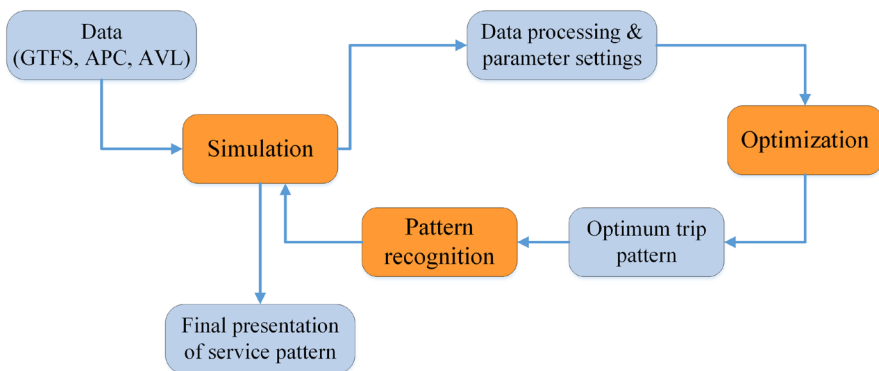


Fig. 3 The limited-stop service modelling framework

incorporating a random-forest speed prediction model which utilizes historical automatic vehicle location (AVL) data available publicly; details on data preparation and speed prediction can be found in Wen et al. (2018). Dwell times are set based on a formula prescribed by Vuchic (2005) with an assumed boarding and alighting time per passenger of 3.0 and 1.5 s, respectively. Transit user agents traverse this surface network in Nexus as they make their way from origin to destination using a quickest-path (based on published schedules with a penalty for transfers) path-finding method.

As an agent-based microsimulation software, Nexus produces highly detailed output for all movements of both surface vehicles and transit user agents, with aggregate values matching closely transit sensor data (AFC, APC, AVL). Logs are kept of vehicle interactions at all bus stops; recorded information includes the vehicle and trip number, arrival and departure times, incoming vehicle load, and the total number of agents (passengers) boarding, alighting, and unable to board due to capacity limitation. Nexus also keeps for each agent a record of its individual trip, with information including departure and arrival times, the boarding and alighting stop of each segment (both in-vehicle and transferring), the actual vehicle trip, and any walking, waiting and in-vehicle duration.

In the simulation, passengers arrive randomly following a Poisson distribution with mean $\lambda_{k,n}$ at each stop. The simulation model considers the vehicle travel times between two adjacent stops as a random variable with a log-normal distribution; $LN(\mu, \sigma^2)$ that has been commonly used in previous studies (Wen et al. 2018; Nesheli and Ceder 2015). Results for ten simulation days (using different random seeds) were produced. This data was processed to act as input for the optimization.

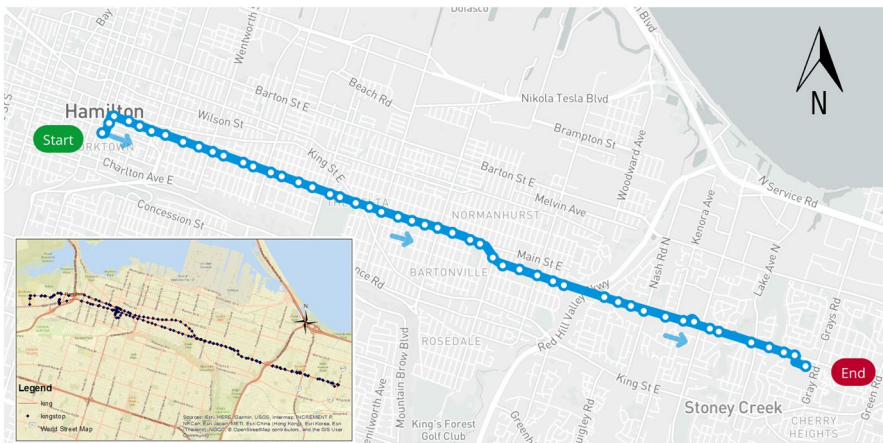


Fig. 4 Bus Route 1 (King), in Hamilton, Canada

5 Case study

An analysis was carried out based on a case study of the Hamilton Street Railway (HSR) transit network to evaluate the application of the developed limited-stop service method. The HSR serves over 80,000 passengers daily. One of the major bus routes of the network, Route 1-King, was selected for our case study (see Fig. 4). The route has two branches with a daily ridership of over 12,000 passengers. The first branch, which is the main one, serves between Hamilton GO Centre and Eastgate Square. The second branch runs between University Plaza and Hamilton GO Centre. For this study, the first branch is used. The route consists of 45 stops and runs in a west-east direction in lower Hamilton from the Hamilton GO Centre in the west to the Eastgate Square in the east. The route length is 12 km in each direction, and the planned headway is 10 minutes during the morning peak.

