Public-Transit Frequency Setting Using Minimum-Cost Approach with Stochastic Demand

Yuval Hadas
Department of Civil and Environmental Engineering
University of Auckland, New Zealand
and
Department of management, Faculty of Social Sciences
Bar-Ilan University, Israel
y.hadas@auckland.ac.nz, yuval.hadas@biu.ac.il

Matan Shnaiderman
Department of management, Faculty of Social Sciences
Bar-Ilan University, Israel
matan.shnaiderman@biu.ac.il

Avishai (Avi) Ceder
Department of Civil and Environmental Engineering
University of Auckland, New Zealand
a.ceder@auckland.ac.nz

Abstract

Common practice in Public-Transit planning is to determine the frequency of service based on accumulated hourly passenger counts, average travel time, given vehicle capacity, and the standard of minimum frequency by time of day. With the increased usage of automatic vehicle location (AVL) and automatic passenger counting (APC) systems, it is possible to construct the statistical distributions of passenger demand and travel time by time of day. This can give rise to improve the accuracy of the determination of frequencies. This study presents a new approach enabling the use of stochastic properties of the collected data and its associated costs. An optimization framework is constructed based on two main cost elements: (a) empty-seat driven (unproductive cost), and (b) overload and un-served demand (increased user cost). The objective function is to minimize the total cost incurred with decision variables of either frequency or vehicle capacity (vehicle size). That is, from the operator perspective it is desirable to utilize efficiently the fleet of vehicles (decisions of the vehicle size). From the authority perspective, the concern is to provide an adequate level of service in terms of frequency. The study contains sensitivity analysis of the cost elements for economic evaluation.

Key words: Public-Transit, Frequency, Optimization.

1 Introduction

Public-Transit planning consists of five major components: network design, setting frequencies, timetable development, vehicle scheduling, and crew scheduling (Ceder and Wilson, 1986). Common practice in Public-Transit planning is to determine the frequency of service based on accumulated hourly passenger counts, average travel time, given vehicle capacity, desired occupancy (load standard) and the minimum frequency permitted by time of day (Ceder, 2007). Other models for frequency settings exists, such an economic costs model which is based on a simple route (Jara-Díaz et al., 2008). Another model introduces performance measures for crowdedness level (COD) and probability of not allowed to alight (POF) (Grosfeld-Nir and Bookbinder, 1995). This model assumes simple stochastic attributes of arrival and departure rates, and treats vehicle capacity as unlimited. An optimization model for headways was presented as part of a more general cost model for public transit route design (Kocur and Hendrickson, 1982), with the assumption of evenly spaced stops and demand.

Automatic Vehicle Location (AVL) and Automatic Passenger counting (APC) are two technologies enabling tracking the location of vehicles en route (AVL), and collecting the number of passengers alighted and boarded at each stop (APC). The data acquired can be used for analysis, as well as to enhance the performance of public transit systems, and the introduction of advanced models. It was shown that the availability of bus locations, estimated arrival times, number of passengers and their destinations, can open the door to the implementation of bus-dispatching at timed transfer transit stations algorithm (Dessouky et al., 1999). Such an algorithm can intelligently decide whether to hold a bus in order to achieve a transfer with a late bus or not. Based on AVL technology it is possible to forecast accurately the buses estimated arrival times and to use bus holding strategies to coordinate transfers (Dessouky et al., 2003). The use of advanced public transit systems in fixed-route and paratransit operations was found important for improvements in departure times and transfers (Levine et al., 2000). Travel time estimation is also possible (Tétreault and El-Geneidy, 2010), as well as the evaluation of transit operations based on both AVL and APC data (Strathman et al., 2002).

Advanced supply chain models, specifically inventory models (Lee et al., 2000, Zipkin, 2000), are used to optimize the total costs occurs within a time frame, by setting an inventory replenishment strategies (Kogan and Shnaiderman, 2010). The costs are associated with shortage and overage (or surplus). The former relevant when demand is not met, while the later when inventory is higher than the demand. Such models can be used to formulate an optimal frequency setting model, in which shortage and overage costs are transformed to overload and empty-seats costs respectively, and inventory strategy transforms to bus capacity.

This paper presents a new concept for optimal frequency setting, which based on supply chain models that integrates costs, stochastic demand and travel time; it is organized as follows. Section 2.1 introduces a general formulation of the model, followed by deterministic and stochastic variations of the model (Sections 2.2, 2.3, 2.4). Section 3 provides a comparison of the different models. Section 4 presents the conclusions and recommendations of study.

