A stochastic model for bus injection in a public transport service

Diego Morales · Juan Carlos Muñoz

Abstract  Randomness affecting the operation of public transport systems generates significant increments in waiting times. A strategy to deal with this randomness is bus injection, which holds buses in specific points along the route ready to be dispatched when an event such as an extremely long headway occurs. In this work, stochastic models based on the second moment of the headways distribution are developed to determine if bus injection is worth using at a given public transport service, how to operate it optimally and its impact on each stop. A single stop approach is initially used to determine a closed expression for the optimal time threshold for dispatch, then a complete service approach determines the optimal stop to locate the injection fleet and the instant to dispatch it. Simulations with real data are used to test models, proving their accuracy in terms of measuring impact on waiting times. Results show that injection is worth operating in some cases, reducing total waiting times compared to constant circulation.

Keywords  Bus injection · Transit operations · Bus bunching · Stochastic models

1 Introduction and motivation

Randomness in passenger arrival times and vehicle travel times affects the operation of public transport systems significantly. If a vehicle gets delayed, the number of people waiting to ride increases, which reduces its speed even more. The vehicle that follows the delayed one will follow a short interval, so its speed will grow quite fast if no action is made to prevent it. This phenomena, known as bus bunching since it affects buses more than trains, damages waiting time reliability and comfort [Delgado et al (2016)]. Headway regularity has been identified as the key operational attribute besides speed needed by the bus industry to improve the
quality of its service. Headway regularity is important because randomly arriving passengers have proportionally more chance of arriving during a long headway than during a short. It defines the reliability of a bus route and the amount of trust that passengers can have in it [Group et al (2013)].

Bus bunching is a source of concern for both transit riders and providers. It increases passenger waiting time and crowding by augmenting the variability of headways. When buses are traveling together, they operate at the speed of the leading vehicle, which is the slowest. Bunching also forces passengers to arrive early at stations and to budget long travel time Berrebi et al (2015).

Several strategies have been studied for dealing with this variability. Hickman proposed a stochastic model for vehicle holding based on recursive equations for expected values of headways and bus loads Hickman (2001). His control strategy consists of holding operating buses along the road to regularize the system. Delgado et al use continuous variables for a mixed strategy in which vehicles are held and passenger boarding may be limited to improve headway regularity [Delgado et al (2009)]. Adamski and Turnau provided a control strategy in which buses are dispatched at certain moments that will allow a more regular arrival at critical high demand stops along the route [Adamski and Turnau (1998)].

A strategy that has not been studied in detail is bus injection. In this strategy, one or more buses are kept at specific points along the route waiting to be included when a specific operational conditional arises. For example, in some Metro systems a train is injected at a critical combination station when the passenger accumulation reaches a certain threshold. Many operational conditions can be used as a triggering event, including thresholds for headway time, bus load and number of waiting passengers at specific stops. In our case we will consider injecting a bus when a headway exceeds a certain threshold.

This paper proposes a stochastic model that allows us to determine under which conditions a bus should be injected at a stop, and determines the convenience of using this strategy. Thus, the model can be used to determine the best stop to inject the bus. Initially the impact of injecting a single bus is measured only at the stop where it is injected. Then a more general model in which multiple buses may be injected is proposed and then a model measuring the impact at all stops downstream is developed. Finally, the best instant within the headway to dispatch the injected bus is discussed.

2 Model for injection considering a single stop

Consider a stop visited by a single bus service. The headways \( H \) between arriving buses follow a known distribution. Let \( W \) be the waiting time of a random passenger arriving to the stop who always boards the first arriving bus (so bus capacity never binds). Osuna and Newell [Osuna and Newell (1972)] show that if the arrival of passengers is independent to the arrival of the buses, then

\[
E[W] = \frac{E[H^2]}{2E[H]} \tag{1}
\]

We assume that a given fleet is available. If all the buses in the fleet are used for the operation then \( E[H] \) is constant. This means that waiting times of passengers
are directly related to the second moment of $H$, hence to minimize waiting times, $\mathbb{E}[H^2]$ must be minimized.

2.1 Single injection

Define $H$ as the headways sequence of the $N$ next buses arriving to the stop and assume that every element of $H$ is identically distributed according to $H$ (but not necessarily independent). Assume that one bus is kept at the stop ready to be dispatched as soon as the following condition is triggered. The bus will be dispatched at the midpoint of the first headway larger than a given threshold $l$. Note that since we are only interested in this stop, the midpoint is clearly the optimal time to reduce the squared headways, however, since the injected bus is different to the rest, this will not be true once we consider the waiting times at downstream stops. The purpose of this section is to identify the optimal length of $l$ so that the waiting time of all passengers at this stop is minimized.

To implement this strategy, we assume that the controller knows the arrival time of the next bus at all times so the bus can be injected exactly at the middle of the headway. We denote $Z$ as the headway sequence where the bus has been injected. For example, if the bus is injected after the $k$th arrival, the process $Z$ will have 2 headways of length $H_k/2$ instead of a single $H_k$. This will be the only difference between $H$ and $Z$ as the injection won’t affect any other interval at the stop. So, if the $k$th headway is the first one larger than $l$, then the two sequences will be:

$$H = (H_1, \ldots, H_k, \ldots, H_N) \quad \text{and} \quad Z = (H_1, \ldots, \frac{H_k}{2}, H_k, \frac{H_k}{2}, \ldots, H_N)$$

If no headway in the sequence happens to be larger than $l$ then $H = Z$. It is important to note that $Z$ is dependent of the choice of $l$ but $H$ clearly is not. Figure 1 is an example of the sequences $H$ and $Z$ considering $l = 15$ and $l = 20$. Each node is a bus arrival at the stop and each edge represents the headway between the buses. Dotted traces and nodes are injections. Note that in this example, clearly the second sequence of $Z$ yields a lower average waiting time than the first one since the largest headway was covered by the injection. However, the decision of the threshold length has to be made before the headways are observed. The controller does not know the actual sequence $H$, only its distribution. A threshold set too high risks not injecting the bus in the whole period which means not using a valuable resource.

Since the goal is to minimize waiting times at the stop after the bus injection policy has been implemented, we minimize the second moment of $Z$, defined as

$$\mathbb{E}[Z^2] := \sum_{k \in N} (Z_k)^2$$

In order to compute this, let $K$ be a random variable representing the number of headways in $H$ that are larger than $l$. With this, conditioning on the number of headways larger than $l$:

$$\mathbb{E}[Z^2] = \mathbb{E}_K[\mathbb{E}[Z^2 | \text{Exactly } l \text{ headways are larger than } l]]$$

and since $K \in N_0$. 