A weekday morning peak-period (7:00–9:00 AM) Nexus simulation model of the Greater Toronto and Hamilton Area (GTHA) region was utilized. A recent study by Rashedi et al. (2021) used a passenger survey of GO rail transit of GTHA data to investigate the impact of access cost on transit mode switching behavior of commuters in the GTHA. Being a proof-of-concept model, the 2016 GTFS data was used to construct the regional transit system, while transit demand was provided by the 2011 Transportation Tomorrow Survey (TTS). The TTS is a regional travel survey conducted every five years in the GTHA, which in part asks for a working day travel diary, including detailed route information for those using transit. The data retrieval system for the TTS allows for the extraction of O-D matrices in 10 min departure time slices; these matrices were used to generate an agent population for the region. Additionally, APC count data was used to validate the model output at the aggregate level. This larger regional model was used in place of a more focused single-route model. This was done given the prior construction/and availability of the regional model, as well as the transit demand O-D being available only at the regional level;

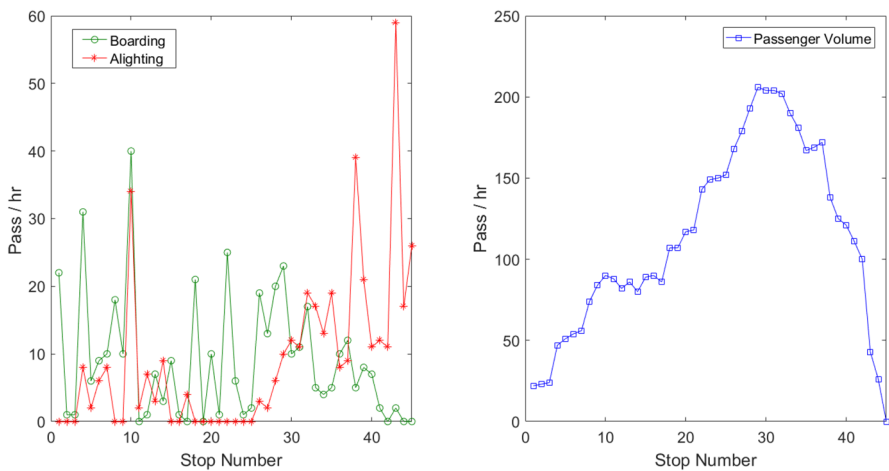


Fig. 5 Passenger boarding, alighting and volume profiles at stops for Bus Route 1

Table 1 Summary of optimization

Objective function	Computational times	Integer gap of convergence
– 15,562.31 (passenger seconds)	173 (s)	Less than 1%

Table 2 Optimal result for LS service

Trip	W-E direction LS pattern	Net time saving (passenger seconds)
1	0000101010100101010000 101 01010000101000010100	1872
2	000000000010010101010 1001 000001001010001001000	3115.54
3	0000101010101000000000 101 00000000100100010000	1710.15
4	0101000001010000010000 1000 100000001000100000	1640.77
5	0000000000101010101010 101 01000010100000001000	4315.02
6	01001000000100010001000 1000 00001000000000100000	516.05
7	0010010101001000101010 101 01010101010101000000	1784.78
8	0101001010100000000000 101 01010101010000000000	312
9	0000000000000000000000 101 00000000000000000000	296

1 = Non-served stop; 0 = served stop; stops 23, 24, and 25 are highlighted in bold

this allowed for the avoidance of assumptions to be made for travel to/from Hamilton. Figure 5 presents the passenger demand profile of one day for Route 1-King, at each stop in the Eastbound direction produced from the simulation output. The framework was executed for the specified peak-time period of 10 working days in order to accumulate enough data required for the analysis.

