

**ABSTRACT**

One of the continual challenges faced by transit managers is how to provide better service in a dynamic operating environment. In the context of high-frequency services, in an ideal situation, all buses would be equally spaced along their line, resulting in even intervals between successive bus arrivals to a certain stop.

However, in reality, there are many factors that hinder buses from keeping constant headways, such as irregular passenger demand and variations in traffic conditions. One of the extreme consequences of an irregular transit service is bus bunching (BB), in which two or more buses that should be evenly running along the same route arrive simultaneously at the same stop.

This paper explores the potential uses of automatic vehicle location (AVL) systems to better understand the operation of bus routes and the dynamics of service reliability. It creates visualization tools that allow the analysis of the regularity of the transit service in line 758 detailed down to the stop level, as well as the detection of BB occurrences, based on the data provided by Carris from the month of May of 2018. Also, the developed model can be easily applicable to other routes.

For the total of 4180 trips that were investigated in the regular weekdays, the model detected BB frequencies of 20 to 40% of the trips for each day. Several recommendations are made to the operator, including the tracking of the deviations of the departure headway from the terminal and the implementation of proactive measures.

**Keywords:** bus bunching; regularity; service reliability; headway control; public transport; punctuality

## 1 Introduction

One of the continual challenges faced by transit managers is how to provide better service in a dynamic operating environment. In many areas, services are constantly subject to delays and disruptions due to traffic congestion, weather, vehicle breakdowns and other events [1].

The service delays and disruptions adversely affect transit ridership [2]. In fact, transit passengers consistently rank on-time performance and schedule reliability as one of the most critical factors affecting their use of transit [3].

The match between planning and operations is the basis of service reliability: transit operators offer a network and a timetable, which is their promise to the customer and the extent to which they keep to this promise in all its components defines how reliable the transit service they provide is [4].

In an ideal situation, all buses would be equally spaced along their line, resulting in even intervals between successive bus arrivals to a certain stop. However, this rarely happens, even if no headway deviations are seen in the origin terminal.

One of the extreme consequences of an irregular transit service is bus bunching (BB), in which two or more buses that should be evenly running along the same route arrive simultaneously at the same stop. This BB phenomenon can be simply explained as follows: if a bus is slightly delayed (due to traffic congestion, for instance), that bus will have to pick up more passengers than expected at the next bus

stop (given a certain passenger arrival rate to the stops). This extra passenger demand further aggravates the bus delay, as the boarding and alighting takes more time than planned. On the other hand, the following bus will experience lower passenger boarding and alighting rates, consequently saving more dwell time and catching up in time and space with its preceding bus [5]. This outcome can quickly propagate to the following stops and eventually yield little or zero headway between the two buses.

The bunching problem is an example of a positive feedback loop that reinforces itself: if no measures are taken, it will escalate along the line until bus platoons are ultimately formed [6]. This will result in an increase on the average waiting time for passengers and in a deterioration of in-vehicle comfort levels due to overcrowding in the delayed buses.

The uneven passenger loads mirror a poor capacity utilization, which also represents a decrease on public transport performance on the operators' side: when bunching occurs, an overcrowded bus is followed by a near-empty bus, which leads to a waste of the already limited resources and to a less efficient crew management, resulting in an increase of the operating costs for the transit companies.

The uneven passenger loads also results in a smaller chance of having a seat, and some passenger may see their boarding denied, which is extremely frustrating for transit riders [4]. Transit agencies may also lose loyal customers with such

unreliable transit service causing revenue reduction [1].

In fact, a healthy and efficient public transit system is indispensable to reduce congestion, emissions, energy consumption, and car dependency in urban areas [7].

The performance indicators of Carris provide concerning results on the way the operation of the company has evolved throughout the years. The total demand continuously decreased from 256.6 to 140.6 million passengers between 2004 and 2016 (which represents a variation of 45%) while the supply (*veic x Km*) was reduced by the company in about 32%.

In order to increase the ridership levels of Carris, it is important to understand the underlying causes of the decreasing reliability of the transit service. Being that the prevalence of BB is one of the most visible proofs of the unreliable service [8], this paper aims at developing a method to detect BB occurrences and computing key performance indicators of the regularity of the system. I explore the potential uses of automatic vehicle location (AVL) systems to better understand the operation of bus routes and the dynamics of service reliability.

The objective of this research is to provide Carris (and possibly other transit operators) with a straightforward framework to study the reliability of their services, the conditions that lead to poor regularity and the improvement of service planning and operations monitoring through the available AVL datasets. The levels of complexity surrounding service reliability and the stochasticity of the operation of transit systems do not allow for a clear and simple solution for BB. Therefore, I developed a framework in which the implementation of several strategies can be tested.

## 2 Service reliability

Service reliability can be defined in terms of the variability of service attributes and its effects on traveler behavior and on agency performance [2] and it can be considered either in terms of punctuality or regularity.

In the context of high-frequency and high-demand services (with headways smaller than 10 minutes [9]), regularity is the main determinant of passenger waiting time and reliability needs to be interpreted in terms of regularity rather than punctuality. Therefore, measures to improve reliability must focus on keeping even headways between buses rather than adhering to the schedule [6].