Sequence $H$ (no injection):

$H = 6 \quad H = 4 \quad H = 16 \quad H = 9 \quad H = 22 \quad H = 7$

Sequence $Z$ with threshold $l = 15$:

$H = 6 \quad H = 4 \quad H = 8 \quad H = 9 \quad H = 22 \quad H = 7$

Sequence $Z$ with threshold $l = 20$:

$H = 6 \quad H = 4 \quad H = 16 \quad H = 9 \quad H = 11 \quad H = 11 \quad H = 7$

Fig. 1: Example of $H$ and $Z$ with different thresholds

$$\mathbb{E}[Z^2] = \sum_{i=0}^{N} \mathbb{E}[Z^2|K = k]p[K = k] \quad (2)$$

Defining $p_k = p[K = k]$, the expected value of $Z$ can be calculated as follows for the range of $K$:

- $K = 0$:
  If all headways are shorter than $l$, then $Z = H$ (no injection is made) which results in
  $$\mathbb{E}[Z^2|K = 0] = \mathbb{E}[H^2|K = 0] = N\mathbb{E}[H^2|H \leq l] \quad (3)$$

- $K = 1$:
  The only headway larger than $l$ will be the one where a bus is injected, so $Z$ results in:
  $$Z = (H_1, \ldots, H_{k-1}, \frac{H_k}{2}, \frac{H_k}{2}, H_{k+1}, \ldots, H_N)$$

Where $H_k$ is larger than $l$ (a bus was injected) and every other headway is smaller than $l$, so

$$\mathbb{E}[Z^2|K = 1] = \mathbb{E}[H^2/4|H > l] + \mathbb{E}[H^2/4|H > l] + (N-1)\mathbb{E}[H^2|H \leq l] \quad (4)$$

- $K = k > 0$:
  $Z$ will have the same form as the last case, except now there are $k-1$ headways larger than $l$ besides the injection and $N - k$ headways where $H \leq l$. Using this, the expression becomes

$$\mathbb{E}[Z^2|K = k] = (k-1/2)\mathbb{E}[H^2|H > l] + (N-k)\mathbb{E}[H^2|H \leq l] \quad (5)$$
Notice that this expression is also correct for the $K = 1$ case.

If we define $p_k = P[K = k]$, replacing (3) and (5) in (2) results in

$$
\mathbb{E}[Z^2] = NE[H^2|H \leq l]p_0 + \sum_{k=1}^{N} \left( (k - 1/2)\mathbb{E}[H^2|H > l] + (N - k)\mathbb{E}[H^2|H \leq l] \right) p_k
$$

(6)

This is an expression for the second moment of the resulting headways $Z$ for a given threshold $l$, which depends exclusively on the distribution of $H$ and the probabilities $p_k$ of having $k$ headways larger than $l$. This is very helpful for minimizing the waiting times since it grants an expression which can be minimized as a function of the threshold $l$. This means that (6) is an appropriate expression for the second moment we are interested in. However, two problems arise when attempting to minimize it:

1. The probabilities $p_k(l)$ of exactly $k$ headways being greater than $l$ are required.
2. It is required to calculate the second moment of the headways conditioned to the case where the headways are greater than $l$, where $l$ is modelled as a continuous variable. Since there are no assumptions on the distribution of $H$ this second moment might not be given by analytic functions and this can lead to serious numerical errors.

The first problem can be solved easily if there is good historical data of the system. Since $p_k$ doesn’t depend on the headways themselves but on the correlation between them, a frequentist approach can be used to count the number of headways larger than $l$ in the data and get an estimate for the probabilities.

If historical data isn’t available, assuming independence makes the number of headways larger than $l$ follow a Binomial distribution. Note that this is clearly not a realistic assumption, but it simplifies the expressions significantly and as will be observed in the next sections, it does not cause significant differences in the results.

The second one is solved thanks to the work of Munkhammar et al [Munkhammar et al (2017)]. They propose a method to approximate any density function (in our case, $H$) to a polynomial. This is particularly useful because the second moment is calculated by multiplying by $x^2$ and integrating, which leads to another polynomial, resulting in a very simple numerical calculation.

2.1.1 Application to real data

In order to test the accuracy of this model, a simulation was made using data from the service 216 of Transantiago. Buses were dispatched every 10 minutes. When a bus reaches a stop all waiting passengers board it and a subset of the ones aboard alighted, each taking a fixed constant amount of time. Travel time between departing a stop and arriving at the next one is modelled as a random variable independent to anything else, but dwell time considers the amount of passengers boarding and alighting. The injection was arbitrarily located in the 25th stop out of 51. Figure 2 shows the histogram of the headways at the stop alongside a polynomial fit of the density function. Using the empirical headway distribution for $H$ in (6) and assuming independence of the headways, figure 3
uses (6) to display the second moment of $H$ as a function of the threshold $l$ for different values of $N$. Since $E[Z^2]$ is directly linked to waiting times at the stop, we observe the following:

- A large value of $l$ implies the injection is (almost) never made, hence we obtain the same $E[Z^2]$ for any value of $N$. For low values of $l$ the bus is injected within the very first headways so it is clearly not optimal.
- The incentive to have a large $l$ is to cover a large headway. The bigger $l$ is, the better the injection becomes. However, a large $l$ risks that the bus is never
injected, so it is not optimal for it to be arbitrarily high. As the figure shows, there is an optimal $l$ for each value of $N$.

- As $N$ grows so does the optimal $l$. This is expected because $N$ is the number of opportunities for injecting, so more opportunities means a riskier threshold can be chosen.

In this example, the optimal threshold varies between 16 and 20 minutes depending on the number of buses available on the period of interest, which is roughly double the dispatch rate.

Figure 4 displays the value of $\mathbb{E}[Z^2]$ with fixed $N = 20$ under both suggested methods for $p_k$: Assuming headway independence and based on simulation historic information. Note that although both curves are not identical, the optimal threshold is roughly the same and the minimum value of $\mathbb{E}[Z^2]$ is slightly higher assuming independence. This is expected, as the probability of all headways being smaller than $l$ is much lower in a real service. In this example the optimal implementation threshold is found at 19 minutes, reducing the waiting times at the stop by almost 9%. If the operation of the service didn’t consider injecting a bus and all buses were continuously operating the waiting time gain at this stop would be only 5%. Note that although the reduction in this stop is significant, a more thorough analysis should consider the waiting time impact of both operational strategies in all stops of the route. This analysis will be presented in the next chapter of this thesis.

2.2 Multiple injections

Expression (6) can be adapted for the case of multiple injections. The process follows the same logic: every time a headway is larger than the threshold a bus is injected, but now more buses are available. In order to obtain an expression for
Let $b$ be the number of vehicles separated for eventual injection. Similar to the previous section, we need to consider two cases:

1. There are $b$ or less headways larger than $l$.
2. There are more than $b$ headways larger than $l$.

Note that the main difference between the two cases is that in the first one every single 'large' headway is covered by an injection, while on the second one some are not.

Just like in the previous section, let $K$ be a random variable representing the number of headways in the sequence larger than $l$ and consider a fixed realization of it, $k$. Using this, in the first case we have that $k \leq b$ and all $k$ large headways are covered by an injection. Mathematically:

$$E[Z^2 | K = k] = k \left( E[(H/2)^2 | H > l] + E[(H/2)^2 | H > l] \right) + (N - k)E[H^2 | H \leq l] \quad (7)$$

Here, the first term represents the $k$ larger than $l$ headways which are split in half by the injections, while the second term corresponds to the $N - k$ headways that were shorter than $l$ and therefore no injection was made.