6 Analysis and results

Modifications of the simulation model for consistency with the initial simulation model are made for this study, with the implementation of the aforementioned random forest speed prediction model for more realistic movements of buses in the case study area. As a baseline, the in- and out-of-vehicle wait times contribute the same marginal time; however, different relative wait ratios (in-vehicle over out-vehicle times) are tested in the sensitivity analysis. The average walking speed of a pedestrian is 5 km/h, and the average bus speed is 25 km/h, according to the study data. The service schedule headway is fixed at 10 minutes over the time period. The bus capacity is 60 pass/veh, with 40 seated and 20 standing.

The optimization problem is coded in Python, and “Bonmin”—an open-source optimization solver (Bonami and Lee 2007)—is used to solve the problem. Bonmin features several algorithms, such as the branch-and-bound algorithm. The branch-and-bound allows to efficiently solve the problem to determine the value of each decision variable and find an optimal solution. The computations are carried out on

a Mac OS X machine with a 2.4 GHz Intel Core i5 processor and 8 GB of RAM. Table 1 summarizes the total objective function value and the computational times for the case study. The solution is obtained after applying the branch-and-bound method within an optimality gap of solution (i.e., less than 1% gap).

The detailed results obtained by the developed methodology are shown in Table 2. Each trip shows the optimum service pattern. In particular, it can be seen which stops are the most appropriate candidates for applying the LS strategy. The patterns illustrate that Stops 23, 24, and 25 (highlighted in bold in Table 2), which are located in the middle of the route, are the most non-served stops for the majority of the trips. The benefit of using this strategy yields saving times for each trip, as indicated in Table 2. As can be seen, the results also show a somewhat superior outcome for trip #2 and #5 than for other trips, with a 20% and 28% improvement in total passenger travel time, respectively. This is because most of the passengers of these two trips board at the start of the route and alight near the end of the route. In other words, the number of onboard passengers in the middle of the route is large and very few passengers board or alight at those stops, which are located in the middle section of the route. It is increasingly evident that due to special passenger O-D patterns, the method was able to customize special service patterns accordingly. It is to note that in this case study, the vehicle capacity is not reached. It is also worth noting that the proposed framework is capable of capturing the LS effects on user behavior, specifically regarding route choice. However, based on the scope of this study we assumed no change to travel behavior (i.e., no mode or route shift). Further analysis and investigation of the route choice and mode choice behavioral changes to the LS strategy are recommended for future study.

6.1 Stop pattern recognition

The results of Table 2 demonstrate that a group of trips (i.e., branch) may follow a similar pattern. As implementing a trip-by-trip skip-stop operation would not be practically feasible, finding a small number of LS branches is necessary. Consequently, further analysis using a pattern recognition approach for the binary result data set is carried out. We select a binary pattern [0 1] (meaning a served stop followed by a non-served stop). The following algorithm 1 illustrates the process.

In this study, the frequency of occurrence (μ) of the defined binary pattern for each trip is depicted in Fig. 6. In the figure, those stops that follow the defined pattern are selected, meaning that the x-axis shows stops where the 0 of the pattern occurs, implying that the indicated stop will not be served (or is equal to 1 in the pattern). This will lead to reducing the search space needed to find the best combination of stops and the number of branches by removing those stops that are not candidates for an LS strategy. As the figure illustrates, there are two different patterns of using stops that are recognizable from the peak points. Peaks with a value of three or more are considered in this study, since the most frequent situation happens in the case of three or more trips. The first pattern is mostly shared among odd-numbered trips, while the second pattern is mostly shared among even-numbered trips. We name these two patterns as “Branch A” and “Branch

