Transportation (2011) 38:169–189
DOI 10.1007/s11116-010-9287-8

Optimal design and benefits of a short turning strategy for a bus corridor

Alejandro Tirachini · Cristián E. Cortés · Sergio R. Jara-Díaz

Published online: 18 June 2010

Abstract We develop a short turning model using demand information from station to station within a single bus line-single period setting, aimed at increasing the service frequency on the more loaded sections to deal with spatial concentration of demand considering both operators' and users' costs. We find analytical expressions for optimal values of the design variables, namely frequencies (inside and outside the short cycle), capacity of vehicles and the position of the short turn limit stations. These expressions are used to analyze the influence of different parameters in the final solution. The design variables and the corresponding cost components for operators and users (waiting and in-vehicle times) are compared against an optimized normal operation scheme (single frequency). Applications on actual transit corridors exhibiting different demand profiles are conducted, calculating the optimal values for the design variables and the resulting benefits for each case. Results show the typical demand configurations that are better served using a short turn strategy.

Keywords Public transport · Short turning · Users' costs · Operators' costs · Frequency

Introduction

Daily operations of urban public transport systems have to deal with both spatial and temporal peak periods of demand that result in inefficient operation schemes when using

A. Tirachini (✉)
Institute of Transport and Logistics Studies, Faculty of Economics and Business,
The University of Sydney, Sydney, NSW 2006, Australia
e-mail: alejandro.tirachini@sydney.edu.au

C. E. Cortés · S. R. Jara-Díaz
Civil Engineering Department, Universidad de Chile, P.O. 228-3, Santiago, Chile
e-mail: ccortes@ing.uchile.cl

S. R. Jara-Díaz
e-mail: jaradiaz@ing.uchile.cl

the same vehicle supply along the entire route and over the whole day. With regard to the spatial dimension, there are several strategies to better assign the available fleet, by increasing the service frequency on the most demanded route sections in order to adjust the demand with the vehicles' supply.

Among the vehicles' assignment strategies, the most studied in the specialized literature are expressing, deadheading and short turning. The vehicles that perform expressing (Jordan and Turnquist 1979; Furth 1986; Eberlein et al. 1999) serve only one section of their route, and then they proceed without stopping until reaching either the terminal or a pre-specified zone where the service is reestablished. The deadheading strategy (Eberlein et al. 1998, 1999; Furth 1985, Ceder and Stern 1981; Ceder 2003a, 2004) is usually considered for transit corridors that present high demand on one direction of operation and low demand on the other. It consists in increasing the frequency in the peak direction by suppressing some services in the off-peak direction, that is, vehicles not in service are deadheaded back to the initial terminal of the peak direction.

In the present paper we focus our analysis on a very flexible and popular strategy called short turning, expanding its scope to account for all costs involved, users' and operators', considering station to station demand information. It consists in selecting part of the fleet to serve short cycles or loops on those route segments exhibiting high demand. This strategy has been studied in several forms by Furth (1987), Ceder (1989, 2003b), Vijayaraghavan and Anantharmaiah (1995) and Delle Site and Filippi (1998).

Furth (1987) assumes a scheduled operation scheme, in which the frequency of vehicles performing the short turn (hereafter, fleet B) is a multiple n of the frequency of those vehicles serving the entire route (hereafter, fleet A); the parameter n is known as the "scheduling mode". As many demand patterns can be satisfied by bus fleets adopting either full-length or short-turn services, schedule coordination of the two operation schemes is essential. He describes possible schedule coordination modes, and proposes algorithms for finding the schedule offset between the schemes that will balance loads. The optimization variables are the bus headway associated with fleet A, and the schedule offset between the vehicles of fleet A and those of fleet B. The process assumes as external parameters the number of cycles, the limit stations for the short turn and the bus capacities. The author focuses on several problems. First, the fleet size is minimized, which is equivalent to maximizing the vehicle headway. Next, a number of solutions depending on the turn-back points are evaluated, choosing the alternative that provides the lowest waiting time given the fleet size. Savings on fleet size are shown to depend on the offset between vehicles A and B, on whether whole-minute offsets are required and on the possibility of interlining vehicles between fleets A and B.

Ceder (1989, 2003b), by means of an aggregated demand model, proposes a two-stage optimization approach. In a first stage, the fleet size is minimized for a given demand level while in a second stage, the number of trips using the short turn is minimized in order to reduce the effect of the strategy on passenger trip times.

Vijayaraghavan and Anantharmaiah (1995) also pursue the reduction of fleet size through the insertion of express services as well as short turns for the service of some trips, where the headways and speeds of the new services are exogenously introduced and their potential benefits calculated. The authors report savings when applying their approach in terms of fleet and crew utilization, and an associated reduction in passenger travel times as well.

Delle Site and Filippi (1998) develop a more complex model, in which the short turning strategy is applied on a multi-period basis with both elastic and inelastic demand. In addition, the authors extend the work by Furth (1987) to the case of bus arrivals at stations following a Poisson distribution. In the case of elastic demand, the social benefit is maximized, whereas in the case of inelastic demand, the sum of user and operator costs is minimized. The decision variables are the limit stations of the short turn (start and end), the frequency, the capacity of vehicles (treated parametrically) and the service fare. An application that compares the strategy with a given base situation (normal operation with single frequency) shows that the strategy turns out to be beneficial only with demand patterns that exhibit pronounced peaks, reducing both waiting time and fixed operator cost (due to the operation of a smaller fleet size) and increasing the operating variable cost.

In this work we develop a short turning model taking into account that the strategy affects not only the operation costs but also passenger in-vehicle and waiting times. This approach is based on a microeconomic modeling of transit operations, where we clearly establish the users and operators cost components, to investigate how the strategy affects both parties. The analysis is restricted to a single period of operation, under the assumption that this strategy should be useful in most cases during peak demand periods only. The design variables in the formulated problem are frequencies, segment where the short turn strategy should be applied, and vehicle sizes. The model is structural, as we get optimal values of frequencies and capacities and the expected benefits from applying short turning in different cases in order to identify and support potentially effective public transport planning policy. We do not attempt to reach detailed results as, for instance, the timetabling of buses A and B, which can be done once the policy has been deemed useful.

To our knowledge, this is the first work that, under some assumptions, finds analytical expressions for optimal frequencies both inside and outside the short cycle, which are used to analyze the influence of the parameters of the problem in the final solution. Unlike the existing literature that compares the short turning strategy with a given normal operation or base case (no strategy is applied, resulting in a single frequency along the entire route), in our model the situation with normal operation is also optimized, in order to compare both cases on a fair basis. Once the analytical framework has been established, we conduct a number of experiments finding that:

- The short turning strategy can yield benefits for both users and operators at the same time.
- Cost savings are reachable not only when the demand is concentrated in the middle of a route, but also when the most loaded section is on an extreme of the line.
- Cost savings depend largely on the imbalance between trips inside and outside the short cycle.
- When the strategy yields benefits, vehicles may be smaller than those resulting for the single frequency case.
- The concentration of demand plays a fundamental role: the more concentrated the trips (the shorter the cycle is), the larger the benefits.

In Sect. 2 the short turning strategy is presented and an optimization problem is formulated, solved and compared analytically against an also optimized normal operation. In Sect. 3 we show several applications to measure the comparative benefits of the strategy under different situations. This section is crucial to better understand the behavior of the key elements in the definition of the strategy, such as optimal frequencies, capacity of vehicles, features of the demand and position and length of the short turn. Finally, in Sect. 4 the main findings are summarized and extensions of this approach are proposed.

Modeling and optimization of the short-turning strategy

Without loss of generality, we model a single linear transit line. Fleets A and B are defined as detailed above. The system contains N stations in one direction (N–1 segments), as shown in Fig. 1. The operation directions are denoted direction 1 (from station 1 to N) and direction 2 (from station N to 1). We develop a model that minimizes total cost for two situations: normal operation, where all vehicles operate along the whole route, and operation with short turning. The model is used to measure potential benefits of short turning under several demand configurations. Following Fig. 1, the decision variables are the start and end stations associated with the short turn, denoted by s_0 and s_1, the frequency f_A of those vehicles serving the whole route (fleet A) and the frequency f_B of those vehicles serving the short cycle (fleet B). Under normal operation the decision variable is the single frequency f.

The known parameters are L, length of the corridor (km); R_k, bus running time under normal service between stations k and $k+1$ including acceleration and deceleration at bus stops (min); b, marginal passenger boarding time (seg/pax); and λ_{kl}, the trip rate between stations k and l (pax/h). This disaggregated demand is assumed fixed (steady state) over the studied period, defining a trip matrix.

