TÉCNICO LISBOA

Bus bunching: the case of Carris' transit line 758

João Maria Petinga de Almeida

Dissertação para obtenção de Grau de Mestre em:

Engenharia Civil

Orientadores

Prof. Doutor Filipe Manuel Mercier Vilaça e Moura

Prof. Doutor João António de Abreu e Silva

Júri

Presidente: Prof.ª Doutora Maria do Rosário Mauricio Ribeiro Macário

Orientador: Prof. Doutor Filipe Manuel Mercier Vilaça e Moura

Vogal: Prof. Doutor Amílcar José Martins Arantes

Junho de 2019

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

Firstly, I would like to thank my advisors, Professor Filipe Moura and Professor João de Abreu e Silva, not only for their guidance and support throughout the elaboration of this dissertation but also for their valuable insights on this topic's development.

I am also very thankful to all my friends and colleagues for allowing me to grow personally and professionally throughout my faculty years. Heartfelt thanks go to João Serralha and Édi Gonçalves.

I would also like to thank Pedro Rodrigues, for making it easier to depict the configuration of the route of line 758, and the remaining colleagues of the research group, for their valuable inputs concerning the presentation of the dissertation.

I am profoundly grateful to CARRIS, S.A. for providing me with the location records of the company's vehicles for one month of operation, which enabled the development of the detection model of bus bunching in a real scenario in Lisbon. My sincere thanks to Alexandra Marcelino and Nuno Fonseca, from CARRIS, S.A. for trying to answer all my questions and for facilitating the data collection.

My research would have been impossible without the aid and support of Professor Guilherme Ramos, whose knowledge on *Mathematica* software was essential to fasten the process of data mining. Thank you, Guilherme!

Lastly, I would like to thank my brother Pedro and my parents, Júlio and Lurdes, for always being there and for always believing in me. Thank you!

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 dissertation 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 (and validation) of preventive and proactive measures.

Key words: bus bunching; regularity; service reliability; transit; headway control; punctuality

Resumo

Um dos desafios que os operadores de transportes públicos (TP) rodoviários enfrentam é o de oferecer um serviço fiável num ambiente de operação dinâmico. No contexto de serviços de alta frequência, e num cenário ideal, todos os autocarros estariam espaçados uniformemente na sua linha, resultando em intervalos constantes entre chegadas de autocarros sucessivos a uma dada paragem.

No entanto, na realidade, há diversos factores que impedem os veículos de manterem intervalos de tempo constantes entre si, tais como a procura irregular dos serviços ou as variações nas condições de tráfego. Uma das consequências extremas de um serviço de TP rodoviário irregular é o bus bunching (BB), situação em que pelo menos dois autocarros da mesma linha chegam à mesma paragem em simultâneo.

Nesta dissertação são explorados os usos potenciais dos sistemas de localização automática de veículos (AVL) para melhor compreender a operação nas linhas de autocarros e a dinâmica da fiabilidade do serviço. Foi criado um modelo de visualização que permite a análise da regularidade do serviço de TP rodoviário na linha 758 (que é detalhado ao nível das paragens deste percurso), assim como a detecção de ocorrências de BB. O modelo baseia-se em dados que foram fornecidos pela Carris, referentes ao mês de Maio de 2018, sendo que o algoritmo é generalizável à restante rede de serviços da Carris.

Tendo em conta a amostra de 4180 viagens analisadas para dias de semana regulares, o modelo detetou frequências de BB entre os 20 e os 40% em relação às viagens efetuadas em cada dia. Foram recomendadas diversas medidas ao operador, incluindo a monitorização cuidadosa dos desvios dos intervalos de tempo nas partidas do terminal, assim como a implementação e posterior validação de medidas preventivas e proactivas.

Palavras-chave: bus bunching; regularidade do serviço; fiabilidade; controlo de headways; transporte público; pontualidade

Table of Contents

List of figures

List of tables

List of acronyms

ADC - Advanced data collection
AFC - Automatic fare collection
ANN - Artificial neural network
APC - Automatic passenger counts
APTS - Advanced public transport systems
AVL - Automatic vehicle location
BB - Bus bunching
BBS - Bunching black spot
BDS - Bus dispatching system
BLIP - Bus lane with intermittent priority
BRT - Bus rapid transit
COV - Coefficient of variation
CTA - Chicago Transit Authority
DBL - Dedicated bus lane
DR - Delta rule
EWT - Excess waiting time
EP - Evening peak
GIS - Geographic information systems
GPR - Gaussian process regression
GPS - Global positioning systems

HA - Headway adherence

HD - Headway deviation

HR - Headway regularity

IBL - Intermittent bus lane

ITS - Intelligent transportation systems

KNN - K-nearest neighbor

LF - Load factor

LOS - Level of service

LS-SVM - Least square support vector machine

LTT - Link travel time

ML - Machine learning

MSL - Minimum service level

MP - Morning peak

MPS - Minimum performance standards

PBC - Performance based contract

PT - Public transportation

RF - Random forest

RTI - Real time information

SVM - Support vector machine

TSP - Transit / traffic signal priority

VMS - Variable message sign

1. Introduction

1.1. Transit service background

The last few decades have shown a substantial increase in personal mobility. Not only in interurban travel but also in the urban environment, traffic and transport volumes have been increasing for years worldwide. However, the share of public transport in this mobility growth did not change much and still remains rather limited. Nevertheless, in order to ensure the accessibility and livability of our cities for future generations, a substantial quality leap in public transport is necessary. This will facilitate a desired modal shift from car traffic towards public transport, which is safer, cleaner and produces less congestion (van Oort, 2011).

Private modes (personal car, bicycle, walking) share similar features such as the fact that the user defines most aspects of a trip and is responsible for the full operational and maintenance cost. On the other hand, in public transportation, the operational and maintenance cost is shared by all users as well as subsidized in most cases by the government. This shared cost creates a situation where the stakeholders have opposing objectives. The question is then how to provide a reasonable transportation service for a diverse set of users and desired trips that is not too expensive. To do so, trips are consolidated both spatially along predefined bus routes and temporally at some frequency.

The time it takes a transit vehicle to travel a route can be broken into two parts: time spent overcoming distance and time spent at stops. The time spent overcoming distance depends on the cruising speed of the vehicle, i.e. the speed at which it can move between stops, and on the acceleration and deceleration rates. These depend on the physical capabilities of the vehicle but also on external factors that can impede its progress, such as traffic congestion and traffic signals. The time spent at stops can be broken into several parts, such as opening and closing doors and allowing passengers to board and alight. There is also an amount of time spent decelerating to a stop and accelerating back to cruising speed.

The average speed at which the vehicle can move, including both the time spent traveling and stopped, is called the commercial speed. This is the speed which transit schedules are based on and which a user experiences when traveling to a destination (Pilachowski, 2009).

Since there is an amount of time lost with each stop, the more densely the stops are located, the lower the commercial speed of the mode will be. However, the total time users spend on their trip also depends on how long it takes them to access a transit stop from their origin, and how long it takes them to progress to their destination from the transit stop they alight. If stops are located too far apart, this time can outweigh the time savings from having fewer stops. This is part of the trade-off that must be taken into account when designing a transit network.

The other part of the trade-off is determining frequency of service. The headway, defined as the time between successive bus arrivals to a point in space, determines the number of users served by each vehicle and how long users must wait for a vehicle to arrive to a transit stop. Vehicles must be dispatched often enough that they do not become overcrowded and users do not have to wait too long for one to arrive. However, more vehicles are required to provide more frequent service and each additional vehicle carries with it the cost of a driver as well as capital and maintenance costs.

All of these components are taken into account during the design process to provide a certain level of service to the user. However, the actual level of service users experience depends on the reliability of the system to operate as designed. This 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 (van Oort et al., 2007).

For example, having a dedicated right-of-way allows a vehicle to travel at a constant speed uninterrupted by externalities between stops, and having a fixed dwell time removes any randomness caused by random passenger arrivals at transit stops. Most heavy rail systems have both of these features and so under ordinary circumstances have very high reliability and schedule adherence. In contrast, most bus lines operate within general traffic and have dwell times based on the number of users at each bus stop (even skipping stops when there is no reason to stop) and because of this, bus transit can have very low reliability. Reliability is therefore acknowledged as one of the most important performance factors for both passengers and operators (Redman et al., 2013).

1.2. Bus bunching

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 in consecutive runs, even if no uncertainty is associated with the dispatching from the origin terminal. These factors include variations in traffic conditions, heterogeneous driver behavior, irregular passenger demand and unreasonable bus bay layouts, among others (Haiyang Yu et al., 2016).

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 is presented on the right side of figure 1 and 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 (Fonzone et al., 2015). This outcome can quickly propagate to the following stops and eventually yield little or zero headway between the two buses.

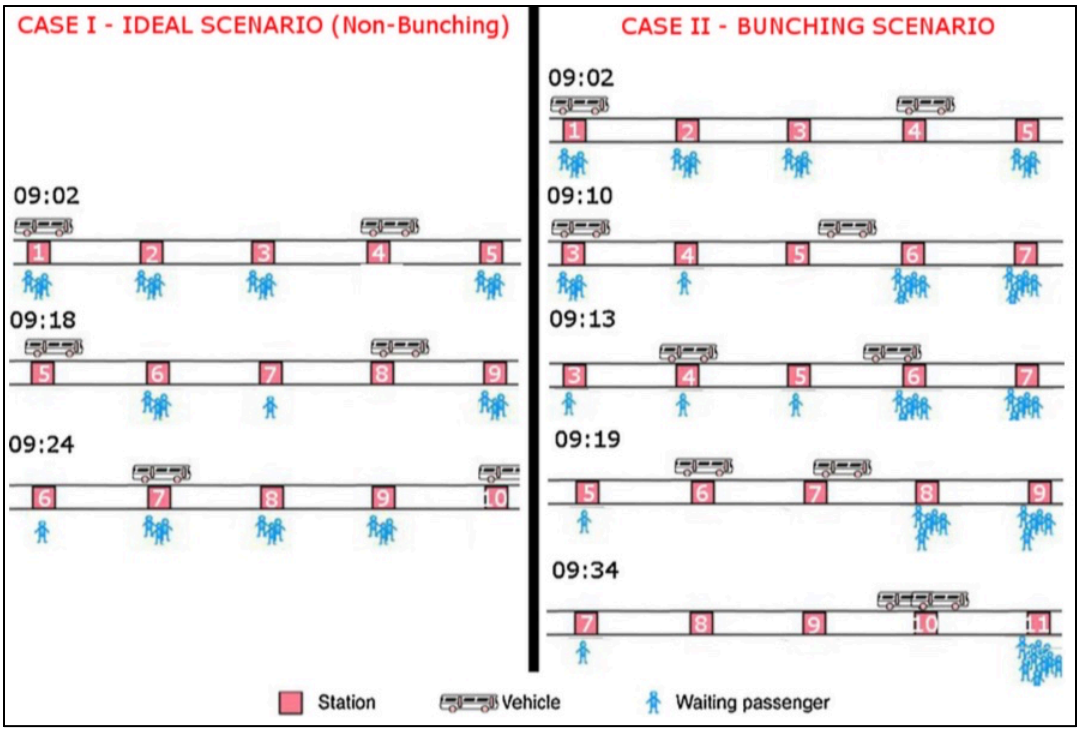


Figure 1 - Bus bunching illustration (Moreira-Matias, 2016)

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 (Cats, 2014). 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 in 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. Transit agencies may also lose loyal customers with such unreliable transit service causing revenue reduction (Haiyang Yu et al., 2016). That is why it is of utmost importance to monitor and predict the level of service reliability, and specifically to study the bus bunching phenomenon.

1.3. Motivation and scope - transit services in Lisbon

As Gershenson (2009) mentions, the prevalence of BB is one of the most visible blueprints of an unreliable service. Two (or more) buses running together on the same route is an undeniable sign that something is going terribly wrong with the company’s service. It is indeed not uncommon to see Carris buses bunched in pairs and even triples, particularly at peak times, which causes great frustration for transit riders. But BB should not be seen as an inevitable phenomenon. There is a significant number of studies that reveal its causes and try to model these bus systems and implement measures to prevent BB from happening, or at least mitigate the consequences of its occurrence, as will be seen in chapter 2. And that's why companies have invested on improving the reliability of their system, bearing in mind the direct impacts not only on the number of passengers they can attract to transit systems but also in the perspective of lowering their operation cost.

Since this dissertation will build on the Lisbon bus operator, Carris, it is also important to understand the spatial and legal context in which the company is embedded and have a first look at the main performance indicators of the company in the last years.

PAMUS-AML (2016) defines the strategy on the regional transport system of the metropolitan area of Lisbon and, among the points that are outlined, one can highlight two strategic objectives that are particularly related to the topic of this work: the surpassing of inefficiencies that affect the performance of the system and the transition to a more energetically efficient and environmentally sustainable system. Bus bunching occurrences prevent any of these objectives from being achieved.

In fact, a healthy and efficient public transit system is indispensable to reduce congestion, emissions, energy consumption, and car dependency in urban areas (Feng & Figliozzi, 2011). If there are areas where bus routes chronically underperform, transit becomes more unattractive for private modes users and transit becomes consequently incapable of having more ridership.

According to INE database (INE, 2001 and 2011), the daily commuting trips in Portugal have seen significant differences between 2001 and 2011: while there was a decrease from 15.9% to 11.8% of bus trips and from 25.1% to 16.4% on the trips made walking, the trips made by private car have increased from 45.7% to 61.6%, which mirrors the growing dependency on private modes across the country. Further fluctuations can be seen in annex 1. This trend applies in particular to the metropolitan area of Lisbon, as shown in figure 2. The histograms show that the reduction on the public transport trips is roughly replaced by the growing of the trips made with private modes.

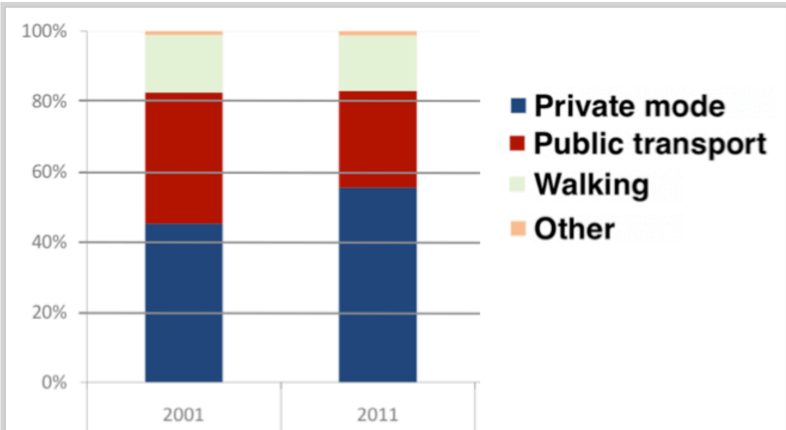


Figure 2 - Distribution of transportation modes in the commuting trips for students and workers living inside the Metropolitan Area of Lisbon. Source: INE, Recenseamento Geral da População, 2001 and 2011

Furthermore, the Statistical Yearbook of Portugal (INE, 2017) states that, in 2016, the providers of public passenger road transport services (in heavy motor vehicles) in mainland Portugal provided 17.8 million transport services and transported 478.9 million passengers. The 25.2 billion seat-kilometers had a demand of 6.8 billion passenger-kilometers, i.e. only 26.9% of total capacity was actually used. In particular, in domestic scheduled transport, the average utilization rate of road transport was 18.8%, a value that can compromise the environmental advantages of transit services, considering that diesel and gas accounted for the majority of the fuel consumed by Carris buses up until this moment.

The performance indicators of Carris also provide important information 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% in twelve years in the transit service). Likewise, the supply $(veic \times Km)$ was reduced by the company in about 32%. The average commercial speed of buses also decreased between 2004 and 2016 in about 0.5 km/h and the utilization rates varied between 22% and a minimum value of 15.5% (2014). In the years that followed 2014, the company was able to obtain bigger utilization rates, but mostly at the cost of having offered less available $seats \times Km$. This can be seen in detail in figure 3 where the supply and demand data of the company throughout the years is presented. In these twelve years, the number of $passengers \times Km$ decreased 48% (on the demand side) and the number of available $seats \times Km$ offered by the company was reduced in 47%. At the first sight, it seems like the company adjusted the levels of the supply to the levels of demand, with further negative impacts on the demand. These performance indicators are shown in detail in annex 2.

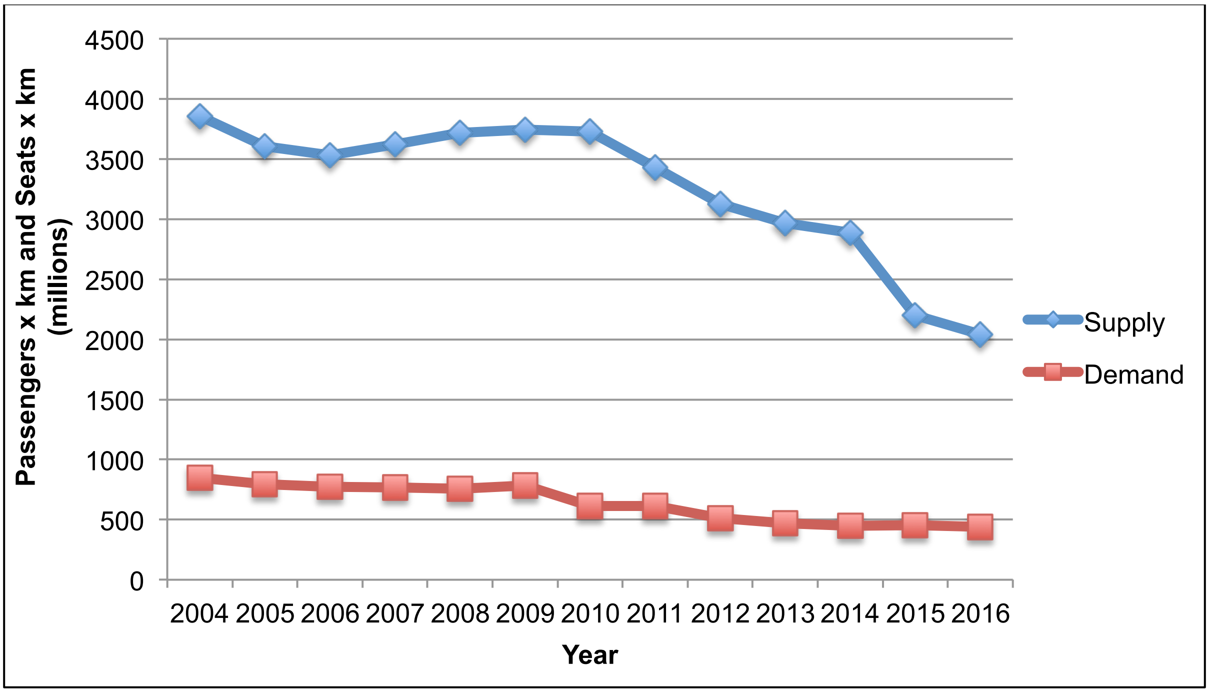


Figure 3 - Carris supply (Seats x km) and demand (Passenger x km) evolution between 2004 and 2016

The dissatisfaction of Carris customers is also mirrored in their complaints to the company. In 2017 alone, more that 2800 written complaints were handed over to the company concerning the irregularity of the service (which represents more than 7 complaints a day). These complaints covered problems such as breakdowns, punctuality issues, delayed departures from the terminal and other service failures.

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, this dissertation aims at studying the causes and effects of BB, namely by developing a method to detect BB occurrences and by computing key performance indicators of the regularity of the system, which the transit operator can use to compare its performance before and after the implementation of any corrective strategy.

This dissertation explores the potential uses of automatic vehicle location (AVL) systems to better understand the operation of bus routes and the dynamics of service reliability. With the availability of larger AVL datasets, there is an opportunity for transit agencies to answer these questions:

- How can service reliability be assessed based on the available AVL datasets?
- What conditions lead to service reliability problems?
- What strategies can be applied to avoid or reduce poor service regularity?
- How can AVL data be used to improve service planning and operations monitoring?
- How can severe regularity deviations such as BB be detected and predicted based on the AVL data?

The objective of this research is to provide Carris (and possibly other transit operators) with a straightforward framework to address them.

By better understanding the dynamics of transit service, transit agencies would be able to improve the service through improved delivery, management and planning (Cham, 2006).

This study recognizes that 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 the causes of bus bunching and its consequences. However, it aims at developing a framework in which the implementation of several solutions can be tested.

1.4. Structure of the document

The remaining chapters of this document are organized in the following way: in chapter 2, I present a literarature review on how reliability and BB have been studied so far; in chapter 3, I present the BB detection model, where the methodology for data analysis is explained - the data acquisition, cleaning, the transformation and visualization processes and how the AVL data was modeled in order to evaluate the occurrence of BB - and I introduce the case study (bus line 758); in chapter 4, I perform the analysis of the results (both on the BB detections and on the new key performance indicators for regularity, in order to support decisions regarding the reliability of the system). Finally, in chapter 5, I draw the main conclusions of the dissertation and explain how this research can be implemented and further developed.

2. State of the art

2.1. Framework and methodology

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. It is well known that transit passengers consistently rank on-time performance and schedule reliability as one of the most critical factors affecting their use of transit (Hickman, 2001). The inevitable service delays and disruptions do adversely affect transit ridership and the overall level of service (Abkowitz et al. 1978 and Abkowitz 1980). Moreover, from the point of view of the transit operator, service delays and disruptions have a real monetary cost, in terms of lower utilization of vehicles and drivers.