### 2.1 Regularity indicators

Several authors have proposed a large range of measures to assess service regularity and they can be divided in two categories: the indicators that are based on headway distribution and the

indicators that focus on the passenger waiting time distribution.

The first category includes the headway coefficient of variation,  $COV(h)$  [6] and the standard deviation of the recorded headways [10,11]. Other measures refer to the share of headways that are within a certain time interval [9].

The  $COV(h)$  represents the ratio between the standard deviation of the observed headways and the mean actual headway (equation 1):

$$COV(h) = \frac{\sigma_{h_{k,s}}}{\left[ \frac{\sum_{k \in K} \sum_{s \in S} (h_{k,s})}{(|K| \times |S|)} \right]} \quad (1)$$

where  $\sigma_{h_{k,s}}$  is the standard observed headways. This is a normalized measure of headway variability, which takes the value of zero in the ideal case that all headways are equal. The more irregular the service is the higher the  $COV(h)$ . This is a robust statistical measure that provides a direct indication of service variability yet it is not intuitive and may not be fully representative of users' experience.

[9] also established headway adherence levels of service (LOS) grades based on both the  $COV(h)$ , as presented in table 1:

**Table 1 - Fixed-route headway adherence level of service. Adapted from [9]**

Level of service (LOS)	Coefficient of variation of headway (COV)
A	0.00 - 0.21
B	0.21 - 0.30
C	0.31 - 0.39
D	0.40 - 0.52
E	0.53 - 0.74
F	$\geq 0.75$

The second category contemplates measures such as the average passenger waiting time,  $E(\text{waiting time})$ , and the excess waiting time [6], which requires boarding information.

The  $E(\text{waiting time})$  was shown to be the sum of one-half of the average headway with one-half of the ratio between headway variance and the average headway [12], i.e.:

$$E(\text{waiting time}) = 0.5 E(\text{headway}) + 0.5 \frac{V(\text{headway})}{E(\text{headway})} \quad (2)$$

### 2.2 Bus bunching

As explained by [13], the most important variable regarding the BB events is the distance (in time) between two consecutive buses running on the same route, i.e., the headway between the

vehicles. Let the trip  $k$  of a given bus route be defined by  $T_k = \{T_{k,1}, T_{k,2}, \dots, T_{k,s}\}$  where  $T_{k,j}$  stands for the arrival time of the bus running the trip  $k$  to the bus stop  $j$  and  $s$  denotes the number of bus stops defined for such trip. Consequently, the headway between two buses running on consecutive trips  $k, k + 1$  can be defined as follows (equation 3):

$$H = \{h_1, h_2, \dots, h_s\}; h_i = T_{k+1,i} - T_{k,i} \quad (3)$$

A common acknowledgement to all the bibliography is that BB occurs not when a bus platoon is formed but sooner, when the headway becomes unstable, i.e., if the headway is shorter than a certain limit, it can already be defined as BB. According to [9], the threshold that defines instability is the quarter of the planned headway.

It is logical that to efficiently implement strategies to avoid BB or reduce its impacts, the fundamental causes of BB and unreliability must be understood. Common (external) factors for all transit operations include general traffic conditions and congestion; the presence of signalized intersections along the route (route configuration); demand variations; vehicle and driver availability and weather. One can also identify significant inherent (internal) causes of an unreliable service such as: late departures from origin points; unrealistic scheduled running and recovery times; driver behavior or poor performance [14].

Strategies to deal with BB have also been analyzed through the years, whether they are preventive (aimed at reducing the likelihood of BB) or corrective (directed at avoiding further propagation of problems and restoring normal operations) [14]. Nonetheless, preventive and corrective measures may not be enough and taking proactive measures is of extreme importance in the context of intelligent transportation systems (ITS), where real time operations (RTO) and the dissemination of real-time information (RTI) to passengers are essential [15].

### 3 Bus bunching detection

In this section, I will briefly explain the framework that I followed in order to detect bus bunching in the transit service of Carris. Firstly, I analyzed the reliability indicators used by Carris (comparing the *offered km\*vehicles* with the *lost km\*vehicles*). Then, I selected the bus line that I would focus on and proceeded with the data collection in cooperation with Carris. After the data was made available, I used the *Mathematica* software to process all the information, finally leading to a BB detection method for different days and periods of the day.

The stops and the schedule of line I will study can be found in (<http://www.carris.pt/pt/autocarro/758/ascendente/>).

The current control strategy in Lisbon is designed to improve service punctuality - the bus drivers of each vehicle are given a certain time to leave the two extreme stops (terminals) and they are expected to follow the scheduled arrivals to selected intermediary stops (that can be seen as regulation stops). The company provides the clients with a schedule that presents the planned departures from the origin points and the expected time between these regulation stops. However, bus drivers are not expected to change their driving behavior to comply with these arrival times, unless they receive some instruction from the operational control center.

The controllers follow the positions of the vehicles in each transit line through simplified diagrams that are updated every 30 seconds. One of the shortcomings in the current practice is that, even though AVL units allow for real time monitoring of the buses, there seems to be little use of the data for reliability improvement. The two main uses of the data are informing users of estimated bus arrivals and allowing agencies to know the on-time performance of their buses. The actual controls tend to be implemented only when BB has already occurred. These circumstances of Carris set grounds for methodological improvements based on available AVL data, which is explored in the forthcoming sections of this article.