Additionally, the following functions and quantities are defined:

- Passenger boarding rate at station k, whose destination is among stations l_1 and l_2 inclusive (pax/h): $\lambda_k^+(l_1, l_2) = \sum_{l=l_1}^{l_2} \lambda_{kl}$
- Passenger alighting rate at station k, whose origin is among stations l_1 and l_2 inclusive (pax/h): $\lambda_k^-(l_1, l_2) = \sum_{l=l_1}^{l_2} \lambda_{lk}$
- Passenger boarding rate at station k, direction 1:$\lambda_k^{1+} \equiv \lambda_k^+(k+1, N) = \sum_{l=k+1}^{N} \lambda_{kl}$
- Passenger alighting rate at station k, direction 1:$\lambda_k^{1-} \equiv \lambda_k^-(1, k-1) = \sum_{l=1}^{k-1} \lambda_{lk}$
- Passenger boarding rate at station k, direction 2:$\lambda_k^{2+} \equiv \lambda_k^+(1, k-1) = \sum_{l=1}^{k-1} \lambda_{kl}$
- Passenger alighting rate at station k, direction 2:$\lambda_k^{2-} = \lambda_k^-(k+1, N) = \sum_{l=k+1}^{N} \lambda_{lk}$

We assume that at stations boarding and alighting process are simultaneous and that boarding dominates alighting (i.e. the boarding process is slower than the alighting process); therefore, total dwell time at stops is considered through the boarding parameter b.

Optimal design of a bus system should include the operator cost (fuel consumption, crew costs, lubricants, tires, maintenance, etc.) and the user costs (access, waiting and in-vehicle times). Both are affected when a short turning strategy is applied; the former

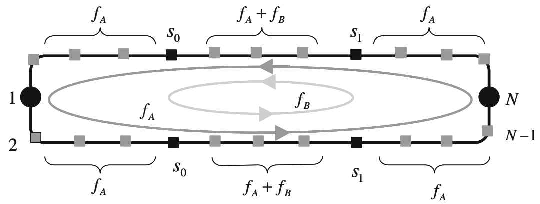


Fig. 1 Short turning strategy

because the fleet size can be adjusted when applying the strategy, and the latter because users' waiting and in-vehicle times vary due to the adjustment of frequency and changes in dwell time at stations with respect to the normal case.

In what follows we will find the operators and users cost components as functions of the design variables for two scenarios: normal and strategy-based operations. Then we will find the analytical expressions for the optimal frequencies, which in the case of the strategy will depend parametrically on the stations that define the boundaries of the short turn service. This will be the basis for a two-stages process in order to find overall optimal frequencies, optimal vehicle sizes and optimal short-turn segment.

The analytical expression for the waiting time depends on the vehicle and passenger arrival processes. Regarding bus arrivals, as in Delle Site and Filippi (1998) we consider two opposite cases: a highly controlled service where headways are kept regular along the line, and a poorly controlled system where headways vary randomly along the line due to traffic and demand variations. In the latter case (random arrivals), we will make the usual assumption that vehicle arrivals are Poisson in which headways follow a negative exponential distribution, noting that this is a simplification as the headways between successive buses are likely to be correlated due to the influence of dwell times, which in turn depends on the previous headway.

In the regular service scheme, we impose regular headways within the short turn operation as well, although this could result in larger loads observed in buses belonging to fleet A when compared with buses of fleet B. This differs from the work of Furth (1987), who specifically determines the offset between fleets A and B in order to balance passenger loads. Note that waiting times are proportional to the headway variance (Welding 1957; Osuna and Newell 1972), which is minimized when keeping regular headways as the possibility of vehicle bunching is reduced. Therefore, it is assumed that potential negative effects of operating with uneven loads are outweighed by the benefits of a regular headway scheme. Regarding passengers, we assume they arrive at stations uniformly at a fixed rate, which is a reasonable assumption in cases of high frequency transit lines. Combining passengers and bus arrival schemes, the average waiting time at each station would be equal to half of the headway under regular bus arrivals, and equal to the whole expected headway for random (Poisson) bus arrivals.

In summary, we can compute the waiting time cost (C_{w}) component associated with normal service, as follows:

$$C_{\mathrm{w}} = P_{\mathrm{w}} \frac{1+x}{2} \left\{ \sum_{k=1}^{N} \frac{\lambda_k^{1+}}{f} + \sum_{k=1}^{N} \frac{\lambda_k^{2+}}{f} \right\} \tag{1}$$

In this expression, the two bus arrival schemes discussed above are captured by the auxiliary binary variable x, which is equal to 1 if buses arrive Poisson and 0 if they arrive regularly spaced; P_{w} is the waiting time value. For the short turning strategy, we will assume that bus arrivals at both inside (frequency $f_{\mathrm{A}} + f_{\mathrm{B}}$) and outside the short turn (frequency f_{A}) follow a Poisson distribution. Then C_{w} becomes:

$$\begin{aligned} C_{\mathrm{w}} = P_{\mathrm{w}} \frac{1+x}{2} \Bigg\{ & \sum_{k=1}^{s_0-1} \frac{\lambda_k^{1+}}{f_{\mathrm{A}}} + \sum_{k=s_0}^{s_1-1} \left(\frac{\lambda_k^{+}(k+1, s_1)}{f_{\mathrm{A}} + f_{\mathrm{B}}} + \frac{\lambda_k^{+}(s_1+1, N)}{f_{\mathrm{A}}} \right) + \sum_{k=s_1}^{N} \frac{\lambda_k^{1+}}{f_{\mathrm{A}}} \\ & + \sum_{k=s_1+1}^{N} \frac{\lambda_k^{2+}}{f_{\mathrm{A}}} + \sum_{k=s_0+1}^{s_1} \left(\frac{\lambda_k^{+}(1, s_0-1)}{f_{\mathrm{A}}} + \frac{\lambda_k^{+}(s_0, k-1)}{f_{\mathrm{A}} + f_{\mathrm{B}}} \right) + \sum_{k=1}^{s_0} \frac{\lambda_k^{2+}}{f_{\mathrm{A}}} \Bigg\} \end{aligned} \tag{2}$$

Inside the key brackets in expression (2) we have split the expected waiting times of users per direction and per origin and destination, because passengers boarding and alighting inside and outside the short turn face different frequencies, and therefore, different waiting times. For example, in direction 1 (first line in Eq. 2) passengers boarding before (first term) and after (third term) the short turn are served by a frequency f_A, the same as those passengers who board inside but have destination after the short turn (second summation in second term). Only passengers whose origin and destination are inside the short cycle observe a frequency $f_A + f_B$ (first summation in second term). The terms in the second line (direction 2) have an analogous interpretation. This assumes that passengers do not transfer between vehicles A and B.

The in-vehicle time between stations corresponds to the sum of the bus running time (R_i) and the time that passengers wait for others to board the bus ($b\lambda_i^+/f$). Under normal operation, the cost associated with in-vehicle time (C_v) is,

$$C_v = P_v\left\{\sum_{k=1}^{N}\sum_{l=k+1}^{N}\left[\sum_{i=k}^{l-1}\left(R_i + b\frac{\lambda_i^{1+}}{f}\right)\right]\lambda_{kl} + \sum_{k=1}^{N}\sum_{l=1}^{k-1}\left[\sum_{i=l+1}^{k}\left(R_{i-1} + b\frac{\lambda_i^{2+}}{f}\right)\right]\lambda_{kl}\right\} \tag{3}$$

where P_v is the in-vehicle time value. When a short turning strategy is introduced, the analytical expression for C_v becomes more complicated, as a trip can encompass areas in which users board vehicles at rates λ_i^+/f_A or $\lambda_i^+/(f_A + f_B)$. Using the definitions of functions $g_i(s_0, s_1)$ in Appendix A, the total in-vehicle time cost can be expressed as

$$C_v = 2\sum_{k=1}^{N-1} R_k + b\frac{g_5(s_0, s_1)}{f_A + f_B} + b\frac{g_6(s_0, s_1)}{f_A} \tag{4}$$

where functions $g_5(s_0, s_1)$ and $g_6(s_0, s_1)$ group second order terms of the trip rates λ_{kl}, as a function of the start and end stations s_0 and s_1.