A review of previous studies reveals that considerable research already exists on service reliability and bus bunching in particular, on the effects of unreliability on passengers and transit agencies, on the identification of the causes of unreliable systems and on the application of strategies to improve the level of service. However, the complexities of transit service and the limitations of manual data collection have made it difficult to fully understand service reliability and to identify the relationship between service attributes and reliability in the past. In recent years, the development and implementation of automated vehicle location (AVL) monitoring and the automated passenger counts (APC) data collection systems have sparked the opportunity for more detailed analyses and for an improvement on transit operation monitoring.

Consequently, it is logical that Bartholdi and Eisenstein (2012) state that literature can be divided into two groups: one prior to AVL systems, such as global positioning systems (GPS), and the other after. Before GPS, approaches assumed that not much information was available. Barnett (1974), for instance, computed bus delays based on the distribution of observed headways. In these papers, the object was generally to reduce variation.

Subsequent to GPS, models have commonly assumed accurate and real-time knowledge of the locations of all buses on a route, not to mention bus velocities and even instantaneous arrivals of passengers to each bus stop (Eberlein et al., 2001). Many also assume communication amongst buses (Daganzo and Pilachowski, 2011) or between buses and bus stops (Zhao et al., 2003). And while earlier papers concentrated purely on adjusting the bus delays at control points, more recently authors have considered additional means of control, including adjusting the velocities of buses (Daganzo and Pilachowski, 2011), vehicle overtaking (Hickman, 2001), skipping some stops, or even refusing to allow some passengers to board (Delgado et al., 2009).

In this chapter, an extensive literary review on bus bunching is provided. Before doing so, it is important to explain the methodology that was followed on the selection of the articles. The aim was to review all the main literature on bus bunching and service reliability in transit services and the article that served as source of inspiration for the development of this dissertation was (Cats, 2014). The sources to find the reviewed articles were widely available databases such as SCOPUS and Google

Scholar and the search was done with keywords such as "bus bunching", "bus reliability" and "transit reliability". Only articles written in English were covered and the time frame included articles written between 1964 and 2016, with more emphasis on more recent articles. The selection of the papers was based on their impact level (based on the number of citations each paper had) and backward snowballing was often used (i.e., finding citations in a paper that is being reviewed and review it as well). No geographical areas were excluded in this review, even though there might be differences in the public transport systems from country to country. The analysis of those differences may also be interesting for the topic and may also allow the definition of possible gaps in the literature.

In section 2.2, I begin by providing a review on how reliability measures have been studied and how reliability is achieved in various jurisdictions worldwide. It should be noted that bus bunching is an underlying issue of an unreliable service and it is consequently impossible to study one topic detached from the other. Then, on section 2.3, a reference is made on the role of new technologies such as automated data collection and real time information (RTI) on transit operation. Moving on to section 2.4, a deeper analysis is made on the origins of the bus bunching phenomenon and its consequences, while in section 2.5, I analyze how different researchers have addressed this topic throughout the years and on their recommendations for transit operators. Finally, in section 2.6, I draw the main trends in the research.

2.2. Service reliability

Transit authorities around the world have been striving to improve transit service quality to attract more transit riders (Haiyang Yu et al., 2016). The level of service enhancement largely relies on whether the transit system can be operated as it was planned (Bellei and Gkoumas, 2010).

Service reliability is indeed one of the most important level-of-service determinants for both public transport users as well as for attracting car users (Redman et al., 2013). Service reliability can be defined in terms of the variability of service attributes and its effects on traveler behavior and on agency performance (Abkowitz et al. 1978) and it can be considered either in terms of punctuality or regularity. That leads to one of the key aspects of transit reliability: the distinction between low-frequency and high-frequency routes. The attributes of transit reliability differ based on the frequency of the route because of the difference in passenger behavior and passenger perception of an unreliable service. Passengers on low-frequency routes are inclined to study the schedule and coordinate their arrivals at the bus stops with the scheduled bus arrival time in order to minimize waiting time. As a consequence, punctuality is fundamental in these services.

In contrast, in the context of high-frequency and high-demand services, passengers typically arrive randomly at stops without referring to the timetable. Hence, service 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 the statistics on passenger waiting time distribution must focus on keeping even headways between buses rather than adhering to the schedule (i.e., giving preference to a headway-based operation over a schedule-based operation (Cats, 2014)). However, agencies often see schedule adherence as an important and easier-to-collect measure of the

effectiveness of service delivery. As a reference, TCQSM 2nd Ed. (2003) considers the threshold of 10 minutes headways above which the passengers check the posted schedules (low-frequency routes) or below, when they are assumed to arrive randomly (high-frequency routes).
Abkowitz et al. (1978) present the first comprehensive assessment of transit service reliability. The study explored the impacts of transit unreliability on both travelers and transit agencies and the main conclusions of this study are presented, compared with the ones seen in more recent papers.

2.2.1. Traveler behavior

Abkowitz et al. (1978) firstly reviewed the relationship between reliability and travel behavior, in particular mode choice and departure time decisions. Travelers choose a travel alternative and departure time based on their perception of travel time variability and on their risk aversion. This choice reflects the assumptions that travelers associate a certain loss in value with being late at their destination. It is assumed that travelers aim to minimize total travel time, accounting for the variability of in-vehicle travel time and waiting time. For the latter, it was found that travelers do not perceive a reliability problem if bus arrivals are predictable to a certain degree. Travelers tended to adjust their arrival time at bus stops to reduce their expected waiting times, on infrequent services. The theoretical view of departure time behavior was based on the travelers' need for on-time arrival at the destination, their familiarity with the system and bus arrival patterns.
The authors state that a number of benefits to travelers can result from improved service reliability. The most important is the increase in the probability of on-time arrival, as the variability in total travel time, especially waiting time, is reduced. This probability increase may decrease the disutility of transit relative to other modes and thus lead to more frequent use of the system. It may also attract new riders whose transit disutility is now lower than their prior travel alternative (Cham, 2006). Reducing travel time variability may allow travelers to choose a later departure time with the same probability of on-time arrival.

2.2.2. Transit operators perspective

The theoretical discussion presented in Abkowitz et al. (1978) examines agency behavior as it affected transit service reliability. To transit operators such as Carris, reliability is most often defined in terms of schedule adherence. In systems that keep a tight scheduled running time, with no slack built in, companies rely on recovery time to account for the variability in running times.
Decisions of service operations managers involve a trade-off in resource allocation and operating policies such as setting frequency, running times and recovery times. To the agency, reliability strategies consist of maintaining schedules as reflected through on-time performance measures and ensuring an adequate availability of spare vehicles and drivers. The decisions on the number of spare vehicles and drivers and the strategies for assigning them also affect the ability to cover vehicle breakdowns and additional service disruptions.
The authors state that improving service reliability may yield benefits in reduced capital and operating costs as a result of reduced travel time variability. Schedule adjustments to improve service may reduce the size of the spare fleet required. Improving service reliability also has the potential to increase ridership, and therefore, revenues (Cham, 2006).

2.2.3. How to measure reliability - worldwide overview

Recognizing the centrality of this aspect of service quality in passenger experience, industry regulators around the world have introduced various service reliability frameworks in their performance monitoring regimes. In order to measure and monitor service reliability, it is essential to select key performance indicators that are both comprehensive and concise, which facilitate a consistent and understandable comparison among various lines and operators. In the first study of Abkowitz et al. (1978), the authors developed criteria to address a set of service measures that characterize reliability and are useful for problem identification and strategy implementation that focus on three main aspects:

- The compactness of the distribution of running times;
- The likelihood of extreme delays (which cause great dissatisfaction for travelers);
- The normalization of measures (using standard indicators to allow for the comparative analysis between different time periods and routes and to assess the impacts of different service attributes).

The authors also recognized the importance of time-of-day and day-to-day variability. Based on the presented arguments, Abkowitz et al. (1978) recommend the measures presented in table 1:

Table 1 - Measures of reliability of transit service (Adapted from Abkowitz et al. (1978))

Measures of reliability	Indicators
Distributions of travel time (total, in-vehicle or waiting times).	1. Mean; 2. Coefficient of variation (COV); 3. Percent of observations "N" minutes greater than the mean value.
Schedule adherence, measured at any point along the route.	1. Average deviation from schedule; 2. COV (from the average deviation, not of the schedules); 3. Percent of arrivals "N" minutes later than the schedule.
Distribution of headways	1. Mean; 2. COV; 3. Percent of headways greater than X (or lower than Y) percent of the average or scheduled headway, with $X \geq 1, Y \leq 1$.
Seat availability	1. Passenger loads (demand and capacity)

Forty years later, this dissertation revisits this framework to assess which features are still appropriate and applicable in light of the improved understanding based on more available data. In more recent years, and throughout the world, a large number of entities has developed and tested enhanced measures of service reliability that build on these.

TriMet in Portland, Oregon, uses the Bus Dispatching System (BDS) to monitor public transport reliability (Feng & Figliozzi, 2011). The BDS combines AVL and APC data to provide detailed information on bus service performance. The two performance measures are headway deviation and actual headway spatial distribution. Headway deviation looks at the difference between actual headway and scheduled headway in each stop while actual headway spatial distribution depicts the proportion of actual headways deviating from scheduled headway against different stops along the route. High and low deviations on the distribution would suggest congestion or disrupted areas that require improvements.

In Shanghai and Jiangyin City, a normalized average headway index is used to determine the actual headway deviation from the scheduled headway (Leong et al., 2016). An index below 100% indicates that the bus is earlier than scheduled, while an index above 100% indicates that the bus is behind schedule. However, averaging headways may not be the best method to study headway deviations since when a bus is very close to its precedent, it is usually further away from the following bus, and the average of these two deviations may come across as null.

In Chicago, the AVL data is used to determine running time adherence and headway regularity (Leong et al., 2016). Running time adherence measures the average difference between actual and scheduled run times, while headway regularity measures the average difference between actual and scheduled headways. A high metric value for these two indicators will indicate irregular bus services and poor reliability.

In Sydney, Transport for New South Wales measures punctuality of buses at the beginning of trip, mid-point of trip and at the last transit stop, requiring at least 95% of the trips to be between 2 min early and 6 min late. It is important to note that this on-time performance measure depends on the selected punctuality window, and there is no common standard among different bus operators. More importantly, if this is applied as a binary score: 1 in a region predefined as punctual and 0 otherwise, it penalizes equally all headways that dissatisfy the threshold requirement regardless of the extent of the deviation.

Transport for London (TfL) characterizes London bus services depending on whether they are high or low frequency. The reliability of high frequency services (defined as those with headways of less than 15 minutes) is assessed based on average excess waiting time (EWT) experienced by commuters. Unreliable bus services, as evidenced by irregular spacing of buses, will result in high EWT. On the other hand, low frequency services are assessed on percentage of buses departing on time according to bus schedules. The most relevant metrics are further developed in section 2.3.

2.2.4. Regularity metrics

As already mentioned, in the scope of frequent services, service regularity is the main determinant of passenger waiting time and consequently reliability needs to be interpreted in terms of regularity rather than punctuality. A large range of measures used to assess service regularity was proposed by the previously presented studies, but they can be divided in two categories: 1st - the indicators that are based on headway distribution and its relation to the planned headway and 2nd - the indicators that focus on the passenger waiting time distribution.

The first category includes the headway coefficient of variation (Cats, 2014), the headway adherence (TCQSM 2nd Ed., 2003), an index based on the Gini ratio and an irregularity index that particularly penalizes long headways (Cats, 2014) or the standard deviation of the recorded headways (Bartholdi and Eisenstein (2012) or Cats et al. (2010)). Other measures refer to the share of headways that are within a certain time interval, similarly to their on-time performance counterparts. Strathman et al.

(1999) used the ratio between observed and scheduled headway as an instantaneous measure for identifying bus bunching.

Besides the above-mentioned metrics of the 1st category, Lin and Ruan (2009) propose a probability-based metric that is meaningful to both passengers and transit managers, focusing on frequently serviced urban bus routes. While arrival time information is certainly valuable to passengers, the reliability of the information (defined as the probability of the value within a predefined acceptable range) is equally desirable to passengers. Lin and Ruan (2009) first present the mathematical formulation of the metric and then apply it to evaluate service reliability of bus route monitored by Chicago Transit Authority (CTA). Using variables such as bus dwell time, number of stops into the trip, passenger activities (i.e. arrival, boarding and alighting) and expectation (or tolerance) of bus headways, they define headway regularity (HR) in terms of a probability of a bus arriving within a certain maximum anticipated headway (H) in each specific stop k:

$$HR = P(headway_k \leq H) \tag{1}$$

The proposed measure is derived from data which is available from automatic vehicle location (AVL) and automatic passenger counter (APC) and thus it provides an appropriate way for transit agencies to improve service and meet passengers' expectations and thus increase ridership. It is found that headway regularity during a bus trip is closely impacted by dispatching headway and that headway reliability decreases as the number of stop increases, as echoed by other studies. Furthermore, the time-point level service reliability declines as passenger activity levels increase or as the maximum passenger anticipated headway decreases (i.e. passengers become more demanding of frequent bus services).

Even though this approach is practical and intuitive, equation 1 does not constrain the lower bound of the headway. This is consistent with passengers' preference in frequent bus service, which value small headways but, from the agency's point of view, frequent bus service means high operating costs, which may lead to much less desirable policies and operation. Such constraints are not considered in (Lin and Ruan, 2009), nor any penalty for the scenarios where bus bunching occurs. Therefore, the headway regularity metric won't be used in this study.

As for the second category, Osuna and Newwell (1972) first established the relationship between headway variation and average passenger waiting time, based on the assumption that passengers arrive uniformly at stops. Average passenger 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, i.e.:

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

London buses use a variation of this indicator which computes the difference between the actual and scheduled waiting times at the disaggregate level based on individual headways (Cats, 2014), calculating the scheduled waiting time as half of the planned headway. This measure indicates how much longer than intended passengers are waiting on average - for example, a value of 1.5 indicates that passengers wait 50% longer than planned (Transport for London, 2012).

Cats (2014) specifically analyses a new regularity-driven operation scheme, tested in two field experiments in Stockholm, in order to demonstrate its capability to mitigate bus bunching. In this paper, the three following measures are used (the first two are from the 1st category and the last one is from the 2nd category):

- Headway coefficient of variation, $COV(h)$ - the ratio between the standard deviation of the observed headways and the mean actual headway:

$$COV(h) = \frac{\sigma_{h_{k,s}}}{\left[\frac{\sum_{k \in K} \sum_{s \in S} (h_{k,s})}{(|K| \times |S|)}\right]} \tag{3}$$

where $\sigma_{h_{k,s}}$ is the standard deviation of the 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.

- Headway adherence, HA - the share of buses that arrive with a headway that does not deviate from the planned headway by more than a certain percentage:

$$HA = \frac{\sum_{k \in K} \sum_{s \in S} (\delta_{k,s})}{|K| \times |S|} \tag{4}$$

where $\delta_{k,s} = 1$ if $\left|\frac{h_{k,s} - h_p}{h_p}\right| < \alpha$, where α is a pre-defined threshold, and $\delta_{k,s} = 0$ otherwise. Depending on the service and the level-of-service standards, α may vary between 0.1 and 0.75, although for most urban services this range can be truncated to from 0.2 to 0.5 (Cats, 2014). This measure is easy to communicate, as it is percentage-wise and comparable across lines and systems. However, it penalizes equally all headways that dissatisfy the threshold requirement regardless of the extent of the deviation.

- Average excess waiting time, EWT - the additional waiting time that passengers experience due to irregular bus arrival:

$$EWT = \sum_{k \in K} \sum_{s \in S} (b_{k,s} \frac{h_{k,s}}{2} - \frac{h_p}{2}) \tag{5}$$

where $b_{k,s}$ is the number of passengers boarding trip k at stop s. The fraction $\frac{h_{k,s}}{2}$ represents the average waiting time assuming that passengers arrive randomly during the headway, from which one subtracts $\frac{h_p}{2}$, the theoretical waiting time that would have occurred if buses followed the planned headway. In case passenger counts are not available for each trip, the number of boarding passengers could be estimated as linearly proportional to the corresponding headway based on the average of historical or sample passenger counts. A perfectly regular service – where all buses are evenly spaced – implies that $h_k = h_p \; \forall k$ and thus yields $EWT = 0$. Even though the calculation is not intuitive, the measure itself is easy to communicate as it is measured in time units. Moreover, it is comparable across lines and systems. This measure also incorporates the impact of missed trips, which is also true for the headway adherence (HA).

Cats (2014) defines headway adherence in the form of a probability, whereas (TCQSM 2nd Ed., 2003) previously established headway adherence levels of service (LOS) grades based on both the COV(h) and their definition of the probability of bunching, as presented in table 2:

Table 2 - Fixed-route headway adherence level of service. Source: TCQSM 2nd Ed. (2003)

Level of service (LOS)	Coefficient of variation of headway (COV)	Probability [%]	Comments
A	0.00 - 0.21	≤ 1	Service provided like clockwork
B	0.21 - 0.30	≤10	Vehicles slightly off headway
C	0.31 - 0.39	≤20	Vehicles often off headway
D	0.40 - 0.52	≤33	Irregular headways
E	0.53 - 0.74	≤50	Frequent bunching
F	≥ 0.75	>50	Most vehicles are bunched

In this manual, the probability of bunching is defined as the probability that a vehicle's headway will differ from the scheduled headway by more than 50%. Besides headway adherence LOS, (TCQSM 2nd Ed., 2003) also proposed levels of service for on-time performance, which is a critical reliability characteristic for low-frequency routes. This punctuality estimator is presented in annex 3.

Passenger loads can also be seen as a measure for regularity. Service frequency is set to ensure that the demand is evenly distributed and that buses do not exceed a given load factor (LF). This benchmark is a percent of the number of seats:

$$LF = \frac{occupancy\ of\ the\ vehicle}{seating\ capacity} \tag{6}$$

Most transit agencies nowadays use a benchmark maximum number of standees as criteria in schedule design standards. When passenger loads are high, passengers experience a lower level of comfort, and boarding and alighting times generally increase, along with the dwell time at stops. Overcrowding eventually leads to a more severe problem when passengers are not able to board the first bus that arrives at a stop because it is full and have to wait for the second vehicle. This obviously increases the waiting time and user's frustration.

2.2.5 How to improve reliability

Abkowitz et al. (1978) categorize strategies to improve reliability as follows:

1) Priority strategies, where transit vehicles receive special treatment to reduce the influence of external factors.
2) Control strategies, which involve direct handling of active service operations.
3) Operational strategies, that relate to changes in route, schedule and resource allocation.

The authors also group strategies into two other categories, according to their application:

a) Preventive strategies, aimed at reducing the likelihood of reliability problems developing.
b) Corrective or restorative strategies, directed at avoiding further propagation of problems and restoring normal operations.

The identified strategies are summed up in table 3:

Table 3 - Strategies to improve reliability. Source: Abkowitz et al. (1978)

		Preventive	Corrective
Priority	Exclusive lanes	✔	
	Signal Priority	✔	✔
Control	Holding		✔
	Passing / Overtaking		✔
	Turning back		✔
	Stop skipping		✔
	Speed modifications	✔	✔
Operational	Reserving vehicles and drivers	✔	✔
	Schedule adjustments	✔	✔
	Express service	✔	
	Improving vehicle access (fare collection, boarding/alighting)	✔	

Several studies evaluated the effectiveness of these strategies on improving reliability, through empirical analyses and simulation models, as will be further investigated in section 2.4.

One of the approaches to improve reliability of a bus system is to reduce the severity of the perturbations that can affect the components of travel time. There are several ways to do this (as seen in table 3). One of the ways of doing so is implementing Bus Rapid Transit (BRT) lanes. Giving a bus a dedicated lane allows it to travel without being delayed by general traffic. Also, installing Transit Signal Priority (TSP) systems allows buses to avoid stopping at traffic signals. Furthermore, pre-boarding ticketing systems and aligned platforms reduce the amount of time each passenger takes to board a bus. All of these components reduce the amount of randomness that can be added to a bus trip, which results in less control needed to provide on-time performance and service regularity. Well-known examples of BRT are Curitiba, Bogata, and Seoul. While many of the abovementioned principles are already in use (or being tested) by Carris, full implementation of BRT in Lisbon is not yet in place. These measures will be detailed in section 2.4.4.

2.2.6. Improving reliability in strategic and tactical design

Usually, reliability issues are analyzed and improved at the operational level of public transport, which is the main focus of this study. The hypothesis followed in the research by van Oort et al. (2007) is that already during the strategic and tactical planning phases reliability can be taken into account and be improved. The authors make a separate analysis of several components of travel time: access time; waiting time; in-vehicle time; transfer time (optional) and egress time to better understand the problem. One of the impacts of variations in system's performance on travellers is their travel times becoming longer. For instance, irregularity on headways causes an increase on waiting times. Research on the appreciation of different components of travel time (van der Waard, 1988) shows that one minute of waiting time is perceived equally to one and a half minutes of in-vehicle time or more, which depicts the bigger dissatisfaction associated with an increase on waiting time and the resulting importance of this parcel. According to Wardman (2004), the disutility is even higher, since the ratio between waiting time and in-vehicle time is in the range of 1.5 to 2.0.