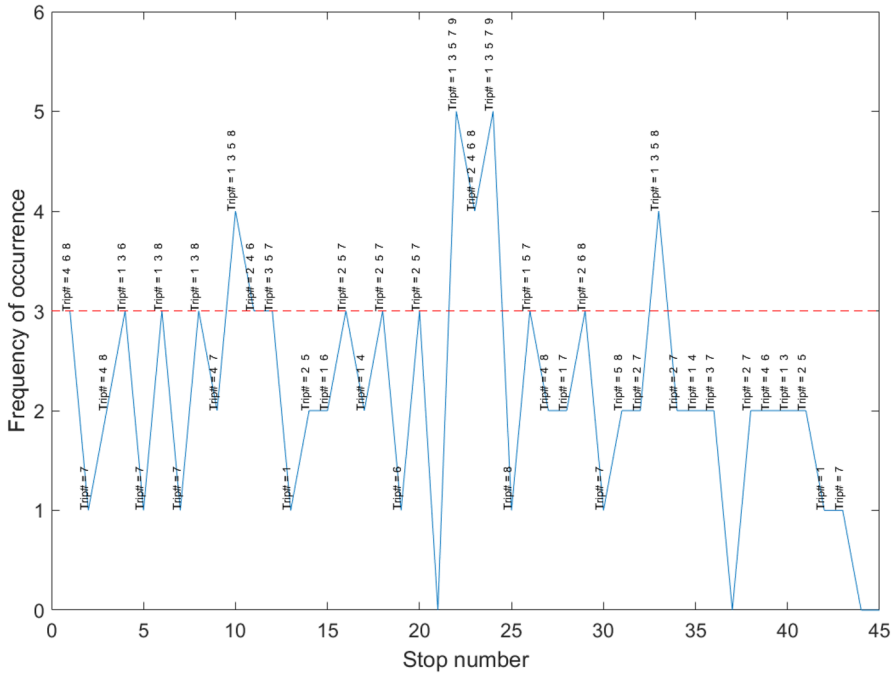


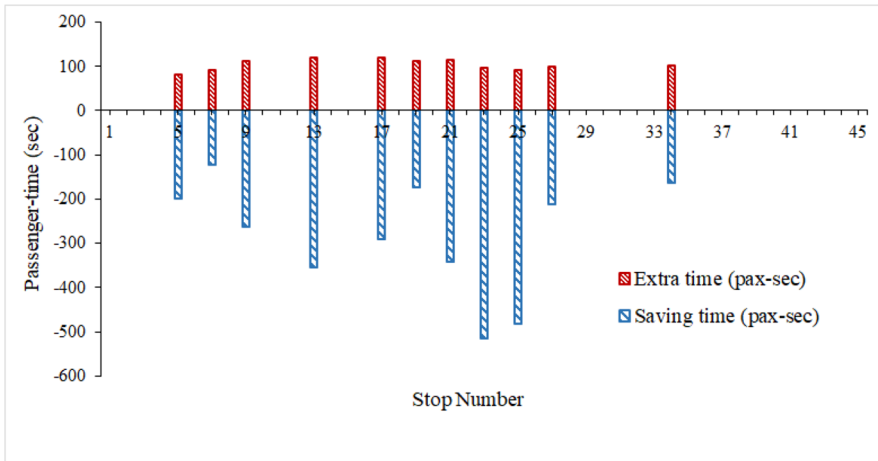
Fig. 6 Stop pattern recognition according to the frequency of occurrence of the binary pattern [0 1] for the given bus trips

B”. Now, we should assign suitable stops to each branch utilizing the information of this figure. As an example, suppose we assign the most frequent stops (22, and 24) to Branch A and the second frequent stop (23) to Branch B, respectively. This indicates that Branch A serves stops 22 and 24 while skipping stops 23 and 25.

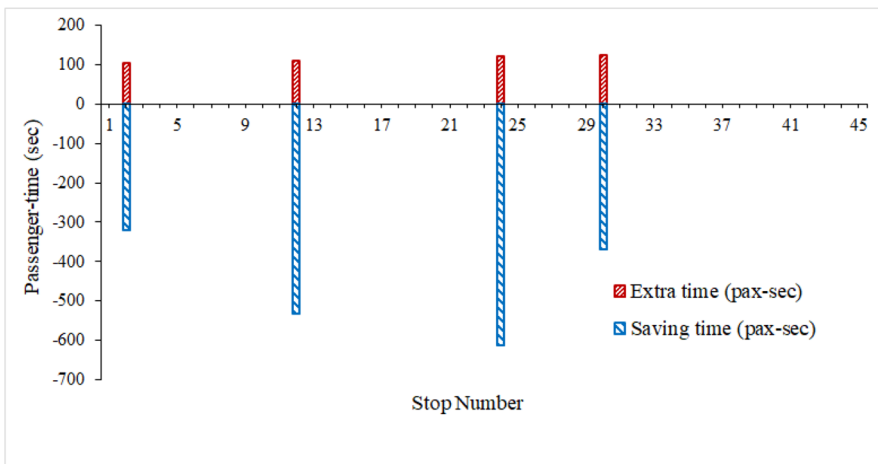
Similarly, Branch B serves stop 23 and skips stop 24. We then continue the same procedure for the other peaks (i.e., frequency ≥ 3) with careful attention to stop assignment being due to the highest number of odd or even trip numbers that are sharing the stop where the pattern occurs. Consequently, the service and stop pattern assigned to both Branch A and B is obtained as follows, as shown in