#### 3.1 Data collection

The data used in this study was obtained from Carris AVL system, which ensures the vehicle-to-infrastructure communication and the data storage at the Carris control center (at least every 30 seconds).

The AVL data is stored on a folder that contains as many files as the number of operating vehicles in that day. Each vehicle has two different identifiers: 1) the radio number - related to the components that gather the location of the vehicle; 2) the ID number - the number that Carris uses for fleet identification. These identifiers are unique and there is a one-to-one correspondence between the two identifiers.

Inside the folder for each day, one can find the txt. files of the AVL information for each radio number. Each file has a number of entries (lines) that is equal to the amount of times when the location of that vehicle was recorded by the system. Each line has information on several variables (rows), such as: 1) the time stamp; 2) the bus line; 3) the location (longitude and latitude); 4) the direction; 5) the plate number.

Besides the AVL data information, the following data sources were used in this study:

- The ID numbers of the vehicles that were used in bus line 758 in each day;
- The information on the stops for bus line 758;

- The scheduled trips for each plate number and direction for bus line 758.

### 3.2 Data processing

I started by processing the AVL information for two days only (the 1st and the 6th of February of 2018), having in mind the future replication of that same methodology to a larger dataset (namely, the complete month of May of 2018). Each AVL folder for a specific day represents about 1 GB of storage.

The data processing was developed using the Wolfram Mathematica software, due to both its technical power and ease of use. I performed the following actions:

- 1) Importing the information on the ID numbers of the vehicles that were used on each day

- 2) Converting the ID numbers into radio numbers (through a conversion table)

- 3) Accessing the AVL information of each vehicle: I created a function that imports the AVL data file of each radio among all the files on the daily folders (there are about 1200 AVL files for each day, which represent the number of operating vehicles from Carris fleet).

- 4) Data cleaning: Each AVL file contains all the location records of a vehicle throughout the complete day of operation. I eliminated erroneous or mismatch records, through a series of criteria (for instance, eliminating the entries that had missing location records) and I kept only the variables that would be necessary for the BB detection model.

- 5) Computing the distance between points: I computed the distance between (latitude, longitude) records, using the simplified *Haversine* method, due to the lack of information on altitude. This method is used to calculate the shortest distance over the earth's surface between two points, ignoring any hills between the points. This is a valid approximation, given the frequency in which the records are collected (at least every 30 seconds).

- 6) Visualizing the daily journey of a vehicle: I plotted the distance over time to the terminal (Linda-a-Velha). By looking at the daily journeys of several vehicles, I confirmed that all the records are aggregated for the complete day (no separation is made when the vehicle starts or ends a trip) and that, as expected, the same vehicle can be allocated to different bus lines, which implies that the data will need to be filtered in order to consider the trips on bus line 758 only.

- 7) Data disaggregation: I performed the following actions in order to separate different trips inside the daily journeys of each vehicle:

- Identifying the beginning and the end of each trip, through the identification of the variation in the direction, trip number and type of record saved on the AVL files;

- Separating the AVL data between records of an ending trip (turning point);

- Grouping the trips from each direction for each vehicle;

- Eliminating the trips from bus lines other than 758 (through the plate number identification);

- 8) Data aggregation: Then, I created a function that uploads and allows for the visualization of the AVL files for a specific direction and timeframe.

- 9) Detecting the passing times in each stop: The next goal was to detect the times in which each vehicle passed through each stop, combining the data on the location of the stops with a proximity criteria, since the location records from the vehicles are gathered every 30 seconds.

Therefore, I created a routine that looks into every record on each trip and searches for the line that has the closest location to each stop. The searching function only looks at records that are less than 200 meters away from the stop.

- 10) Detecting bus bunching: In order to identify the occurrences of BB in line 758, the final stage is to compare the timestamps in every stop for consecutive vehicles (i.e. to look at the headways between vehicles at each stop) and decide if these are smaller than the assumed threshold. I compared the several trips, without repeating comparisons, and identified the headways that are smaller than the assumed threshold value.

One of the most commonly used thresholds for instability is the quarter of the planned headway (see section 2.2). The planned headway for bus line 758 for the morning peak (MP) period and for the evening peak (EP) period varies between 6 and 10 minutes. Therefore, I tested three different thresholds (1, 2 and 3 minutes) to identify BB occurrences. The lower the threshold, the lower the number of occurrences is expected.

The output is the number of trips in which BB was detected, for each considered threshold, together with the radio numbers of the pairs of buses that were bunched.

- 11) Headway calculation: Besides the BB detection, I also sorted the different trips in each direction by departure time in order to compute the headways for consecutive trips.

The sorting function looks at the passing times of each trip in each stop and it sorts the trips based on the passing times on the first stop (departure time from the terminal). If there is a missing record for the passing time on the first stop (due to failures in the GPS records), then the function looks at the first valid passing time and performs binary search to order that trip (with incomplete records) with the already ordered trips, based on the passing time on that specific stop. This step will be essential for the computation of headways' statistics for consecutive trips.