Following the same logic, in the second case there are $k > b$ large headways, this means that $b$ large headways will be covered by injections, $k - b$ large headways will not be able to be covered, and $N - k$ short headways. This is modelled by

$$E[Z^2 | K = k] = b \left( E[(H/2)^2 | H > l] + E[(H/2)^2 | H > l] \right) + (k - b)E[H^2 | H > l] + (N - k)E[H^2 | H \leq l] \quad (8)$$

Again, the first term represents the $b$ covered headways, the second one the $(k - b)$ large headways that are not covered and third are the $(N - k)$ short headways that do not need to be covered.

With these expressions, using the Total Law of Probabilites over the number of headways larger than $l$ and using linearity of the expectation, similar to the previous section:

$$E[Z^2] = \sum_{k=0}^{b} \left( k \frac{1}{2} E[H^2 | H > l] + (N - k)E[H^2 | H \leq l] \right) p_k + \sum_{k=b+1}^{N} \left( (k - b) \frac{1}{2} E[H^2 | H > l] + (N - k)E[H^2 | H \leq l] \right) p_k \quad (9)$$

Now, we can compute the second moment as a function of two variables, $l$ and $b$, where $f(l, b) = E[Z^2](l, b)$ (considering $N$ as a fixed value). Here, $f(l, b)$ is a representation of the waiting times if the threshold is set as $l$ and a fleet of $b$ buses is reversed for eventual injections. It is obvious that $f$ is nonincreasing with respect to $b$. If more buses are added for injection while keeping the operative fleet constant, then the expected waiting time must drop or in the worst case, remain the same. Another interesting result is obtained by computing the reduction in expected waiting times by adding another bus reserved to be injected:
\[
f(l, b) - f(l, b + 1) = \left( -(b + 1) \frac{1}{2} E[H^2|H > l] + \left( \frac{b}{2} + 1 \right) E[H^2|H > l] \right) p_{b+1} + \sum_{k=b+2}^{N} \frac{1}{2} E[H^2|H > l] p_k
\]

\[
= \frac{1}{2} E[H^2|H > l] p_{b+1} + \sum_{k=b+2}^{N} \frac{1}{2} E[H^2|H > l] p_k
\]

\[
= \frac{1}{2} E[H^2|H > l] \sum_{k=b+1}^{N} p_k
\]

\[
= \frac{1}{2} E[H^2|H > l] \mathbb{P}[K \geq b + 1]
\]

(10)

This shows that increasing \( b \) by one reduces waiting times by the product of half the conditional second moment and the probability of there being at least \( b+1 \) large headways, this is, the probability of the additional bus being actually used multiplied by the benefit of it being used. If we consider this probability decreases as \( b \) increases, we have a sort of concavity in \( f \) with respect to \( b \).

### 2.2.1 Application to real data

Similar to the previous section, data from service 216 of Transantiago was used to test this result.

Figure 5 shows \( E[Z^2] \) for \( N = 20 \) assuming headway independence, and with up to 4 buses reserved for injection. We can observe the following:

- Every additional bus reserved for injection has a lower marginal impact in the optimal value of \( E[Z^2] \). This is expected since the probability of using the extra bus is strictly lower than the previous one, while its expected impact is exactly the same.
- The first bus reduces the waiting times by 8.5%, while the fourth one reduces them by 4.5%.
- The optimal threshold drops when the number of reserved buses grows. Note that in order to fully use a reserve fleet, multiple headways must be large, and this is less likely as the fleet size grows. This means that the risk of wasting a bus increases with the fleet size and therefore the optimal threshold is reduced.

It is important to remark that this analysis compares fleets of different sizes because the non reserved fleet is the same for every case. Since every additional injected bus is less impactful, these results suggest that a large injection fleet is generally not optimal. To determine the optimal size of the reserve fleet, its benefits identified in (10) should be balanced with the cost of reserving an extra bus for this purpose.
In order to correctly evaluate the impact of bus injection, the results from the previous chapter must be adapted to consider the effect in waiting times at all downstream stops along the service. To do this, we must evaluate the evolution of a headway sequence at downstream stops. The work by Marguier, later improved by Hickman [Hickman (2001)] is particularly useful as a starting point.

3.1 Hickman’s work and Marguier’s Model

In 1985, Marguier proposed a stochastic model for transit operations on a single route that predicts the trajectories of vehicles. Later, in 2001, Hickman improved this model by correcting some assumptions and adding new computations including the variance of the headways. In order to develop a framework for modelling bus injections, Hickmans’ model will be used with some slight variations. Our interest is not headways’ variance, but their second moment, so new computations must be made. Also, boundary conditions must be defined to represent bus injections.

The following assumptions will be made, which are almost identical to Marguier’s (and Hickman’s).

1. Passengers arrive randomly as a homogeneous Poisson process at each stop. We will assume that the arrival rate does not change over the time period of interest.
2. Fleet size will be constant, so as in the previous chapter, our measure of performance will be given by the second moment of the headways.
3. When a vehicle arrives at a stop, a subset of arriving passengers will alight, and then every passenger at the stop will board. This means that these events

Fig. 5: Second order Moment of headways for different injection fleet sizes
are done in series and both times will be added to determine dwell time at each stop. Also, each passenger takes the same amount of time boarding, regardless of how many passengers are boarding or how full the bus is. The same logic applies to alighting time.

4. Every time a passenger boards a bus, they have a fixed probability distribution for their alight stop. More specifically, this means that the destination stop of each passenger follows a multinomial distribution. This is equivalent to saying that the number of passengers alighting from each bus at each stop follows a Binomial distribution depending on the stop and the current load of the bus. This assumes that passengers’ destination are independent of each other (they do not travel in groups).

5. Travel time of a bus between stops has finite second moment and its distribution does not change over the period of interest. Also, they are all independent and therefore their distribution does not change between different buses over the same two stops.

6. Passengers that arrive at the stop while passengers are boarding and alighting is negligibly small.

We use the following notation based on Marguiers’ model, remarking that capital letters represent a random variable, while lower case and greek letters represent fixed (and known) parameters of the system.

- $s$ is a subscript denoting the stop, $s = 1, \ldots, S$.
- $i$ is a subscript denoting the bus number, $i = 1, \ldots, I$. Note that in this situation $I$ is not necessarily equal to $N$ (the number of headways we are observing).
- $L_{i,s}$ is a random variable denoting the load on vehicle $i$ as it departs stop $s$. Note that this means that the load on vehicle $i$ as it enters stop $s$ is $L_{i,s-1}$.
- $R_{i,s}$ is a random variable denoting the travel time of bus $i$ between departing stop $s - 1$ and arriving at stop $s$.
- $H_{i,s}$ is a random variable denoting the headway between the bus $i - 1$ and bus $i$ as they depart stop $s$.
- $\lambda_s$ is the passenger arrival rate at stop $s$ (passengers per unit time).
- $p_s$ denotes the probability that a passenger who is on board the vehicle entering stop $s$ will alight at stop $s$. Note that by this definition, $\sum_{s=1}^S p_s$ is not 1. In fact, $p_S = 1$.
- $a$ denotes the time lost by accelerating and decelerating at a stop.
- $b_A$ and $b_B$ denote the time a passenger takes alighting and boarding respectively.