To learn about actual service reliability and its effects, van Oort et al (2007) observed the excess waiting time on several stops of a high-frequency tramline of a PT network of The Hague, The Netherlands and concluded that unreliability of the service (in terms of irregularity) had a large effect on the travel times.

In (van Oort et al, 2007), reliability is presented as the match between planning and operations: operators offer a network and a timetable, which is their promise to the customer and reliability defines to what extent they keep to this promise. Hence, two ways of improving reliability are possible: by adjusting practice to planning on one hand and by fitting planning to practice on the other hand.

One can divide the planning of PT in 3 levels. At the strategic level the network is designed, with the public transport lines and indicative frequencies. At the tactical level, the indicative frequencies are extended to a detailed timetable, which is used by both travelers and operators (for the planning of the crew and the fleet). The last level is the operational level, in which the LOS is measured by several quality indicators. By using feedback loops to the upper levels these results can be used to improve the system's performance and reliability. Ibarra-Rojas et al. (2015) elaborated an extensive review of the literature on transit network planning problems and real-time control strategies suitable to bus transport systems.

van Oort et al. (2007) give examples on how the operation itself can be used at the first two levels in order to improve reliability: at the tactical level - by using accurate driving times to design achievable timetables, also by testing the impact of using different percentile values of actual driving times on the punctuality, the extra waiting time for the passengers and the number of vehicles running ahead of the designed schedule; at the strategic level, where reliability is not usually one of the design criteria - by studying the impacts of parameters such as the line length, the coordination of several bus lines in the same infrastructure and the stop spacing choices on reliability (in terms of regularity or punctuality).

2.2.7. Improving reliability through the regulator's policies

From the regulator's perspective, one way of ensuring improvements on the reliability of public transport is through the use of performance based contracts (PBCs) in the bus service provision in particular. Hensher and Houghton (2002) proposed a system that takes into account various external costs such as costs of congestion, among many others, with social surplus maximization as the underlying motivation in order to ensure that bus operators deliver the optimal service level that is consistent with the needs of stakeholders, especially the government.

Working along similar lines, Hensher and Stanley (2003) highlighted the importance of PBCs as a crucial factor that aligns commercial objectives with social objectives by rewarding operators for achieving a minimum service level (MSL) and for an increase in ridership. Social benefits such as reduced waiting time, reduced number of transfers and transfers of riders from car to public transport can be internalized into the operators' remuneration contracts based on revenue kilometers and passengers, for instance.

Cats (2014), among other studies, have also highlighted the need for transport companies to place greater emphasis on bus reliability and to incorporate it into an incentive scheme: besides being an important factor in determining people's behavior choices, reliability is also found to be a main reason behind users' dissatisfaction towards public bus services. In addition, service reliability may be a decisive factor in determining people's tolerance level in other areas such as comfort and ventilation (Disney, 1999).

The incorporation of bus service reliability into a PBC regime for operators has been widely implemented around the world, such as in New Zealand, in Santiago, Chile and in London (Transport for London, 2015). In London's case, an incentive mechanism embedded in its quality incentive contract is used to encourage bus operators to provide reliable service, with a bonus of up to 15 percent or deduction of up to 10 percent of the contract price relative to the required standards. These standards are known as the Minimum Performance Standards (MPS) and are a crucial part of London's scheme as they act as a reliability benchmark for operators. Operators are not paid for any mileage not operated for reasons within the operator's control, such as staff shortages or mechanical issues. Setting MPS properly is a very important part of the contracting process, as failure to do so could mean bus operators passing the risk of poor performance to the public authority by putting the anticipated penalty in the contract bid.

Since the implementation of the Quality Incentive Contracts in 2001, EWT of high frequency services has fallen from 2.2 min to 1.0 min and the percentage of on-time low frequency services has risen from 68 percent to 83 percent in 2014 (Transport for London, 2014). TfL reports that service improvements are estimated to have accounted for some 30 percent of the growth in demand for bus services between 1997 and 2012 (Transport for London, 2014).

In particular, in the context of high-frequency systems, a shift in the business models is also needed (Cats, O., 2014) since, in most of the current incentive schemes, public transport agencies often determine the bonuses or penalties given to service operators based on the punctuality. Instead, the agreements should contemplate the regularity of the system, through indicators like the ones presented in section 2.2.4, which reflect the passenger perception of service reliability much more accurately.

Furthermore, measuring the performance only at a small subset of stops along the line through the computation of punctuality related to the schedule contributes to abrupt performance patterns.

2.3. The role of automated data collection and real time information

In recent years, intelligent transportation systems (ITS) have been extensively deployed in planning, control, and management of modern transportation systems. Public transport agencies have either successfully implemented or are in the process of transforming their systems by adapting ITS applications for planning, maintenance, and operational purposes (Cats and Loutos, 2015).

Public transport systems are increasingly equipped with information and communication technologies in order to improve passenger level of service and facilitate fleet management. Advanced data collection (ADC) techniques, such as automatic vehicle location (AVL) and automatic passengers counts (APC), were first used in order to improve operations and management. Later on, they were also utilized to disseminate real-time information (RTI) to passengers (Cats and Loutos, 2015).

Extensive research exists on the potential of both AVL and APC systems to improve operations, performance monitoring, scheduling and planning of transit systems (Wilson et al. 1992; Furth et al. 2003; Kimpel et al., 2004; Hammerle, 2005; Lin and Ruan, 2009; Cats and Loutos, 2015).

With the development of such systems, public transport agencies obtain details regarding fleet positioning, travel time and speed, passengers on board, and dwell time at stops. These sources are then processed and integrated in order to generate RTI concerning present and future states of the public transport system. RTI could refer to information on service disruptions, crowding conditions, and prescriptive journey planners. However, the most commonly provided as well as the main focus of research is information concerning the remaining waiting time, since it constitutes an important and uncertain part of the trip (Cats and Loutos, 2015).

Most AVL systems are also used for real-time operations control and emergency response (the most recent systems transmit data to a server in real-time). However, this is often labor-intensive, as traffic dispatchers need to take numerous decisions in response to rapidly evolving conditions, which applies to Carris control center. These systems are usually not designed to provide archived data due to the huge data storage requirements that would be needed to upload this information. On the other hand, APC systems are mostly designed for off-line analysis and have been used to evaluate performance.

Among the technological advances outlined in the studies are the development of location-based tools such as global positioning systems (GPS) and geographic information systems (GIS); the improvements on communication with wireless network technologies and more efficient radio communications; the integration of systems and "smart bus" design, which combines the functionalities of various systems such as passenger information, fare collection and scheduling.

For example, knowing the headways and the passenger loads on a bus allows decision-makers to assess whether it is beneficial to hold a bus, what time point to hold at, for how long, and approximately how many passengers on-board will be impacted. Off-line data from these systems may also be useful in helping decision-makers evaluate the effectiveness of holding strategies. Archived data may be used to analyze the impacts of control strategy decisions, and to develop decision-support tools that enable supervisors to estimate the conditions under which holding and other strategies are most appropriate. Boarding and alighting profiles generated from historical data would be helpful to approximate the number of passengers, both on-board and down the route, who would potentially benefit from holding a bus.

As mentioned in (Haiyang Yu et al., 2016), the recent advent of automatic fare collection (AFC) (i.e., smart card) can reduce the costs of data collection, as well as accelerate the passenger boarding and alighting process through contact-less cards and card readers. The data gathered by the AFC system contains abundant temporal and spatial information of individual passengers and can be utilized to quantify transit performance, improve transit operation, and understand passenger behavior.

Furth et al. (2003) outlines common issues of matching captured data with the corresponding schedule. Matching captured data with bus stop locations has been difficult because many agencies do not have an accurate "stops" database. However, the introduction of stop announcements systems tied to GPS systems has increased the need for more accurate stop matching because errors are very obvious in real-time announcement systems (which is also true in the case of Carris). Problems also arise when trying to match data to schedule files for the current timetable. An automated process is needed which requires compatibility between the different systems. Matching is also limited by the need for valid entries for driver, run and trip numbers. Issues also exist with end-of-line data records because passenger counts may be skewed by confusion surrounding the end of the trip and the beginning of the next trip.

The study also states that good analysis using archived data from APC systems depends on the availability of other useful data items and databases. The capabilities of AVL and APC systems are enhanced when other related data items are integrated. Potentially valuable data items include: door open and close times, start and stop times, time stamps on passenger entries and exits, off-route events, mechanical and security alarms, communications to and from the control center, traffic control messages, farebox transactions, and annunciation and destination signs. Related databases include schedule data, GIS, payroll, farebox, maintenance, weather or special events, and customer satisfaction. To sum up, the study notes the opportunities presented by the transition to a data-rich environment and its resulting analysis possibilities, highlighting the need for good integration with related databases of stop locations, schedule information and fare collection data.

The fact that more and more cities now provide real-time information (RTI) for passengers also leads to new directions and extensions to the work developed so far by researchers on bus bunching. The increasing presence of service schedule information to passengers due to online availability of journey planners changes the ‘‘visibility of the network’’ and hence passenger behavior, even for passengers that are unfamiliar with the network. For instance, it is reasonable to expect that permanent RTI on departure time (accessed by internet and/or mobile phone apps) induces non-uniform passenger arrivals also for short headways and irregular services.

Cats and Loutos (2015) develop on the added value of RTI on its implications on passengers' waiting time uncertainty. The prediction produced by the RTI generation method performs better than the timetable in predicting actual waiting times and hence contributes to a reduction in waiting-time uncertainty, an important level-of-service determinant. In other words, the prediction scheme in this study yields a waiting time estimate that is more than twice as close to the actual waiting times than the timetable is. This implies that RTI enables passengers to shift their expectations considerably closer to their experienced waiting time. This is especially important since passengers perceive waiting time as longer than the actual waiting time they have experienced and the authors show that this overestimation of waiting times did not occur when RTI was available. Hence, the accuracy of the information provided to the customers plays a bigger role on their perception of the reliability of the system.

In fact, the hypothesis of random (uniform) passenger arrivals over time and space in models of public transport with headways shorter than ten minutes might change with the availability of information on bus arrivals to stops. Actually the threshold between schedule-dependent and uniform passenger arrival can be lower than the conventional one (Fonzone et al., 2015). These authors develop a ‘‘reliability-based arrival pattern model" which is applicable to situations in which passengers consider timetables in deciding their arrival at stop. They develop a "bus propagation model" - an algorithm to derive the dwell time of subsequent buses serving a stop in order to illustrate when bus bunching might occur.

One should also note further the effect of electronic ticketing. Through smart card technologies, boarding has become faster which possibly reduces the bus bunching effect (Fonzone et al., 2015). At the same time, more efforts are made by bus operators to further cater to the needs of population groups with special needs such as wheelchair users. These require additional boarding times, leading to more variability in the boarding rates.

2.4. Bus bunching

It is well known from experience and theory that collective bus motion within high frequency transit lines is unstable i.e., that even if buses start with perfectly even headways, they invariably become irregular, and if enough time passes, they bunch up (Daganzo, 2009). This has been known for more than fifty years: as explained in Newell and Potts (1964), a reason for this instability is that if a disruption causes a bus to slow up relatively to the bus it follows, the bus encounters more passengers along the way, and these extra passengers delay it further on its line. On the other hand, the next bus tends to catch up, causing a positive feedback loop that results in collective arrivals of buses to the same stop.

2.4.1. Designations and measurement

As summed up in (Moreira-Matias et al., 2012), this phenomenon has several designations: the Bangkok effect, bus platooning, vehicle pairing, bus clumping, bus cluster or bus bunching (BB). Nevertheless, it's important to note that they all refer to the same effect of headway instability. In this study, I will use the most commonly used designation (bus bunching).

As explained by Moreira-Matias et al. (2014), the most important variable regarding the BB events is the distance (in time) between two consecutive buses running on the same route. Such distance is denominated as time headway (or simply headway). 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 headways between two buses running on consecutive trips $k, k+1$ can be defined as follows (equation 7):

$$H = \{h_1, h_2, \dots, h_s\}: h_i = T_{k+1,i} - T_{k,i} \tag{7}$$

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., the headway between two consecutive buses doesn't need to equal zero for operators to acknowledge an occurrence as BB. If the headway is unstable (shorter than a certain limit), it can already be defined as BB. What is different among the reviewed articles is the way they define the system as unstable.

According to Moreira-Matias et al. (2014), the threshold that defines instability is when the headway between two consecutive buses is shorter than a quarter of the planned headway. As for Cham (2006), the Chicago Transit Authority (CTA) defines irregular headways as headways below an absolute value of 1 minute (buses with headways lower than 1 minute are considered bunched). And on the other extreme, headways that are above 150% of the scheduled headway are considered gaps in service.

2.4.2. The causes

It is logical that to efficiently implement strategies to avoid bus bunching or reduce its impacts, the fundamental causes of BB and unreliability must be understood. However, it is also clear that poor reliability can be triggered by many different and interrelated factors and so it is difficult to fully understand and address the underlying causes of BB. In many cases, the effects of one cause yield the occurrence of another cause.

Abkowitz et al. 1978 classify basic causes of unreliability as environmental or inherent. The authors consider environmental factors to be those resulting from the surrounding of the system, generally random in nature, while inherent factors are those associated with the transit service itself. The authors separate the causes of unreliability common to all transit systems, as well as those specific to fixed route bus system. That same framework will be followed in this document. Common (external) factors for all transit operations include:

- **General traffic conditions and congestion**: they affect transit service, since transit vehicles typically operate in mixed-traffic. Variations in running time result from interactions with other vehicles, including accidents, turning movements, illegal parking or speed changes.

- **The presence of signalized intersections along the route (route configuration)**: it interrupts the free flow of traffic and increases the probability of delays. Running times increase due to stop-and-go operations and stop times at red lights. Shi-Gong et al. (2016) elaborate on the influence of synchronized traffic light on the states of a bus operating system, focusing on the relationship between the system states or the average speed and the signal period while Albright and Figliozzi (2012) present a regression model to analyze the impact of Transit Signal Priority (TSP) strategies specifically on bus bunching.

- **Demand variations by day-of-week and season-of-year:** agencies account for systematic or known fluctuations in demand by adjusting schedules accordingly, but random variations in demand cause variability in dwell times and travel times.

- **Vehicle and driver availability**: it is related to the operating policies of the agency, but affects reliability in terms of the driver being able to cover all scheduled trips and having spares in case of breakdowns, accidents or driver absences.

- **Weather**: lower visibility levels, cautious drivers or poor road condition during adverse weather conditions can reduce the speed of vehicles on the road and increase the probability of accidents occurring. Boarding and alighting times also tend to be above average. According to Hofmann and O'Mahony (2005), rain has a negative impact on travel times and their variability (worsening on-time performance) but it seems to contribute to a more regular service with a slight trend to less bunching, which is surprising. This may be due to the increased traffic congestion, which prohibits the buses to move more freely. Also, the size of the samples taken for this study in Ireland may have been too small, as mentioned by the authors.

These causes are generally environmental in nature. While the availability of vehicles and drivers tends to be driven by operating policies and can be controlled to some extent, random externalities still exist, such as vehicle accidents and driver's sickness. One can identify the following as significant inherent (internal) causes of an unreliable service:

- **Late departures from origin points (garage or terminals):** such initial deviations tend to propagate, creating unbalanced passenger loads and may worsen conditions downstream. Increased boardings at further stops result in higher dwell times, which increase total running times. Besides the possible occurrence of BB, if not enough recovery time is available at the end of trips, late arrivals of previous trips might carry over to the next trip and cause further late departures. Emphasis is given to schedule deviations at the terminals because drivers and supervisors have more control at these points than at intermediate points along the routes. Hammerle et al. (2005) analyzed two causes of bus bunching: on-route effects and at the terminal or route start. By analyzing a two-day sample of archived operations data, they found that most bus bunching was the result of irregular departure headways at the terminals instead of on-route effects (Figliozzi et al., 2012). This is further supported by the work of Cham (2006), who analyzed the Silver Line on Washington Street in Boston (the first BRT system and flagship bus service operated by the Massachusetts Bay Transportation Authority), and concluded that the main cause of service reliability problems were initial deviations from the terminal.
- **Unrealistic scheduled running times and recovery times that buses are unable to follow**: when actual running times are lower than scheduled, buses will tend to run "early" for most of the route. On the other hand, if too little running time is scheduled, buses will not run as scheduled. The need to fulfill the scheduled arrivals at the stops itself may cause drivers to adjust their speed, possibly causing BB. Transit agencies typically deal with this problem by inserting slack into their schedules. Slack, however, reduces the commercial speed (this reduction is quantified in Daganzo (1997)).
- **Driver behavior or poor performance**: some employees tend to drive faster and more aggressively, while others might be slower and more cautious. This causes a variation in travel times between consecutive buses. Drivers may also report late to work or take long personal breaks. In the worst case scenario, they might also disregard scheduled times and headways when they deliberately catch up with their preceding bus and ride along bunched in order to pick up fewer passengers and have an easier trip. Previous studies (for instance, Kimpel et al., 2004) identified some driver characteristics as most affecting their behavior and performance.

Feng & Figliozzi (2011) provided a method to identify the temporal and spatial location of bus bunching events and analyze the causes of those events in a bus route in Portland. This study looks at deviations from bus states along their routes and determines the frequency in which those deviations happen for all BB occurrences. The study looks at seven attributes for the front bus and six attributes for the following bus and the results indicate that late departures of the front bus and on-time or early departures of the following bus are the major triggers of BB events, as expected. They found that short departure headway at a stop is mainly due to irregular departure from an upstream stop instead of irregular passenger demand or uncertain travel times between each two consecutive stops. The authors suggest that the agency moves from a schedule-based service to a headway-based service during the high frequency service hours. This recommendation is common to several other articles. The authors also highlight that an increase on dwell time may happen not only due to passenger boarding and alighting movements, but also due to the lift use, to the existence of other buses at the same stop (possibly of other operators), or due to the poor stop location relative to a signalized intersection.

2.4.3. The consequences

The consequences of bus bunching on users, on the provider and on society are broad and they have already been mentioned within previous sections of this work, at least partially, but they are summarized here.

Bus bunching is associated with longer average waiting times for some riders and an uneven passenger distribution along the buses (Feng & Figliozzi, 2011), which causes a deterioration of in-vehicle comfort levels due to overcrowding in the delayed buses. 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 (van Oort et al. 2007). Also, many passengers find it uncomfortable to stand for a long time or not to be able to use their travel time productively (TCQSM 2nd Ed., 2003). Transit utility decreases even further when passengers compare feeling cramped in a bus with the privacy and comfort of a single seat ride on a private vehicle. Besides that, higher crowding levels make it more difficult to move around inside the bus, making the boarding and alighting processes more stressful. This added interference increases the dwell time at bus stops but also in-vehicle driving time.

Passenger will experience longer driving times and dwell times, leading to deviations on the departure and arrival times at stop. The commercial speed of the bus will decrease as it falls behind schedule and must bear more passengers, resulting in longer travel times and waiting times at the stops. Furthermore, intermodal transfers will become less smooth - transfers can be delayed or made unviable.

If the number of buses on a route is fixed, for every group of bunched buses, there will be locations along the route not served by buses for long periods of time, which mirrors a poor capacity utilization. On the other hand, if additional resources are applied by transit companies to cope with the decrease on the system's performance, the operation cost will increase and the resource allocation and crew management will be less efficient.

All these factors will reduce the level of service, the quality and reliability of the transit service. This combined with traveler's uncertainty during a trip will lower customer satisfaction and discourage riders from using the transit system (van Oort et al. 2007). Transit agencies may end up losing loyal customers resulting in a decrease on ridership, leading to revenue reduction (Haiyang Yu et al., 2016).

As more and more people move towards urban areas, overloading the existing infrastructure, the role of public transportation as an efficient means of mass transportation needs to increase. Additionally, with political pressure towards promoting more environmentally friendly transportation options there is a growing awareness of public transportation as a means to travel. There is a portion of the population that is captive and relies on public transportation, regardless of the level of service. However, beyond this, agencies that provide public transportation are not guaranteed a mode share. In order to get people to ride, transit must be fast, convenient, and reliable. Bus bunching negatively affects all three of these qualities (Pilachowski, 2009). If a bus line has a reputation for being unreliable, that is a disincentive for users to ride the bus.

If not enough people ride buses, two things can happen: either the buses operate with lower occupancy and higher fares, or service can be reduced. Since buses output more emissions than private vehicles, they will be more polluting per user per km than private vehicles if they run at a low enough occupancy. This could cancel out the environmental benefits of transit. On the other hand, if fares are raised or service is reduced, those with no other option than to ride the bus are negatively impacted and those who have a choice are more likely to choose a different mode. Because transit is funded in part by the government, there is also the issue of wasting resources to provide a poorly performing and largely unused service. Finally, there is an effect on the bus drivers. Depending on how far behind schedule a bus is running, a driver may find their break time cut short or missed entirely. In addition, there is often hostility towards drivers when a bus arrives to a bus stop late. These can lead to dissatisfaction among the drivers and a hostile work environment.

2.4.4. Preventive measures

In section 2.4.2, the internal and external causes of the variability on the scheduled headway that may lead to BB were defined. In sections 2.4.4 and 2.4.5, the strategies to deal with BB are now detailed. Following the same categorization seen on section 2.2.5, preventive and corrective strategies can be distinguished: the preventive aim at reducing the likelihood of BB from occurring while the corrective ones try to restore normal service and to minimize the effects of BB, once it has occurred. I will firstly analyze the preventive strategies and then focus on the corrective ones (section 2.4.5).