## 4 Analysis of the results

Once the Mathematica script was developed, I draw the approach to study the BB occurrences on both directions of line 758 in the month of May of 2018 that is divided in two main areas: the bus bunching detection results and a boarder analysis of the regularity of the service.

### 4.1 Bus bunching detection

#### 4.1.1 Data visualization for each day

One of the most useful tools to visualize large amounts of tabular information is to use time-space diagrams. They are very helpful to identify bus operations and scheduling problems and evaluate the effectiveness of management interventions [16]. The data processing framework was developed in a way that users (the transit agency, Carris, for instance) can select the travel direction, the time period and the dates and simply click a button to obtain the time-space diagrams along the route.

An example of a time-space diagram is shown in figure 1. The x-axis represents time and the y-axis shows the distance to the terminal. The lines represent the actual trajectories and different colors represent different trips. According to the script, there were 9 bunching detections (following the threshold of 2 minutes), which represent a 30% rate of occurrence (out of the 30 trips in this period). The visualization of the time-space diagram can help us see those events more clearly.

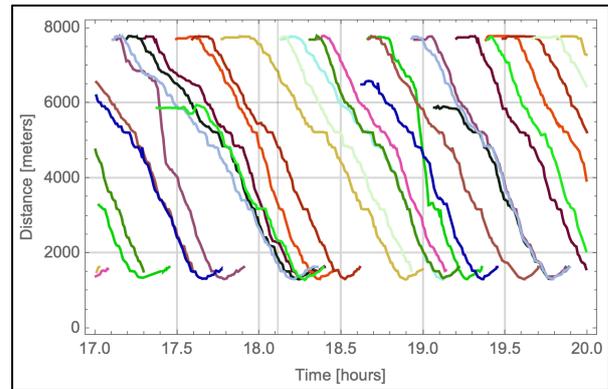


Figure 1 - Time-space diagram example: 22nd of May, evening peak, direction 1

#### 4.1.2 Overall bus bunching detection results

After running the script for every regular weekday of May of 2018, rather than looking at all the details of the BB events of each day, I gathered the information for each weekday. A total number of 4180 trips were computed (2053 for direction 1 and 2127 for direction 2). For direction 1, 604 occurred during the morning peak (MP) and 587 occurred in the evening peak (EP) while, for direction 2, 684 occurred during the MP and 532 occurred in the EP.

Figure 2 depicts the percentage of trips in which BB was detected (according to the threshold of 2 minutes), compared to the total number of trips in each day.

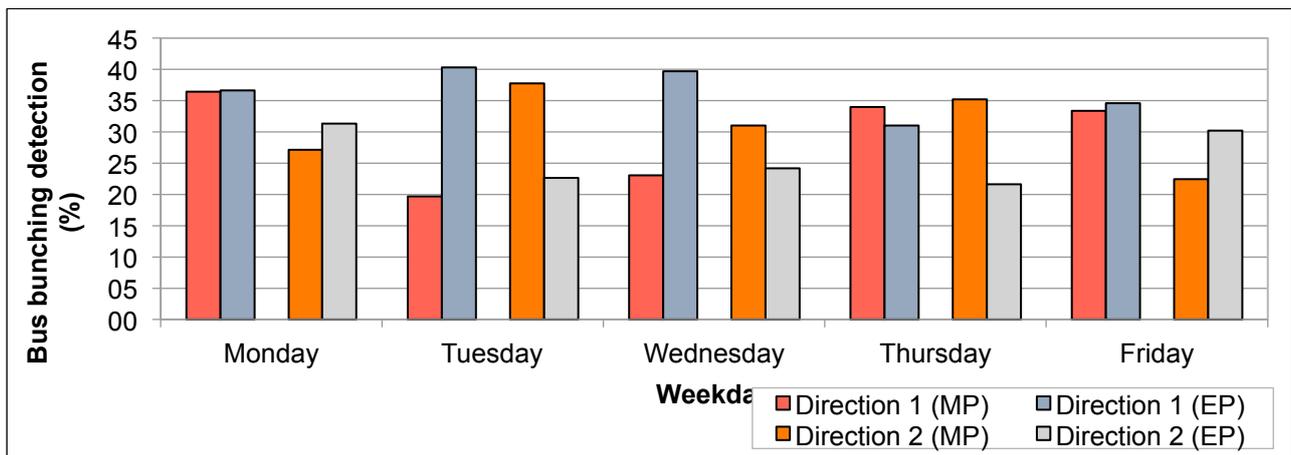


Figure 2 - Percentage of trips with BB detection by weekday and peak period (morning, MP or evening, EP)

From figure 2, we can conclude that:

- The minimum average BB detection frequency was slightly under 20% and the maximum was slightly above 40% (this confirms the importance of studying this phenomenon in this bus line);

- There isn't a specific day in which the average BB detection frequency was particularly higher or a clear trend between the BB detections in the MP and the detections in the EP;