With this set of assumptions and notation, the first thing to note about the model is that each bus spends time both travelling between stops and dwelling at them to allow passengers to board and alight. Time spent travelling is represented by $R_{i,s}$. Note that we are assuming that travel time distribution does not change between stops, so while $R_{i,s} \neq R_{j,s}$ for $i \neq j$, $E[R_{i,s}] = E[R_{j,s}]$ for all $i, j \in \{1, \ldots, I\}$ and we will abuse notation defining this expected value as simply $E[R_s]$. We will define dwell time of bus $i$ at stop $s$ as the sum of the time spent accelerating, decelerating and the time the passengers board and alight. Mathematically, this is represented by

$$D_{i,s} = a + b_A A_{i,s} + b_B B_{i,s}$$
Where $A_{i,s}$ and $B_{i,s}$ are the number of passengers alighting and boarding bus $i$ at stop $s$ respectively. The number of passengers alighting, $A_{i,s}$, distributes $\text{Bin}(L_{i,s-1}, p_s)$. To obtain its expected value, we must recognize that this random variable depends on a parameter which is another random variable, thus we must use the Law of Total Probabilities:

$$E[A_{i,s}] = E_L[E[A_{i,s}|L_{i,s-1}]]$$

$$= \sum_{n=0}^{\infty} np_s P[L_{i,s-1} = n]$$

$$= p_s \sum_{n=0}^{\infty} n P[L_{i,s-1} = n]$$

$$= p_s E[L_{i,s-1}] \quad (12)$$

The number of people boarding a bus at a stop corresponds to the number of people who have arrived at the stop during the interval that the bus. We assume that the passengers arrive to each stop according to a Poisson process with a constant rate. Thus, $B_{i,s} \sim \text{Pois}(\lambda_s H_{i,s-1})$. Note that again, the parameter of this distribution is a random variable so we follow the same previous logic.

$$E[B_{i,s}] = E_H[E[B_{i,s}|H_{i,s-1}]]$$

$$= \int_{R^+} \lambda_s h f_{H_{i,s-1}}(h) dh$$

$$= \lambda_s E[H_{i,s-1}] \quad (13)$$

Although both results seem obvious, on principle they are not. In the next section we will compute the second moment of both $A_{i,s}$ and $B_{i,s}$ and we will observe that because they rely in random parameters, multiple terms are added in the results.

Using (12) and (13) in (11) we obtain

$$E[D_{i,s}] = a + b_A p_s E[L_{i,s-1}] + b_B \lambda_s E[H_{i,s-1}] \quad (14)$$

Note that (14) assumes that buses stop and open doors at every stop in the system. Although this is a strong assumption, it is made in order to simplify the expressions. A more complete expression would include the probability of no passengers interacting at the stop.

Now, Marguier identified the most important dynamics of the system: the headway of a vehicle $i$ at a stop $s$ corresponds to the observed headway at the previous stop $s - 1$ plus the difference in travel and dwell times between the vehicle $i$ and the previous one $i - 1$. Mathematically, this is

$$H_{i,s} = H_{i,s-1} + R_{i,s} - R_{i-1,s} + D_{i,s} - D_{i-1,s} \quad (15)$$
Again, using the linearity of the expected value operator, along with (14) and the previous assumptions, we get

\[
E[H_{i,s}] = E[H_{i,s-1}] + E[D_{i,s}] - E[D_{i-1,s}]
\]

\[
= E[H_{i,s-1}] + b_A p_s (E[L_{i,s-1}] - E[L_{i-1,s-1}]) + b_B \lambda_s (E[H_{i,s-1}] - E[H_{i-1,s-1}])
\]

(16)

Here, the expected values of \(R\) get cancelled out because of assumption (v). Note that this equation allows us to obtain an estimate for the headway in a stop if we know the headway in the previous one and the load of the vehicles. Similarly, the load on a vehicle leaving stop \(s\) corresponds to the load the vehicle had when it left the previous stop \(s-1\), plus the number of people who boarded on stop \(s\) minus the number of people who alighted on stop \(s\). This is,

\[
L_{i,s} = L_{i,s-1} + B_{i,s} - A_{i,s}
\]

(17)

Again, we can apply expected value and use (12),(13) to get

\[
E[L_{i,s}] = E[L_{i,s-1}] + B_{i,s} - E[A_{i,s}]
\]

\[
= (1 - p_s)E[L_{i,s-1}] + \lambda_s E[H_{i,s-1}]
\]

(18)

This allows Marguier to represent system dynamics with a matrix notation, but it is not particularly useful to us. Note that given complete information of a system at any fixed instant of time, this model allows to get estimates for the headways and loads of every bus in the future. Also, this model allows overtaking, as \(H_{i,s}\) can be negative. This usually happens when boarding times are significant and there is high variance in travel times. When overtaking occurs, we will swap the index \(i\) of the buses so that the bus \(i-1\) will always be ahead of the bus \(i\).

Hickman then formulated expressions for the variance of these variables but we are not particularly interested in those. Now that the base model is defined, we will define boundary conditions that allows us to utilize it in bus injection situations. However, as we will see, this model does not capture variability of the system so it will not suffice and we will develop a more detailed one.

3.2 Applying the base model to real data and defining boundary conditions

Before presenting how to represent a bus injection with the previous model, it is important to understand how the recursive model works. As can be seen from (16) and (18), the expected value of the headways and loads of any bus at any stop depends exclusively on

(a) The headways and loads of the buses on the previous stop.

(b) The system parameters.

Hence, if we know the values for the expected load and headways at any stop, then we can know the expected values for the next stop. This means that once the expected state of the system is known at a stop, then the next one is completely determined, and recursively, the entire downstream system can be determined. Our goal will be to determine the expected state of the system in an injection, and using the described model we will be able to compute the expected state of all future stops. This structure of previous-stop and parameter
Bus arrivals at injection stop $s'$:

![Diagram of bus arrivals at injection stop $s'$]

Fig. 6: Example bus arrivals

dependance is what we will denominate as our recursion rule for the model to work.

The same data from chapter 1 was used to test this model for our purpose. In order to do so, first we need to define boundary conditions that represent a bus injection.

Let $l$ be the chosen threshold for injection (using the method from chapter 1), $s'$ the stop where the bus injection will be made and $i'$ the bus right before the injected one. Also, we will use the index $e$ to refer to the injected bus.

Figure 6 illustrates the notation used. It is the same example as the previous chapter, where it shows arrivals at a single stop, and each node is a bus arriving. Note that Once a headway is larger than the threshold, a new decision must be made: the exact instant along the headway where the additional bus must be dispatched. In the previous chapter it was exactly at half the headway since we were not concerned about downstream stops. Now, however, we want to minimize the waiting times for the entire system, and dispatching at the middle might not be optimal for the next stops. We will define $\pi$ as the fraction of the headway where the injection is made, where $\pi = 0$ means that the bus is dispatched together with the previous one, $\pi = \frac{1}{2}$ means it is dispatched exactly at the middle, $\pi = 1$ means it is dispatched with the 'late' bus, and so on.

With this, the following observations arise:

(a) Since the bus $i'$ is the one where an injection was decided to be used, every single bus before it had a headway shorter than the threshold $l$.
(b) Since the injected bus is empty at stop $s'$ and the rest of the buses are not, it is expected to travel faster.
(c) The ideal scenario is perfectly regular headways, so the optimal $\pi$ is expected to be larger than 0.5 so the bus on average travels at the middle point between $i'$ and $i'+1$.