2 Cost-based approach of frequency setting

Consider a single route in which several vehicles of the same capacity, are serving the route at a fixed frequency (or headway). Let *i*, *i*=1..*I*, be the vehicle index, and let *k*, *k*=1..*K*, be the stop index. *CP* will denote the capacity of each vehicle. Let μ_k be the average travel time between stops *k* and *k*+1. Also, the passenger demand (the unconstrained load) from stop k will be denoted by d_{ik}. For simplicity, the model uses headways instead of frequencies, thus for a frequency *f* (vehicles per hour), the equivalent headway *H* (minutes between consecutive departures) is $H = 60/f$.

The study begins with a general formulation of the model including stochastic demand and distributed travel time. The study continuous with a deterministic model, stochastic demand model, and stochastic demand and distributed travel time.

2.1 Optimal cost-based frequency setting model

The study uses two common frequency setting methods. The first is based on a point check and called hourly max-load-point method and the second is based on ride check and called weighted average load profile method (Ceder, 2007); the latter method is introduced for comparison.

Assume that for a time period T (which is measured in minutes, and usually referred to one hour), the average demand at stop *k*, for every $1 \le k \le K-1$ is $\overline{D}_k$, and L_0 is the desired occupancy.

According to the hourly max load point method, the headway is determined by the highest loaded stop. Generally, the headway is equal to $L_0 T \big/ \max_k \overline{D}_k$. However, there are two constraints which refer to the lower and upper bounds of possible headways. The lower bound is equal to T/N, where N denotes the number of vehicles available to serve during time period T. The upper bound is determined by the service level, and is denoted by H_0.

Taking into account the constraint

(1) $$\frac{T}{N} \le H \le H_0$$

the following headway is obtained

(2) $$H = \min\left\{ \max\left(\frac{L_0 T}{\max_k \overline{D}_k}, \frac{T}{N} \right),\ H_0 \right\}.$$

For the weighted average load profile method, we've substituted the distance travelled by the travel time, so a comparison with the general model is possible.

(3) $$L_0 T \sum_{k=1}^{K-1} \mu_k \Bigg/ \sum_{k=1}^{K-1} \left(\mu_k \sum_{i=1}^{I} d_{ik} \right).$$

with the constraint

(4) $$H \le \frac{CP \cdot T}{\max_k \overline{D}_k}.$$

Taking into account (3), (1) and (4), the following headway is attained:

(5) $$H = \min\left\{ \max\left(L_0 T \sum_{k=1}^{K-1} \mu_k \Big/ \sum_{k=0}^{K-1} \left(\mu_k \sum_{i=1}^{I} d_{ik} \right),\ \frac{T}{N} \right),\ H_0,\ \frac{CP \cdot T}{\max_k \bar{D}_k} \right\}.$$

Formulas (2) and (5) are deterministic in nature; moreover, overload and empty-seats costs are not included. Hence, a new approach, based on supply chain principles (Kogan and Shnaiderman, 2010) is introduced.

Let c^+ be the empty seat overage cost per time unit, let c^- as the unserved passenger shortage cost per time unit, and let R_{ik} be the running time between stop $k-1$ to stop *k*, for vehicle *i*.

If the load is smaller than the capacity CP, then overage cost of vehicle i at stop k is equal to

(6) $$R_{ik} c^+ \max(CP - d_{ik}, 0).$$

On the other hand, if the load is higher than the capacity, then the shortage cost of vehicle i at stop k is equal to

(7) $$C_{ik}^- = R_{ik} c^- \max(d_{ik} - CP, 0).$$

The total cost of vehicle i at stop k is therefore the sum of (6) and (7):

(8) $$C_{ik} = R_{ik} c^+ \max(CP - d_{ik}, 0) + R_{ik} c^- \max(d_{ik} - CP, 0).$$

Proposition 1: Assuming that either a) R_{ik} are deterministic or b) The probabilities $\left\{ \Pr(L_{ik} = n) \right\}_{n=0}^{\infty}$ do not depend on the exact distributions of R_{ik}, but only on their expected values, then the expected cost C_{ik} depends on the distribution of L_{ik} as follows:

(9) $$\mathrm{E}[C_{ik}] = \mu_k \sum_{n=0}^{CP} c^+ (CP - n) \Pr(L_{ik} = n \mid H) + \mu_k \sum_{n=CP+1}^{\infty} c^- (n - CP) \Pr(L_{ik} = n \mid H).$$

The objective function is the expected total cost for all vehicles and stops:

(10) $$\text{Minimize } \mathrm{ETC}(\mathrm{H,CP}) = \sum_{i=1}^{I} \sum_{k=1}^{K-1} \mathrm{E}[C_{ik}].$$

with decision variables H and CP.