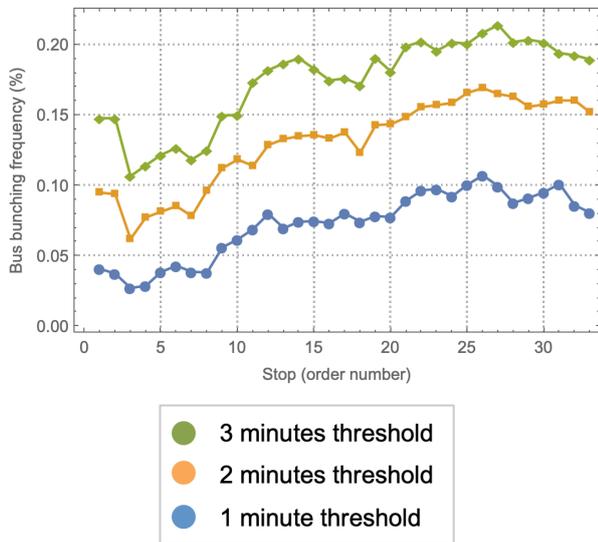
- There isn't a clear trend between the BB detections in the two directions in the morning peak period, however, in the evening peak, in average bunching is more frequent in direction 2 than in direction 1 for every weekday.

#### 4.1.3 Bus bunching along the route

In this section, I assess how the BB events evolved along the route on both directions. As an

example, I depict the results for direction 1 in figure 3. The same representations were obtained for direction 2 and for the peak periods and the results were very similar (they are not depicted due to space limitations).

The x-axis represents the ordered stop numbers of each direction and the y-axis shows the frequency in which the headways on those stops were under the BB threshold. Three lines are shown for the three possible thresholds. As expected, the lower the threshold, the less BB is detected in each stop.



**Figure 3 - Bus bunching frequency - evolution along the route (direction 1)**

In figure 3, we can see that the BB frequency has an overall growth throughout the line. This is expected and confirms that the BB phenomenon propagates along the line and, once two buses are bunched, if they run freely until the end of the route, the headways will not become regular again. Also, the closer the buses are to the end of the line, the more likely it is for them to have faced irregularity causes throughout their service (because they have travelled a longer distance).

#### 4.1.4 Where does bus bunching start?

Besides the identification of the BB occurrences, it is worth knowing where buses get bunched for all pairs of bus trips where BB was detected. I identified those stops for the thresholds of 1 and 2 minutes for each direction and then I analyzed the two peak periods for both directions. In this section, I will sum up the most significant results.

In direction 1, more than 15% of the BB trips bunched in Sete Rios. Also, the majority of the BB trips were formed between stops Cais do Sodré and Estrada de Benfica (Furnas). This matches my expectations: Sete Rios is an expected bottleneck since it represents an important transfer stop. Therefore, a peak in demand is expected, with consequences in the boarding and alighting times, and in the dwell times. Moreover, it is also known that Sete Rios is a driver relief point. If a driver

arrives late to the station, that directly affects the departure time from the stop.

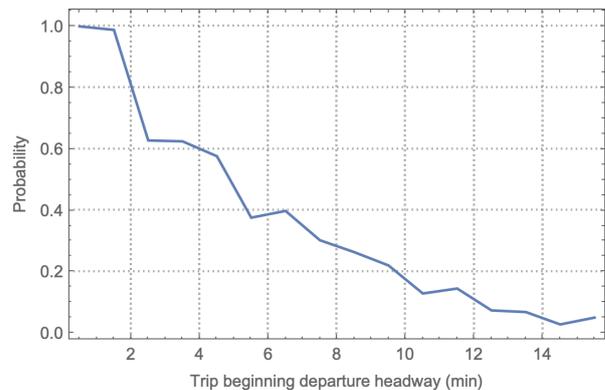
Besides, the fact that most BB trips have origin between Cais do Sodré and Estrada de Benfica is not surprising, since this is the part of the line where the route configuration is more challenging: the narrow structure of the city itself closer to the more historical neighborhood (between Cais do Sodré and Rato), the fact that the bus operates almost always in mixed-traffic conditions, the presence of signalized intersections and the interaction with other vehicles including turning movements, illegal parking and speed changes cause a variation in the cruising speed of the bus, possibly causing the formation of more BB events.

Looking at direction 2, it can be seen that no BB trips have origin in the first 12 stops of the route (between Portas de Benfica and Estrada de Benfica - Furnas) and then, in the stops that follow Sete Rios, the amount of formation of BB trips is more considerable. Also, more than 20% of the BB trips are originated in Sete Rios only. These results are once again consistent with the differences in the route configuration along the line.

#### 4.1.5 The influence of the departure headway

One of the conclusions of the previous section was that some of the BB trips started in the first stops of the trip, which may indicate that some of the trips start with irregular headways from the beginning. To assess this possibility, I estimated the probability of downstream bunching based on the BB detections for both directions for the complete period of analysis. Consequently, I established a relation between the headway at the departure terminal between every two consecutive trips and the existence of BB is any stop in the route for each of those trips.

Figure 4 shows the probabilities of downstream BB trips (considering the 2 minutes threshold) for different departure headway bins for direction 1 (a similar curve was obtained for direction 2).



**Figure 4 - Probability of downstream bunching with varying trip beginning departure headways for direction 1**

Looking at direction 1, if the departure headway is between 0-1 or 1-2 minutes, the probability of a

downstream BB is 100%; if the departure headway is between 2-3 or 3-4 minutes, the probability of BB is about 60% while if the departure headway is between 8-9 minutes, the probability of BB is about 30%.