At the strategic level of PT planning, the choices on route design, lane priority and signal priority may influence the occurrences of BB:

- **Route length** is an important variable since studies have shown that the variability in running time is expected to increase with distance since there is more chance for triggering events to occur and for initial deviations to propagate (Cham, 2006).

- **The number of stops** also increases the variability in running times due to the variability on stop-and-go operations (deceleration, boarding and alighting of passengers, acceleration). However, fewer stops as improvement strategies are limited by network design issues and service standards, such as maximum walking distance to bus service or maximum number of transfers.

- The implementation of **dedicated bus lanes** (DBLs) reduces the variability in bus running times due to traffic conditions and allows for greater control of speeds. However, their introduction is costly and not always feasible - its implementation is constrained by current street dimensions, surrounding development and traffic volumes. Also, DBLs remove one lane from general use and therefore reduce capacity. In addition, exclusive lanes, unless fully protected or grade separated, are still subject to intersections, turning movements and illegally parked vehicles.

- Besides DBLs, the route design may also consider **intermittent bus lanes** (IBLs), in which the status of a given section changes according to the presence or not of a bus in its spatial domain: when a bus is approaching such a section, the status of that lane is changed to BUS lane, and after the bus moves out of the section, it becomes a normal lane again, open to general traffic. Therefore when bus services are not so frequent, general traffic will not suffer much, and bus priority can still be obtained. This measure can be operating at a single city block, but if all related control parameters along bus lines are considered together, more time gains can be obtained. Viegas and Lu (2004) firstly developed on the structure and operation of IBLs on a single intersection and then along one bus line and among several bus lines within an area.

- **Traffic signal priority** (TSP) may also reduce running time variability by reducing delays at signalized intersections. Without signal priority, buses may be "fortunate" enough to catch all the green phases at intersections and go through intersections without having to stop, while "unlucky" buses may arrive at intersections during the red-phase of the signal and be forced to stop and wait. With TSP, buses can extend the green phase of traffic signals to claim the right-of-way and proceed unimpeded through an intersection.

There are two types of traffic signal priority: *1) absolute priority,* where a green phase is given to all buses; and *2) conditional priority,* providing a green phase only to buses that are running behind schedule.

Muller and Furth (2000) examined the impacts of TSP on average delay, measured as the time difference between the actual crossing time and the time it would take a typical unimpeded vehicle to go through the intersection, for both transit buses and private vehicle traffic in Eindhoven, The Netherlands. The results showed conditional priority did not cause significant changes to private vehicle traffic. For buses, average delays were lower than without priority. Another study (Kimpel et al. 2004) examined running times, on-time performance and passenger excess waiting time with data from before and after the implementation of traffic signal priority. The analysis results were mixed with an overall improvement in average running times, but mixed outcomes at the individual route level and by direction and time-of-day. On-time performance decreased as the tendency for early arrivals increased, and headway variability and excess wait time increased.

- Eichler and Daganzo (2006) expand on the work presented in (Viegas and Lu, 2004) by introducing **bus lanes with intermittent priority** (BLIPs), a variant of IBLs in which traffic is forced out of the lane reserved for the bus with variable message signs (VMS). BLIPs do not require changes to the signal settings, and they can be combined with TSP. A BLIP is essentially a set of rolling spatial cocoons (bus-lane sections) in which buses travel to the exclusion of other traffic. Each cocoon starts at the rear bumper of its bus and extends a fixed distance ahead. This zone is kept clear of non-bus traffic to ensure that the bus does not experience any delay. For practical reasons, the exclusion zone is assumed not to travel continuously along the roadway, but to advance discretely one block at a time. VMSs, possibly combined with in-pavement lights, would announce the changes.

- **Yield-to-Bus Law** is the regulation that requires motorists to yield to buses that are reentering the traffic from the bus stop. Similar to the boarding islands and curb extensions method, these laws also aim to reduce the clearance time of the bus. However, car users do not always give priority to buses, as they should.

- **Parking restrictions** are also very important for the bus priority operations (TCQSM 2nd Eds, 2003). Parking restriction is mainly applied near a curbside bus stop where the transit agency has to free up some space for the bus to pull out of the traffic and up to the curb for dwelling. The inspection on car irregular parking is also extremely important. Among the transit network operated by Carris, the coordination with EMEL (the municipal company for mobility and parking in Lisbon) and the municipal police is essential for their quick response.

Tactical and real time operations level**:** besides strategic planning, there are other decisions that should be taken on the tactical level, as well as on real time operations:

- **Reserving drivers and vehicles** to avoid missed trips or avert large schedule deviations: the absence of a scheduled driver or vehicle creates a gap in service, which can significantly decrease service quality. Extraboard personnel cover for drivers who are on vacation, who call in sick or who are absent without leave while a reserve fleet of vehicles is maintained to replace a missing bus resulting from vehicle breakdowns, accidents or other emergencies. Standby buses can also be used to service a route where peak demand asks for extra capacity. These on-demand buses help avoid overloads and schedule deviations that can propagate unreliability throughout the route. The number of reserve drivers and vehicles is important but there must be a trade-off between the service quality improvement and the extra cost of the reserve policy and the underutilization of resources.
- **Drivers training** and implementing incentives/penalties for drivers help reducing the variability on drivers' work performance and its impact on the service.
- **Service supervision** serves as a strategy to both prevent and correct service reliability problems. With the introduction of automated vehicle location systems, RTI allows supervisors to better monitor bus operations and to better evaluate the route and take control decisions.
- **Adjusting schedules** to reflect actual conditions also serves as a strategy to maintain reliability. The common practice among bus operators is to use the 85th or 90th percentile of trip travel time distribution when constructing vehicle schedules, according to (Cats et al.,2011). However, passenger demand and traffic conditions will gradually change over time, causing the running time distributions to deviate from those used during schedule planning when vehicles and drivers were assigned. Inserting extra time, or slack, into a schedule as a buffer to prevent the propagation of disturbances is one of the most widely used strategies to increase reliability. The slack is calculated so that buses can make up time lost due to random recurrent travel disruptions between control points, but slack reduces the commercial speed of the buses, so it cannot be too big. To limit the total amount of bus delay, control points are spaced widely so that typical routes include only a few (Cats, 2014).

2.4.5. Corrective measures

Moving on to the corrective strategies, the following will be reviewed: holding, expressing, short-turning and deadheading.

- **Holding** is the control strategy of purposely delaying a bus at a time point (or control point) for a certain amount of time. It aims at correcting early buses or at preventing buses from bunching. Holding can be schedule-based (to ensure on-time performance), or headway-based (to maintain even headways between consecutive buses). Holding strategies can also be preventive rather than corrective and they are the ones that have been most widely used, therefore they will be treated in detail in section 2.5.1.

- Expressing or stop skipping involves sending a bus to a stop further downstream and skipping (not serving) some, or all, intermediate stops. The objective of this strategy may be either to split bunched buses or to close a service gap further downstream, both in an attempt to balance headways and improve service past the end of the express segment. One can distinguish three types of expressing: *1) full expressing*, in which the bus is expressed to a point downstream without serving any intermediate stops; *2) limited stops*, when the expressed bus serves only a limited number of intermediate stops; *3) alighting-only*, in which the expressed bus does not pick up any additional passengers and only drops off at intermediate stops.

An expressing strategy involves selecting the intermediate stops that are skipped or served with limited service, considering the ability of a bus to express over a segment and pass regular-service buses, and informing passengers of the change in service. Decision-makers must weigh the benefits to downstream passengers and to the overall performance of the route, considering the negative impacts on passengers at intermediate stops and those on-board affected by the extra transfer.

Again, RTI enables decision-makers to better evaluate the trade-offs regarding the decision to express a bus. The location and load data from automated systems provide more detailed information to help supervisors decide whether expressing is appropriate, which of the three expressing strategies to apply, which bus is the most fitting to express and how long should the express segment be. Archived data may also support decisions by studying the impacts of previous expressing decisions and assessing passenger demand estimates through load profiles.

- Short-turning involves instructing a bus to end its current trip before it reaches the terminal, and serve the route in the other direction. This strategy is employed to return a late bus to schedule, or when extra service is needed in the opposite direction, whether it is due to higher passenger demand or large gaps in service. Route design, street layout and the vehicles themselves may limit the number of control points at which buses are able to switch directions.

- **Deadheading** involves pulling a bus from service and running it empty for a segment of the route. Deadheading is one of the most employed control strategies and if applied appropriately, it can be effective in reducing service irregularity (Eberlein et al. 1998). It is similar to expressing except that the vehicle runs without any passengers on-board.

2.5. Modeling / Proactive measures

In this section, I will look at the different modeling techniques that were developed throughout the years to study the bus operation environment concerning service reliability and bus bunching. I will start by reviewing the commonly applied holding techniques, focusing on the most recent analytical models. Then, I will review the most up-to-date methods to predict the occurrences of bus bunching. Furthermore, I will examine the work on simulation and control methods and address their drawbacks when compared to the analytical models.

2.5.1. Holding strategies and control models

Holding strategies are among the most widely used transit control methods aimed at improving service regularity. The implementation of holding strategies involves two key design decisions: selecting the set of time/control point stops and the holding criteria (Cats et al., 2011).

As for the holding criteria, holding strategies are commonly classified into two categories: schedule-based and headway-based strategies (van Oort et al., 2010).

A **schedule-based holding** strategy defines the earliest time that a bus can depart from a time point stop relative to the schedule. Scheduled-based holding is more appropriate for low-frequency routes (Cats et al., 2011) and is also appropriate for routes with important transfer connections: by holding a bus until its scheduled departure time, passengers are less likely to miss connections because the second bus departed earlier than scheduled. If schedules are slack, a large percent of the buses will spend too much time idling, as they wait for their scheduled departure time, and passengers will get frustrated with the increase in travel time. If schedules are tight, buses will be running late most of the time and the holding strategy will be ineffective in achieving on-time performance or maintaining even headways.

On the other hand, **headway-based holding** strategies use headways between consecutive vehicles as their criterion for regulating departure times from time point stops. These strategies require real-time AVL information and they can take into account only the headway from the preceding vehicle or both the headway from the preceding and to the succeeding bus. When headway-based holding is applied, the bus in front continues service down the route and moves away from the held bus, while the bus behind also continues regular service and moves closer to the held bus, evening out the space (and time) between consecutive buses when applied to buses with short headways. This prevents the occurrence of bus bunching and helps balance passenger loads. As a result, headway-based holding is more suitable for high frequency routes, where bus bunching is more likely to occur.

Research into holding strategies started with single control point models. Hickman (2001) provides an extensive literary review of these models. The investigation on the topic started with the analytical model of Osuna and Newell (1972), with further refinements and extensions by Newell (1974), Barnett (1974), Barnett and Kleitman (1978), and Barnett (1978). These models assume fairly simple transit networks, typically some kind of shuttle service or a simple service loop, with a limited number of vehicles (one or two) and they include some probability distribution of vehicle running times for the service. Using this framework, these models derive an optimal value of a threshold, corresponding to a threshold-based holding policy. Because of the limitations on the route structures, these models are generally not directly applicable to a more typical fixed-route transit service.

The analytical work of Turnquist and Blume (1980) examines conditions under which a threshold model is likely to produce benefits by reducing passenger waiting and travel time. Using an analytic model, the research found that the benefits of holding are higher when the headway coefficient of variation $COV(h)$) is considerable and/or when the ratio of on-board passengers to expected downstream passengers is low. They note that this is somewhat conflicting: $COV(h)$) tends to increase along the route, while the ratio of on-board to downstream waiting passengers tends to decrease. Hence, the choice of a holding station is not entirely obvious. The research concludes that control should be enacted at a stop where there is already substantial variation in headways, the vehicle load is light, and the number of downstream passengers is significant. These results on the selection of control sites were echoed in subsequent studies by Turnquist (1981).

A much larger literature on transit vehicle holding has developed using heuristics and Monte Carlo simulation. This approach is appealing because vehicle operations on a transit route are inherently stochastic and hence difficult to describe analytically (Xuan et al., 2011). Hickman (2001) provides a detailed review of the authors that focused on these route-level simulations in the 1980s. These models typically assume a random pattern of passenger arrivals at each stop (i.e., Poisson arrivals) and a binomial distribution for the number of alighting passengers at each stop. Generally, these studies have concluded that headway-based holding is more effective than schedule-based holding, while both methods offer substantial improvements over no holding at all (Abkowitz et al., 1986).

More recent studies present more general analytic models for the transit vehicle holding problem, formulating the vehicle holding problem as a quadratic programming problem to solve for the optimal set of vehicle headways at a dispatch point. Eberlein (1995) provides an iterative heuristic to determine the optimal set of headways, while the models of Furth (1995) and O'Dell and Wilson (1999) solve these problems using nonlinear optimization methods. Shen and Wilson's (2000) model includes integer variables to accommodate short-turning and expressing of vehicles; their solution method uses a piecewise linear objective and linear constraints within a branch-and-bound technique. In each of these formulations, the transit service model includes more realistic characteristics of fixed-route service, such as minimum acceptable vehicle spacing and the dynamics of vehicle running and dwell times along a transit route.

Optimization models with the objective to minimize the passenger waiting time are quite popular, even though this type of model is generally difficult to solve for complex problems (Xuan et al. 2011). Some studies used heuristics with less restrictive assumptions that may not render near-optimal but fast enough solutions for real-time application while others make stronger assumptions. Boarding while holding has been ignored in optimization modeling. Otherwise, the problem would become non-convex and with no closed form solution. While that simplification has greatly reduced the complexity of the bus holding problem, the literature often possesses an inconsistency on how to treat holding time in relation to dwell time, having an impact on the estimation of the number of on-board passengers and the non-consideration of waiting time savings (Chen et al., 2013).

The most recent models also include vehicle capacity constraints, although this requires integer variables in the formulation. The drawback to these analytic models, however, is that they use deterministic running and dwell times - there are no stochastic service attributes in these models.

Hickman (2001) provides a first step for greater realism by presenting an analytic model that determines the optimal vehicle holding time at a control stop that explicitly accounts for route-level operational dynamics, for stochastic travel times and passenger boarding and alighting times and for real-time knowledge of vehicle status. As it is formulated, the single vehicle holding problem is a convex quadratic program in a single variable, and is easily solved using gradient or line search techniques. However, it is important to keep in mind that although this holding model is useful for decision support, the model presented by Hickman (2001) has several critical assumptions. One of them is the scope of the holding decision: in only solving for a single vehicle, the model is necessarily "myopic" in that it optimizes only a single headway, independently of other control actions, instead of solving for all subsequent holds (i.e., all subsequent vehicle headways). Besides that, vehicle capacity is assumed to be infinite and consecutive vehicle running times are considered independent.

A somewhat separate literature formulates the transit vehicle holding and dispatching problem as an optimal control problem, using techniques such as half-wave and shifted linear rectifiers, polynomial rectifiers, and step function controls (Hickman (2001) develops on these topics). While the modeling construct is different, the resulting solution methods and properties are similar to the optimization-based methods cited above.

Unfortunately, single-point control cannot succeed for long routes with frequent service and a strong bunching tendency (Daganzo, 2009). In this case one needs to figure out how many control points to provide, examining the whole route as a system. Analytical models that partly address this question have been provided in Daganzo (1997) and Zhao et al. (2006). The former assumes that just enough slack is added at the control points to avoid with a high probability the occurrence of bunching before the next control point, and then examines the system-wide trade-off between in-vehicle riding time and waiting time as one varies the number of control points. The latter reference focuses on waiting times and ignores the commercial speed issue.

As pointed out in Newell (1974), even considerable amounts of schedule slack cannot guarantee on-time performance all the time because a disruption large enough to spread system-wide can never be ruled out, such as a mechanical malfunction. Generalized disruptions, such as heavy traffic or adverse weather conditions, are even more problematic because they can push large groups of vehicles behind schedule. If this happens, the delayed vehicles would try to go as fast as they can to catch up with the schedule and in doing so would possibly tend to bunch again, disrupting themselves and the rest of the system. In this case, the schedule loses its utility (Daganzo, 2009).

In order to alleviate these problems and improve resiliency, Daganzo (2009) proposes using dynamic holding times based on RTI on headways instead of the schedule. They also propose introducing enough control points to ensure that headways can be kept close to the ideal everywhere along the route, thus reducing the need for large corrective actions. The proposed model recognizes that vehicle travel times, stop times and passenger demand are random.

This dynamic headway-based problem is considerably more complex than the schedule-based problem. Prior studies such as the one conducted by Eberlein et al. (2001) have addressed the inclusion of the available RTI but only in an idealized case with known deterministic demand and bus travel times, not exploring the consequences of the chosen control points on the commercial speed and not considering random variations explicitly.

Daganzo (2009) proves that an adaptive control strategy can produce regular headways without significantly slowing buses, using an idealized model. To demonstrate this potential analytically, Daganzo (2009) shows that the model requires less holding time than schedule-based control to keep headways within any desired tolerance from a target, thus reducing in-vehicle time. In addition, the model takes into account random effects, such as the random variations in bus travel time, bus dwell time and passenger demand, making it resemble more realistic, real-life situations.

Daganzo (2009) defines the schedule in terms of bus arrivals to a series of control points and different bus runs are modeled by the dynamic equations of travel time between different control points. Even though this is one of the most up-to-date analytical models to study bus motion, the framework is still simplified: in reality, random disturbances due to traffic, passenger demand, passenger needs and the behavior of bus drivers give rise to errors between the actual bus arrival times and those scheduled. As a result, actual headways also differ from the target, and this affects the travel times since longer headways imply more passengers to be served - the author roughly assumes that the resulting average uncontrolled time is approximately linear. The analytical equations are also incapable of incorporating situations such as the boarding of passengers arriving at the stop while the bus is stopped and other assumptions are made, such as the average marginal delay per boarding move.

In order to model the BB phenomenon, the author states that headways increasingly deviate from the target as time passes until buses bunch up, because there are forces that attract the pair of buses on opposite sides if the headway is shorter than the planned headway and repel them when the headway is longer, which is in line with previous studies. His strategy to prevent BB from happening is introducing a compensating force that would attract buses when they are too far and repel them when they are too close. The simplest policy of this type would act only on the following bus of each pair, speeding it when it lags and retarding it when it closes. The authors propose adding a term to the dynamic equations: a headway-dependent delay to the time that the bus on each run would otherwise spend traveling uncontrolled between two stops. An appealing feature of the proposed strategy is that it increases bus speed at the first hint of an unanticipated long headway, before the problem grows to unmanageable proportions. In so doing, it acts as a robust mechanism that compensates for the type of recurrent disruptions that tend to spread with schedule-based control (Daganzo, 2009).

However, the introduced added delay term may sometimes imply that the bus would drive faster than it actually can. Therefore, Daganzo (2009) proposes asking bus drivers to speed up by not picking up passenger until the added delay becomes feasible again, which may be difficult to implement in reality and frustrating from the passengers' perspective, even though we are only dealing with systems operated with small headways. Thus, even though the results are encouraging, the application of this control theory based model in a real situation seems quite difficult.

(Xuan et al., 2011) also builds on the model of Daganzo (2009) to develop a new formulation for the bus motion and to extend the control law allowing not only the maintenance of regular headways but also the compliance of the schedule, which adds value on the application of these methods to bus lines with both short and long headways.

Bartholdi and Eisenstein (2012) propose a model that not only abandons the notion of schedule but also the attempt to follow a pre-specified target headway as presented in (Daganzo, 2009). By doing so, the system is free to express the natural headway, which may change over time. The authors argue that the method will converge to the smallest common headway possible given the current conditions on the bus line and that, even after a large disruption of service, the system will spontaneously reposition buses to achieve a new, yet necessarily larger, common headway.

The approach of Bartholdi and Eisenstein (2012) is close to the work of Daganzo (2009), in the way that its data requirements are minimal and it simply corrects headways of buses departing from control points. The scheme makes use of AVL, but as little as possible, due to the practical limitations of current GPS (for example: inaccuracy during heavy rain or amongst tall buildings). Also, the model is independent of demand - the scheme makes no assumptions on passenger arrival rates or utilities for waiting, and therefore has no need to measure them. The authors argue that rather than trying to exercise more control over bus drivers, the presented scheme exercises less, leaving them free to focus on driving. The main goal is simply to reduce both the mean headway and the variation among headways (rather than to achieve a pre-specified target).

The authors indicate that under this self-equalizing headways scheme, the positions of buses are collected every 15 seconds and the equilibrium for the system is computed, which will influence the holding time for each bus at each stop. After departing from each stop, drivers are asked to drive with the flow of traffic, reducing the work for management on building schedules or target headways. Also, since the model is built to respond to changes in the system, operators can also implement or test actions (for example, a new process for improved boarding or a forced temporary re-rerouting) and the headway will consequently adapt to those conditions.

The work of Liang et al. (2016) develops on the self-coordinating bus route model proposed by Bartholdi and Eisenstein (2012), establishing new dynamic equations which are said to make the headways self-equalize quicker (the converging rate guarantees that a bus passes through less than eight control point until the headways settled to the optimum value).