Up to this point, in order to determine the expected behaviour of the system once the injection is made and to find the optimal $\pi$, the operator needs information about the load and headway on each vehicle of the system at the moment the injection is going to be made.

All of this information can be obtained (and should be available for operators) in real-time once the threshold has been surpassed, but in order to test this model, we will use the expected values with historical data for estimates. We will assume a set of historical information (or simulation results) about the system without injection is available and we will call $K$ the number of observations in the set, as well as $h_{k}^{s}$, $l_{k}^{s}$ the headway and load respectively at stop.
s in observation number \( k \) (each observation corresponds to a bus arriving at a stop in any observed day).

For the load on each vehicle, note that using the first observation, we can estimate

\[
E[L_{i,s'}|H_{i,s'} \leq l]
\]

from historical data. This is, out of all loads of vehicles departing at stop \( s' \), we take the first empirical moment (expected value) only of those where the headway preceding it was shorter than the threshold \( l \). The only exception is the load of the bus \( i' + 1 \) because it is actually preceded by a long headway and therefore will be given by

\[
E[L_{i',s'}|H_{i,s'} > l]
\]

To simplify notation, we will define \( \mathcal{K}_\alpha \) as the set of observations that satisfy condition \( \alpha \). For example, \( \mathcal{K}_{H > l} \) will be the set of all observations of buses from historical data whose preceding headway is larger than \( l \). Note that the cardinality of set \( \mathcal{K} \) (without a condition, meaning all observations) is \( K \).

Now, if an injection is made with a fixed \( \pi \), then the injected bus is expected to absorb a \( \pi \) fraction of the load the bus \( i' + 1 \) was going to receive originally, which, as given by (13), corresponds to

\[
\pi E[H_{i',s' - 1}] \lambda_{s'}
\]

This means that the expected initial load on bus \( e \) is given by

\[
E[L_{e,s'}] = \pi E[H_{i',s' - 1}] \lambda_{s'}
\]

and the expected load on bus \( i' + 1 \) will be the estimate from historical data minus \( E[L_{e,s'}] \).

As for the headways preceding each bus, we use the same information we used for the load, so that for \( i = 1, \ldots, i' \) we can estimate them from historical information as the expected value of only the observations where they are shorter than \( l \), this is

\[
E[H_{i,s'}|H_{i,s'} \leq l]
\]

and for the bus \( i' + 1 \) it is

\[
E[H_{i,s'}|H_{i,s'} > l]
\]

With this, the boundary conditions for the model to represent our situation of interest are given by the estimators:

(a) \( \hat{E}[H_{i,s'}] = \frac{1}{|\mathcal{K}_{H \leq l}|} \sum_{k \in \mathcal{K}_{H \leq l}} h_{s'}^k \quad \forall i = 1, \ldots, i' \)

(b) \( \hat{E}[H_{e,s'}] = \frac{\pi}{|\mathcal{K}_{H > l}|} \sum_{k \in \mathcal{K}_{H > l}} h_{s'}^k \)

(c) \( \hat{E}[H_{i'+1,s'}] = \frac{1 - \pi}{|\mathcal{K}_{H > l}|} \sum_{k \in \mathcal{K}_{H > l}} h_{s'}^k \)
Basically, the first four conditions represent the headway each bus on the system has, where (a) corresponds to the observed mean at stop $s'$ from historical observation where the headway was shorter than the threshold, (b) corresponds to the additional bus, obtained as the expected headway of a bus with a large threshold multiplied by the fraction of the headway where the bus will be dispatched, (c) is the remaining fraction of that same headway and (d) is the classic mean of the observations since we have no additional information about those buses at the time of injection.

The last four conditions represent the load on each bus, where (e) and (h) are totally analogous to (a) and (d) respectively. (f) is simply the expected amount of passengers to arrive at the injection headway multiplied by the fraction $\pi$, and (g) is the load obtained from historical data, subtracting the load that the injected bus will absorb.

Note that $\pi$ is a variable in this framework. With this, we can obtain through the recursive set of equations (16) and (18) the expected headways for the entire system in the stops downstream $s'$ and then square them as a measure of how effective different values of $\pi$ are (remember we want them to be as regular as possible, and perfectly regular headways means minimum squared values).

Figure 7 shows the expected trajectories of buses $i$ and $i'$ according to the model if there was no injection. The arbitrarily chosen stop for injection in this example is stop number 25 out of 51. Note that stop 25 is a critical point in the route, having a high arrival rate (it is a combination station to Metro) and marking the end of the bus corridor. This high arrival rate damages the performance of the bus $i' + 1$ considerably because since it has such a high headway preceding it, the number of people boarding it will be very high.

Figure 8 shows the expected trajectories computed by the model for the same buses in case the injection was dispatched with $\pi = 0.5$ (at the middle of the headway). We can observe that since initially the bus $e$ is empty, it travels considerably faster than bus $i' + 1$ which is expected to be heavily loaded and slightly faster than bus $i'$, which is expected to be slightly loaded. Note that although bus $i' + 1$ improves greatly compared to the previous case, it is clearly not an optimal use of the injection, as the bus travels much more closer to bus $i'$ than to $i' + 1$ and therefore not reducing bus bunching efficiently.

Another important point to note is how much faster bus $i' + 1$ runs with an injection. The time it reaches the final stop is reduced from 6400 seconds to
3400. This excessive reduction is caused because we are not assuming any limit on the buses’ capacity so on the case with no injection, buses are loaded with up to 170 passengers, therefore the injection is overestimating the impact on reduced dwell time. Even though this reduction in load is not realistic, we must consider that excessive load can be interpreted as passengers who will board the next bus and therefore as additional waiting times.

Running the model leaving $\pi$ as a variable, we found that the value of $\pi$ that balances the bus optimally between $i'$ and $i' + 1$ (by minimizing squared expected headways) is $\pi = 0.546$. Figure 9 illustrates the situation with this
optimal $\pi$. In this situation, the additional bus is much more aligned with the others along the route, reducing the load on bus $i' + 1$ considerably.

Even though this model gives a reasonable approach for the problem, there is a particularly worrying assumption. We are using $\mathbb{E}[H^2]$ as the performance indicator, when the real one is $\mathbb{E}[H^2]$. These are equal if and only if $\text{Var}[H] = 0$ (i.e. the headways $H$ are not random variables, but fixed values). This can also be seen if we note that we are obtaining the expected values for each stop once at a time using the expected value of the previous one and thus we are not including variability in their values, only taking the expected value and using it to compute the next expected value. A more realistic approach would include in the recursive model itself the second moment of the headways and loads so that the variability is captured. This is what will be presented in the next section of this chapter.