Let L_{ik} be the load in vehicle i after departing from stop k:

(11) $$L_{ik} = \min(d_{ik}, CP).$$

Let A_{ik} denotes the number of passengers alighting from vehicle i at stop k, and B_{ik} the number of passengers arriving at stop k *after* the previous vehicle ($i-1$) has left. Also, let r_{ik} be the number of un-served passengers by vehicle i at stop k, that is

(12) $$r_{ik} = \max(d_{ik} - CP, 0).$$

Thus, the following formula is derived for the demand d_{ik}:

(13) $$d_{ik} = L_{i,k-1} - A_{ik} + r_{ik} + B_{ik} = \min(d_{i,k-1}, CP) - A_{ik} + \max(d_{i-1,k} - CP, 0) + B_{ik}$$

$2 \le i \le I$ and $2 \le k \le K-1$

with the initial conditions:

(14)
$$\begin{cases} d_{1k} = L_{1,k-1} - A_{ik} + B_{ik} = \min(d_{i,k-1}, CP) - A_{ik} + B_{ik}, \quad k \geq 2 \\ d_{i1} = r_{i1} + B_{i1} = \max(d_{i-1,1} - CP, 0) + B_{i1}, \quad i \geq 2 \\ d_{11} = B_{11}. \end{cases}$$

Proposition 2: The probability of d_{ik} to be n, $n \geq 0, n \in \mathbb{Z}$ is:

(15)
$$\begin{cases}
\Pr(d_{ik} = n) = \\
\sum_{\ell=0}^{n-1} \Pr(d_{i,k-1} = \ell) \left[\sum_{m=0}^{\ell} \Pr(A_{ik} = m \mid L_{i,k-1} = \ell) \left(\begin{array}{l} \sum_{j=0}^{CP-1} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n - \ell + m) + \\ \quad + \sum_{j=CP}^{CP+n-\ell+m} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = CP + n - \ell + m - j) \end{array} \right) \right] + \\
\quad + \sum_{\ell=n}^{CP} \Pr(d_{i,k-1} = \ell) \left[\sum_{m=\ell-n}^{\ell} \Pr(A_{ik} = m \mid L_{i,k-1} = \ell) \left(\begin{array}{l} \sum_{j=0}^{CP-1} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n - \ell + m) + \\ \quad + \sum_{j=CP}^{CP+n-\ell+m} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = CP + n - \ell + m - j) \end{array} \right) \right] + \\
\quad + \sum_{\ell=CP+1}^{\infty} \Pr(d_{i,k-1} = \ell) \left[\sum_{m=CP-n}^{CP} \Pr(A_{ik} = m \mid L_{i,k-1} = CP) \left(\begin{array}{l} \sum_{j=0}^{CP-1} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n - CP + m) + \\ \quad + \sum_{j=CP}^{n+m} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n + m - j) \end{array} \right) \right], \\
\hfill \text{if} \quad n \leq CP \\
\sum_{\ell=0}^{CP} \Pr(d_{i,k-1} = \ell) \left[\sum_{m=0}^{\ell} \Pr(A_{ik} = m \mid L_{i,k-1} = \ell) \left(\begin{array}{l} \sum_{j=0}^{CP-1} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n - \ell + m) + \\ \quad + \sum_{j=CP}^{CP+n-\ell+m} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = CP + n - \ell + m - j) \end{array} \right) \right] + \\
\quad + \sum_{\ell=CP+1}^{\infty} \Pr(d_{i,k-1} = \ell) \left[\sum_{m=0}^{CP} \Pr(A_{ik} = m \mid L_{i,k-1} = CP) \left(\begin{array}{l} \sum_{j=0}^{CP-1} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n - CP + m) + \\ \quad + \sum_{j=CP}^{n+m} \Pr(d_{i-1,k} = j) \Pr(B_{ik} = n + m - j) \end{array} \right) \right], \\
\hfill \text{if} \quad n \geq CP + 1
\end{cases}$$