Operator cost (C_o) is divided into temporal and spatial components. The former includes items as personnel costs (crew), while the latter comprises running costs, such as fuel consumption, lubricants, tires, maintenance, etc. Following Oldfield and Bly (1988) and Jansson (1980), the unit operator costs are expressed as a linear function of the vehicle capacity K, where $c(K)$ is the vehicle–hour cost (expressed in \$/vh) and $c'(K)$ corresponds to the vehicle–kilometer cost (expressed in \$/vkm). Analytically,

$$c(K) = c_0 + c_1 K \quad c'(K) = c'_0 + c'_1 K \tag{5}$$

$$C_o = c(K)F + c'(K)vF \tag{6}$$

In the previous expressions, v is the bus system commercial speed (including dwell times) and F is the fleet size needed for a design occupancy rate (define below in Eq. 11), obtained as the product of the frequency and the cycle time t_c (which by simplicity assumes no layover time or slack in the schedule), $F = ft_c$. Since $v = 2L/t_c$, C_o can be rewritten as follows for the case of normal operation as (7) and (8)

$$C_o = c(K)ft_c + 2c'(K)fL \tag{7}$$

$$C_o = f\left\{c(K)\left[\sum_{k=1}^{N-1}\left(R_k + b\frac{\lambda_k^{1+}}{f}\right) + \sum_{k=2}^{N}\left(R_{k-1} + b\frac{\lambda_k^{2+}}{f}\right)\right] + 2c'(K)L\right\} \tag{8}$$

On the other hand, for the short turning strategy, the operator cost must include the costs of both fleets, A and B. Operator cost of Fleet A, which operates normally between terminals, is

$$C_{oA} = f_A\left\{c(K)\left[\sum_{k=1}^{s_0-1}\left(R_k + b\frac{\lambda_k^{1+}}{f_A}\right) + \sum_{k=s_0}^{s_1-1}\left(R_k + b\left\{\frac{\lambda_k^+(k+1,s_1)}{f_A+f_B} + \frac{\lambda_k^+(s_1+1,N)}{f_A}\right\}\right)\right.\right.$$
$$+ \sum_{k=s_1}^{N}\left(R_k + b\frac{\lambda_k^{1+}}{f_A}\right) + \sum_{k=s_1+1}^{N}\left(R_{k-1} + b\frac{\lambda_k^{2+}}{f_A}\right)$$
$$+ \sum_{k=s_0+1}^{s_1}\left(R_{k-1} + b\left\{\frac{\lambda_k^+(1,s_0-1)}{f_A} + \frac{\lambda_k^+(s_0,k-1)}{f_A+f_B}\right\}\right)$$
$$\left.\left. + \sum_{k=2}^{s_0}\left(R_{k-1} + b\frac{\lambda_k^{2+}}{f_A}\right)\right] + 2c'(K)L\right\} \quad (9)$$

while vehicles belonging to fleet B travel a shorter distance and then their cost is:

$$C_{oB} = f_B\left\{c(K)\left[\sum_{k=s_0}^{s_1-1}\left(R_k + b\frac{\lambda_k^+(k+1,s_1)}{f_A+f_B}\right) + \sum_{k=s_0+1}^{s_1}\left(R_{k-1} + b\frac{\lambda_k^+(s_0,k-1)}{f_A+f_B}\right)\right]\right.$$
$$\left. +2c'(K)L\frac{s_1-s_0}{N-1}\right\} \quad (10)$$

Using (9) and (10), the total operator cost is $C_o = C_{oA} + C_{oB}$.

Once waiting time, in-vehicle time and operator costs have been analytically found, it is possible to minimize total cost for both normal operation (sum of expressions 1, 3 and 8) and for the system with short turning strategy (sum of 2, 4, 9 and 10). Moreover, for the waiting time calculations, in Eqs. 1 and 2 passengers are assumed to board the first bus that arrives at their station; to fulfill this condition, bus capacity is set to accommodate the expected demand in the most loaded segment along the line, namely q_{max}. Under normal operation:

$$K = \frac{q_{max}}{\eta f} \quad (11)$$

where $q_{max} = \max_k\left\{\sum_{i=1}^{k}\sum_{j=k+1}^{N}\lambda_{ij}, \sum_{i=k+1}^{N}\sum_{j=1}^{k}\lambda_{ij}\right\}$ is obtained from the OD matrix. Factor η defines the maximum design occupancy rate (for example, $\eta = 0.85$), whose purpose is to keep extra capacity for absorbing the intrinsic randomness of the demand. Note that if demand is assumed to follow a random distribution, such as Poisson, this capacity and any other finite value can be theoretically superseded, case which is not considered in this model.

After introducing (11) in (5), the optimal value of frequency f is obtained applying first order conditions (FOC), obtaining

$$f^* = \sqrt{\frac{P_w\frac{1+x}{2}\left(\sum_{k=1}^{N}\lambda_k^{1+} + \sum_{k=1}^{N}\lambda_k^{2+}\right) + P_v b\left(\sum_{k=1}^{N}\sum_{l=k+1}^{N}\lambda_{kl}\sum_{i=k}^{l-1}\lambda_i^{1+} + \sum_{k=1}^{N}\sum_{l=1}^{k-1}\lambda_{kl}\sum_{i=l+1}^{k}\lambda_i^{2+}\right) + c_1\frac{q_{max}}{\eta}b\left(\sum_{k=1}^{N}\lambda_k^{1+} + \sum_{k=1}^{N}\lambda_k^{2+}\right)}{2\left(c_0\sum_{k=1}^{N-1}R_k + c_0'L\right)}} \quad (12)$$

Expression (12) corresponds to the classical "square root formula", derived for a single bus corridor with fixed demand (Mohring 1972; Jansson 1980; Jara-Díaz and Gschwender 2003 among others), with the difference that in this formulation, (12) is expressed in terms of the disaggregated OD demand between stations (Jara-Diaz et al. 2008).

For the short turning strategy, total cost can be expressed more concisely as

$$\begin{aligned}C_t(f_A,f_B,s_0,s_1,K)=&f_A\left\{c(K)\left(g_0+b\frac{g_1(s_0,s_1)}{f_A+f_B}+b\frac{g_2(s_0,s_1)}{f_A}\right)+2c'(K)L\right\}\\&+f_B\left\{c(K)\left(g_3(s_0,s_1)+b\frac{g_1(s_0,s_1)}{f_A+f_B}\right)+2c'(K)\frac{s_1-s_0}{N-1}L\right\}\\&+P_w\frac{1+x}{2}\left(\frac{g_1(s_0,s_1)}{f_A+f_B}+\frac{g_2(s_0,s_1)}{f_A}\right)+P_v\left(g_4+b\frac{g_5(s_0,s_1)}{f_A+f_B}+b\frac{g_6(s_0,s_1)}{f_A}\right)\end{aligned} \tag{13}$$

Functions $g_i(s_0, s_1)$ are defined in Appendix A; g_0 is the total running time, g_1 is the demand benefited by the strategy (origin and destination inside the short turn), g_2 quantifies the passengers whose origin or destination are outside the short turn, g_3 is the running time of vehicles inside the short turn, g_4 is the total in-vehicle time experienced by passengers, and g_5 and g_6 are quadratic factors on demand used to calculate the dwell times of passengers whose origin and destination is inside the short turn (g_5) and of all other passengers (g_6), respectively. The maximum load occurs for vehicles of fleet A, as these enter the short turn section already with passengers on board (fleet B vehicles start service at stations s_0 or s_1). Using recursively the boarding and alighting rates at stations λ_k^+ and λ_k^-, it is possible to obtain the load of the bus along the route for each segment and from that, the capacity of vehicles can be set as

$$K=\frac{1}{\eta}\left(\frac{\vartheta_0(s_0,s_1)}{f_A}+\frac{\vartheta_1(s_0,s_1)}{f_A+f_B}\right) \tag{14}$$

where $\vartheta_0(s_0, s_1)$ and $\vartheta_1(s_0, s_1)$ represent the terms that yield the maximum load, which depends on the selection of stations s_0 and s_1 (see Appendix B for details). Therefore, by introducing (14) in (13), we can see that the total cost depends only on f_A, f_B, s_0 and s_1, namely $C_t(f_A, f_B, s_0, s_1)$. The problem is solved in two stages. First, parametric in the value of discrete variables s_0 and s_1, we apply the proper FOC to find $f_A^*(s_0, s_1)$, $f_B^*(s_0, s_1)$ and $C_t(f_A^*(s_0,s_1),f_B^*(s_0,s_1),s_0,s_1)\equiv\bar{C}_t(s_0,s_1)$. Next, in a second stage, the feasible limit stations s_0 and s_1 are explored to select the ones that minimize $\bar{C}_t(s_0,s_1)$.

Let us first assume that bus arrivals are distributed Poisson both for the normal operation and for the short turning strategy, not only inside (frequency $f_A + f_B$) but also outside the short cycle (frequency f_A). This is equivalent to set $x = 1$ in expressions (12) and (13).