The general trend for both directions is the probability of BB in the line decreasing as the departure headway bins at the beginning trip increase up to a value that is closer to the scheduled headway (10 minutes for the peak periods and about 15 minutes for the remaining of the day). This indicates that the closer the buses depart from each other in the first stop, the more likely BB is to happen in the line, which is consistent with the expectations, and with the work presented by [16]

This trend outlines the importance of keeping track of the departing headways from the terminals in Cais do Sodré and Portas de Benfica.

## 4.2 Service regularity assessment

### 4.2.1 Headway distribution

Service regularity was primarily analyzed by constructing the distribution of all headways upon arrival to all stops. The narrower the distribution is around the planned headway, the more regular the service is [17]. I started by computing the headway distribution for the complete days (6h00-20h00) and then focused on the peak periods. As an example, I depict the results for direction 2 for the evening peak (17h00-20h00), which had the worst performance (figure 5). The x-axis represents headway bins of 2 minutes and the y-axis represents the frequencies in which the headways among consecutive buses at all stops were within each bin.

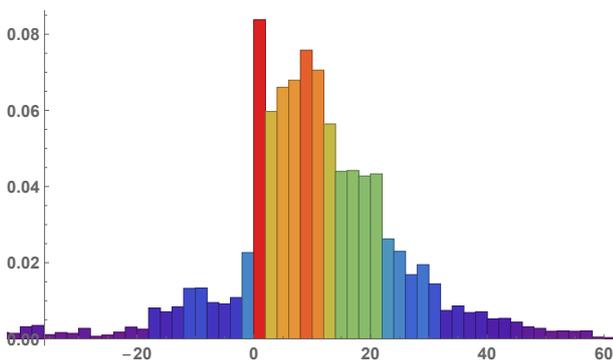


Figure 5 - Headway distribution for direction 2 (17h00-20h00)

The disperse shape of the headway distribution outlines the poor performance in terms of regularity. The most frequent bin is the one that represents the BB situations (0-2 minutes), which is alarming. The share of very short or very long headways severely increases when I compare this period to the morning peak period, and on the other hand, the share of regular headways decreased. There is a big range of headway values, with no clear focus on the planned headway (around 10 minutes). We can also see

that overtaking took place for some pairs of buses since there are records of negative headways.

### 4.2.2 Headway coefficient of variation - fluctuation over time

In order to gain a better understanding on how regularity evolved over the experiment period, I computed the headway coefficient of variation,  $COV(h)$ , for each day. The smaller the values of  $COV(h)$ , the more regular the service is and the less likely BB is to occur (table 1).

I analyzed the variation of the  $COV(h)$  for both directions. Concerning the extent to which regularity fluctuates from one day to the other, no clear patterns were seen. I concluded that the LOS in terms of regularity during the MP period ranged from E to F (the two worst levels, where BB is frequently expected), while during the EP period, the LOS was always F, the worst category. It should be noted that only the positive values of the headways were considered for the computation of the  $COV(h)$  for each day. Otherwise, there could have been a tendency for negative values to cancel positive values and change the average headway. I also compared the fluctuations of regularity in direction 2 with the ones on direction 1 and no pattern was observed. The irregularity problems are common to both directions and periods of the day.

### 4.2.3 Headway coefficient of variation - fluctuation by stop

I also investigated the propagation of irregularity along the route, by computing the values of  $COV(h)$  for each stop individually.

In figure 6 we can see the variation of the  $COV(h)$  along the line for direction 1 (the figure for direction 2 is not included due to space restrictions). In both directions, the headway variability increases considerably along the line and buses arrive very irregularly at the last stops. However, the departures are already very irregular from the first stop (in both directions, the LOS in terms of headway adherence only scores E, the second lowest level, and tends to aggravate for level F along the route). Also, one should note that, in general, the service is more irregular in the evening peak than in the morning peak, supporting the previous results.

### 4.2.4 Headway coefficient of variation - fluctuation by stop

A second category of measures of regularity focuses on the indicators that assess the passenger waiting time distribution. In fact, passenger waiting times are determined by service regularity (for instance, gaps in service and unusual headways have significantly negative impacts on passenger waiting time).

I evaluated the effects of BB on the waiting times distribution in both directions through the computation of the average passenger waiting time