$\forall\ 2 \leq i \leq I,\ 2 \leq k \leq K-1$

(where the probabilities $\{\Pr(d_{i-1,k} = j)\}$ are under the condition $d_{i,k-1} = \ell$, and the probabilities $\{\Pr(B_{ik} = J)\}$ are under the conditions $d_{i,k-1} = \ell$ and $d_{i-1,k} = j$).
The initial conditions are:

$$\Pr(d_{1k}=n)=$$

(16)
$$\begin{cases}\sum_{\ell=0}^{n-1}\Pr(d_{1,k-1}=\ell)\left[\sum_{m=0}^{\ell}\Pr(A_{1k}=m\,|L_{1,k-1}=\ell)\Pr(B_{1k}=n-\ell+m|d_{i,k-1}=\ell)\right]+\\ \sum_{\ell=n}^{CP}\Pr(d_{1,k-1}=\ell)\left[\sum_{m=\ell-n}^{\ell}\Pr(A_{1k}=m|L_{1,k-1}=\ell)\Pr(B_{1k}=n-\ell+m|d_{i,k-1}=\ell)\right]+\\ \sum_{\ell=CP+1}^{\infty}\Pr(d_{1,k-1}=\ell)\left[\sum_{m=CP-n}^{CP}\Pr(A_{1k}=m|L_{1,k-1}=CP)\Pr(B_{1k}=n-CP+m|d_{i,k-1}=\ell)\right], & \text{if} \quad n\le CP\\ \\ \sum_{\ell=0}^{CP}\Pr(d_{1,k-1}=\ell)\left[\sum_{m=0}^{\ell}\Pr(A_{1k}=m|L_{1,k-1}=\ell)\Pr(B_{1k}=n-\ell+m|d_{i,k-1}=\ell)\right]+\\ \sum_{\ell=CP+1}^{\infty}\Pr(d_{1,k-1}=\ell)\left[\sum_{m=0}^{CP}\Pr(A_{1k}=m|L_{1,k-1}=CP)\Pr(B_{1k}=n-CP+m|d_{i,k-1}=\ell)\right], & \text{if} \quad n\ge CP+1\end{cases}$$

$$\forall\ 2\le k\le K-1$$

$$\Pr(d_{i1}=n)=\sum_{j=0}^{CP}\Pr(d_{i-1,1}=j)\Pr(B_{i1}=n|d_{i-1,1}=j)+$$

(17)
$$\sum_{j=1}^{n}\Pr(d_{i-1,1}=CP+j)\Pr(B_{i1}=n-jd_{i-1,1}=CP+j)$$
$$\forall\ 2\le i\le I$$

(18)
$$\Pr(d_{11}=n)=\Pr(B_{11}=n).$$

2.2 Optimal frequency setting with deterministic data

In this case we assume that the travel and dwell times are all deterministic. Assume that the rate in which passengers arrive to stop k (denoted by λ_k) is known, then the variable B_{ik}, which is approximately equal to $\lambda_k H$, is necessarily equal to some $n_k^{(B)}$ which satisfies $\left|n_k^{(B)}-\lambda_k H\right|=\min_{n\in\mathbb{N}\cup\{0\}}\left|n-\lambda_k H\right|$, to obtain

(19)
$$\Pr(B_{ik}=n)=\begin{cases}1, & n=n_k^{(B)}\\ 0, & n\ne n_k.\end{cases}$$

Also, we assume that at stop k it is known that p_k of the passengers alight ($0\le p_k\le 1$), then A_{ik} is necessarily equal to a value $n_k^{(A)}$ such that $\left|n_k^{(A)}-p_k L_{i,k-1}\right|=\min_{n\in\mathbb{N}\cup\{0\}}\left|n-p_k L_{i,k-1}\right|$, that is

(20)
$$\Pr(A_{ik}=n)=\begin{cases}1, & n=n_k^{(A)}\\ 0, & n\ne n_k.\end{cases}$$

Therefore, the value of d_{ik} is well known for every $1\le i\le I$ and $1\le k\le K-1$, and it is calculated from (13)-(14). The expected cost is then found by (8).