Daganzo and Pilachowski (2011) also build on the adaptive control model of Daganzo (2009) and focus on the equalization of headways, proposing a cooperative control method in which bus speed is regulated in real-time. The new model adds on value to the previous one in the way it deals with large disturbances: it proposes a two-way-looking strategy based on the spacing in the front and back of each bus. This bus-to-bus cooperation allows a bus to slow down due to a disturbance that happened behind it, minimizing the overall passenger waiting time. One disadvantage of the model presented in this study is the fact that the speed adjustments are dependent on estimates of passenger demand in each stop, besides the spacing between buses.

Hernández et al. (2015) developed an optimization and simulation model that is pioneer in considering multiple bus lines (managed by different operators) in a corridor. In the experiment carried out with 2 bus lines, as expected, the central operator scenario is the one that brings the best level of service in contrast with the scenarios in which each operator works individually, without considering its competitor's offer. Unfortunately, the model is incapable of assessing the impact of the length and proportion of the common sections to different operators on the results. Besides that, the model is composed by a very large set of cumbersome equations, which makes it not intuitive.

On the other hand, Schmöcker et al. (2016) test the effect of different proportions of common lines on BB. The authors formulated discrete state equations to obtain bus trajectories, introducing in particular the possibility of overtaking among the buses. However, the approach is limited in a way that the scenario tests all assumed that there are no delays except for one initial delay to one of the bus lines. Thus, the model only focuses on the secondary effects that a delay in one bus has on the other buses that use the same corridor, disregarding other random delays occurring to buses between stops (traffic lights, congestion, etc.) or at stops (e.g. passengers requiring additional time for cash handling).

To sum up, the first analytical models started by looking at single control holding methods. Then, new models included several control points along a route rather than a unique control point. After that, it was understood that in the context of high-frequency routes, headway-based control was more effective then scheduled-based control and instead of focusing on a certain threshold for the variation of the headways, systems could also be "let free" to have the best headway that would mirror the system's state. For instance, if a bus would have a mechanical breakdown, the rest of the chain would adapt to that expressing a new wider common headway between all the other buses. The introduction of board computers equipped with GPS and wireless communication devices was fundamental on the development of the research, allowing the transit agency to monitor its bus routes continuously and also making the bus-to-bus cooperation possible. The objective functions of the reviewed models include one or more of these components:

- Waiting time the passengers experience at the bus stops;
- Extra waiting time of passengers who are prevented from boarding a bus because it is at full capacity;
- In-vehicle waiting time for passengers aboard a bus that is being held at a stop;
- Extra waiting time of passengers who are obliged to wait for another bus because the bus skipped their stop;
- Variance in the headway of the buses.

2.5.2. Bus bunching prediction models

As aforementioned in section 2.3.5, transit authorities have developed corrective measures to overcome the negative effects of bus bunching. However, it was also seen that these corrective methods are reactive and may dissatisfy the passengers waiting at stops or sitting in the bus (for example stop-skipping). Adopting proactive strategies to monitor irregular headways and alert possible bus bunching in advance to prevent transit service breakdown is therefore preferable and that is why some researchers have more recently focused on models that predict the occurrence of bus bunching through irregular headway fluctuation.

(Haiyang Yu et al., 2016) provides a review on several complex algorithms and the data they are based on that were developed to predict bus arrival time, which are not limited to the applications on BB studies. Among these models, there are forecasting models based on historical data, time series models, artificial neural network (ANN) models, support vector machine models and regression prediction models. Bin et al. (2006) and Kieu et al. (2012) also resume some of the existing literature on bus arrival time prediction models within the context of intelligent transportation systems (ITS) and they outline their advantages and drawbacks in such advanced systems.

Predicting bus bunching is a more challenging task than predicting single-bus arrival times because more than one bus is involved in BB. Both the dwell times and arrival times of different buses at different stops fluctuate and lead to a highly stochastic process. In this section, the focus will be on the most up-to-date models that focus on the BB phenomenon.

Predicting BB using AVL data: Moreira-Matias et al. (2012) explore time-series of headways based on AVL information to find consistent patterns of frequent headway deviation (HD) events occurring in the same bus stops whenever a BB occurs. Their goal was to point out a route region where an HD event systematically propagates itself along the route, forming a Bunching Black Spot (BBS). The BBS could be specific of a period of the day or continuous along the day. The authors ran a sequence-mining algorithm to explore such data and one of the main finding of this article was that headway deviations in the beginning of the route would propagate faster along the line and would have a bigger impact on the whole system than deviations occurring further along the route, even though there was a high correlation between deviations in one stop and the deviations in the following stops.

Moreira-Matias et al. (2016) present a systematic proactive approach to eliminate BB in real-time: the authors explore online learning techniques to address headway instability by simultaneously considering historical and real-time data concerning service perturbations. They then develop and apply a comprehensive procedure for BB event detection and corrective action deployment.

As already mentioned, most of the research on BB departs from the assumption that the probability of BB events is minimized by maximizing headway stability, which requires multiple corrective control actions which may impose high mental workload for both drivers and control center staff, yielding results with sub-optimal results and low compliance rates. Hereby, the authors propose a proactive rather than a reactive operational control framework, in which the fundamental idea is to estimate the likelihood of occurrence of a BB event further downstream.

The presented modeling framework constitutes a combination of machine learning (ML) methods to extract valuable information from massive sources of continuous data (such as AVL).

The occurrence of BB is triggered by stochastic processes and is thus difficult to predict. Notwithstanding, current system states may allow uncovering such future occurrences, for instance, behavioral patterns such as consecutive headway reductions and travel times longer than expected. To do so, the authors state that it is not sufficient to mine historical AVL data, as there is no obvious trend or a simple static association rule that can explain such events. In (Moreira-Matias et al., 2012) it is postulated that a BB event is usually preceded by a headway deviation prevailing further upstream along the route. However, such a rule cannot handle the random series of events that may affect a given bus trip which can arise sporadically rather than systemically.

Therefore, the authors introduce a stepwise learning methodology to detect and then prevent the BB phenomenon from propagating in real-time. It utilizes simultaneously historical and real-time AVL data. The framework works on two different parts: (I) BB event detection and (II) corrective action deployment. The BB event detection comprises an advanced ML framework that builds on three steps:

1) An offline regression method, which is used to predict link travel times (LTT) for every trip in the following day (the forecasting horizon) using some of the most recent days (the learning period) to train the model;

2) Prediction refinement (i.e. online learning), using trip-level information as well as stop-based information. Both steps are based on the perceptron delta rule by reusing each prediction's residuals to improve successive predictions.

3) Estimation of the likelihood of a BB event to occur at downstream stops, by assuming that headway is normally distributed. Given a certain user-defined threshold, a BB detection alarm is launched.

The authors devise recursive relationships between the arrival times in consecutive stops, considering the dwell times, the riding time between the stops and then define the link travel times (LTT) and compute the headways for consecutive buses and explain how they infer the future values of the headways based on the predictions of the future values of the LTT. The details can be explored in (Moreira-Matias et al., 2016).

The authors formulated the LTT prediction problem as an inductive learning regression problem, combining offline learning and online learning models and using one of the most well known offline learning techniques for regression, Artificial Neural Networks (ANN).

Moreira-Matias et al. (2014) also presents a very similar probabilistic framework to predict BB occurrences in real-time. In both articles, the learning task is performed incrementally by employing the Delta Rule (DR) in two steps: the trip-based refinement, in which the residuals between the predicted link travel times and the real travel times are computed and used; the stop-based refinement, in which the headway residuals from the offline prediction are used to refine the online prediction of the headways in the following stops.

For the third and final step of the model presented in (Moreira-Matias et al., 2016), the bus bunching detection, a probabilistic framework for detecting a BB event at downstream stops is proposed. The likelihood of a BB event to occur at any of the downstream stops between two consecutive trips is computed by inferring the short-term probability distribution function (p.d.f.) of their headways at each stop. This definition allows quantifying the p-value of a BB event to occur at a certain stop. It is then possible to quantify BB likelihood for all downstream stops (and also to update them each time a more up-to-date headway value is obtained).

With such versatile structures and adaptive learning processes, ANNs appear to be a promising approach to describe complex systems where various time and location dependent factors are interrelated. However, it has been commonly reported that ANN models require a large amount of training data to estimate the distribution of input pattern and they have difficulty generalizing the results because of their overfitting nature. In addition, it fully depends on researchers' experience or knowledge to preprocess data in order to select control parameters including relevant input variables, hidden layer size, learning rate, and momentum (Bin et al., 2006). These are in fact the main disadvantages of these ML methods.

Predicting BB using APC data: while Moreira-Matias et al. (2014 and 2016) base their BB prediction models in AVL information, Haiyang Yu et al (2016) present a predictive framework to capture the stop-level headway irregularity based on transit smart card data (AFC or APC). The study aims at leveraging transit smart card data into bus bunching detection by forecasting stop-level headway in varying time horizons. By learning the irregularity of headways from the historical smart card data, the authors will also provide a real-time headway forecasting system for passengers to know where the next bus will arrive. This study mainly focuses on predicting BB in Beijing bus routes, where both boarding and alighting demands for each stops can be estimated. The same doesn't apply to Carris operation scene, where passengers are only required to tap their cards when boarding. The authors assumed that the arrival time at each stop would be approximated by the tapping time of the first boarding passenger on that stop and used this definition for the calculation of the headways between consecutive bus runs.

For the prediction of the occurrence of BB, their methodology involved three steps:

1) The influential factors that are relevant to headway variability are identified and calculated from smart card data (the variable selection);

2) The LS-SVM (least square support vector machine) algorithm is utilized to forecast the bus headways for the next stop on the basis of historical data generated from the first step;

3) Finally, a relationship between predicted headways and bus bunching occurrence is established to detect the irregular fluctuation of headways.

The authors use a LS-SVM algorithm to capture the nonlinear and stochastic nature of bus headway. Different from the traditional SVM algorithm that solves a convex quadratic programming problem, the LS-SVM algorithm solves a set of linear equations to find the global optima, thus demonstrating high efficiency and low computational efforts for large-scale datasets, according to the authors. The mathematical approach is further detailed in the article.

SVM is a very specific type of learning algorithm characterized by the capacity control of the decision function, the use of the kernel functions, and the sparsity of the solution. SVMs are based on the structural risk minimization (SRM) inductive principle, which seeks to minimize an upper bound of the generalization error consisting of the sum of the training error and a confidence level. Unlike ML methods, SVMs are shown to be very resistant to the overfitting problem, eventually achieving high generalization performance in solving various time series forecasting problems. Another key property of SVM is that the solution of SVM is always unique and globally optimal, unlike other networks' training, which requires nonlinear optimization with the danger of getting stuck into local minima (Bin et al., 2006). Another advantage of SVMs is that the solution to the problem is only dependent on a subset of training data points which are referred to as support vectors. Using only support vectors, the same solution can be obtained as using all the training data points.

On the first step - variable selection - bearing in mind the importance of the variations on passenger demand, four candidate variables were identified: the number of boarding and alighting passengers for two consecutive bus runs. Then, two other variables were considered to take into consideration the effects of the traffic conditions and driver behavior: the headway in the previous stop and the link travel time between on the previous run. The input dimensions and the input-output relation are detailed in Moreira-Matias et al. (2014). In the last step, BB is detected according to the same criterion applied in (Moreira-Matias et al., 2016), i.e., if the predicted headway at any given stop is shorter than a quarter of the planned headway.

As part of the algorithm evaluation and the analysis of the results, the authors compared the prediction performance of the LS-SVM algorithm on both headway estimation and BB classification with four well-established algorithms: ANNs, K-nearest neighbor (KNN), Random Forest (RF), and Gaussian Process Regression (GPR). The chosen algorithm seemed to outperform the others in the specified performance metrics for the studied routes.

As aforementioned, this model used passenger demand as one of the influential factors for BB and four variables with the boarding and alighting information are inputted to the LS-SVM regression to forecast the headway deviation at the next stop. In the case of Carris, if the same model were to be tested, the two variables considering the alighting information would have to be disregarded, due to the lack of records. This could have a significant impact on the predictive power of the model. Another drawback of LS-SVM regression is the limited ability to handle outliers and noise, which may be very important if the available data for validation and calibration of the model is limited.

2.5.3. Simulation models and field experiments

Simulation models have also been established as a primary tool for the evaluation of transit performance at the operational level. Transit simulations provide a dynamic perspective on transit operations by enabling the comparison of various scenarios and the representation of complex interactions between the network components: general traffic, transit vehicles, and passengers (Cats et al., 2010).

However, the capability of simulation models to effectively simulate advanced public transport systems (APTS) applications in large networks may be limited. While microscopic models simulate well the local impacts of APTS, these models are inefficient when applied to large-scale applications because of the unnecessary level of detail and extensive computational effort they require. In contrast, mesoscopic simulation models, which represent individual vehicles but avoid detailed modeling of their second-by-second movement, may be as useful for system-wide evaluation of transit operations and APTS as they are for general traffic (combining the advantages of microscopic and macroscopic models). Examples of potential applications include frequency determination, evaluation of real-time control strategies for schedule maintenance, restoration from major disruptions, and assessing the effects of vehicle scheduling on the level of service.

In this section, the most recent simulation models that deal with BB will be revisited and an explanation will be given on why the work in this dissertation doesn't follow a simulation approach.

Cats et al. (2011) use BusMezzo, a transit simulation model developed on the platform of Mezzo (Cats et al., 2010), a mesoscopic traffic simulation model, to evaluate nine scenarios of holding strategies for a high frequency bus line in Stockholm, Sweden, characterized by articulated vehicles, a high level of signal priority, and RTI on arrivals at stops. The analysis of the results considers the implications of holding strategies from both passenger and operator's perspectives. The main service measures of performance utilized were the headway coefficient of variation ($COV(h)$), the average waiting time per passenger, the percentage of on-time arrivals, the percentage of bunched vehicles and the regularity LOS (both according to the definition presented in TCQSM 2nd Eds., 2003). This study also mentions the importance of the value-of-time in the trade-off analysis between passenger in-vehicle delay and the waiting time. The ratio between waiting time and in-vehicle time is in the range of 1.5 to 2.0, according to (Cats et al., 2011). Other authors have confirmed this, sometimes reporting even higher disutility associated with waiting (Wardman, 2004).

The simulation results highlighted substantial potential benefits from implementing an even headway-based strategy, in terms of service reliability improvement, passenger time-savings and reduced operating costs, as well as better schedule adherence at the relief points.

Following the simulation results, Cats (2014) tested the feasibility and implications of the proposed real-time control strategy in a field experiment. A real-time indicator of bus position with respect to the preceding and following buses was displayed on a computer screen that is located in the driver cabin on each bus vehicle. The indicator was constantly displayed and drivers were encouraged to adjust their speed as much as possible according to that indicator and they could decide to hold at stops whenever needed. The field experiments demonstrated benefits in terms of passengers waiting time, capacity utilization and fleet operations: the ($COV(h)$) decreased for the peak period as well as for the entire day along the entire line for both directions by 11–26%; the share of bunched buses decreased by 13–24%. This resulted with a decrease of 38% in the excess waiting time compared to before.

The regularity improvements were achieved without negative implications on vehicle and passenger travel times. Total dwell time along the line decreased by 10% and dwell times were much more evenly spread along the line as each trip involved holding at different stops.

The new control method is based on a decentralized control system, where buses regulate their progress continuously based on the headways from the preceding bus and the following bus, in a cooperative operation as the decisions of each driver are interdependent on other drivers' decisions, allowing for mutual corrections. The headway-based indicator was embedded into the already existing computer display that is located in the driver cabin, supplemented with a proactive fleet management approach by the control center. It should be stressed that the performance and LOS were improved solely by introducing a new control strategy that utilizes the existing equipment, without an investment in infrastructure or technology.

More even passenger loads were indeed obtained when measured through variations in observed loads from automatic passenger counts (APC) data. Hence, there is empirical evidence that a more regular service yields a more even distribution of passengers over buses (Cats, 2014). A more regular service implies therefore that passengers wait for a shorter time and have better on-board conditions. This highlights substantial potential benefits from implementing a decentralized regularity-driven control strategy, which is in line with the results of an analytical study by Daganzo (2009). Besides passenger time savings, a reduction in the production costs as well as better schedule adherence at driver relief point were observed.

The authors also looked into the total travel time distributions under the various holding strategies and into the schedule adherence distribution at driver relief points - some bus operators use driver schedules that include driver replacement at intermediate stops, also known as relief points. The presence of relief points along the line is an additional concern as it is especially important to have high schedule adherence at these stops and that might represent a barrier for the application of holding strategies.

In terms of the number and location of control points, Cats et al. (2014) expand on previous studies by considering the general case of determining both the optimal number and optimal location of the stops where holding takes place, and assessing their impacts on transit performance using Mezzo for the same case study.

A shift from punctuality-focused to regularity-driven service requires the consideration of a series of measures along the service chain in order to facilitate a long-term implementation. Regularity-driven operations imply the removal of as many schedule constraints as possible from the operational routine (turnaround stops, regulation stops and driver relief stops).

In addition to its implications on operational planning, a shift to regularity-driven operations scheme should also be accompanied by consistent real-time fleet management strategies. The common practice among operators is to deploy several reserved buses that are allocated dynamically upon request. These buses should be positioned in proximity to locations where service regularity tends to deteriorate, such as just before a sequence of high-demand stops. Those are also the locations where increased bus capacity is required. An efficient dynamic allocation of buses requires proactive dispatchers at the control center and could be enhanced by a decision support tool for fleet management interventions.

In (Zimmermann et al., 2016), two classes of control methods (predictive control and feedback control) are revisited in a microsimulation environment applied to a Bus Rapid Transit (BRT) corridor in order to point out that allowing greater variance in headways between consecutive buses leads to possible gains in total delay as compared to a prescribed headway-based holding scenario, which is in line with the findings of Bartholdi and Eisenstein (2012). The mathematical programming approach seems rather complex and every control scheme follows a significant number of assumptions. Furthermore, the tests are made in limited scenarios, due to the required computational effort, which represents one of the biggest drawbacks of simulation models.

Pilachowski (2009) also develops a microscopic simulation tool to validate a model of the BB phenomenon: his work presents a continuum approximation framework as a tool to systematically analyze the behavior of the system with discrete bus stops, using GPS data to directly counteract the causes of the bunching by allowing the buses to cooperate with each other and determine their speed based on relative position. The continuous model provided an upper bound for the BB situations, as expected because of the numerous conservative assumptions (for instance, uniform service and demand in space and time). This is an idealized scenario, very far from the reality, highlighting once again the limitations on the applications of simulation models.

2.6. Trends and gaps in the literature

Different measures to assess reliability (and specifically regularity in the context of high-frequency transit services) have been established throughout the years. Also, the strategies to improve the performance of the system in terms of reliability (corrective, preventive or proactive) have been deeply explored for more than 50 years. In the context of proactive measures, the availability of RTI is extremely important since it enables a quicker evaluation of system states and the deployment of control measures in real-time.

Bus bunching is also a strongly studied topic in the context of public transit systems, given its recognized importance in the provided level of service. Several thresholds have been deployed to recognize bus bunching but the focus in headway instability is common to all the reviewed articles. The causes and consequences of the BB phenomenon were also examined throughout the years and the preventive measures in the strategic, tactical and operational levels have been thoroughly investigated. In the scope of corrective and proactive measures, holding strategies are amongst the most widely control models used by transit operators.

The trend in the research in the control models is to disregard schedule and move towards a headway-based operation scheme, aiming at reducing the variation among headways and minimizing the passenger waiting time. The bus-to-bus cooperation and the real-time communication with the control centers allow for more intelligent and integrated transportation systems. Simulation models also played an important role on the evaluation of transit performance and on the validation of models of the BB phenomenon.

In the context of BB prediction models, machine learning models are the most up-to-date ones. However, it is recognized that predicting dwells times and arrival times of several buses is a highly stochastic process, much more complex than the estimation of the arrival times of individual vehicles. Also, the predictive power of the models is very dependent on the selected input variables and on the learning rate, which are based on the experience of the developer of the model. Besides that, the over fitting nature of these models and the difficulty to transpose the results to other contexts and assumptions is also an issue.

The approaches to predict the occurrence of BB are data-driven, whether by AVL or APC databases, which rely on GPS systems or transit smart card data, respectively. The sooner a bus bunching is detected, the more effective is the corresponding reactive control strategies (Haiyang Yu et al., 2016). By providing early warning information on when and where the bus bunching may occur, transit authorities can employ proactive countermeasures to mitigate bus bunching, and therefore improve the quality of transit service for operating cost savings and increasing ridership. From the perspective of passengers, knowing when the next bus will arrive will alleviate their anxiety when waiting at a bus stop and help them choose a less crowded vehicle.

One of the aspects that can be perhaps improved in the current models is the fact that different data sources are usually not integrated (models that include both AVL and APC information are still to be developed). By feeding the control center with RTI on both data types, a much more integrated system can be accomplished, and more reliable dwell times and arrival times can be generated. Also, bus driver behavior and overcrowding are not extensively incorporated in the current prediction models, given their complexity. Besides, the models that contemplate the effect of other bus routes using the same lines are still scarce.

3. Bus bunching detection

3.1. Methodology

In this section, I will briefly explain the framework that I followed in order to detect bus bunching in the transit service of Carris (figure 4). Firstly, I analyzed the reliability indicators used by Carris (comparing the $offered\ km{\times}vehicles$ with the $lost\ km{\times}vehicles$). Then, I selected the bus line that I would focus on and proceeded with the data collection (AVL information; the location of stops, among other) in cooperation with Carris. After the data was made available, I used the Mathematica software to process all the information, finally leading to a bus bunching detection method for different days and periods of the day, which represents the end of section 3. In chapter 4, I proceed with the analysis of the bus bunching events.