In general, FOC do not yield analytical forms for the optimal frequencies and the problem must be solved numerically. Nevertheless, if the unit operator costs do not depend on capacity K, i.e. $c(K) = c$, $c'(K) = c'$ (there are no economies of vehicle size), the optimal frequencies are derived as shown in (15) and (16), conditional on s_0 and s_1. A second order analysis shows that (15) and (16) minimizes total cost.

$$f_A^*(s_0,s_1)=\sqrt{\frac{P_w g_2(s_0,s_1)+P_v b g_6(s_0,s_1)}{c(g_0-g_3(s_0,s_1))+2c'\left(1-\frac{s_1-s_0}{N-1}\right)L}} \tag{15}$$

$$f_B^*(s_0,s_1)=\sqrt{\frac{P_w g_1(s_0,s_1)+P_v b g_5(s_0,s_1)}{c g_3(s_0,s_1)+2c'\frac{s_1-s_0}{N-1}L}}-\sqrt{\frac{P_w g_2(s_0,s_1)+P_v b g_6(s_0,s_1)}{c(g_0-g_3(s_0,s_1))+2c'\left(1-\frac{s_1-s_0}{N-1}\right)L}} \tag{16}$$

Expression (15) has the same analytical form as (12) and corresponds to the optimal frequency for serving the stations outside the short turn (those whose demand is represented by functions g_2 and g_6), while (16), which is the optimal frequency of those vehicles serving only the short turn, is computed as the difference between the optimal frequency to serve only the short turn (first term of 16) and f_A (second term of 16). Note that in the particular case of $f_B = 0$, expression (13) collapses to the total cost under normal operation, and therefore if the solution (16) is positive, the strategy is beneficial (total cost is less than that obtained under normal operation). Therefore, a condition for the strategy to be beneficial is $f_B^*(s_0, s_1) > 0$ or

$$\frac{P_w g_1(s_0, s_1) + P_v b g_5(s_0, s_1)}{P_w g_2(s_0, s_1) + P_v b g_6(s_0, s_1)} > \frac{c g_3(s_0, s_1) + 2c' \frac{s_1 - s_0}{N-1} L}{c(g_0 - g_3(s_0, s_1)) + 2c' \left(1 - \frac{s_1 - s_0}{N-1}\right) L} \tag{17}$$

where the demand benefited by the short turn is represented by functions g_1 and g_5, while demand whose origin or destination is outside the short turn is represented by g_2 and g_6. From (17) we conclude that the greater the demand imbalance between inside and outside the short cycle (g_1 and g_5 against g_2 and g_6), the more likely the strategy results beneficial. Regarding the length of the short turn, diminishing s_1–s_0 reduces the right side of (17) as this implies lower operator costs; this also decreases left side of the expression, since the shorter the short turn, less users are benefited by the strategy. Then, from the equations we can set neither the optimal length of the short turn nor its position. For the case of a scheduled service defined earlier ($x = 0, f_B = nf_A$), the vehicle capacity can be expressed as:

$$K = \frac{1}{\eta}\left[\frac{\vartheta_0(s_0, s_1)}{f_A} + \frac{\vartheta_1(s_0, s_1)}{(n+1)f_A}\right] = \frac{1}{\eta f_A}\left(\vartheta_0(s_0, s_1) + \frac{\vartheta_1(s_0, s_1)}{n+1}\right) \equiv \frac{\vartheta_{\max}(n, s_0, s_1)}{\eta f_A} \tag{18}$$

Thus, it is possible to obtain an analytical expression for the optimal frequency f_A using the full version of the unit operator costs shown in (5). Analytically,

$$f_A^*(n, s_0, s_1) = \sqrt{\frac{P_w\left[\frac{g_1(s_0,s_1)}{2(n+1)} + \frac{g_2(s_0,s_1)}{2}\right] + P_v b\left[\frac{g_5(s_0,s_1)}{n+1} + g_6(s_0, s_1)\right] + c_1 b \frac{\vartheta_{\max}}{\eta}[g_1(s_0, s_1) + g_2(s_0, s_1)]}{c_0[g_0 + n g_3(s_0, s_1)] + 2c_0' L\left(1 + n\frac{s_1 - s_0}{N-1}\right)}} \tag{19}$$

Once expression (19) is introduced in (13), considering $x = 0$ and $f_B = nf_A$, the values for the scheduling mode n and stations s_0 and s_1 can be searched to minimize total cost. Note that in practice, the search procedure of stations s_0 and s_1 is constrained by physical conditions, as only in some spots along a bus route is technically possible to install a short cycle by turning vehicles back.

Applications

Examples

In this section the objective is to quantify the potential benefits of the short turning strategy using the analytical results obtained under the optimization scheme developed above, on a hypothetical bus corridor comprising 10 stations. Further, we perform sensitivity analysis regarding demand volumes and its distribution. We consider two demand scenarios that represent two different demand distribution patterns of particular interest, as shown in the load profiles in Fig. 2. The first case (Example 1, Fig. 2a) represents a situation with the central business district located around one corridor terminus, and is taken from Delle Site

O\D	1	2	3	4	5	6	7	8	9	10
1		29	14	64	4	3	3	1	1	25
2	14		15	70	4	4	3	1	1	27
3	5	5		49	3	3	2	1	0	19
4	8	7	4		18	15	12	4	3	111
5	74	63	35	0		5	4	1	1	37
6	4	4	2	0	0		5	2	1	50
7	1	1	0	0	0	3		20	16	636
8	8	6	3	0	0	26	5		7	262
9	16	14	7	0	0	58	11	0		77
10	13	11	6	0	0	47	9	0	10	

O\D	1	2	3	4	5	6	7	8	9	10
1		5	8	9	13	10	17	41	4	2
2	7		7	8	11	8	15	35	4	2
3	40	35		13	17	13	23	55	6	3
4	20	18	22		24	19	33	77	8	4
5	17	15	19	34		150	265	624	67	35
6	15	13	16	29	348		267	630	67	36
7	20	18	23	40	484	131		529	56	30
8	21	18	23	40	494	134	130		52	28
9	3	3	3	6	72	19	19	5		5
10	1	1	1	2	23	6	6	2	3	

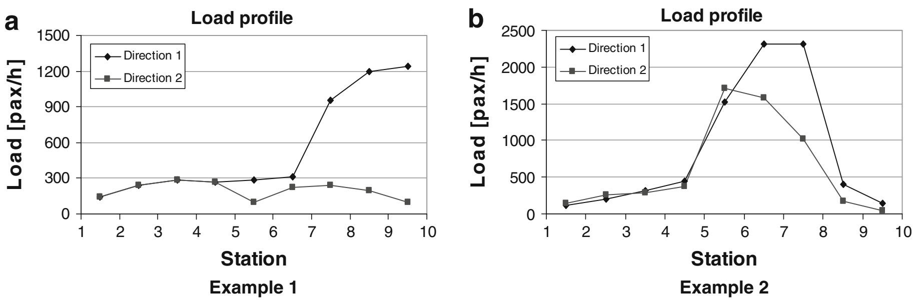


Fig. 2 OD matrices (pax/h) and load profiles, examples 1 and 2

and Filippi (1998). The second case has the same number of stops and comparable demand level as Example 1, but represents a typical cross-town peak-demand route pattern, illustrated by the load profile of Fig. 2b (Example 2). The values assigned to the parameters are: $R = 2.5$ min $\forall k$, $L = 8$ km, $P_{\mathrm{w}} = 2700$ (\$/h), $P_{\mathrm{v}} = 900$ (\$/h), $c_0 = 1800$ (\$/vh), $c_1 = 30$ (\$/h-pax), $c_0' = 400$ (\$/vkm), $c_1' = 1$ (\$/h-pax), $b = 5$ (s/pax) and $\eta = 0.9$ (\$: CLP-Chilean peso, USD 1 $\approx$ CLP 500).

With these data we searched for the optimal values of s_0, s_1, f_{A}, f_{B}, n, K_{A} and K_{B} for each example under the two bus arrival schemes, i.e. regular and random (Poisson). In Tables 1 and 2, a summary of the key results after running our model is provided for Examples 1 and 2. Let us analyze Example 1 first.

In this case, the optimal short turn resulted between stations $s_0 = 7$ and $s_1 = 10$ (benefiting 49.8% of the passengers) which coincides with the most demanded section along direction 2 in Fig. 2a. Regarding the scheduled service case (regular headways), the optimal scheduling mode is $n = 1$, i.e., $f_{\mathrm{A}} = f_{\mathrm{B}}$ and the 25 veh/h frequency is raised to 36 veh/h inside the short cycle, and reduced to 18 veh/h outside the short cycle. The inclusion of the strategy provides savings in both in-vehicle and operator costs, and losses in waiting time cost, with a total saving of 2.6% under random headways (Poisson arrivals) and 3.7% considering regular headways. Therefore, in this example the waiting time savings for passengers inside the short turn are outweighed by the increase in waiting time for passengers whose origin or destination is outside the short turn; nevertheless, there are savings associated with the in-vehicle time since this component depends on quadratic terms of the demand (g_5 and g_6), which amplifies the demand imbalance shown in Fig. 2a. The intuition is the following: when frequency increases for a group of stations, the waiting time is affected only once (in the station where a passenger boards a vehicle), while the in-vehicle time is reduced in all the downstream stations where the frequency has been increased, which implies having less users boarding and alighting each bus.