2.3 Optimal frequency setting with stochastic demand

For this model we assume that demand is stochastic, the variables $\{\lambda_k, p_k\}$ are stochastic and finite-valued. That is, for every $1 \le k \le K-1$, $\lambda_k \in \{\lambda_{k_1}, \ldots, \lambda_{k_{j_1(k)}}\}$, such that for every $1 \le \ell \le j_1(k)$ $\Pr(\lambda_k = \lambda_{k_\ell}) = q_{k_\ell}^{(B)}$, and for every $2 \le k \le K-1$, $p_k \in \{p_{k_1}, \ldots, p_{k_{j_2(k)}}\}$, such that for every $1 \le \ell \le j_2(k)$ $\Pr(p_k = p_{k_\ell}) = q_{k_\ell}^{(A)}$.

In order to calculate the probabilities (15)-(18), we define for every $1 \le \ell \le j_1(k)$ the index $n_{k_\ell}^{(B)}$ which satisfies $\left| n_{k_\ell}^{(B)} - \lambda_{k_\ell} H \right| = \min_{n \in \mathbb{N} \cup \{0\}} \left| n - \lambda_{k_\ell} H \right|$.

In addition, given a value of $L_{i,k-1}$, we define the index $n_{k_\ell}^{(A)}$ which satisfies $\left| n_{k_\ell}^{(A)} - p_{k_\ell} L_{i,k-1} \right| = \min_{n \in \mathbb{N} \cup \{0\}} \left| n - p_{k_\ell} L_{i,k-1} \right|$ for $1 \le \ell \le j_2(k)$.

Therefore, the probabilities (19) and (20) become now

(21) $$\Pr(B_{ik} = n) = \begin{cases} q_{k_1}^{(B)}, & n = n_{k_1}^{(B)} \\ \vdots & \\ q_{k_{j_1(k)}}^{(B)}, & n = n_{k_{j_1(k)}}^{(B)} \\ 0, & \text{else} \end{cases}$$

and

(22) $$\Pr(B_{ik} = n) = \begin{cases} q_{k_1}^{(A)}, & n = n_{k_1}^{(A)} \\ \vdots & \\ q_{k_{j_2(k)}}^{(A)}, & n = n_{k_{j_2(k)}}^{(A)} \\ 0, & \text{else} \end{cases}$$

respectively.

2.4 Optimal frequency setting with stochastic demand and travel time

This section incorporates the general cost-based model with properties from a previous work that modelled stochastic demand and dwell time (Bellei and Gkoumas, 2010, Hickman, 2001).

Let D_{ik} be the dwell time of vehicle i at stop k. Also, let H_{ik} be the headway between the departure times of vehicles $i-1$ and i (for $2 \le i \le I$) at stop k, then:

(23) $$\begin{cases} H_{i1} = H + D_{i1} - D_{i-1,1} \\ H_{ik} = H_{i,k-1} + R_{ik} - R_{i-1,k} + D_{ik} - D_{i-1,k}, \quad k \ge 2 \end{cases}$$

(Hickman, 2001).

Assuming that the headways $\{H_{ik}\}$ for $2 \le i \le I$ and $1 \le k \le K-1$ are stochastic, and that boarding is performed *after* alighting (there are researches which assume that these events are performed in parallel, for example Sun and Hickman, 2008), the stochastic dwell time of vehicle i at stop k is equal to

(24) $$D_{ik} = a + b_A A_{ik} + b_B (r_{i-1,k} + B_{ik}),$$

where a denotes the lost time due to accelerating and decelerating of the vehicles at each stop, b_A denotes the incremental time for a single passenger to alight from the vehicle, and b_B denotes the incremental time for a single passenger to board the vehicle. The travel times between the stops are stochastic as well, satisfying an autoregressive process AR(1)

(25) $$\begin{cases} R_{1k} - \mu_k = \varepsilon_{1k} \\ R_{ik} - \mu_k = \alpha_k (R_{i-1,k} - \mu_k) + \varepsilon_{ik}, \quad 2 \le i \le I \end{cases}$$

for $2 \le k \le K$ (Mishalani et al., 2008). The parameter α_k in (25) denotes the correlation between the travel times of two consecutive vehicles, and $0 \le \alpha_k < 1$. In the special case of $\alpha_k = 0$, the travel times (from stop $k-1$ to stop k) of all the vehicles are identically distributed and independent. Also, the means of the noises $\{\varepsilon_{ik}\}$ are all zero. From (25) we have $R_{ik} = \mu_k + \sum_{j=1}^{i} \alpha^{i-j} \varepsilon_{jk}$, and therefore, while there are no real-time updates, the expected value of R_{ik} is μ_k, namely