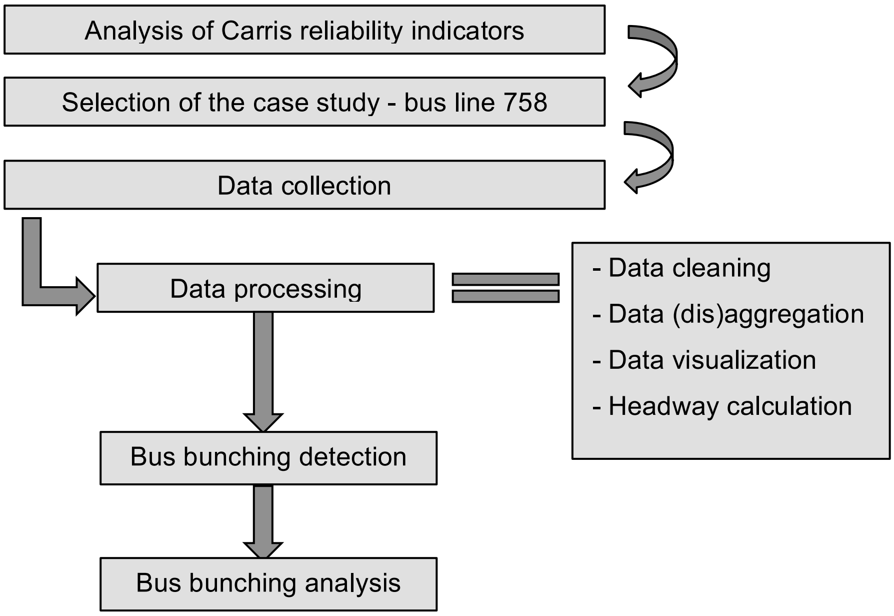


Figure 4 - Framework for the bus bunching detection model

3.2. Case study

3.2.1. Overview on reliability

Transit authorities around the world, just as Carris, have been striving to improve transit service quality to attract more transit riders. It has been seen in section 2.1 that the level of service enhancement largely relies on whether the transit system can be operated as it was planned, i.e., how reliable the transit system is.

There are several approaches to evaluate that match between planning and operations (that were also reviewed in the previous section). Carris, in particular, assesses the gaps in service by looking at the distance that was not covered for each bus line that was initially programmed ($lost\ km{\times}vehicles$). The fact that the offer is not fulfilled may be a consequence of many distinct events such as vehicle breakdowns or driver absence, and Carris gathers the information on the causes that led to each interruption.

This information on the causes of the interruptions in service was made available for bus lines 728, 750, 751 and 758 for the first trimester of 2018. These four bus lines were chosen due to their importance on the network, in terms of $offered\ km{\times}vehicles$, and on indications from the company concerning the bus bunching (BB) occurrences. The results show that 36% of the gaps in service were due to traffic congestion, while 30% of the interruptions were based on "*various conditions*", which may be related to other regularity issues. Also, 18% of the interruptions were caused by vehicle breakdowns while 7% were due to a missing vehicle or driver absence.

The analysis of the failures to match the scheduled service is also valuable for the identification of the bus lines where BB is more likely to occur. For instance, two of the corrective measures applied by Carris to deal with perturbations in service are short-turning and deadheading. The application of these measures implies higher $lost\ km{\times}vehicles$, since the vehicles will skip part of the service to be reallocated in another position on the line. Hence, bus lines with higher $lost\ km{\times}vehicles$ are expected to be more irregular (and consequently less reliable). It was also already discussed that BB is one of the most visible effects of an irregular service. Therefore, one can infer that BB is one of the triggers that makes controllers apply the corrective measures that end up increasing the $lost\ km{\times}vehicles$, in order to improve the overall performance of the lines.

Consequently, it is important to examine which bus lines have worse results in terms of $lost\ km{\times}vehicles,$ when compared to the total $km{\times}vehicles$ they are scheduled to offer. The information on the total offer ($offered\ km{\times}vehicles$) and the part of that offer that wasn't fulfilled ($lost\ km{\times}vehicles$) were made available by Carris to allow for that evaluation. The data was gathered for the first trimester of 2018 for all the bus lines.

With that information, it was possible to calculate the ration between the $lost\ kms$ and the $offered\ kms$ for each route, which I will name non-compliance rate from now on. The mean non-compliance rate for the first trimester of 2018 for each bus line ranged from 0.2% (almost no deviations from the planned service) to 7.2% and for the total network of 64 bus lines and the mean non-compliance rate for that trimester was 2.7% (considering all bus lines in the network). Out of those 64 bus lines, 21 presented a non-compliance rate greater than 3%. The detailed results for those 21 bus lines are shown in annex 4. Amongst those, one can find bus lines 758, 750, 751 and 728, which is consistent with the company's expectations. In fact, bus line 758 is the one with the worst performance (non-compliance rate of 7.2%) and is also one of the bus lines that operates with high frequency in most of the day (with headways lower than 10 minutes). Having that in consideration and also taking into account the considerable data requirements needed to study BB occurrences in each bus line, I decided to study this bus line at first in order to build a detection model for bus bunching, based on the available AVL data. This decision was also justified given the indications from Carris on this specific bus line.

3.2.2. Bus line 758 - route configuration and schedule

This study focuses on Carris' bus line 758, an inner-city transit line with an average planned headway of 6-10 minutes during most of the day (6h00-20h00) that operates in two directions between Cais do Sodré and Portas de Benfica. From 22h00-01h00, the bus operates in a section of the line (between Cais do Sodré and Sete Rios) only, but this part of the day won't be analyzed since the planned headways for this period of the day are between 20-30 minutes, which represents a low-frequency bus route. Hence, the focus of the analysis will be on the operation from 6h00-20h00, when bus bunching is more likely to occur.

The stops for direction 1 (from Cais do Sodré to Portas de Benfica) are shown in figure 5, while the stops for direction 2 are shown in figure 6. The extension of the route is 9.8km in both directions (9816m in direction 1 and 9811m in direction 2). By looking at the schedules provided to the customers, we can see that there are 32 stops in direction 1 and 34 stops in direction 2. However, for direction 1, 33 different stops will be analyzed, since there are 2 different stops called Sete Rios (one of them is the last stop of the service provided at night after 10 p.m.). It should be noted that this doesn't reflect on the results of the bus bunching detection model, because every vehicle goes through the stops, even though they may or may not stop in each stop. Besides, the BB detection model considers the passing times in each stop, and if there isn't a record in a certain stop, then BB is not scanned at that stop (therefore, I am considering all the possible/detectable stops of the route). Likewise, for direction 2, instead of 34 different stops, the analyses will focus on stops with 36 different codes (there are certain Sete Rios and Amoreiras stops that are only used in a part of the day, as will be seen).

Bus line 758 is characterized by long travel distances across congested areas and the route passes by different areas of the city, with implications on the route layout.

Considering direction 1, for instance, the vehicles start their journey in Cais do Sodré (stop 1 in figure 5), and go through the narrow streets of the city until Rato (first 8 stops). The vehicles go through streets such as Rua da Escola Politécnica where the route is shared with other vehicles (cars and motorbikes) as well as trams. It is common to see trams being overtaken by buses but this is not always possible. In the Príncipe Real area, illegal parking is an important issue, causing interruptions on the trams' journeys, with consequences for the other modes. Besides, there are several crossings and the pedestrian walkways are not signalized, further delaying the bus run.

From Cais do Sodré to Rato, the bus encounters 3 signalized intersections (after the departure in Cais do Sodré, before Praça Luís de Camões and before Rato and none of them contemplates priority for the buses). In Rato, the passengers can transfer from and to other buses and also to the metro service. After Rato, the route configuration is not so restricted by the narrow structure of the old part of the city. The buses move to Amoreiras through Rua D. João V and then proceed to Campolide and Sete Rios areas. Amoreiras is also an important stop, allowing transfers to other bus lines and also to other companies' services, such as Vimeca and TST.

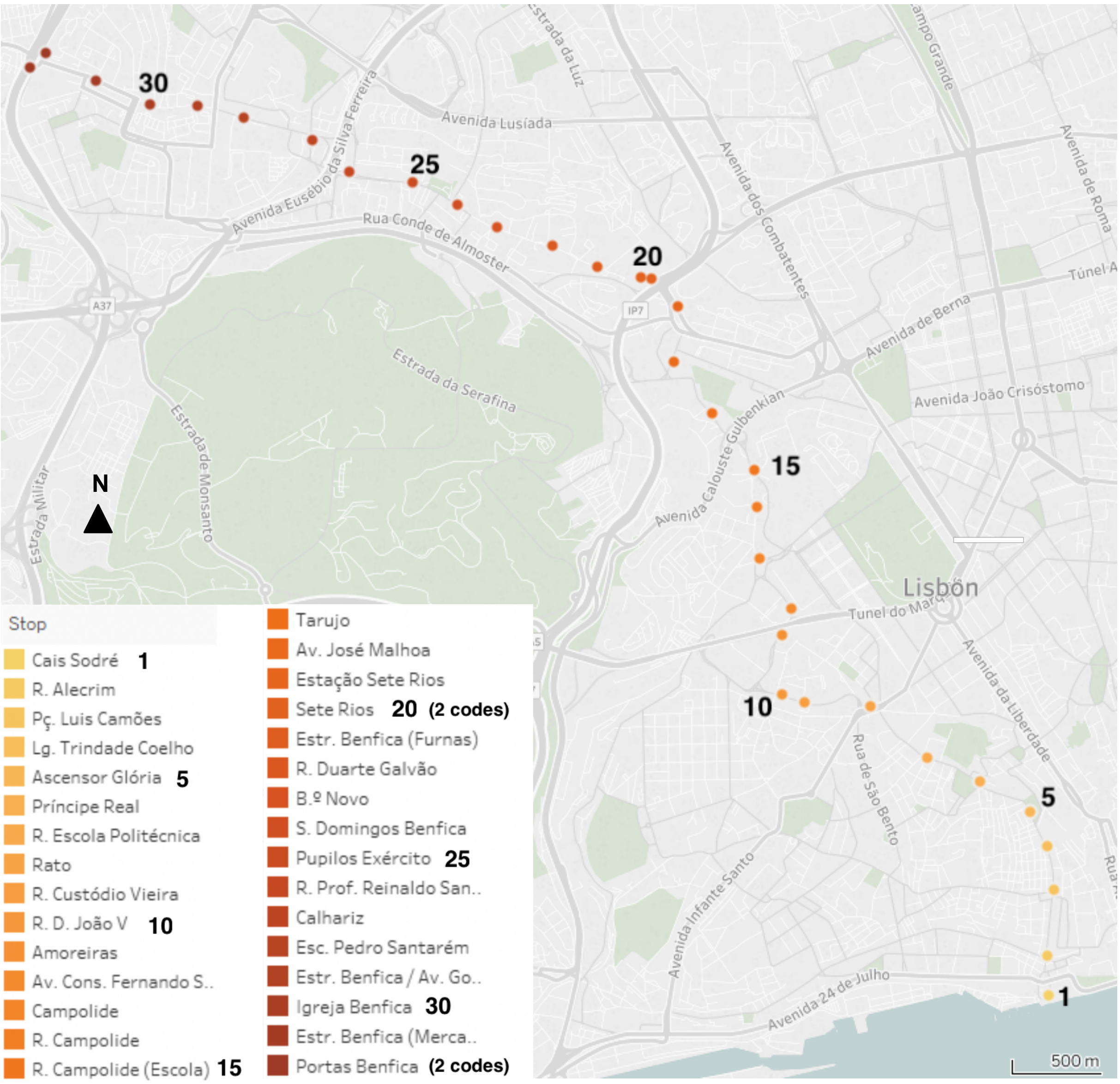


Figure 5 - Stops location in bus line 758 (direction 1, from Cais do Sodré to Portas de Benfica)

Between Amoreiras and Sete Rios, the vehicles encounter a significant number of signalized intersections, which may contribute for the variability of the travel times. Sete Rios (stop 20 in 5) is also an important transfer point - it allows for transfers to and from other bus lines, metro, trains and also intercity buses. From Sete Rios, bus line 758 follows Estrada de Benfica in its almost complete extension until Portas de Benfica. In these last 13 stops of the route, the vehicles find several intersections and pedestrian walkways but it is in this section of the route that more priority is given to buses through dedicated lanes in some sections of the route.

The stops for direction 2 (from Portas de Benfica to Cais do Sodré) are shown in figure 6 and the route restrictions are very similar to the ones found in direction 1.

For winter business days and for the period of the day that we are analyzing (6h00-20h00), a total of 19 vehicles are allocated to this bus line for its complete length and there is an extra vehicle that operates between Sete Rios and Amoreiras for three trips (starting at 7h55, 8h25 and 8h55 in Sete Rios). These trips were inserted in the schedule to cope with higher demand in this section of the route in the morning-peak period. That vehicle only serves the direction from Sete Rios to Amoreiras, running without passengers on the other direction until the next departure from Sete Rios (indicating "Reserved bus" in its variable messaging system). This time window will be particularly important in the analysis of BB occurrences due to the increase on the service frequency in this direction, which decreases the scheduled departure headways of stops within this zone. The extension of this part of the route (between Sete Rios and Amoreiras) is 2056m.

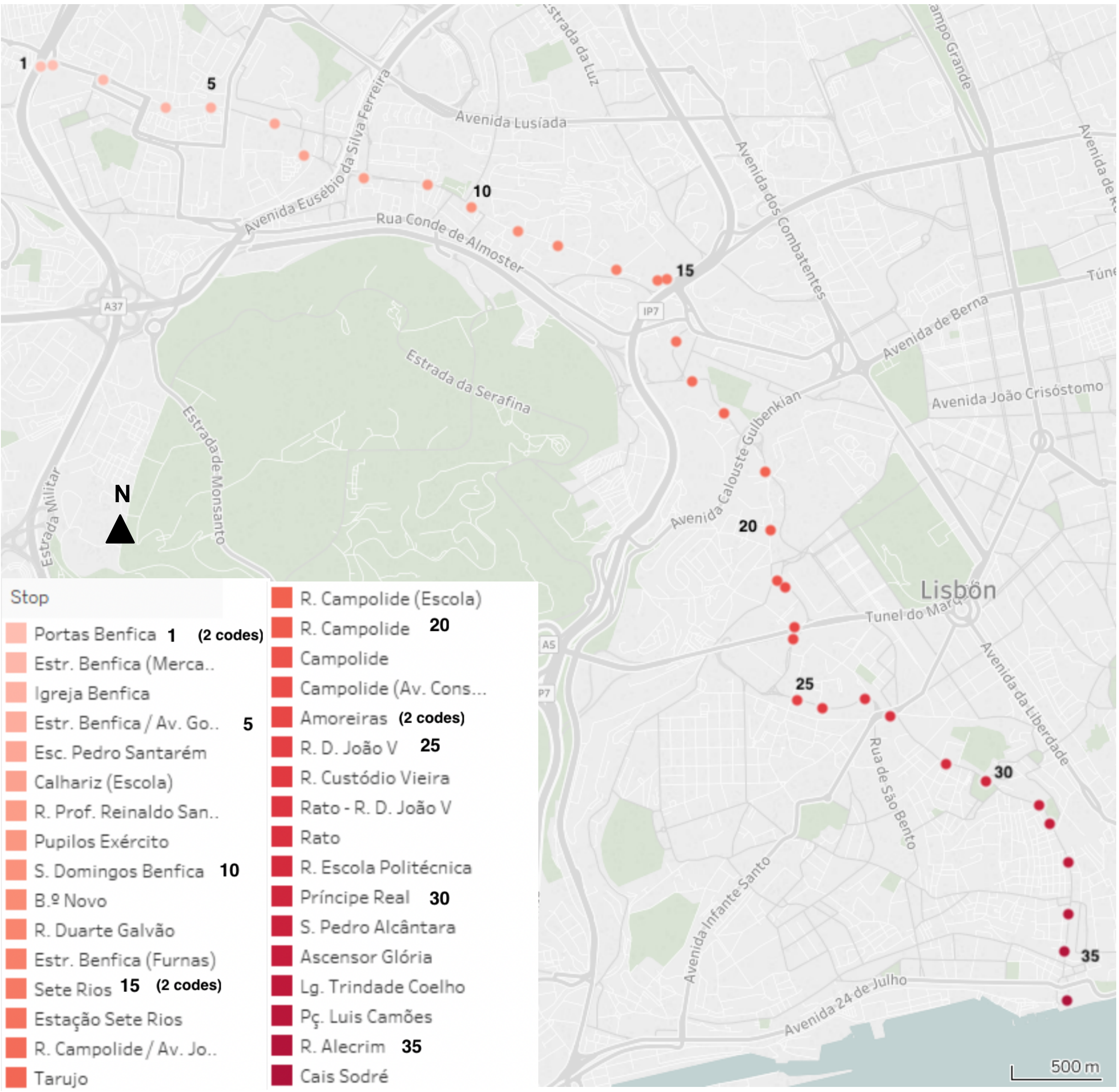


Figure 6 - Stops location in bus line 758 (direction 2, from Portas de Benfica to Cais do Sodré)

After June of 2018, there was a reduction on the frequency of the service, and the number of allocated vehicles diminished from 19 to 18.

All the vehicles start their day at the garage terminal in Linda-a-Velha and travel to the first stop: Portas de Benfica (where the vehicle starts its service for line 758) or another stop (in the case of vehicles that start their service in other lines (for instance 754) and later serve line 758. Each vehicle will do as many trips between the two extreme stops as the schedule indicates and then will return to the terminal from Portas de Benfica or start its service for another line. The entry and exit from the service is always made in Portas de Benfica because it is the closest station to Carris' garage, for a better allocation of the resources.

The analysis of BB will have a focus on two main periods of the day (the morning peak and the evening peak). I will define these two periods based on the schedule that is available to the users of bus line 758 (figure 7). Since I want to compare the frequency in which these events occur among different periods of the day, and also among the two directions, I selected two periods of three hours and used the same period for both directions (the morning peak, 7h00-10h00, and the evening peak, 17h00-20h00). These peaks were chosen because they are the ones when more vehicles are in service in both directions - in figure 7, we can see the scheduled departures from the terminals (in the upper schedule we see the departures from Cais do Sodré, whereas in the lower schedule we see the departures from Portas de Benfica). The chosen peak periods are bordered by red rectangles.

Besides the information on the scheduled departures from the terminal, the customers also have access to an estimate time that the bus takes in among sections of the route. That information is given to the customer for each direction, regardless of the time of the day, which doesn't make much sense because the travel times between stops vary considerably throughout different periods of the day.

4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1
		10	00	06	01	01	02	05	09	01	05	08	05	08	02	05	11	06	09	05	
		20	10	15	08	09	13	16	19	12	15	18	14	17	11	16	31	27	35		
		30	20	24	15	17	23	26	30	22	26	29	23	26	20	28	45	48			
		40	30	33	22	25	34	37	40	33	36	38	32	35	29	39					
		50	39	40	29	33	44	48	51	44	47	47	41	44	38	51					
			48	47	37	41	55	58		54	57	56	50	53	47						
			57	54	45	52							59		56						
					53																

4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	1
	35	07	02	00	00	02	05	09	01	04	03	01	04	07	03	19	22	15	00		
	43	15	09	06	10	13	16	19	12	14	13	10	13	16	15	39	45	35	35		
	51	23	16	13	21	23	26	30	22	24	24	19	22	25	27	59					
	59	31	23	20	31	34	37	40	33	34	34	28	31	34	39						
		39	30	28	41	44	48	51	44	43	43	37	40	43	59						
		47	36	36	52	55	58		54	53	52	46	49	52							
		55	42	44								55	58								
			48	52																	
			54																		

Figure 7 - Schedule for customers: scheduled departures from the terminal for line 758 for a regular winter day: hours (yellow line) and minutes after the hour (each column) - upper: Cais do Sodré to Portas de Benfica; lower - Portas de Benfica to Cais do Sodré. (Source: http://www.carris.pt/)

3.2.3. Current control practices

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. There are also specific stops that serve as driver relief points, where schedule adherence is especially significant. 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.

This control center is located in Carris facilities in Linda-a-Velha, where each controller / dispatcher monitors all the vehicles that are running in a number of transit lines (4 to 10, depending on their frequency and length). These controllers follow the positions of the vehicles in each transit line through simplified diagrams that are updated in real-time. The AVL system tracks the position of each vehicle and uploads that information to the control center. The positions of each vehicle are updated every 30 seconds, which allows for an almost continuous tracing of the vehicles. In extreme situations, such as a vehicle breakdown, controllers can communicate with the drivers if there's a need to shift service from one point of the route to another (due to a gap in service), instructing the drivers to drop the passengers and reallocate service (through short-turning or deadheading). Drivers can also be told to skip parts of their route or suspend service if they get too far behind schedule. However, it seems that drivers are not notified when they are running early or late, or if they are getting too close to the bus that precedes it (approaching a bus bunching situation).

Also, there is a chief-controller in the control center that has information on the complete network and may share information on incidents such as the existence of sports events, riots, or extreme traffic congestion, that result in the need to adjust the routes to avoid those incidents. The chief-controller communicates with the controllers in the control center and these notify the drivers of any modification made to their journeys.