3.3 A model considering the second moment of the headways

In order for the model to capture the variability of the system, we must develop recursive expressions for the second moment of the headways and loads. The goal is to obtain a recursive model similar to the previous sections. Now (12) and (13) become relevant. A similar procedure for their squared forms can be developed using the Law of Total Probabilities:
\[ E[A_{i,s}^2] = E_L[E[A_{i,s}^2|L_{i,s-1}]] \]
\[ = \sum_{l=0}^{\infty} E[A_{i,s}^2|L_{i,s-1} = n] \]
\[ = \sum_{l=0}^{\infty} (n^2 p^2 + np(1-p))P[L_{i,s-1} = n] \]
\[ = p_s^2 E[L_{i,s-1}^2] + p_s(1-p_s)E[L_{i,s-1}] \]  
\[ (19) \]

Now, we perform similar computations for \( B_{i,s} \). Since the number of people who board the bus is identical to the number of people who has arrived to the stop between the previous bus arrived and the current bus arrived, it corresponds to a Poisson process and its parameter is the arrival rate times the headway. Mathematically, this is \( B_{i,s} \sim \text{Pois}(\lambda_s H_{i,s-1}). \) Note that again, the parameter is a random variable so we follow the same previous logic.

\[ E[B_{i,s}^2] = E[H][E[A^2|H_{i,s-1}]] \]
\[ = \int_{\mathbb{R}^+} E[B^2|H_{i,s-1} = h]f_{H_{i,s-1}}(h)dh \]
\[ = \int_{\mathbb{R}^+} (\lambda h + \lambda^2 h^2) f_{H_{i,s-1}}(h)dh \]
\[ = \lambda E[H_{i,s-1}] + \lambda^2 E[L_{i,s-1}^2] \]  
\[ (20) \]

Note that these two expressions follow the recursion rules from the previous model, so although they are considerably more complicated the recursion can still be made. As expected, our final goal will be obtain \( E[H_{i,s}^2] \) and \( E[L_{i,s}^2] \) from (15) and (17). Now, we will compute \( E[D_{i,s}^2] \) as it will clearly be needed. For this, remember that our dwell time is defined as

\[ D_{i,s} = b_A A_{i,s} + b_B B_{i,s} \]

Squaring the expression and applying expected value yields:

\[ E[D_{i,s}^2] = b_A^2 E[A_{i,s}^2] + b_B^2 E[B_{i,s}^2] + 2b_A b_B E[A_{i,s} B_{i,s}] \]
\[ = b_A^2 E[A_{i,s}^2] + b_B^2 E[B_{i,s}^2] + 2b_A b_B E[A_{i,s}]E[B_{i,s}] \]
\[ = b_A^2 \left(p_s^2 E[L_{i,s-1}^2] + p_s(1-p_s)E[L_{i,s-1}] \right) \]
\[ + b_B \left(\lambda E[H_{i,s-1}] + \lambda^2 E[H_{i,s-1}^2] \right) \]
\[ + 2b_A b_B p_s E[L_{i,s-1}]\lambda_s E[H_{i,s-1}] \]  
\[ (21) \]

The importance of (21) is that it allows us to represent \( E[D_{i,s}^2] \) in terms of moments at the previous stop and system parameters, enabling the use of a recursive model. We remark that to obtain (21) we assumed that the covariance between the number of passengers boarding at alighting at the same stop is zero. This is not necessarily true, since both depend on the headway and load.
of the same bus two stops ago. However, real system dynamics do not show a high correlation between these two values, so the assumption is not expected to be impactful.

Now, we will obtain an expression for \(E[H_{i,s}^2]\) in order to complete the recursion for the second order moment. We will get it by squaring (15) and note that as this will result in an exceedingly high number of terms, the expression we will obtain is expected to be particularly long.

\[
E[H_{i,s}^2] = E[H_{i,s}^2] + E[R_{i,s}^2] + E[R_{i-1,s}^2] + E[D_{i,s}^2] + E[D_{i-1,s}^2] \\
+ 2E[H_{i,s}R_{i,s}] - 2E[H_{i,s}R_{i-1,s}] + E[2H_{i,s}D_{i,s}] - 2E[H_{i,s}D_{i-1,s}] - 2E[R_{i,s}R_{i-1,s}] \\
+ 2E[R_{i,s}D_{i-1,s}] - 2E[R_{i,s}D_{i-1,s}] - 2E[R_{i-1,s}D_{i,s}] + 2E[R_{i-1,s}D_{i-1,s}] - 2E[D_{i,s}D_{i-1,s}]
\]

Now, fortunately many terms can be simplified. Note that the travel times \(R\) are independent from any other variable, therefore any terms with a positive \(R\) multiplied by a random variable and then a negative \(R\) multiplied by the same variable will get cancelled out. Also, as previously stated and from assumption (v),

\[
E[R_{i,s}^2] = E[R_{i-1,s}^2] = E[R_s^2]
\]

and

\[
E[R_{i,s}R_{i-1,s}] = E[R_{i,s}]E[R_{i-1,s}] = E[R_s]^2
\]

Hence, terms 2,3 and 10 from (22) form

\[
2E[R_s^2] - 2E[R_s]^2 = Var[R^2]
\]

Not only this allows to simplify (22) significantly, it also means that the variance of the travelling time propagates into the variability of headways on the system. We will also assume that there is no covariance between \(D_{i-1,s}\) with \(H_{i,s}\) and \(D_{i,s}\). Again, these assumptions do not seem to be particularly relevant based on real systems observations. This simplifies considerably our previous expression:

\[
E[H_{i,s}^2] = E[H_{i,s-1}^2] + Var[R_s] + E[D_{i,s}^2] + E[D_{i-1,s}^2] \\
+ 2E[H_{i,s-1}D_{i,s}] - 2E[H_{i,s-1}D_{i-1,s}] - 2E[D_{i,s}E[D_{i-1,s}]]
\]

The first two terms are parameters and the next two are already computed in (21). The last two were already computed in the previous section so all we need to finish this expression is the fifth term. Note that assuming independence on this term would be totally unrealistic because the headway preceding this bus is directly correlated to the number of people boarding it at the next stop and therefore at the dwelling time. However, we can separate this expression,

\[
E[H_{i,s-1}D_{i,s}] = b_AE[H_{i,s-1}A_{i,s}] + b_BE[H_{i,s-1}B_{i,s}]
\]

Here, the first term is a multiplication of independent variables (where both values are already computed in the previous section), and the second one, where we cannot assume independence, is the expected value of a Poisson random variable multiplied by its parameter (which is also a random variable).

We can use the Law of Total Probabilities to obtain an analytic expression for this:
\[
E[H_{i,s-1}B_{i,s}] = E[E[H_{i,s-1}B_{i,s}|H]]
\]
\[
= \int_{\mathbb{R}^+} E[H_{i,s-1}B_{i,s}|H_{i,s} = h]f_{H_{i,s}}(h)dh
\]
\[
= \int_{\mathbb{R}^+} h\lambda h f_{H_{i,s}}(h)dh
\]
\[
= \lambda E[H_{i,s-1}^2]
\]
\hspace{1cm}(24)

Note that with this, all seven terms in (23) have been expressed in expected values of the variables of previous stops or parameters of the system. This allows to compute \(E[H_{i,s}^2]\) enabling us to define a recursive model similar to the one in the previous section.