Table 1 Optimal values for Example 1

Operation		f_A (veh/h)	f_B (veh/h)	F (veh)	K (pax/veh)	C_w ($/pax)	C_v ($/pax)	C_o ($/pax)	C_t ($/pax)	ΔC_w (%)	ΔC_v (%)	ΔC_o (%)	ΔC_t (%)
Poisson arrivals	Normal	31 (*f*)		27	45	87	144	143	375	0.9	−3.9	−3.5	−2.6
	Strategy	23	26	27	36	88	139	138	365				
Regular arrivals	Normal	25 (*f*)		22	57	55	151	119	326	2.1	−4.1	−5.8	−3.7
	Strategy	18	18	22	48	56	145	113	314				

$s_0 = 7$, $s_1 = 10$, $n = 1$

Table 2 Optimal values for Example 2

Operation		f_A (veh/h)	f_B (veh/h)	F (veh)	K (pax/veh)	C_w ($/pax)	C_v ($/pax)	C_o ($/pax)	C_t ($/pax)	ΔC_w (%)	ΔC_v (%)	ΔC_o (%)	ΔC_t (%)
Poisson arrivals	Normal	78 (*f*)		27	34	35	52	72	158	−3.1	−14.5	−11.1	−10.5
	Strategy	40	89	25	31	34	44	64	142				
Regular arrivals	Normal	65 (*f*)		23	39	21	56	62	139	−1.4	−14.2	−13.2	−11.9
	Strategy	34	68	21	37	20	48	54	123				

$s_0 = 5$ and $s_1 = 8$, $n = 2$

It is worth comparing results of Example 1 with those of Delle Site and Filippi (1998, p. 32) as we used the same demand profile (AM peak in their paper), although we included two additional optimization variables (which they treat parametrically: limit stations and bus size) and a different set of parameters. In terms of optimal frequencies, these authors obtain $f_A = 7.6$ veh/h and $f_B = 7.4$ veh/h, which in our model are $f_A = 23$ veh/h and $f_B = 26$ veh/h, i.e. the latter values are roughly three times the former ones, mainly because the assumed unit operator costs c_0 and c_1 are much larger than P_v and P_w in Delle Site and Filippi' example (representative of Rome) than in ours (representative of Santiago), which increases the relative importance of operators cost in their example, reducing frequency. However, in both cases frequencies f_A and f_B are similar to each other, which is an expected qualitative result in relative terms, because both approaches attempt to minimize total cost for the same demand profile. Finally, cost savings obtained from both approaches are not directly comparable as Delle Site and Filippi (1998) consider an optimized short turning situation against a given normal operation case and not against an optimized base situation as in our model.

In Example 2, the relative savings are more important than in Example 1, with total benefits higher than 10%. This is because 73% of the demand occurs between stations $s_0 = 5$ and $s_1 = 8$ (where the short turn is optimally established), i.e. demand within the optimal short cycle is more concentrated than in Example 1. The demand imbalance is enough to even produce benefits in the waiting time cost, since the waiting time loss by users outside the short turn (frequency is reduced from 78 to 40 veh/h with random headways) is compensated by the savings of passengers inside the short turn (frequency goes from 78 to 129 veh/h). On the other hand, operators require a smaller fleet (two vehicles less than under normal operation) to provide the service. The scheduling mode $n = 2$ with regular headways is concordant with the results for Poisson arrival, where f_B is slightly larger than the double of f_A.

Regarding optimal capacity, in both examples K is smaller when the strategy is applied, relative to the base case. This is because a short turn increases the frequency in the segment where the demand is the largest.

In conclusion, both examples show that a rearrangement of frequencies due to the spatial distribution of demand, can be beneficial not only for users, whose cost is reduced on average, but also for operators, since it may be possible to use less and smaller vehicles.

Demand sensitivity

General demand growth

We conduct additional experiments to analyze the changes in the configuration of the strategy (location of the short turn and cost savings) due to variations in the structure of demand, namely total demand growth or change in the trips generation from a single station.[1] The scheduled operation is studied, with emphasis on the resulting scheduling mode n.

First, we study the case in which every OD pair has the same growth rate α for Example 2, amplifying the OD matrix by factors 2, 4 and 8. In the three amplified cases, the load profile has the same shape as that of the base case ($\alpha = 1$, Fig. 2b), as factor α weights loads evenly. In this example the demand is highly concentrated inside the short turn (73% of the trips between stations 5 and 8); therefore, an even growth rate α magnifies the

[1] Results from changes in attraction of trips are analogous to those from changes in generation, as shown in Tirachini (2007).

Table 3 Optimal values after even OD matrix growth in Example 2

α	n	Normal operation				Strategy				Cost savings			
		C_w ($/pax)	C_v ($/pax)	C_o ($/pax)	C_t ($/pax)	C_w ($/pax)	C_v ($/pax)	C_o ($/pax)	C_t ($/pax)	ΔC_w (%)	ΔC_v (%)	ΔC_o (%)	ΔC_t (%)
1	2	21	56	62	139	20	48	54	123	−1.4	−14.2	−13.2	−11.9
2	3	23	119	114	257	25	100	98	223	6.9	−16.1	−14.2	−13.2
4	3	25	248	217	490	27	209	183	420	9.6	−15.6	−15.6	−14.4
8	3	26	506	423	954	29	429	353	810	11.4	−15.3	−16.5	−15.1

Table 4 Cost ratio after doubling total demand in Example 2

Factor $C_i(2\alpha)/C_i(\alpha)$	Normal operation				Strategy			
	C_w	C_v	C_o	C_t	C_w	C_v	C_o	C_t
$C_i(2)/C_i(1)$	1.12	2.12	1.83	1.84	1.22	2.08	1.81	1.82
$C_i(4)/C_i(2)$	1.07	2.08	1.90	1.91	1.10	2.09	1.87	1.88
$C_i(8)/C_i(4)$	1.04	2.04	1.94	1.95	1.06	2.05	1.92	1.93

absolute difference between trips inside and outside the short turn, which is translated into a larger difference between frequencies inside and outside the short turn (growth of n), and in more benefits from implementing the strategy, as shown in Table 3.

The evolution of the different cost components when total demand is doubled deserves some discussion. The frequencies are f for normal operation (expression 12) and f_A or $f_A + f_B$ with the strategy (Eqs. 15 and 16 for constant unit operator costs). If the trips between every OD pair are amplified by a common factor α, the optimal frequencies f, f_A and $f_A + f_B$ are of the form $\sqrt{k_1\alpha + k_2\alpha^2}$ (where k_1 and k_2 are constant values, different depending the case, according to Eqs. 12, 15 and 16). Then, in both normal operation (single frequency) and with the strategy, the waiting time cost is of the form $k_3\alpha/\sqrt{k_1\alpha + k_2\alpha^2}$ (Eqs. 1 and 2), which tends to be constant as α grows. On the other hand, in-vehicle time and operator costs look like $k_3\alpha^2/\sqrt{k_1\alpha + k_2\alpha^2}$, expression that is proportional to α (with k_3 different for every case, obtained from Eqs. 4, 5, 8, 9 and 10). For this reason, as the demand is doubled in Table 3, C_w grows following a factor that tends to one, while C_v and C_o double following demand, as shown in Table 4 where the values of the cost ratios $C_i(2\alpha)/C_i(\alpha)$ are presented for $\alpha \in \{1, 2, 4\}$, subscript $i \in \{w, v, o, t\}$ depending on the case.

Uneven demand growth

Secondly, we amplify all trips generated from a specific station, using again a constant factor α. This is an abstraction of the change produced by the introduction of a trip generator within an area, as a new residential building. For this experiment, two cases are analyzed, depending on where trips grow, i.e. at a station inside or outside an existing short turn.

Considering the results for Example 2 as the base case (short turn between stations 5 and 8) we amplify the trips originated in outside-stations 1, 2, 3 and 4, until finding the minimum factor α that makes the corresponding station part of the short turn (Table 5).

Table 5 Trips generation growth, stations outside original short turn in Example 2

Station	α	n	s_0	s_1	Operation	C_w ($/pax)	C_v ($/pax)	C_o ($/pax)	C_t ($/pax)	ΔC_w (%)	ΔC_v (%)	ΔC_o (%)	ΔC_t (%)
4	4	2	4	8	Normal	21	63	68	153	1.3	−11.9	−9.7	−9.1
					Strategy	22	56	62	139				
3	8	1	3	8	Normal	22	74	76	172	0.3	−6.3	−5.3	−5.0
					Strategy	22	69	72	163				
2	17	2	2	8	Normal	21	80	78	179	3.5	−7.7	−3.2	−4.4
					Strategy	22	74	75	172				
1	19	1	1	8	Normal	21	93	84	197	1.4	−3.9	−1.7	−2.7
					Strategy	21	89	82	192				

Table 6 Trips generation growth, station inside original short turn in Example 2

Station	α	n	s_0	s_1	Operation	C_w ($/pax)	C_v ($/pax)	C_o ($/pax)	C_t ($/pax)	ΔC_w (%)	ΔC_v (%)	ΔC_o (%)	ΔC_t (%)
7	1	2	5	8	Normal	21	56	62	139	−1.4	−14.2	−13.2	−11.9
					Strategy	20	48	54	123				
7	2	3	5	8	Normal	21	69	75	166	2.1	−17.1	−14.5	−13.4
					Strategy	22	58	64	144				
7	4	3	5	8	Normal	22	99	103	224	0.9	−18.3	−17.5	−16.1
					Strategy	22	81	85	188				
7	8	4	5	8	Normal	21	163	162	347	4.7	−21.0	−20.0	−19.0
					Strategy	22	129	130	281				

The values of α show that the farther is a station from the original short turn, the more difficult is that a longer short turn reaches that station, after an increase in the trips generated.[2] If trips from station 4 grow four times, it is optimal to establish a short turn from 4 to 8, but only when the trips from station 1 are amplified 19 times, the optimal short cycle will include such a station. Regarding the benefits, there is a reduction in total cost saving while passengers are more dispersed and the short turn is longer (the influence of the length of the short turn in the results is analyzed in Sect. 3.3).