Table 3 Summary of feasible branches

Branches	LS pattern	Average net time saving (passenger seconds)
A	“00001010100010001010 1010101000000100000 000000”	1958.91
B	“01000000000100000000 0001000001000000000 000000”	13,725.65



(a) Branch "A"



(b) Branch "B"

Fig. 7 The effect of applying an LS strategy on passenger travel time

Table 3. Each trip (of all 9 trips in our case study) then follows patterns in Branch A and B sequentially.

The simulation served as a validation of the optimization results under a stochastic framework. When the optimization framework recommends a solution to deploy in the simulation, it is possible to evaluate the LS strategy’s efficiency and its effects on passengers’ travel experiences. Figure 7 shows the effect of applying the LS strategy on the passengers’ travel times along branches A and B from the simulation run. As explained in the previous section, the benefits of implementing the LS strategy are accrued mainly by the “through” passengers, who are already onboard the bus, and those boarding downstream. On the other

hand, it has an adverse effect on passengers who want to alight or board at the skipped stop, inflicting extra time on such passengers. As Fig. 7a and b show, the maximum saving of passengers' travel times occurs in the middle of the route where Stops 23, 24, and 25 are located. According to the optimization formulation, the optimal solution will naturally tend to maximize the potential saving in passengers' travel times by identifying the stops where the total passenger travel time can be minimized. Conversely, at stops where the extra travel times experienced by the disadvantaged passengers outweigh the travel time savings of the other passengers, the LS strategy will not take place, rendering no change to the passengers' travel times. It is to note that the proposed model is generic and can offer the best solution based on the route characteristics.

Algorithm 1: Pattern recognition algorithm.

Data: Optimum service pattern for each trip

Input: The binary, $[0,1]$, pattern for each stop along the route

Result: Best possible branches for the SSPs

begin

Step 1: Find stop indices for each trip;

For all trips: record the location (i.e., stop) of the binary pattern.

Step 2: Find the frequency of indices;

For all trips: record the frequency of the binary pattern.

Step 3: Assign the trips to the binary pattern;

For n in stop: define vector $X(n)$ and assign the trip # to the stop # on the basis of the same binary pattern recognition.

Step 4: Building new service patterns;

 define branch "A", branch "B", ..., branch "I" **where** $A \cap B \dots \cap I = \emptyset$;

while n in Stops $< |\mathbb{N}|$ **do**

 Find new pattern;

if frequency of $([0, 1]) \geq \mu$ in $X(n)$ **then**

if The cell value in $X(n)$ follows the pattern of branch "A" **then**

 Save stop # in branch "A";

else if The cell value in $X(n)$ follows the pattern of branch "B"

then

 Save stop # in branch "B" ;

 :

else

 Save stop # in branch "I" ;