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 instead of using preventive measures. These circumstances of Carris set grounds for methodological improvements based on available AVL data, which is explored in the forthcoming sections of this dissertation.

3.3. Data mining

3.3.1. 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 back office / control center. The AVL data is gathered through GPS-antennas installed in every vehicle and 2G/3G antennas that enable the real-time data interchange between the vehicles and the control center servers. The geographic location of each vehicle is gathered at least every 30 seconds (the system automatically refreshes every 30 seconds but extra records are gathered upon the occurrence of some events, such as the beginning and the end of each trip).

At the end of every day of operation, the AVL data is stored on a folder that contains as many files as the number of operating vehicles in that same day. In this folder containing the AVL information gathered for the complete day, each file is identified with a radio number, which is unique and is related to the GPS identification system of each vehicle. However, these radio numbers were generated upon the installation of the AVL systems in each vehicle, which caused a mismatch between these numbers and the internal fleet identification numbers that Carris already had in place. Therefore, it is important to remember that 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 already used for fleet identification. Each vehicle has a single radio number and a single ID number and there is a one-to-one correspondence between the two identifiers. Carris also made available the table that allows for this correspondence. One must not mistake these identifiers with the number of the bus route, these identifiers only distinguish the vehicles and each vehicle can be assigned to one or several bus routes. Similarly, many different vehicles serve each bus route.

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), in which we can highlight the following ones: 1) the time stamp; 2) the type of record (whether it was a regular record happening every 30 seconds or an extra one); 3) the bus line to which that vehicle was allocated at that time stamp; 4) the location (longitude and latitude); 5) the direction in which the bus was travelling on that time stamp; 6) the plate number.

Alongside with the AVL information, Carris also made available the essential information on the stops for every bus line. Carris provided us with an Excel file containing the stops identification for each bus route, direction and variant inside that route, with the distance between stops, the stops ID (which is their unique identification number), their order in each direction and variant of the route. For each stop, there is also information on the name and location of the stop (latitude and longitude).

Furthermore, for each time-series that was analyzed in this work, Carris provided an Excel file with the information on the identification of the vehicles (ID numbers) that were used in each specific day for the bus line of the case study (758). Without that information, it would have been impossible to identify the AVL files that should be studied in order to study bus line 758 in those days. The schedule for each plate number for bus 758 was also made available, for a regular winter day.

To sum up, the following data sources will be used in this study:

- The AVL information (merged in files with the daily records for each radio number);
- The conversion table (ID number to radio number and vice-versa);
- The information on 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.

In section 3.3.2, I will try to better explain how all these data sources were combined in order to assess the occurrence of bus bunching in this bus line.

3.3.2. Data processing

In order to study the bus bunching phenomenon on bus line 758, Carris started by providing the AVL information for two days only (the 1st and the 6th of February of 2018). Therefore, I created a methodology to identify the BB occurrences on those days, having in mind the future replication of that same methodology to a larger dataset. Later on, Carris provided the information for 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. Mathematica provides a computation environment that uses the Wolfram language, which integrates many aspects of statistical data analysis, from getting and exploring data to building high-quality models and deducing consequences. The Wolfram language provides multiple ways to get data, starting with built-in data sources, importing from a variety of file formats, or connecting to databases, which is useful bearing in mind the different data sources. Basic processing of data, including computing statistical quantities, smoothing, testing and visualizing give a first level of analysis on this software. By adding models such as regression models, one can answer a wider range of analysis questions or even provide predictive capabilities. Besides the data analysis capabilities, the symbolic architecture and dynamic interface of the language allow for a flexible and convenient approach to charting and data visualization, which is another advantage of this interface.

I will now explain the methodology that I followed to process the AVL information available for the 1st and the 6th of February:

1) Importing the information on the ID numbers of the vehicles that were used on each day: In first place, I imported the excel file that had the information on the ID numbers of the vehicles that were operating in bus line 758 on the two studied days.

2) Converting the ID numbers into radio numbers: Then, I imported the conversion table from radio numbers into ID numbers and vice-versa and used that table to save the radio numbers of the same vehicles. Both the 1st and the 6th of February had 20 different radio numbers which is the expected number of vehicles, i.e., there are 20 plate numbers, according to the schedule for regular winter days.

3) Accessing the AVL information of each vehicle: Having the information for every radio for the two analyzed days, I started by creating 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, that 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, where each line of the file represents a new entry from the AVL system. I started by uploading the information into Mathematica and then I had to eliminate erroneous or mismatch records, through a series of criteria, and to keep only the information that would be necessary for the bus detection model:

- Selecting the rows that had the information on the timestamp and the location (hour, latitude and longitude);
- Eliminating the entries / lines that had no information on latitude or longitude (missing location);
- Removing all the non-numerical information (labels inside the cells on each line, punctuation, etc.);
- Converting the time-stamp format "00:00:00" to the number of seconds after 00:00;
- Eliminating the entries that had erroneous location records (zero latitude or longitude);
- Converting the format of the coordinates (the gathered GPS coordinates are hexadecimal and they are converted to decimal degrees coordinates) using the "CoordinateConverter", written based on the conversion method applied by Carris (annex 5).

5) Computing the distance between points: Having performed this series of actions, it was now possible to see all the accurate GPS records of a vehicle throughout a complete day. To visualize the different positions of a vehicle, I computed the "Dist" function (annex 5) that calculated 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: Using the "Dist" function to compare the location of each record with the location of the terminal in Linda-a-Velha, I plotted the distance over time to the terminal in order to have a first look at the provided data. For illustration purpose, I printed this information for the vehicle with radio number 503 on the 1st of February between 7 a.m. and 8 p.m. (figure 8). In this figure, we can see that the vehicle started its day at the terminal and then started its service on bus 758 where it traveled between Portas de Benfica and Cais do Sodré twice, before returning to the terminal around 10.30 a.m.. Cais do Sodré is the most distant point to the terminal in the figure (almost 8 kilometers away). Then, around 4 p.m., the vehicle started its service again, but this time it was allocated to another bus line (754), which explains, the difference in the patterns of the distance to the terminal after this time in the day when compared to the first part of the day.

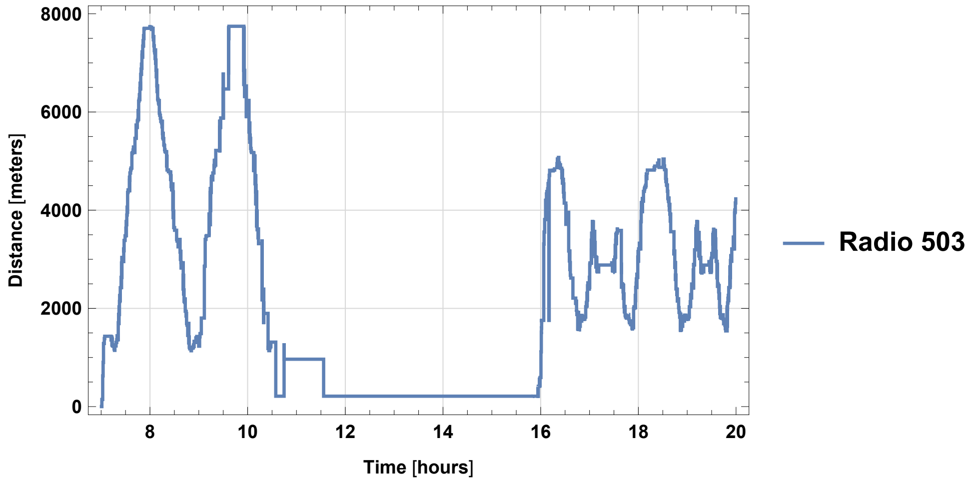


Figure 8 - Daily journey of the vehicle with the radio number 503

There are two important observations that were extracted from the visualization of the daily journeys of several vehicles:

- All the records are aggregated for the complete day (no separation is made when the vehicle starts or ends a trip);
- 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.

After having a closer look at the AVL records, I also noticed that there were entries that had the same time-stamp and different locations: there were records of a vehicle being in locations that were 2 kilometers apart from each other at the same time. Therefore, using the already explained "CoordinateConverter" and "Dist" functions, I computed a function that compared all the records that had the same timestamp and discarded those who were at a distance larger than 20 meters from each other. By simply eliminating both entries, it is true that I may be eliminating a correct and an incorrect entry. However, it is also true that there will be a new record in less than 30 seconds and this method guarantees than no incorrect entries are considered. This is computed through the "RemoveLeiturasRepetidasErradas" function, which is displayed in annex 5.

7) Data disaggregation: So far, I have shown that I am able to import and process the daily journey of a vehicle knowing its radio number. But in order to study the BB phenomenon, rather than having the information for the complete daily journey of a vehicle, one needs to have the location records from each separate trip of all the vehicles running on line 758 in order to compare the headways between consecutive vehicles on each direction. Therefore, besides the information on the timestamp, latitude and longitude, I will need to import more data in order to find the criteria to separate different trips inside the daily journeys of each vehicle. To do so, I created the "LongParseTXT" function that imports the AVL files saving further information from each record, namely the type of record, the plate number, the trip number, the direction, the direction variant and the bus line (annex 5).
Then, I created a function ("posSTOP", annex 5) to identify the records that matched the end of each trip by looking at the "type of record" column of each record and searching for the extra type of record "STOP_VOYAGE_REQ", that is shown each time the driver reaches the last stop of the trip (in the case of bus line 758 this will happen when the driver reaches Portas de Benfica or Cais do Sodré). In the case of vehicles with the plate number 21, that only run between Sete Rios and Amoreiras, the records that had the "START_VOYAGE_REQ" reference were also saved, otherwise the separation of the trips records from the deadheading records wouldn't be possible. In order to assess the validity of the method, I plotted journeys of different vehicles throughout the day. As an example, figure 9 shows the journey of the vehicle with the radio number 134 in, where the "trip ends" records are identified.

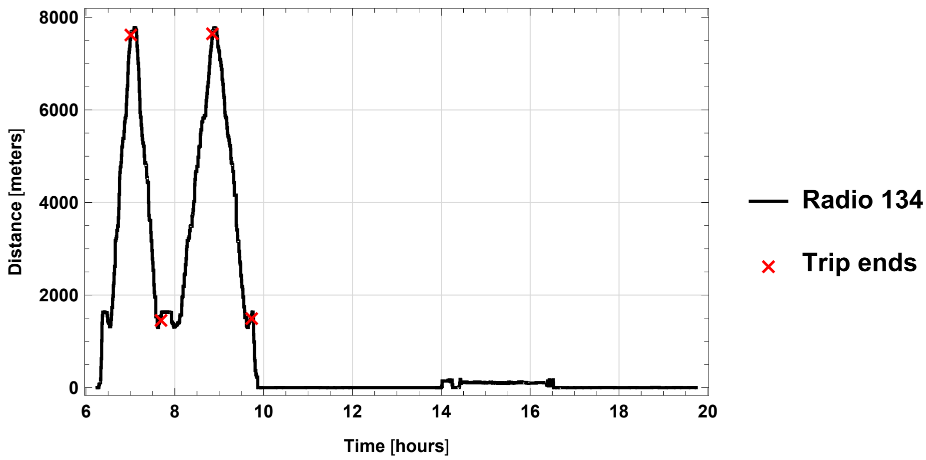


Figure 9 - Daily journey of the vehicle with the radio number 134, before the disaggregation of the trips

After identifying the positions where the vehicles end their trips in one direction of bus line 758, I created the function "SplitPerStop" (annex 5) that separates the AVL data between records of an ending trip. This method looks into all the location records of a vehicle and creates a division every time it finds a turning point, identified by the "posSTOP" function. Consequently, the data is disaggregated in sets between these turning points.
In order to analyze separately the information on each direction and to compare the headways between consecutive vehicles, it was then necessary to identify the trips among these sets from each direction. Taking that into consideration, I created the "CheckSentido" function (annex 5) that looks into the "direction" column of the AVL files (the direction can be 0 - when the bus is not in service; 1 - when it is travelling in the direction from Cais do Sodré to Portas de Benfica; and 2 - when it is

travelling from Portas de Benfica to Cais do Sodré) and defines the direction of each trip based on the counts of this column. Then, I grouped all the trips for each vehicle for each direction.

8) Eliminating the trips from bus lines other than 758: Following the disaggregation of each trip, I developed the "CheckCarreira" function (annex 5) that scans the "plate number" column and identifies the bus line to which each trip is allocated. Then, only the trips from line 758 are considered.

Then, I created the function "AllBusFromDaySentidoCarreira" (annex 5) that uploads and allows the visualization of the AVL files for a day for as many radio numbers as the user wants. By doing so, the user can see the trips of each bus for a specific direction and timeframe. For these visualizations, the functions "DistByTimetoCoord" and "DistByTimeToCoordProg" were developed in order to measure the distances and the cumulative distances run in each trip. The first function calculates the distance from each record to the terminal while the second function (cumulative) calculates the distance between the first record on each trip and the terminal and then adds the distance run between the following records to that one (annex 5). As an example, we can see in figure 10 all the buses running in direction 2 on the 1st of February from 7h30 to 9h30. We can accurately see the 3 trips from the vehicle with radio 394 (plate number 21, that only runs between Sete Rios and Amoreiras). Also, in figure 10, we can already acknowledge that the headways from the consecutive vehicles are very irregular, but I will explore that in detail in section 4.2.1, where a time-space diagram is explored.

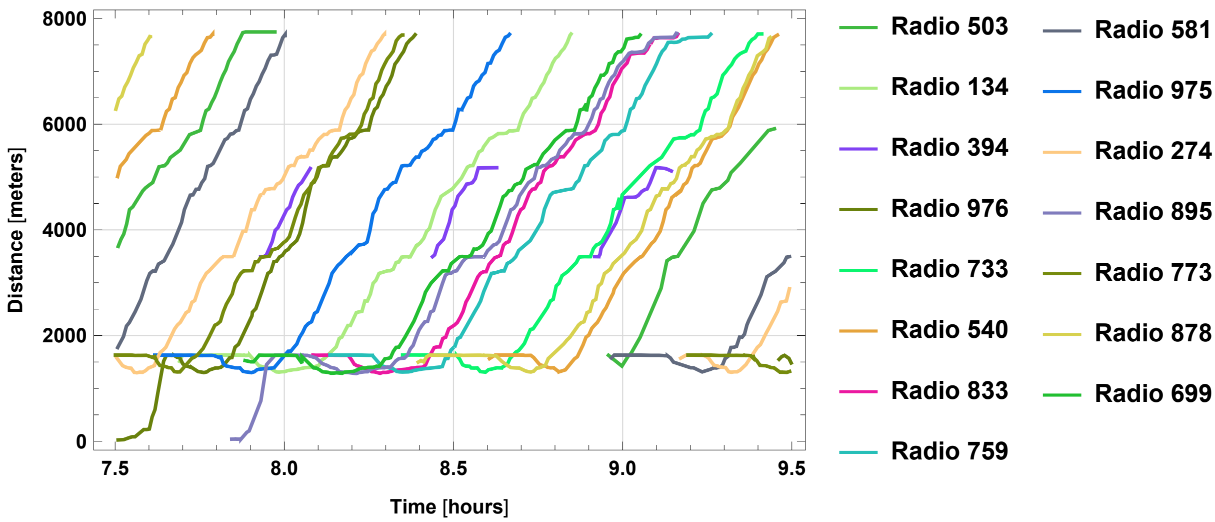


Figure 10 - Time-space diagrams of the 15 vehicles running between 7h30 and 9h30 in direction 2

9) Detecting the passing times in each stop: After creating the routine to upload the AVL information for any service day for bus line 758, the next goal was to detect the times in which each vehicle passed through each stop. Hence, I had to combine the data on the location of the several stops with a proximity criterion to the stops, since the location records from the vehicles are gathered every 30 seconds, regardless of their speed and position.

Consequently, I started by importing the location of all the stops of the bus line (in both directions) from the Excel file provided by Carris with the stops for all the routes. Then, I created a routine that looks into every record on each trip and searches for the location record (line) that has the closest location to each stop. One should note that it is possible that a vehicle doesn't go through all the stops (for instance due to a breakdown). As a consequence, I had to establish a threshold for the maximum distance that would be acceptable on the routine, i.e., I considered that the searching function would only search for records that were less than 200 meters away from the stop. If the method doesn't find any record on the list for that stop, then I will assume that the bus doesn't pass through that stop (it does not save a bus-stop time record). This won't be a problem since it is very unlikely that the method is unable to find closer records to the stop, for the following reasons:

- The buses run with a commercial speed of 15 km/h, which represents an average travelled distance of 125 meters between records (30 seconds).
- The average distance between stops for bus line 758 is about 300 meters (according to the data provided by Carris).

Consequently, it is very unlikely that we can't trace a record on a bus trip that is inside a radius of 200 meters around the stop. However, even if the method doesn't save a record for a certain stop, that won't interfere severely on the detection of BB because the headways between consecutive vehicles are computed for all the stops. The functions that were developed for this task were "PosMinDistToStop" and "LocateStops" (annex 5).

10) Detecting bus bunching: The output of "LocateStops" is a timestamp for every stop covered for each trip and vehicle considered for a certain day. In order to identify the occurrences of BB in line 758, the final stage is to compare these timestamps 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. Consequently, I developed the "DetetaBunching" function (annex 5) that receives as an input a list of all the time passages of each trip on the several stops and compares the passages of that trip with the passages of all the other trips for each stop. The function compares the trips inside a radio with the other trips on the same radio and with all the trips from the other radios, without repeating comparisons, and identifies the headways that are smaller than the assumed threshold value.

As mentioned in section 2.3.1, the approach on the definition of unstable headways varies among authors. One of the most commonly used thresholds for instability is perhaps the quarter of the planned headway (see Moreira-Matias et al. (2014), for instance). The planned headway for bus line 758 for the morning peak period and for the evening peak period varies between 6 and 10 minutes. Therefore, I tested three different thresholds (1, 2 and 3 minutes between consecutive buses) to identify BB occurrences. The lower the threshold, the lower the number of occurrences is expected.

The output of "DetetaBunching" 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. For instance, analyzing the direction from Cais do Sodré to Portas de Benfica for the 6th of February, and using the threshold of 2 minutes, the model detected 38 BB occurrences out of the 111 trips, which represents an incidence rate of about 34% (this means that 1 out of 3 buses bunches with its predecessor, which is very significant). When computing this function, I assumed that:

- Only trips with more than 10 valid records (out of the 33 or 36 stops) would be considered;
- There must be at least 2 stops where BB is detected in order to recognize that BB occurs.

The first assumption is made to prevent the consideration of trips that have insufficient (and possibly erroneous GPS) information and the second assumption is due to the fact that, sometimes, the drivers would mark their departure from the terminal ahead of the real departure, leading to inaccurate detections. Hence, this methodology prefers false negative to false positive detections.

I also computed a dynamic visualization process that allows us to look at each pair of buses that bunched in order to see the evolution of their timestamps along the route. For example, in figure 11, I depict the 27th pair of bunched buses out of the 38 occurrences. One can see that after 9h00 overtaking took place and the second bus (in blue) reached the end of the line prior to the bus it was initially following (in orange).

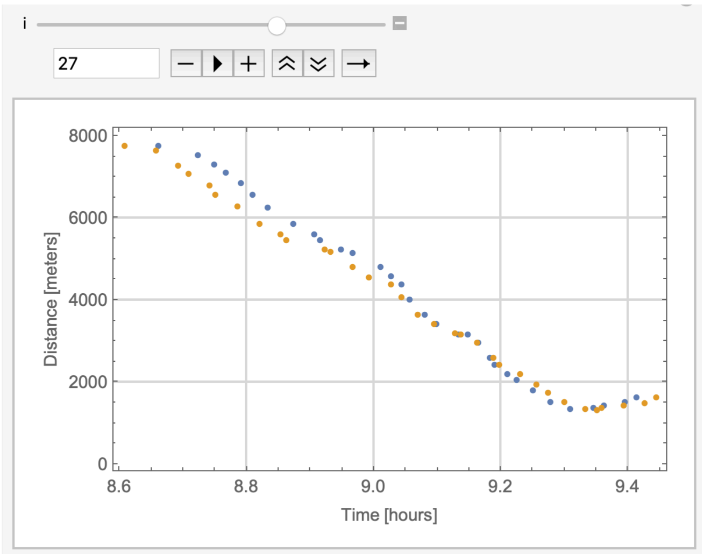


Figure 11 - Position detections (at stops) of two bunched buses after 8.30 a.m. on the 6th of February (the y-axis represents the distance to the garage

Furthermore, I verified that the model identifies a bigger number of occurrences as the threshold for headway instability gets bigger and, using the visualization process, I saw that the model is able to trace the BB occurrences involving vehicles that only cover the segment from Sete Rios to Amoreiras.

11) Headway calculation: Even though the "DetetaBunching" function already looks at the headways among pairs of buses to evaluate the existence of BB, I also developed a function to sort the different trips in each direction by departure time. The sorting function is called "SortByFirstRegister" (annex 5). It looks at the passing times of each trip in each stop (33 stops or 36 stops, depending on the direction) 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 will be essential for the computation of headways' statistics for consecutive trips.