(26) $$\mathrm{E}[R_{ik}] - \mathrm{E}[R_{i-1,k}] = 0.$$

The variable B_{ik} is distributed according to a Poisson process with mean $\lambda_k \mathrm{E}[H_{ik}]$ (Hickman, 2001). Therefore, the probability (21) becomes

(27) $$\Pr(B_{ik} = n) = \frac{\left(\lambda_k \mathrm{E}[H_{ik}]\right)^n}{n!} e^{-\lambda_k \mathrm{E}[H_{ik}]}.$$

Given the value of the load $L_{i,k-1}$, the variable A_{ik} is binomially distributed such that $A_{ik} \sim \mathrm{B}(L_{i,k-1}, p_k)$ for some $0 \le p_k \le 1$. Thus, (22) is now equal to

(28) $$\Pr(A_{ik} = n) = \binom{L_{i,k-1}}{n} p_k^n (1-p_k)^{L_{i,k-1}-n}.$$

In order to calculate the probabilities (15)-(18) the calculation of the expected headway H_{ik} ought to take place (see (27)) for $2 \le i \le I$ and $1 \le k \le K-1$. The following proposition is then used:

Proposition 3: The expected headways at stop 1 are

(29) $$\begin{cases} \mathrm{E}[H_{21}] = \dfrac{H + b_B\left(\mathrm{E}[r_{11}] - \lambda_1 H\right)}{1 - \lambda_1 b_B} \\ \mathrm{E}[H_{i1}] = \dfrac{H + b_B\left(\mathrm{E}[r_{i-1,1}] - \mathrm{E}[r_{i-2,1}] - \lambda_1 \mathrm{E}[H_{i-1,1}]\right)}{1 - \lambda_1 b_B}, \quad i > 2. \end{cases}$$

The expected headways at stop $k \ge 2$ are

(30) $$\begin{cases} \mathrm{E}[\mathrm{H}_{2k}] = \dfrac{\mathrm{E}[H_{i,k-1}] + b_A p_k\left(\mathrm{E}[L_{2,k-1}] - \mathrm{E}[L_{1,k-1}]\right) + b_B\left(\mathrm{E}[r_{1k}] - \lambda_k H\right)}{1 - \lambda_k b_B} \\ \mathrm{E}[\mathrm{H}_{ik}] = \dfrac{\mathrm{E}[H_{i,k-1}] + b_A p_k\left(\mathrm{E}[L_{i,k-1}] - \mathrm{E}[L_{i-1,k-1}]\right) + b_B\left(\mathrm{E}[r_{i-1,k}] - \mathrm{E}[r_{i-2,k}] - \lambda_k \mathrm{E}[H_{i-1,k}]\right)}{1 - \lambda_k b_B}, \quad i > 2. \end{cases}$$

Furthermore, it is required that

(31) $\lambda_k b_B < 1$.

Note that from (11) and (12) we respectively obtain:

$E[L_{ik}] = \sum_{n=0}^{CP} n \Pr(d_{ik} = n) + CP \sum_{n=CP+1}^{\infty} \Pr(d_{ik} = n)$ and $E[r_{ik}] = \sum_{n=CP+1}^{\infty} (n - CP) \Pr(d_{ik} = n)$.

3 Evaluation of Models

Evaluating the performance and the benefits of the models was carried out with the following simple example. Consider a route of five stops, with vehicle capacity $CP = 80$, the arrival rates (per one minute) to the stops are $\lambda_1 = 3.1, \lambda_2 = 1.4, \lambda_3 = 0.65, \lambda_4 = 0.35$, the alighting passengers fractions are $p_2 = 0.09, p_3 = 0.19, p_4 = 0.42$, and the average travel times are $\mu_2 = 10, \mu_3 = 6, \mu_4 = 20, \mu_5 = 3$. According to the Ceder's model (2007), the average loads at stops 1,2,3 and 4 (during 3 hours) are 558, 760, 733 and 488. The suggested headway, according to the weighted average load profile(5), is 17 minutes. A comparison between the four models is summarized in Table 1. Also, a comparison of the expected costs obtained from each model to those obtained from the weighted average load profile is presented.