end

end

end

end

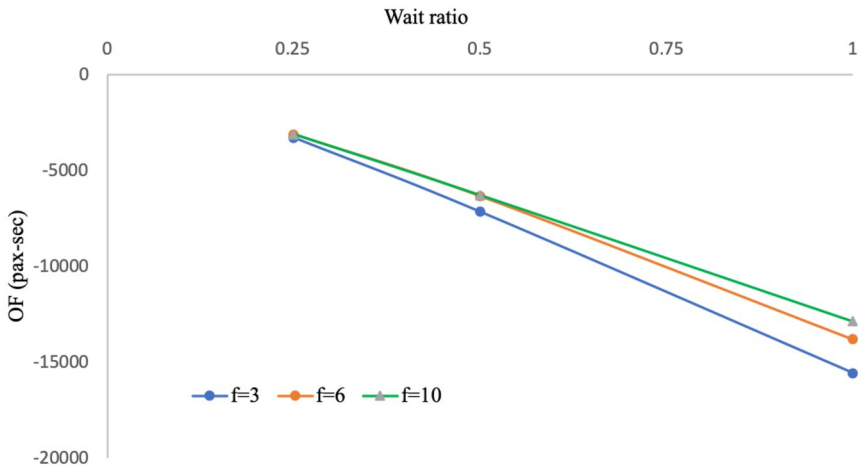


Fig. 8 Sensitivity analysis of wait time factor and speed ratio

6.2 Sensitivity analysis

A sensitivity analysis of the objective function (OF) value with weightings of passenger wait time and speed ratio is shown in Fig. 8. The marginal burden or disutility of out-of-vehicle waiting time is perceived to be more burdensome than in-vehicle travel time (Reed 1995). Thus, the wait time factor is computed as in-vehicle over out-of-vehicle times (δ_2/δ_1). A wait time factor less than 1 means the passengers who are not through passengers are likely to experience a higher disutility of wait time with less saving time for the system.

The other factor when considering the application of an LS strategy is the ratio of the average speed of a bus and the average walking speed of pedestrians (f), which takes the walking penalty time into consideration. An adjustment of this ratio means higher or lower passenger travel time than the base data. The bus speed variation also leads to a vehicle running time variation. That is, a higher ratio means higher bus speed and less walking time penalty. The results of Fig. 8 reveal a somewhat superior saving time for a lower speed ratio with a higher wait ratio. In particular, the higher saving time in the objective function is attained in a lower walking time penalty and greater value of in-vehicle wait time. In other words, the sensitivity analysis depicts that the effectiveness of the proposed optimization model is high when the speed of the bus system is low (e.g., because of traffic congestion or fluctuations of passenger demand), and the in-vehicle travel time ratio is high. This is especially the case when the large group of passengers travels a longer distance, and due to the low speed of the system, they will experience large in-vehicle travel time.

7 Conclusions

This paper presented a new mathematical programming model which aims at finding the best candidate route stops for a limited-stop strategy. The developed model is formulated in two stages, solving the optimization problem, and post-optimization analysis. Within this study, we analyzed the benefits of implementing such a strategy for the purpose of minimizing the total passenger travel time. The adopted approach consists of three steps: optimization, post-optimization, and simulation. An agent-based simulation framework is built to represent a real-life example and to generate reasonable input data for the proposed optimization model. Subsequently, the post-optimization procedure using the optimization outputs enabled us to propose an appropriate service pattern at the branch level. Finally, the simulation platform tests the quality of solutions on passengers' travel experiences. Based on the formulations for optimization and the post-optimization process, we developed an efficient algorithm for finding the best candidate for the route service pattern. In order to furnish a practical solution for selecting a feasible service pattern, a new pattern recognition method as part of the post-optimization procedure is developed. The model was applied in a case study of one route in Hamilton (Ontario), and a detailed discussion regarding the optimum service stop pattern is provided.

A sensitivity analysis of the objective function value with weighting passenger wait time and speed ratio is conducted. The computational results demonstrate the effectiveness of saving time for the greater value of in-vehicle wait time.

The main contribution of the proposed framework is to offer PT agencies a ready-to-use tool for designing and evaluating the limited-stop strategy to real-world routes. For example, the cost per kilometer drops since buses reduce the time they require to accelerate, decelerate, and open their doors. Also, the service cycle time would be shorter, resulting in reducing the fleet size. However, in this study, we investigated the LS strategy to improve the quality of existing services by matching supply and demand in a better way.

Ongoing research related to this study includes validating some of the assumptions, analyzing travel behavior (route and mode choices), developing a family of stop and service pattern strategies (such as semi-express service), and utilizing more accurate simulation models (including learning-based transit assignment with a more realistic passenger arrival distribution at stops) to measure the transit system's key performance indicators.

Acknowledgements This research was supported by NSERC, OCE, Trapeze and SOSCIP. We also gratefully acknowledge the data support from the Trapeze group and the Hamilton Street Railway service.