Now we proceed to compute \(E[L_{i,s}^2]\). The process will be similar to that of the headways. Again, we start by squaring (17) and applying expected value operator to the equation:

\[
E[L_{i,s}^2] = E[L_{i,s-1}^2] + E[B_{i,s}^2] + E[A_{i,s}^2] - 2E[L_{i,s-1}A_{i,s}] + 2E[L_{i,s-1}B_{i,s}] - 2E[A_{i,s}B_{i,s}]
\]
\hspace{1cm}(25)

Since we are assuming \(A_{i,s}\) and \(B_{i,s}\) are independent, the final term is not a problem. Also, we will assume \(L_{i,s-1}\) and \(B_{i,s}\) are independent. Again, this assumption is not necessarily true but based on experience of real system dynamics the covariance between the load of a bus and the number of people boarding it in the next stop is not particularly strong so it is not expected to be a strong assumption. Since the first term is dependent of the previous stop and we have an expression that follows the recursion rules in (19) and (20) for the next two, all that is left to compute is the fourth term. Note that assuming independence here would be totally unrealistic, since the number of people alighting is directly related to the load of the bus, so we must compute this term analytically. First, we observe that the number of people alighting from bus \(i\) at stop \(s\), \(A_{i,s}\) follows a Binomial distribution with parameters \(p = p_s\) and \(n = L_{i,s-1}\) and therefore the term we are interested in is the multiplication of a binomial random variable with one of its parameters which is also a random variable. To obtain an analytic expression for this, we resort one last time to the Law of Total Probabilities:

\[
E[L_{i,s-1}A_{i,s}] = E[E[L_{i,s-1}A_{i,s}|L_{i,s-1}]]
\]
\[
= \sum_{n=0}^{\infty} E[nA_{i,s}]P[L_{i,s-1} = n]
\]
\[
= \sum_{n=0}^{\infty} n^2pP[L_{i,s-1} = n]
\]
\[
= p \sum_{n=0}^{\infty} n^2P[L_{i,s-1} = n]
\]
\[
= pE[L_{i,s-1}^2]
\]
\hspace{1cm}(26)
With this expression, all terms in (25) follow the recursion rules and therefore the model for second moments is complete. This model will allow us to compute analytically the second moment of the headways at any stop downstream the injection one once we define boundary conditions to represent it. As a way to enclose this section, a summary of the steps for computing the second moments and build the model is presented:

(a) Use the equations for the base model along the boundary conditions presented in the previous section to obtain the expected value (first moment) for all headways and loads of interest.

(b) Obtain boundary conditions for a stop (in our case, the injection one) either from historical data, simulations or estimations.

(c) Compute recursively the second moments for the next stops using equations (23) and (25) (which depend on equations (21) (24) and (26)).

3.4 Defining second order boundary conditions and using the model on real data

Now we present an extension to the previous sections for the second order boundary conditions along with real data tests.

In order to begin the recursive process, we need to obtain \( E[H_{i,s'}^2] \) and \( E[L_{i,s'}^2] \) for all buses at the system. We use the same definitions and observations as in section 3.2, so the expressions obtained will be nearly identical.

We start by studying the second order moment of the buses load. Again, for any bus prior to the injection we know the preceding headway is shorter than the threshold \( l \) and therefore we can estimate \( E[L_{i,s'}^2 | H_{i,s'} \leq l] \) from historical data using an expression equivalent to the one for the first moment, this is, out of all buses departing from stop \( s' \), we take the empirical second moment of the load only for those with a preceding headway shorter than \( l \). This gives a good estimate for any bus preceding the injection.

For the injected bus, note that the initial load will be identical to the number of people boarding it at the injection stop. This means that we just need to obtain an expression for \( E[B_{i,s'}^2] \). Since the boardings distribute according to a Poisson distribution with parameter equal to the multiplication of the boarding rate and the headway the bus faces. We know the bus is preceded by a large headway multiplied by \( \pi \), so using (20) we obtain

\[
E[L_{i,s'}^2] = \lambda_{s'} \pi E[H_{i,s'-1} | H_{i,s'-1} > l] + \lambda_{s'-1} \pi^2 E[H_{i,s'-1}^2 | H_{i,s'-1} > l] \quad (27)
\]

For the bus after the injection, \( i' + 1 \), we also know it is preceded by a headway longer than \( l \) and we can do an analogous procedure, but in this case we want to estimate

\[
E[L_{i,s'}^2 | H_{i,s'} > l]
\]

However, the injected bus will absorb a fraction of the load that bus \( i' + 1 \) was originally going to receive. To obtain a good estimate of what we want, note that the load of the bus is composed of two factors:
The 'base' load the bus will have independent of the choice of \( \pi \) (this is, the load it carried from the previous stop minus the alights at \( s' \))

The passengers boarding in \( s' \) (which do depend on \( \pi \)).

If we call the fixed load \( T \) and the boardings \( B \) and assume independence between \( T \) and \( B \), note that the second moment without injection is obtained from

\[
L_{i,s'} = T_{i,s'} + B_{i,s'}
\]

\[
\mathbb{E}[L_{i,s'}^2] = \mathbb{E}[T_{i,s'}^2] + 2\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}] + \mathbb{E}[B_{i,s'}^2]
\]

Which is what we can estimate with the historical data. As for the case with an injection with a chosen \( \pi \), we have

\[
L_{i,s'} = T_{i,s'} + (1 - \pi)B_{i,s'}
\]

\[
\mathbb{E}[L^2] = \mathbb{E}[T_{i,s'}^2] + 2(1 - \pi)\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}] + (1 - \pi)^2 \mathbb{E}[B_{i,s'}^2]
\]

Therefore, the difference in the second moment generated by the injection is

\[
\Delta \mathbb{E}[L_{i,s'}^2] = \mathbb{E}[T_{i,s'}^2] + 2\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}] + \mathbb{E}[B_{i,s'}^2] - \mathbb{E}[T_{i,s'}^2] - 2(1 - \pi)\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}] - (1 - \pi)^2 \mathbb{E}[B_{i,s'}^2]
\]

\[
= 2\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}](1 - (1 - \pi)) + \mathbb{E}[B_{i,s'}^2](1 - (1 - \pi)^2)
\]

\[
= 2\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}] \pi + \mathbb{E}[B_{i,s'}^2](1 - (1 - \pi)^2)
\]

This can be reformulated as

\[
\mathbb{E}[L_{i,s'}^2]_{\text{inj}} = \mathbb{E}[L_{i,s'}^2]_{\text{no inj}} - 2\mathbb{E}[T_{i,s'}] \mathbb{E}[B_{i,s'}] \pi - \mathbb{E}[B_{i,s'}^2](1 - (1 - \pi)^2)
\]

Here, we can obtain all expressions separately from historical data in order to estimate the difference between the two cases.

As for the headways, it is much simpler. Again we use the same procedures as previous sections but for second order moments. The only notable difference is for the injected bus and the next one, since the headway will be split in two.