Table 1 - Frequency setting models' comparison

Model	Variance	$C^+=C^-$	$C^+<C^-$	$C^+>C^-$
Weighted ave. load profile	-	17	17	17
Deterministic	-	19 (26%)	18 (9%)	26 (79%)
Stochastic demand	Low	20 (36%)	18 (9%)	27 (77%)
	High	21 (11%)	15 (20%)	30* (60%)
Stochastic demand and travel time	Low	20 (33%)	17 (0%)	27 (75%)
	High**	7 (68%)	6 (60%)	7 (78%)

* The optimal headway is higher, but due to service level we assumed an upper bound of 30.

** The use of Poisson and Binomial distributions affected the result of the weighted average load profile, thus the stochastic demand and travel time model were compared with 3 minutes.

The results show that both the ratio between the costs and the variability of demand and travel time affect the optimal headway. High variability cause extreme headway, while a high empty-seats cost tends to increase the headway, in contrast to high overload cost which obviously decrease headway (and increase the level of service).

4 Conclusions

An optimal model was developed taking into account the costs associated with running empty-seats and passenger overload, and considering stochastic demand and distributed travel time. The model provides optimal frequency, based on quantitative costs. Moreover, the model reflects the necessary adjustments of the frequency because of the fluctuation of travel time and demand.

The availability of AVL and APC data across transit agencies and operators, justifies the development of models that estimate the demand and travel time (ride time and dwell time) statistical distributions. These data are key input to the optimal frequency setting model.

The sensitivity analysis of the costs introduced portrays the avenues for the authorities and operators to attain a better decision-making process to reach an improved service and more efficient use of resources.

As both frequency and vehicle capacity are decision variables, an optimal capacity variation of the model was developed.

References

G. Bellei, Gkoumas, K. 2010. Transit vehicles' headway distribution and service irregularity. *Public Transport* 1-21.

A. Ceder. 2007. *Public Transit Planning and Operation: Theory, Modeling and Practice*. Butterworth-Heinemann, Oxford, UK.

A. Ceder, Wilson, N.H.M. 1986. Bus network design. *Transportation Research B*. **20**(4) 331-344.

M. Dessouky, Hall, R., Nowroozi, A., Mourikas, K. 1999. Bus dispatching at timed transfer transit stations using bus tracking technology. *Transportation Research Part C: Emerging Technologies*. **7**(4) 187-208.

M. Dessouky, Hall, R., Zhang, L., Singh, A. 2003. Real-time control of buses for schedule coordination at a terminal. *Transportation Research Part A: Policy and Practice*. **37**(2) 145-164.

A. Grosfeld-Nir, Bookbinder, J.H. 1995. The planning of headways in urban public transit. *Annals of Operations Research*. **60**(1-4) 145-160.

M.D. Hickman. 2001. An Analytic Stochastic Model for the Transit Vehicle Holding Problem. *Transportation Science*. **35**(3) 215.

S. Jara-Díaz, Tirachini, A., Cortés, C.E. 2008. Modeling public transport corridors with aggregate and disaggregate demand. *Journal of Transport Geography*. **16**(6) 430-435.

G. Kocur, Hendrickson, C. 1982. Design of Local Bus Service with Demand Equilibration. *Transportation Science*. **16**(2) 149.

K. Kogan, Shnaiderman, M. 2010. Continuous-Time Replenishment Under Intermittent Observability. *Automatic Control, IEEE Transactions on*. **55**(6) 1460-1465.

H.L. Lee, So, K.C., Tang, C.S. 2000. The Value of Information Sharing in a Two-Level Supply Chain. *Management Science*. **46**(5) 626.

J. Levine, Hong, Q., Hug, G., Jr, Rodriguez, D. 2000. Impacts of an Advanced Public Transportation System Demonstration Project. *Transportation Research Record*(1735) 169-177.

R.G. Mishalani, McCord, M.R., Forman, S. 2008. *Schedule-based and autoregressive bus running time modeling in the presence of driver-bus heterogeneity*. City.

J.G. Strathman, Kimpel, T.J., Dueker, K.J., Gerhart, R.L., Callas, S. 2002. Evaluation of transit operations: data applications of Tri-Met's automated Bus Dispatching System. *Transportation*. **29**(3) 321-345.

A. Sun, Hickman, M. 2008. *The Holding Problem at Multiple Holding Stations*. Springer Berlin Heidelberg, City.

P.R. Tétreault, El-Geneidy, A.M. 2010. Estimating bus run times for new limited-stop service using archived AVL and APC data. *Transportation Research Part A: Policy and Practice*. **44**(6) 390-402.

P.H. Zipkin. 2000. *Foundations of inventory management*. McGraw-Hill, Boston.