References

- Bonami P, Lee J (2007) Bonmin user's manual. *Numer Math* 4:1–32
- Cao Z, Ceder A, Li D, Zhang S (2020) Robust and optimized urban rail timetabling using a marshaling plan and skip-stop operation. *Transportmetrica A Transp Sci* 16(3):1217–1249

- Chen J, Liu Z, Zhu S, Wang W (2015) Design of limited-stop bus service with capacity constraint and stochastic travel time. *Transp Res Part E Logist Transp Rev* 83:1–15
- Chiraphadhanakul V, Barnhart C (2013) Incremental bus service design: combining limited-stop and local bus services. *Public Transp* 5(1):53–78
- Eberlein XJ (1995) Real-time control strategies in transit operations: models and analysis. PhD thesis. Department of Civil and Environmental Engineering, Massachusetts Institute of Technology
- Eberlein XJ, Wilson NHM, Bernstein D (1999) Modeling real-time control strategies in public transit operations. *Computer-aided transit scheduling*. Springer, Berlin, pp 325–346
- Furth PG (1986) Zonal route design for transit corridors. *Transp Sci* 20(1):1–12
- Hart N (2016) Methodology for evaluating potential for limited-stop bus service along existing local bus corridors. *Transp Res Rec J Transp Res Board* 2543(1):91–100
- Jordan WC, Turnquist MA (1979) Zone scheduling of bus routes to improve service reliability. *Transp Sci* 13(3):242–268
- Kaerouean K, Phithakkitnukoon S, Demissie MG, Kattan L, Ratti C (2020) Analysis of demand-supply gaps in public transit systems based on census and GTFS data: a case study of Calgary, Canada. *Public Transp* 12(3):483–516
- Larrain H, Muñoz JC (2016) When and where are limited-stop bus services justified? *Transportmetrica A Transp Sci* 12(9):811–831
- Leiva C, Muñoz JC, Giesen R, Larrain H (2010) Design of limited-stop services for an urban bus corridor with capacity constraints. *Transp Res B Methodol* 44(10):1186–1201
- Nesheli MM, Ceder AA (2015) A robust, tactic-based, real-time framework for public-transport transfer synchronization. *Transp Res C Emerg Technol* 60:105–123
- Nesheli MM, Ceder AA (2017) Real-time public transport operations: library of control strategies. *Transp Res Rec J Transp Res Board* 2647(1):26–32
- Rashedi Z, Hasnine MS, Habib KN (2021) Modelling second-best choices from the choice-based sample: revelation of potential mode-switching behaviour from transit passenger surveys. *Public Transp*. <https://doi.org/10.1007/s12469-021-00275-z>
- Reed TB (1995) Reduction in the burden of waiting for public transit due to real-time schedule information: a conjoint analysis study. In: *Vehicle navigation and information systems conference, 1995. proceedings. In conjunction with the Pacific Rim TransTech Conference. 6th International VNIS: 'A Ride into the Future'*, IEEE, pp 83–89
- Soto G, Larrain H, Muñoz JC (2017) A new solution framework for the limited-stop bus service design problem. *Transp Res B Methodol* 105:67–85
- Srikukenthiran S, Shalaby A (2017) Enabling large-scale transit microsimulation for disruption response support using the Nexus platform. *Public Transp* 9(1):411–435
- Sun A, Hickman M (2005) The real-time stop-skipping problem. *J Intell Transp Syst* 9(2):91–109
- Tétreault PR, El-Geneidy AM (2010) Estimating bus run times for new limited-stop service using archived AVL and APC data. *Transp Res A Policy Pract* 44(6):390–402
- Ulusoy Y, Chien S, Wei CH (2010) Optimal all-stop, short-turn, and express transit services under heterogeneous demand. *Transp Res Rec J Transp Res Board* 2197:8–18
- Vuchic VR (2005) *Urban transit: operations, planning, and economics*. Wiley, New York
- Wang DZ, Nayan A, Szeto W (2018) Optimal bus service design with limited stop services in a travel corridor. *Transp Res E Logist Transp Rev* 111:70–86
- Wen B, Srikukenthiran S, Shalaby A (2018) Data-driven mesoscopic simulation models of large-scale surface transit networks. *Transportation Research Board 97th Annual Meeting Transportation Research Board*
- Wirasinghe S, Vandebona U (2011) Route layout analysis for express buses. *Transp Res C Emerg Technol* 19(2):374–385