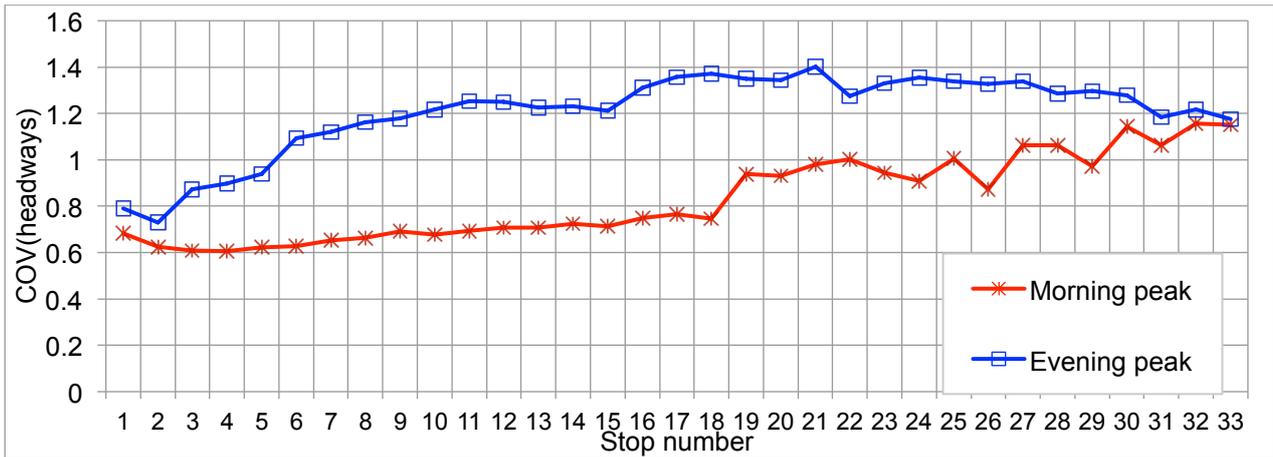


Figure 6 - Headway coefficient of variation (fluctuation along the line for direction 1)

$E(\text{waiting time})$ . I present the results that were obtained along the route in direction 1 (figure 7).

In an ideal scenario, all the buses would be evenly spaced, and considering a scheduled headway of 10 minutes, the average theoretical passenger time would be 5 minutes.

For direction 1 (figure 7), there is a 3 minutes difference between the theoretical and the real value at the first stop (that is about 8 minutes for both peak periods). For both directions, it was seen that the propagation of the irregularity is felt more severely by passengers in the EP than in the MP, even though the  $E(\text{waiting time})$  deteriorates in both periods along the route. Considering the

period when the  $E(\text{waiting time})$  increased less along the route (which was the MP for direction 1), the  $E(\text{waiting time})$  in the stops in the middle of the route is about 10 minutes, which represents a deviation of 5 minutes to the schedule. It may seem like not too much, but if I assume that the average penalty of passenger waiting time is 10€/hour, and that passenger demand is 500 people per hour along the line in this peak period, the daily cost due to passenger waiting time in this peak period of 3 hours will be 1250€. And this is only for one day, for this route and this direction in the morning peak, not to mention the afternoon peak hours, the other direction, the other routes, and that potential ridership may decrease due to irregularity.

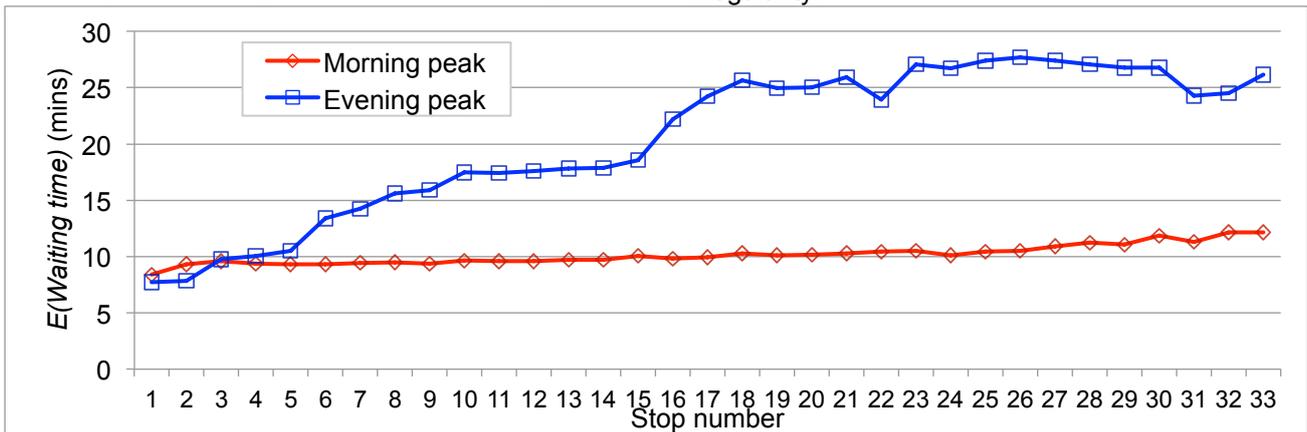


Figure 7 - Average passenger waiting time along the route (direction 1)

Lastly, I observed the total travel time of the observed trips (month of May of 2018) and concluded that the average commercial speed is under the average commercial speed announced by Carris for the complete network: the measured commercial speed varies between 11.4 and 12.4 km/h, while Carris indicates 14 km/h.

## 5 Conclusions

The work developed in this paper is of great value for Carris (and possibly for other transit

operators) and the main contributions are the following:

- The conversion of bus operational records (AVL data) into visible outputs (such as time-space diagrams);
- The computation and visualization of reliability (regularity) performance measures for high frequency services;
- The development of a BB detection model, based only on the archived AVL data according to user-defined parameters (direction, time-window, detection threshold).

The developed script can be easily applicable to other routes, by uploading the information on the ID numbers of the vehicles that were affected to that bus line, as well as the information on the stops.

Based on the results, I also draw some recommendations on how Carris can take advantage of the work here presented:

- Testing strategies that are related to the preventive measures (such as the implementation of dedicated bus lanes, traffic signal priority, parking restrictions) and assessing the differences in the BB detection results and in the key performance indicators before and after the implementation.