For a single station inside the short turn in Example 2 (say station 7) the generation rate is doubled until reaching $\alpha = 8$. Results are shown in Table 6.

As expected, total cost savings are larger when trip generation is augmented inside the short cycle, which remains unchanged between stations 5 and 8 for all cases, as concentration of trips inside the short turn is actually more intensive. Thus, the model increases the value of f_B relative to f_A, which is reflected in Table 6 through the increment in the optimal scheduling mode from 2 to 3 and 4.

[2] This result depends on the fact that the load in the segment after station 1 is lower than after stations 2, 3 and 4 (Fig. 2b). In general, if demands outside the short turning (stations 1–4) were similar, the distance to the short turn would play a role, as the farther a station to a existing short turn is, the higher the operator cost to include that station.

Influence of the demand concentration

In real public transport systems, spatial demand peaks show different shapes, dispersion, load level and length; for example, cases with very concentrated demand around a few stations or more even load profiles, with demand showing peaks but over a wider section of the route, involving more stations. Thus, it is important to understand the way the length of the demand peak (and therefore the length of the short turn) affects the results of the strategy in terms of cost savings. To isolate this effect, we now conduct a different experiment in which three load profiles are considered, all of them showing a peak load of 1,000 pax/h across 2, 3 and 4 consecutive segments in the middle of the route, respectively. In all cases, the extremes show a sharp reduction of the passenger load, to 100 pax/h, (Fig. 3). The matrices were generated such that both the total demand (2520 pax/h) and the trips inside the most loaded section (81% of total) are the same in the three cases.

From Table 7, we can see that the narrower the demand peak is (equivalent to say the shorter the short turn is), the more beneficial the strategy results. This phenomenon is observed in the three components of the total cost. When the same demand level is more concentrated in some section of the route, it is easier for the operators to serve it (they run a shorter distance). On the other hand, Eqs. 15 and 16 show that a shorter cycle increases f_B and decreases f_A, but the former effect is larger than the latter. Therefore the gain of the users that are benefited by the strategy overcomes the loss of users outside the short turn (whose frequency, f_A, is lower if the demand is more concentrated, *ceteris paribus*).

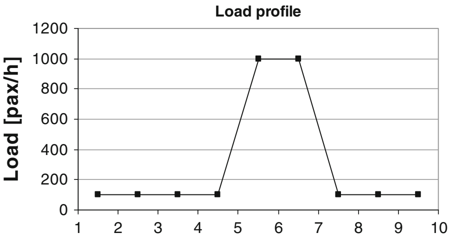
Comments and conclusions

In most cities in the world, the demand for urban transit systems shows clear spatial concentrations over certain segments of bus routes. Then, it is reasonable to offer higher frequency in the most loaded sectors. If the demand is concentrated around a specific sector of the route (for example the *Central Business District* or CBD), a short turn emerges as a suitable strategy to better serve the resulting uneven load of passengers. This strategy aims at reducing the cost of the operator on the one hand, and decreasing the waiting and in-vehicle times that users spend while traveling (with respect to the standard single frequency service along the entire route), on the other.

In this paper, we have introduced a model of a short turn strategy to set the optimal values of frequencies (inside and outside the loop), capacity of vehicles and the position of the short turn limit stations. These values and the corresponding cost components for operators and users (waiting and in-vehicle times) are compared against an optimized normal operation scheme (single frequency). The model was used to obtain results for different examples, and several experiments were conducted to analyze their sensitivity. Let us highlight the most important findings of this work.

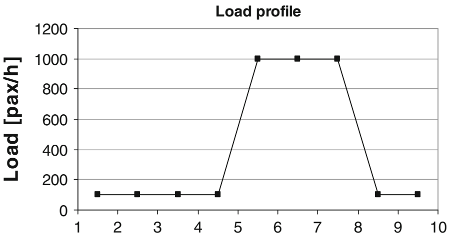
- The short turning strategy can yield benefits for both users and operators at the same time. Note that among passengers there are winners (origin and destination inside the short turn) and losers (origin and/or destination outside the short turn), who observe a lower frequency than in the single frequency case; but overall users save time.
- Cost savings are reachable not only when the demand is concentrated in the middle of a route (a bus line crossing the CBD), but also when the most loaded section is on an extreme of the line (radial line to the CBD).

Station	Direction 1 Load [pax/h]	λ_1^+ [pax/h]	λ_1^- [pax/h]	Direction 2 Load [pax/h]	λ_1^+ [pax/h]	λ_1^- [pax/h]
1	100	100	0	0	0	100
2	100	10	10	100	10	10
3	100	10	10	100	10	10
4	100	10	10	100	10	10
5	1000	910	10	100	10	910
6	1000	190	190	1000	190	190
7	100	10	910	1000	910	10
8	100	10	10	100	10	10
9	100	10	10	100	10	10
10	0	0	100	100	100	0
Total		1260	1260		1260	1260



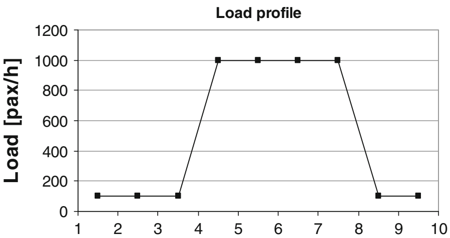
Demand load concentrated between stations 5 and 7

Station	Direction 1 Load [pax/h]	λ_1^+ [pax/h]	λ_1^- [pax/h]	Direction 2 Load [pax/h]	λ_1^+ [pax/h]	λ_1^- [pax/h]
1	100	100	0	0	0	100
2	100	10	10	100	10	10
3	100	10	10	100	10	10
4	100	10	10	100	10	10
5	1000	910	10	100	10	910
6	1000	100	100	1000	100	100
7	1000	100	100	1000	100	100
8	100	10	910	1000	910	10
9	100	10	10	100	10	10
10	0	0	100	100	100	0
Total		1260	1260		1260	1260



Demand load concentrated between stations 5 and 8

Station	Direction 1 Load [pax/h]	λ_1^+ [pax/h]	λ_1^- [pax/h]	Direction 2 Load [pax/h]	λ_1^+ [pax/h]	λ_1^- [pax/h]
1	100	100	0	0	0	100
2	100	10	10	100	10	10
3	100	10	10	100	10	10
4	100	10	10	100	10	10
5	1000	910	10	100	10	910
6	1000	100	100	1000	100	100
7	1000	100	100	1000	100	100
8	100	10	910	1000	910	10
9	100	10	10	100	10	10
10	0	0	100	100	100	0
Total		1260	1260		1260	1260



Demand load concentrated between stations 4 and 8

Fig. 3 Boarding and alighting rates and load profiles, different width of demand concentration

Table 7 Short turning results, different width of demand concentration

s_0	s_1	Normal operation C_w (\$/pax)	C_v (\$/pax)	C_o (\$/pax)	C_t (\$/pax)	Strategy C_w (\$/pax)	C_v (\$/pax)	C_o (\$/pax)	C_t (\$/pax)	Cost savings ΔC_w (%)	ΔC_v (%)	ΔC_o (%)	ΔC_t (%)
5	7	58	50	93	201	49	38	73	160	−15.5	−23.3	−21.9	−20.4
5	8	58	58	94	209	52	49	80	181	−10.1	−15.2	−14.5	−13.5
4	8	58	67	94	218	55	60	86	200	−5.3	−10.1	−8.5	−8.1

- Cost savings depend largely on the imbalance between trips inside and outside the short cycle.
- When the strategy yields benefits, vehicles may be smaller than those resulting for the single frequency case. This is because a short turn increases the frequency where the demand is larger. The size of the vehicles is determined by those serving the whole route (fleet A), since they circulate with more passengers when crossing the short cycle. A corollary of this result is that a model with constant unit operator costs would underestimate the benefits of the strategy compared with our model, where the operator cost increases linearly with capacity K (Eq. 5).
- When facing an even growth of the OD matrix, the strategy results more beneficial. This also happens when the trip generation or attraction from (to) a station inside a short turn is increased. A station outside a short turn can become part of it after a raise in the number of trips it generates or attracts, notwithstanding it produces smaller benefits than in the original case in which the short turn comprised fewer stations.
- The concentration of demand plays a fundamental role: the more concentrated the trips (the shorter the cycle is), the larger the benefits of a short turning strategy.
- From the modeling results, it is interesting that the analytical form for the optimal frequencies with the strategy follows the structure of the square root formula (see Eqs. 15, 16 and 19), even though a more complex account of the costs and benefits has been undertaken, considering the distinct terms for users benefited by the strategy (trips inside the short turn) and users whose situation is worse with the strategy (trips with either origin or destination outside the short cycle), relative to the single frequency case.