For the injected bus, the expression will be

\[
\mathbb{E}[(\pi H_{i,s'})^2 | H_{i,s'} > l] = \pi^2 \mathbb{E}[H_{i,s'}^2 | H_{i,s'} > l]
\]

And for the next one it will be the same except for the following remaining fraction:

\[
\mathbb{E}[(1 - \pi)H_{i,s'}^2 | H_{i,s'} > l] = (1 - \pi^2) \mathbb{E}[H_{i,s'}^2 | H_{i,s'} > l]
\]

With this, our set of equations for the estimators of the second order moments will be:

(a) \( \mathbb{E}[H_{i,s'}] = \frac{1}{|K_H \leq l|} \sum_{k \in K_{H \leq l}} (h_{i,k})^2 \) \quad \forall i = 1, \ldots, s'

(b) \( \mathbb{E}[^{\text{inj}} H_{e,s'}] = \frac{\pi^2}{|K_H > l|} \sum_{k \in K_{H > l}} (h_{k})^2 \)

(c) \( \mathbb{E}[H_{i,s'} + 1,s'] = \frac{(1 - \pi)^2}{|K_H > l|} \sum_{k \in K_{H > l}} (h_{k})^2 \)
(d) \( E[H_{i,s'}] = \frac{1}{K} \sum_{k=1}^{K} (h_{i,s'}^{(k)})^2 \quad \forall i > i' + 1 \)

(e) \( \hat{E}[L_{i,s'}] = \frac{1}{|K_{H \leq l}|} \sum_{k \in K_{H \leq l}} (l_{i,s'}^{(k)})^2 \quad \forall i = 1, \ldots, i' \)

(f) \( \hat{E}[L_{e,s'}] = \lambda_{s'} \pi \hat{E}[H_{e,s'}] + \lambda_{s'}^2 \pi^2 \hat{E}[H_{e,s'}^2] \)

(g) \( \hat{E}[L_{i'+1,s'}] = \frac{1}{|K_{H \geq l}|} \sum_{k \in K_{H \geq l}} (l_{i,s'}^{(k)})^2 - 2\pi((1 - p_s)\hat{E}[L_{i,s'-1}])\lambda_{s'} \hat{E}[H_{i,s'-1}] - (\lambda_{s''} \hat{E}[H_{i,s''-1}] + \lambda_{s'}^2 \hat{E}[H_{e,s''-1}]}(1 - (1 - \pi)^2) \)

(h) \( \hat{E}[L_{i,s'}] = \frac{1}{K} \sum_{k=1}^{K} (l_{i,s'}^{(k)})^2 \quad \forall i > i' + 1 \)
Just like in the first order case, the first four conditions represent the headway preceding each bus on the system, with (a)-(e) and (h) being completely analogous to the previous case. (f) corresponds to an estimator for the second order of the injected bus’ load based on (27) and (g) estimates the second order load for the following bus based on (31). Although computations are not as straightforward as the previous model, the procedure for using the model is completely identical, as was described at the end of the previous section. This set of equations completes the framework needed for the final model to be used. The main difference with the previous model and the reason this one was developed is to include variability in the headways as according to (1) the waiting times are directly related to the second moment, which the previous model was not capturing correctly. Again, data from service 216 of Transantiago was used to test the model.

Figure 10 shows the sum of the second order moment of the headways in the entire system as a function of the fraction before injection $\pi$. It can be seen how the optimal value of $\pi$ is slightly higher than the previous model, being 0.57 in this case. Figure 11 shows that this higher value does not make the injected bus stay in the middle of the headway, however this is still the optimal $\pi$ according to the model because:

(a) It reduces variability in the system, reducing the passengers’ waiting times.
(b) It considerably reduces the total travel time of bus $i' + 1$, which in turn helps reducing waiting times too.

Also, note that lower values of $\pi$ behave worse than higher values, since dispatching an empty bus earlier will instantly join the earlier bus, but dispatching it later will help the overloaded bus coming behind.

Tables 1 and 2 show the load and headways respectively in the case of no injection, naive injection ($\pi = 0.5$) and optimized injection ($\pi = 0.57$). It is clearly seen how poorly the no injection scenario performs in both load and
headways. This is a particularly bad scenario because we are already assuming a long headway occurs (which is not necessarily bound to happen). A naive injection clearly helps bus $i' + 1$ considerably both in terms of load and headways. Note that the load the late bus had with no injection is considerably higher than the sum of the injected bus and the late bus with injection. This is because as the late bus is less crowded, it travels faster, and therefore not capturing as much passengers as originally. These passengers not captured are going to board the next bus, which was originally not crowded because the bus
As for the headways, the same phenomena occurs but even more notably. The preceding headway of bus $i' + 1$ in the case of no injection is considerably higher than the sum of the two buses in the case of injection. This is because the late bus travels faster and now a closer bus exists to reduce the headway.

Also, note how the case with optimized injection performs better in terms of passenger and headway distribution along the route: the maximum load falls from 83.5 to 68.8, which corresponds to 17%. The largest headway lowers around 45% because of optimized injection. Although this is a remarkable reduction, we must remember earlier stops perform slightly worse in order to achieve this. If squared loads are used as a metric of bus bunching, the overall reduction from naive injection to optimized injection is over 20%.

Finally, the model results for the headways second moment were tested with a simulation of service 216, which runs with a fleet size of 20. The simulation was programmed in Python using a Discrete event simulation model. The simulation model represents the exact same situation as the one presented in this chapter, and is used as an alternate way to measure the impact an injection will have on the system. Since the simulation model requires considerably less assumptions (particularly, no independence assumptions are required to be made) it is assumed to be a reliable source of data.

Figure 12 compares the second moment of the headways for each stop at service 216 in three cases:

- No injection is made
- With injection according to our stochastic model
- With injection in the simulation model
We remark that the stop chosen for injection is the 25th one out of 51 (in the figure, 0 is a stop) and it corresponds to an important point in the route, where a Metro station is located and therefore the passenger arrival rate is very high. The figure shows that the stochastic model is slightly but consistently above the simulation results. This is explained due to the independence assumptions mentioned in previous sections of this chapter. However, the difference is small enough for the model to be considered valuable in terms of recommending an injection. Using bus injection in the service would provide a 6.8% reduction in wait times. Since using the added vehicle as a normal circulation would yield about a 5% reduction, the injection would be preferred.

4 Conclusions

A model for analyzing the convenience of injecting a bus incorporating stochastic evolution is provided. Examples and comparison with real-data simulations show that this model, even with independence assumptions provides close to simulated results. In fact, since the model underestimates benefits it can be considered a conservative approach. Some conclusions can be made from results:

- Optimal threshold can be easily computed based on headways historical information. In our example it is almost twice the average headway.
- Since each additional bus in the injection fleet is strictly less beneficial than the previous one, the injection fleet is likely to be small.
- The optimal dispatch time is slightly after half the injecting headway has passed.
- Even though the presented model does not give a method to choose the injection point, calculations are easy enough to repeat at every stop to find the optimal one.
- Injection appeared to be beneficial to the system compared to having the same fleet constantly circulating. Waiting times improve at stops downstream the injection, while stops upstream would be better off without injecting. This means that passenger arrival at each stop is important to determine the optimal injecting stop.

Although multiple-size injections were implemented for a single stop analysis, the generalized model only accounts for single injection. Although the single stop analysis suggested that larger fleets will generally not be optimal, analyzing this for all stops in future work might prove to be beneficial.

References

Osuna E, Newell G (1972) Control strategies for an idealized public transportation system. Transportation Science 6(1):52–72