- Keeping track of the deviations of the departure headways from the terminals (Cais do Sodré and Portas de Benfica).

- Using the developed framework to assess the success of the changes in the route configuration in process between Portas de Benfica and Estrada de Benfica (Furnas).

A more ambitious recommendation would be for Carris to test a new regularity-driven operation scheme and assess its capability to mitigate BB. This would involve a new real-time control strategy that would have to be included in the business models in order to improve bus service regularity by supporting its full-scale implementation. The development of a service focused on regularity involves a paradigm shift in production planning, operations, control centre and performance monitoring. This could be implemented through bus-to-bus cooperation, as reported by [18].

To conclude, I present the main directions in which the work presented in this paper can be extended:

- The dataset - by including a larger dataset (including a set of bus lines that is representative of the entire network, enabling the study of the interaction with other bus lines sharing the same route or stops);

- Extension of the developed key regularity indicators with the inclusion of smartcard information;

- Assessing the remaining causes of BB (besides the headway deviations at the terminals);

- Developing a bus bunching prediction model in real-time that would recognize patterns in headway deviations and would indicate the probability of BB for consecutive buses with some stops in advance. This would also have to be included in the dynamic visualization available for controllers, so that proactive measures could be taken in real-time.

## References

[1] Haiyang Yu, Dongwei Chen, Zhihai Wu, Xiaolei Ma, Yunpeng Wang, 2016, Headway-based bus bunching prediction using transit smart card data, *Transportation Research Part C: Emerging*

*Technologies*, Volume 72, November 2016, Pages 45-59

- [2] Abkowitz, M., H. Slavin, R. Waksman, L. English, N. Wilson. 1978. *Transit Service Reliability: Final Report*. Technical Report MA-06-0049-78-1, Urban Mass Transportation Administration.
- [3] Hickman, Mark D., 2001, An Analytic Stochastic Model for the Transit Vehicle Holding Problem. *Transportation Science* 35 (3):215-237.
- [4] Oort, N. van, Nes, R. van, 2007, Improving Reliability in Urban Public Transport in Strategic and Tactical Design, Prepared for the 87th Annual Meeting of the Transportation Research Board 2008 Waard J. van der. 1988
- [5] Fonzzone, A., Schmöcker, J.D., Liu, R., 2015. A model of bus bunching under reliability-based passenger arrival patterns. *Transport. Res. Part C: Emerg. Technol.* 7, 164–182.
- [6] Cats, O., 2014. Regularity-driven bus operation: Principles, implementation and business models, *Transport Policy* 36, pages 223–230
- [7] Feng, W., Figliozzi, M., 2011. Empirical findings of bus bunching distributions and attributes using archived AVL/APC bus data. *Int. Conf. Chin. Transport. Prof.*, 4330–4341
- [8] Gershenson C, Pineda L.A., 2009, Why Does Public Transport Not Arrive on Time? The Pervasiveness of Equal Headway Instability. *PLoS ONE* 4(10).
- [9] TCQSM 2nd Ed. (2003) *Transit Capacity and Quality of Service Manual (TCQSM)*, Second Edition, Transportation Research Board, TCRP Report 100, Washington, DC.
- [10] Bartholdi III, J.J., Eisenstein, D.D., 2012. A self-coordinating bus route to resist bus bunching. *Transportation Research Part B* 46 , p. 481–491
- [11] Cats, O., Burghout, W., Toledo, T., Koutsopoulos, H., 2010. Mesoscopic Modeling of Bus Public Transportation, *Transportation Research Record: Journal of the Transportation Research Board*, No. 2188, pp. 9–18.
- [12] Osuna, E.E., Newell, G.F., 1972. Operational strategies for an idealized public transportation system. *Trans. Sci.*6 (1), 52–72.
- [13] Moreira-Matias, L., Cats, O., Gama, J., Mendes-Moreira, J., Sousa, J.F., 2016, An online learning approach to eliminate Bus Bunching in real-time, *Applied Soft Computing* 47, Pages 460–482
- [14] Cham, L., 2006, *Understanding Bus Service Reliability: A Practical Framework Using AVL/APC Data*, Submitted to the Department of Civil and Environmental Engineering in partial fulfillment of the requirements for the degree of Master of Science in Transportation at the Massachusetts Institute of Technology
- [15] Cats, O. & Loutos, G. (2015). Real-Time Bus Arrival Information System – An Empirical Evaluation. *Journal of Intelligent Transportation Systems Technology Planning and Operations*.
- [16] Figliozzi, M., Feng, Wu-Chin, Gerardo Lafferriere, and Feng, Wei. 2012 A Study of Headway Maintenance for Bus Routes: Causes and Effects of “Bus Bunching” in Extensive and Congested Service Areas. OTREC-RR-12-09. Portland, OR:

Transportation Research and Education Center  
(TREC)

- [17] Cats, O., 2013 RETT3—Final Report, A Field Experiment for Improving Bus Service Regularity. Available at: <http://www.ctr.kth.se/research.php?research=rett>
- [18] Daganzo, C.F., Pilachowski, J., 2011. Reducing bunching with bus-to-bus cooperation. Transportation Research Part B 45 (1), 267–277.