Future research can include some elements not considered here, such as:

- crowding inside vehicles (which affects waiting time, as shown by Tirachini and Cortés 2007, and in-vehicle time value); as with a short turning strategy we increase the frequency on the most loaded part of the line and decrease it on the less demanded (less crowded) section, and therefore occupancy rates and crowding factors would change;
- vehicle capacities according to commercial buses (and with this, to consider the case of active capacity constraint affecting the waiting time). The capacity of available commercial buses is highly discrete (e.g. 40, 60 and 120 pax/bus) which adds a constraint that would limit the benefits of the approach;
- extension of the model to a network (or at least, two routes connected by a transfer station); in which case there are more degrees of freedom as interlining may be possible (vehicles changing lines on terminals or transfer stations)
- although this model has been conceived for a bus system application, it can be extended to rail lines, taking into consideration some characteristic aspects of rail systems, such as dwell time at stations (that in many cases is fixed and independent of the demand) and the operators' cost structure, among others.

Acknowledgements This research was partially financed by grants 1061261 and 1080140 from Fondecyt, Chile, and the Millennium Institute "Complex Engineering Systems" (ICM: P-05-004-F, CONICYT: FBO16).

Appendix A. Definition of g_i functions

- g_0 is the total running time:

$$g_0 = 2\sum_{k=1}^{N-1} R_k$$

- g_1 and g_2 are in the waiting and cycle times, g_1 is the demand benefited by the strategy (origin and destination inside the short turn), while g_2 are the passenger whose origin or destination are outside the short turn.

$$g_1(s_0, s_1) = \sum_{k=s_0}^{s_1-1} \lambda_k^+(k+1, s_1) + \sum_{k=s_0+1}^{s_1} \lambda_k^+(s_0, k-1)$$

$$g_2(s_0, s_1) = \sum_{k=1}^{s_0-1} \lambda_k^{1+} + \sum_{k=s_0}^{s_1-1} \lambda_k^+(s_1+1, N) + \sum_{k=s_1}^{N} \lambda_k^{1+} + \sum_{k=s_1+1}^{N} \lambda_k^{2+} + \sum_{k=s_0+1}^{s_1} \lambda_k^+(1, s_0-1) + \sum_{k=1}^{s_0} \lambda_k^{2+}$$

- g_3 is the running time for vehicles inside the short turn

$$g_3(s_0, s_1) = 2\sum_{k=s_0}^{s_1-1} R_k$$

- g_4 is the in-vehicle time experienced by passengers

$$g_4 = \sum_{k=1}^{N}\sum_{l=1}^{N} \lambda_{kl} \sum_{i=k}^{l-1} R_i$$

- g_5 and g_6 are the factors to calculate the total dwell time of passengers benefited by the strategy and the others, respectively.

$$g_5(s_0, s_1) = g_5^1(s_0, s_1) + g_5^2(s_0, s_1)$$

$$g_6(s_0, s_1) = g_6^1(s_0, s_1) + g_6^2(s_0, s_1)$$

where

$$g_5^1(s_0, s_1) = \sum_{k=1}^{s_0-1}\sum_{l=s_0+1}^{s_1} \lambda_{kl} \sum_{i=s_0}^{l-1} \lambda_i^+(i+1, s_1) + \sum_{k=1}^{s_0-1}\sum_{l=s_1+1}^{N} \lambda_{kl} \sum_{i=s_0}^{s_1-1} \lambda_i^+(i+1, s_1) + \sum_{k=s_0}^{s_1-1}\sum_{l=k+1}^{s_1} \lambda_{kl} \sum_{i=k}^{l-1} \lambda_i^{1+} + \sum_{k=s_0}^{s_1-1}\sum_{l=s_1+1}^{N} \lambda_{kl} \sum_{i=k}^{s_1-1} \lambda_i^+(i+1, s_1)$$

$$
\begin{aligned}
g_5^2(s_0,s_1) = & \sum_{k=s_1+1}^{N}\sum_{l=s_0}^{s_1-1}\lambda_{kl}\sum_{i=l+1}^{s_1}\lambda_i^+(s_0,i-1) \\
& + \sum_{k=s_1+1}^{N}\sum_{l=1}^{s_0-1}\lambda_{kl}\sum_{i=s_0+1}^{s_1}\lambda_i^+(s_0,i-1) \\
& + \sum_{k=s_0+1}^{s_1}\sum_{l=s_0}^{k-1}\lambda_{kl}\sum_{i=l+1}^{k}\lambda_i^{2+} + \sum_{k=s_0+1}^{s_1}\sum_{l=1}^{s_0-1}\lambda_{kl}\sum_{i=s_0+1}^{k}\lambda_i^+(s_0,i-1)
\end{aligned}
$$

$$
\begin{aligned}
g_6^1(s_0,s_1) = & \sum_{k=1}^{s_0-1}\sum_{l=k+1}^{s_0}\lambda_{kl}\sum_{i=k}^{l-1}\lambda_i^{1+} + \sum_{k=1}^{s_0-1}\sum_{l=s_0+1}^{s_1}\lambda_{kl}\left[\sum_{i=k}^{s_0-1}\lambda_i^{1+} + \sum_{i=s_0}^{l-1}\lambda_i^+(s_1+1,N)\right] \\
& + \sum_{k=1}^{s_0-1}\sum_{l=s_1+1}^{N}\lambda_{kl}\left[\sum_{i=k}^{s_0-1}\lambda_i^{1+} + \sum_{i=s_0}^{s_1-1}\lambda_i^+(s_1+1,N) + \sum_{i=s_1}^{l-1}\lambda_i^{1+}\right] \\
& + \sum_{k=s_0}^{s_1-1}\sum_{l=s_1+1}^{N}\lambda_{kl}\left[\sum_{i=k}^{s_1-1}\lambda_i^+(s_1+1,N) + \sum_{i=s_1}^{l-1}\lambda_i^{1+}\right] + \sum_{k=s_1}^{N}\sum_{l=k+1}^{N}\lambda_{kl}\sum_{i=k}^{l-1}\lambda_i^{1+}
\end{aligned}
$$

$$
\begin{aligned}
g_6^2(s_0,s_1) = & \sum_{k=s_1+1}^{N}\sum_{l=s_1}^{k-1}\lambda_{kl}\sum_{i=l+1}^{k}\lambda_i^{2+} + \sum_{k=s_1+1}^{N}\sum_{l=s_0}^{s_1-1}\lambda_{kl}\left[\sum_{i=s_1+1}^{k}\lambda_i^{2+} + \sum_{i=l+1}^{s_1}\lambda_i^+(1,s_0-1)\right] \\
& + \sum_{k=s_1+1}^{N}\sum_{l=1}^{s_0-1}\lambda_{kl}\left[\sum_{i=s_1+1}^{k}\lambda_i^{2+} + \sum_{i=s_0+1}^{s_1}\lambda_i^+(1,s_0-1) + \sum_{i=l+1}^{s_0}\lambda_i^{2+}\right] \\
& + \sum_{k=s_0+1}^{s_1}\sum_{l=1}^{s_0-1}\lambda_{kl}\left[\sum_{i=s_0+1}^{k}\lambda_i^+(1,s_0-1) + \sum_{i=l+1}^{s_0}\lambda_i^{1+}\right] + \sum_{k=1}^{s_0}\sum_{l=1}^{k-1}\lambda_{kl}\sum_{i=l+1}^{k}\lambda_i^{2+}
\end{aligned}
$$

Appendix B. Load of vehicles between stations

1. Load of vehicles serving the entire corridor (fleet A):

   Direction 1

$$
\pi_k^1 = \begin{cases}
\pi_{k-1}^1 + \frac{\lambda_k^{1+}}{f_A} - \frac{\lambda_k^{1-}}{f_A} & \text{if } 1 \le k \le s_0 - 1 \\
\pi_{k-1}^1 + \frac{\lambda_k^+(k,s_1)}{f_A+f_B} + \frac{\lambda_k^+(s_1+1,N)}{f_A} - \frac{\lambda_k^-(1,s_0-1)}{f_A} - \frac{\lambda_k^-(s_0,k-1)}{f_A+f_B} & \text{if } s_0 \le k \le s_1 - 1 \\
\pi_{k-1}^1 + \frac{\lambda_k^{1+}}{f_A} - \frac{\lambda_k^{1-}}{f_A} & \text{if } s_1 \le k \le N
\end{cases}
$$

   Direction 2

$$
\pi_k^2 = \begin{cases}
\pi_{k+1}^2 + \frac{\lambda_k^{2+}}{f_A} - \frac{\lambda_k^{2-}}{f_A} & \text{if } s_1 + 1 \le k \le N \\
\pi_{k+1}^2 + \frac{\lambda_k^+(s_0,k-1)}{f_A+f_B} + \frac{\lambda_k^+(1,s_0-1)}{f_A} - \frac{\lambda_k^-(s_1+1,N)}{f_A} - \frac{\lambda_k^-(k+1,s_1)}{f_A+f_B} & \text{if } s_0 + 1 \le k \le s_1 \\
\pi_{k+1}^2 + \frac{\lambda_k^{2+}}{f_A} - \frac{\lambda_k^{2-}}{f_A} & \text{if } 1 \le k \le s_0
\end{cases}
$$

2. Load of vehicles performing short turning (fleet B):

Direction 1

$$\tilde{\pi}_k^1 = \begin{cases} 0 & \text{if } 1 \le k \le s_0 - 1 \\ \tilde{\pi}_{k-1}^1 + \frac{\lambda_k^+(k,s_1)}{f_A+f_B} - \frac{\lambda_k^-(s_0,k-1)}{f_A+f_B} & \text{if } s_0 \le k \le s_1 - 1 \\ 0 & \text{if } s_1 \le k \le N \end{cases}$$

Direction 2

$$\tilde{\pi}_k^2 = \begin{cases} 0 & \text{if } s_1 + 1 \le k \le N \\ \pi_{k+1}^2 + \frac{\lambda_k^+(s_0,k-1)}{f_A+f_B} - \frac{\lambda_k^-(k+1,s_1)}{f_A+f_B} & \text{if } s_0 + 1 \le k \le s_1 \\ 0 & \text{if } 1 \le k \le s_0 \end{cases}$$

After applying the short-turning strategy, the maximum load will correspond to some segment, which can belong to either direction 1 or 2. However, by observing the recursive form of the load equations above, we can establish the following:

- If the maximum load occurs along direction 1, then we can say that

$$\eta K = \pi_{\max}^1 = \frac{\varphi_0^1(s_0, s_1)}{f_A} + \frac{\varphi_1^1(s_0, s_1)}{f_A + f_B}$$

- Otherwise, if the maximum load occur along direction 2, then we can say that

$$\eta K = \pi_{\max}^2 = \frac{\varphi_0^2(s_0, s_1)}{f_A} + \frac{\varphi_1^2(s_0, s_1)}{f_A + f_B}$$

Thus, a generic relation between the design capacity of buses and the frequencies f_A and f_B can be established in the way summarized in Eq. 14:

$$K = \frac{1}{\eta}\left(\frac{\vartheta_0(s_0, s_1)}{f_A} + \frac{\vartheta_1(s_0, s_1)}{f_A + f_B}\right)$$

where, if the maximum load happens on a segment that belongs to direction 1, then $\vartheta_0(s_0, s_1) = \varphi_0^1(s_0, s_1)$ and $\vartheta_1(s_0, s_1) = \varphi_1^1(s_0, s_1)$. Otherwise, if the maximum load occurs on a segment that belongs to direction 2, then $\vartheta_0(s_0, s_1) = \varphi_0^2(s_0, s_1)$ and $\vartheta_1(s_0, s_1) = \varphi_1^2(s_0, s_1)$.

References

Ceder, A., Stern, H.I.: Deficit function bus scheduling with deadheading trip insertion for fleet size reduction. Transp. Sci. **15**(4), 338–363 (1981)

Ceder, A.: Optimal design of transit short-turn trips. Transp. Res. Rec. **1221**, 8–22 (1989)

Ceder, A.: Public transport timetabling and vehicle scheduling. In: Lam, W., Bell, M. (eds.) Advanced Modeling for Transit Operations and Service Planning, pp. 31–57. Pergamon Imprint, Elsevier Science Ltd. Pub, Amsterdam (2003a)

Ceder, A.: Designing public transport network and route. In: Lam, W., Bell, M. (eds.) Advanced Modeling for Transit Operations and Service Planning, pp. 59–91. Pergamon Imprint, Elsevier Science Ltd. Pub, Amsterdam (2003b)

Ceder, A.: Improved lower-bound fleet size for fixed and variable transit scheduling. 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT), August 9–11, 2004 San Diego, California, USA (2004)

Delle Site, P.D., Filippi, F.: Service optimization for bus corridors with short-turn strategies and variable vehicle size. Transp. Res. A **32**(1), 19–28 (1998)

Eberlein, X.J., Wilson, N.H.M., Barnhart, C., Bernstein, D.: The real-time deadheading problem in transit operations control. Transp. Res. B **32**(2), 77–100 (1998)

Eberlein, X.J., Wilson, N.H.M., Bernstein, D.: Modeling real-time control strategies in public transit operations. In: N.H.M. Wilson (Ed.), Computer-Aided Transit Scheduling, Lecture Notes in Economics and Mathematical Systems 471, pp. 325–346, Springer-Verlag, Heidelberg (1999)

Furth, P.: Alternating deadheading in bus route operations. Transp. Sci. **19**(1), 13–28 (1985)

Furth, P.: Zonal route design for transit corridors. Transp. Sci. **20**(1), 1–12 (1986)

Furth, P.: Short turning on transit routes. Transp. Res. Rec. **1108**, 42–52 (1987)

Jansson, J.O.: A simple bus line model for optimization of service frequency and bus size. J. Transp. Econ. Policy **14**, 53–80 (1980)

Jara-Díaz, S.R., Gschwender, A.: Towards a general microeconomic model for the operation of public transport. Transp. Rev. **23**(4), 453–469 (2003)

Jara-Diaz, S.R., Tirachini, A., Cortés, C.E.: Modeling public transport corridors with aggregate and disaggregate demand. J. Transp. Geogr. **16**, 430–435 (2008)

Jordan, W.C., Turnquist, M.A.: Zone scheduling of bus routes to improve service reliability. Transp. Sci. **13**(3), 242–268 (1979)

Mohring, H.: Optimization and scale economies in urban bus transportation. Am. Econ. Rev. **62**, 591–604 (1972)

Oldfield, R.H., Bly, P.H.: An analytic investigation of optimal bus size. Transp. Res. **22B**, 319–337 (1988)

Osuna, E.E., Newell, G.F.: Control strategies for an idealized public transportation system. Transp. Sci. **6**, 52–72 (1972)

Tirachini, A.: Estrategias de asignación de flota en un corredor de transporte público. MSc Thesis, Universidad de Chile (2007)

Tirachini, A., Cortés, C.E.: Disaggregated modeling of pre-planned short-turning strategies in transit corridors. CD Proceedings, Transportation Research Board (TRB) 86th Annual Meeting, Washington DC, USA (2007)

Vijayaraghavan, T.A., Anantharmaiah, K.M.: Fleet assignment strategies in urban transportation using express and partial services. Transp. Res. A **29**(2), 157–171 (1995)

Welding, P.I.: The instability of a close-interval service. Oper. Res. Q. **8**(3), 133–148 (1957)

Author Biographies

Alejandro Tirachini born in Chiloé, Chile, is beginning his journey in the world of transport research. He is a PhD candidate at The University of Sydney and holds a MSc in Transport Engineering from Universidad de Chile. His main research interests are public transport economics, sustainable transport and externalities.

Cristián E. Cortés is Assistant Professor at the Civil Engineering Department, Universidad de Chile. Holds a PhD in Civil Engineering from University of California and a MSc in Transport Systems Analysis from Universidad de Chile. Author of more than 20 articles in ISI indexed journals, on public transport optimization, network optimization and equilibrium, logistics, simulation, control applied to dynamic transport systems, among his topics of major interest. Dr. Cortés is currently Associate Editor of Transportation Science, member of the Directory of the Chilean Society in Transport Engineering, and leader of applied research projects in public transport planning.

Sergio Jara-Díaz is Professor of Transport Economics at Universidad de Chile. Holds a PhD and MSc from MIT, where he has taught during various terms. Author of Transport Economic Theory (Elsevier, 2007) and more than 80 articles in journals and books, on the microeconomics of transport demand, multioutput analysis in transport industries, public transport modelling and pricing, and time allocation. He teaches frequently in Spain. Resides in Ñuñoa, Santiago, with his only wife. http://www.cec.uchile.cl/~dicidet/sergio.html.