12) Spatial bus bunching patterns: After the identification of the BB occurrences, I developed the "FindFirstBunchingStop" function (annex 5) that only examines the trips in which BB took place and depicts the first stop in the route in which the headway was identified as unstable. In this way, for a certain day-direction-threshold trio, the function investigates every pair of vehicles that were bunched in that day and outputs the first stop where the headway between each pair of vehicles goes under the threshold value. Therefore, this function outputs as much stops as the number of BB events that were identified.

4. Bus bunching analysis

4.1. Methodology

Once the Mathematica script was developed, I draw the approach to study the bus bunching occurrences in both directions of the line 758 during May 2018. This approach involved the following steps:

- Running the script for each direction and separate day for the bus bunching threshold of 2 minutes for three different periods of the day (6h00-20h00; 7h00-10h00; 17h00-20h00);
- Analyzing the time-space diagrams for the different periods and exploring the results for the bus bunching detections;
- Analyzing the results concerning the stops where bus bunching is more likely to be triggered for different periods of the day and estimating the probability of BB starting at each stop and the probability of BB occurrence for each stop (irrespectively of where BB started);
- Assessing the evolution of the probability of bus bunching along the route and depending on the different bus bunching thresholds;
- Evaluating the statistics for the headways between consecutive bus trips and the total travel times for different trips;
- Estimating measures of the passenger waiting times based on the results for the headways and assessing the influence of the headway in the first stop of the line 758 on the probability of bus bunching along the route;

Section 4.2. focuses on the BB itself while section 4.3 looks at further performance indicators that can be extracted from the developed work (that are also related to the BB occurrences).

4.2. Bus bunching detection results

4.2.1. 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 (Figliozzi et al., 2012). They are mainly used to analyze particular problems in a specific time of the day, just as the case that I am considering. Therefore, 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 (as well as the other visualizations of the statistics for the headways). This is valid as long as the information is already within the archived database and it should also be noted that the integration of other information such as the identification of the vehicles and radio numbers that were allocated to this line on each specific day was also necessary.

An example of a time-space diagram is shown in figure 12. To generate the diagram, the user only has to select the date, the 22nd of May, the period, from 17h to 20h, the direction (1) and then run the script. 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. These diagrams can be improved in the future by including the scheduled trajectories, in order to assess deviations from the schedule. 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.

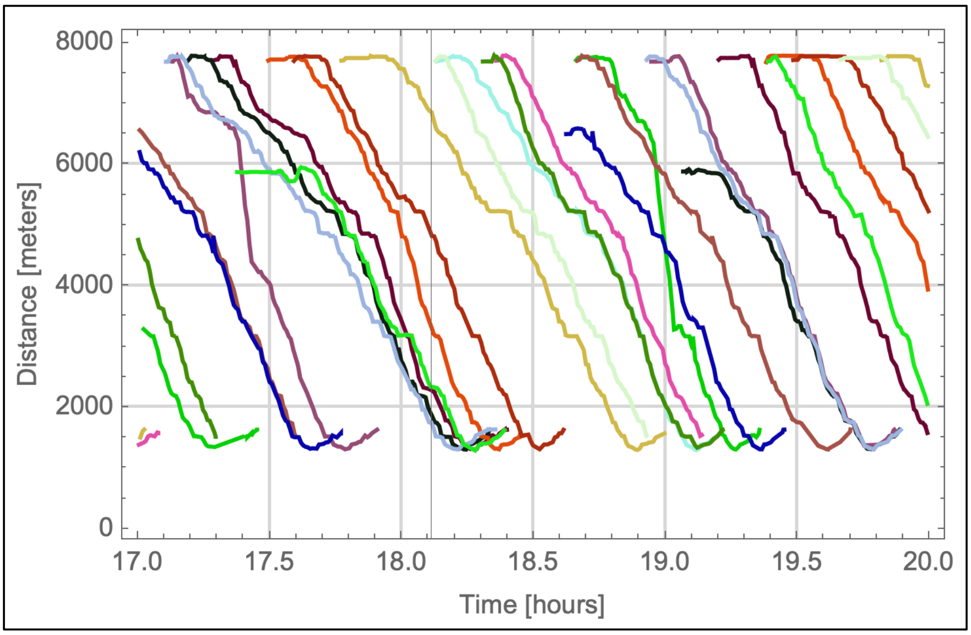


Figure 12 - Time-space diagram example: 22nd of May, evening peak, direction 1

4.2.2. Overall bus bunching detection results

Although the time-space diagram can show movements of all buses along a route in one direction in a certain time of day, detailed information about each bus is not available. And that is why I created the visualization tool that allows us to look at each pair of bunched vehicles individually as was shown in point 10 of section 3.3.2. There we can see at exactly what time each event started and ended or if it was propagated until the end of the route. However, it is not practical to look at each BB event one by one due to time restrictions. Also, the visualization of these results doesn't answer other questions, for instance: how have the two buses become bunched or what were their schedule adherences and departure times or headways at each stop?

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 started by gathering the information for each weekday and for the most critical periods of the day in terms of bus trips frequency (the morning peak, MP, from 7h00 to 10h00, and the evening peak, EP, from 17h00 to 20h00) to assess if there was any clear relation among these factors and the bus bunching occurrences. A total number of 4180 trips were computed. Among these, 2053 were from direction 1 and 2127 were from direction 2. For direction 1, out of the 2053 trips, 604 occurred during the MP and 587 occurred in the EP while, for direction 2, out of the 2127 trips, 684 occurred during the MP and 532 occurred in the EP.

Figure 13 depicts the percentage of trips in which BB was detected (according to the threshold of 2 minutes) compared to these totals, analyzing the peak periods of every weekday of the month.

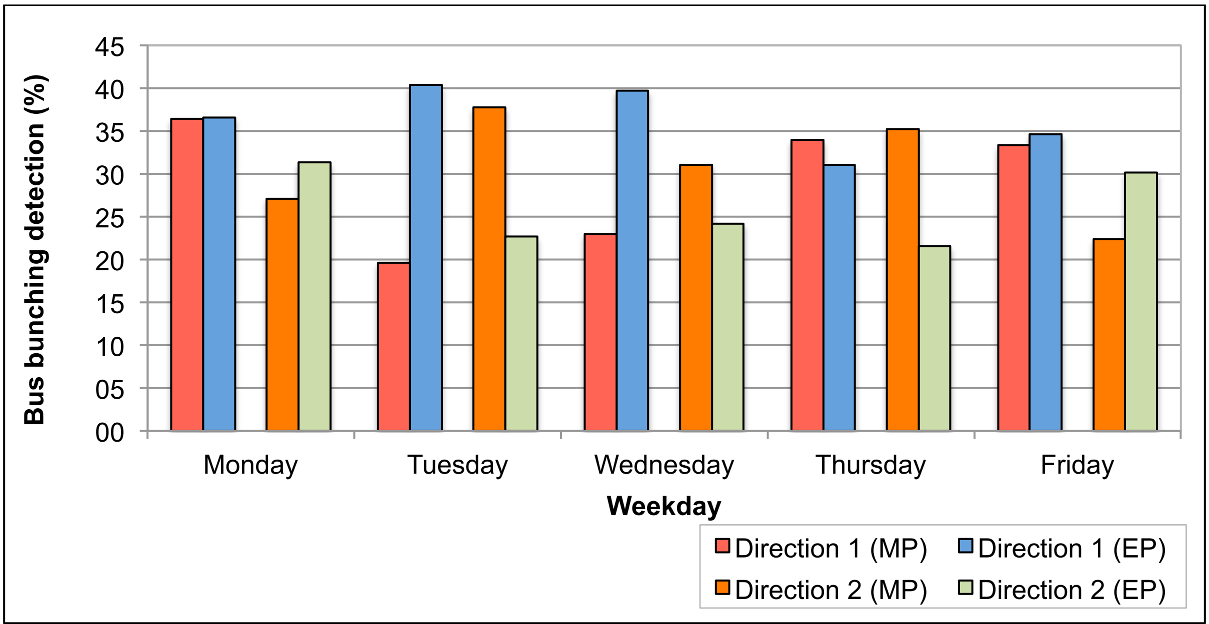


Figure 13 - Percentage of trips with BB detection by weekday and peak period (morning, MP or evening, EP)

We can draw some conclusions from this figure:

1) 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, considering that it represents an extreme irregularity issue);

2) There isn't a specific day in which the average BB detection frequency was particularly higher than in others;

3) There isn't a clear trend between the BB detections in the MP and the detections in the EP (which tells us that this reliability problem doesn't propagate from one peak period to the other);

4) 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 (from Portas de Benfica do Cais do Sodré) than in direction 1 for every weekday. This is against the expectations since bus trips are supposed to be more frequent in the other direction, as can be seen in the schedule presented in figure 7. However, it may be the case that the irregularity in direction 2 propagates to the operation in direction 1. This will be considered in section 4.3.

4.2.3. Bus bunching along the route

In the previous section, I analyzed the frequency in which bus bunching occurred in each trip, regardless of their extension in the route, i.e., BB detections were triggered whether bus bunching occurred in only two stops or if that event was spread over the full route. There, I saw that, in the peak periods, in average there were between 20 and 40% of the trips where bus bunching took place.

By contrast, in this section, I look at all the headways for consecutive buses in each stop for the 4180 trips that were computed (2053 for direction 1 and 2127 for direction 2) and I try to assess how the bus bunching events evolved along the route in both directions. Namely, in figure 14 we can see how the frequency of BB occurrences develops from the initial terminal to the followings stops, for different BB thresholds (in the upper part we see the results for direction 1 while in the lower part we see the results for direction 2).

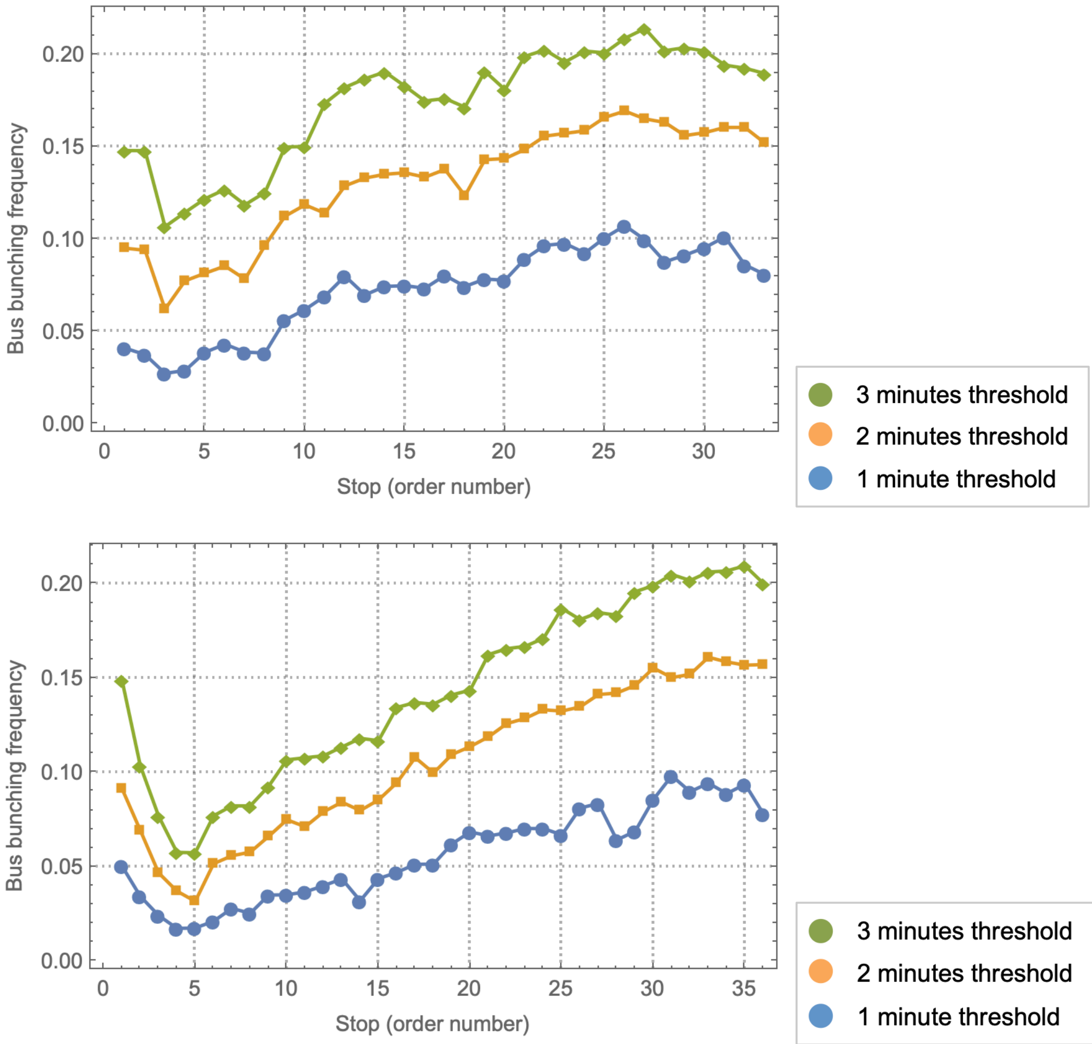


Figure 14 - Bus bunching frequency - evolution along the route (upper - direction 1; lower - direction 2)

These frequencies were computed based on the trips for the complete days (between 6 a.m. and 8 p.m.), so more than 2000 trips for each direction. The same representations were obtained for the peak periods and the results were very similar (annex 6). 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 bus bunching limit/threshold.

Three lines are shown for the three possible thresholds. As expected, the lower the threshold, the less BB is detected in each stop, since having buses only one minute apart from each other represents a more extreme and less likely situation than having those buses separated by three minutes.

In figure 14, we can see that in both directions the bus bunching frequency has an overall growth throughout the line: in the first stops, the likelihood of two buses being bunched is low and this likelihood grows from one stop to the following ones. This is expected and confirms that the bus bunching 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 causes of irregularity throughout their service (because they have travelled a longer distance), resulting in deviations to their scheduled arrivals to the stops that may end up in a BB detection further on the line.

However, figure 14 shows that in the first 3 to 5 stops of the line, the BB frequency tends to decline, which wasn´t expected. This fact suggests that some drivers may commence their journey earlier than scheduled (or they can set the start of their trips earlier on the radio of the bus). This feature can also be recognized when looking at figure 12 (section 4.2.1), where in fact the trips of consecutive buses are closer to each other in the first stops.

Despite this, one should note that, regardless of the direction we are considering, there are more than 10% of the trips between 6 a.m. and 8 p.m. where for more than half the route the buses are spaced less than 2 minutes between each other. When only the morning peak period is considered (annex 6), if we look at direction 2, for instance, the frequency of bus bunching for the 10 final stops of the route (which represent around 1/3 of the route length) is bigger than 20%, for the same threshold. These findings demonstrate once again the severity of these occurrences in bus line 758.

4.2.4. Where does bus bunching start?

The previous sections show the distributions of bus bunching counts for several periods of time and also along the two directions of the bus line. However, these bus bunching events are identified independently without knowing which trips they belong to. A methodology was also developed to look at each pair of bunched vehicles individually but apart from that, it is worth knowing where buses get bunched and separated for all pairs of bus trips. Therefore, I used the "FindFirstBunchingStop" function to identify the stops where BB started in each trip whenever BB was detected.

Identifying the stops where BB started while considering different threshold values and looking at different days may be helpful to identify hypothetical points of the route where the infrastructure itself may play a role on these occurrences (for instance, the effect of signalized intersections, the existence of bus lanes or other aspects of route configuration) so that particular control strategies can be implemented at those stops.

Firstly, I ran the model 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 look more closely to the results for the complete day (between 6 a.m. and 8 p.m.) since they represent a bigger dataset, possibly more conclusive in terms of spatial relations. In annex 7, I present the results for the peak periods for the threshold of 2 minutes, to which the same kind of analysis could be performed.

Following the application of the model, figure 15 shows the results of the BB trips formation for direction 1. In this figure, the y-axis represents the ordered stops from Cais do Sodré to Portas de Benfica while the x-axis represents the frequencies of BB beginnings for each stop. The orange bar chart illustrates the 2 minutes threshold and the blue one shows the 1 minute threshold.

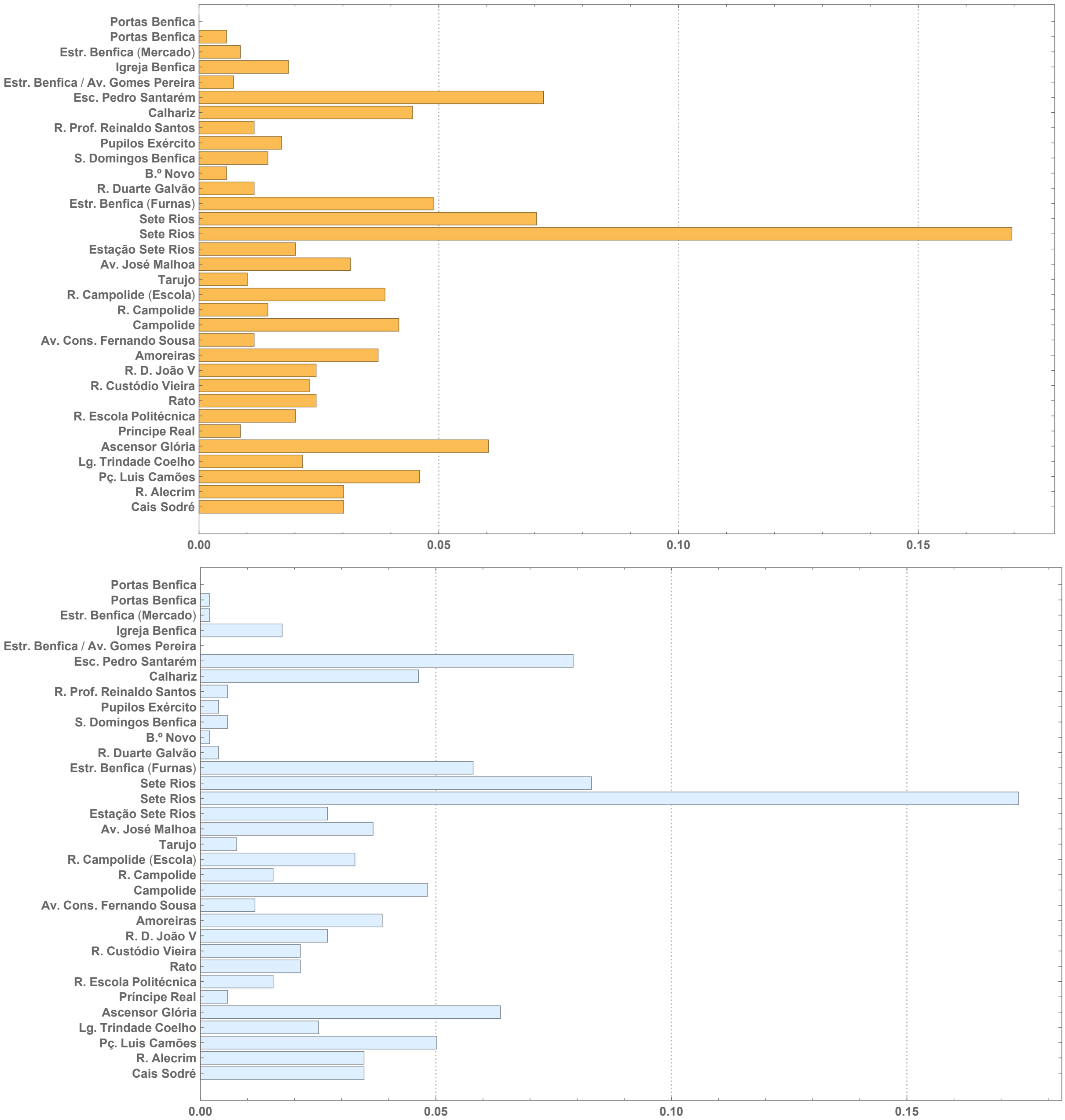


Figure 15 - Frequency of bus bunching formations at each stop in direction 1 (6h00-20h00) for the 2 minutes (orange) and 1 minute (blue) thresholds

Figure 15 illustrates that in direction 1, more than 15% of the BB trips were formed in Sete Rios. Also it is clearer from the blue bar chart that, overlooking stops Calhariz and Escola Pedro Santarém, the majority of the BB trips were formed between stops Cais do Sodré and Estrada de Benfica (Furnas). This matches our 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. Also, it is possible that some drivers are instructed to leave Sete Rios station when a certain train arrives to the station. Moreover, it is also known that Sete Rios is a driver relief point. If a driver arrives late to the station, it will directly affect the departure time from that 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 (also between Rato and Campolide) cause a variation in the cruising speed of the bus, possibly causing the formation of more BB events.

The considerable amount of BB trips in the initial stops of direction 1 is alarming since we have already seen on section 4.2.3. that once the BB starts, it is extremely uncommon for the bunched vehicles to become separated again and recover their regular headways. Therefore, the effects of this irregular service will be felt by the customer throughout a bigger extension of the route. While this pattern of having more BB trips starting in this section of the route (between Cais do Sodré and Sete Rios) is already noticeable in direction 1, it is even more obvious when analyzing the trips from direction 2 (figure 16).

In this figure, the y-axis represents the ordered stops (Portas de Benfica to Cais do Sodré) while the x-axis represents the frequencies in which the BB started in that specific stop (the orange bar chart for the 2 minutes threshold and the blue one for the 1 minute threshold). Looking at the blue chart, we can see that almost 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 then 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 (presented in section 3.2.2.) and also with the fact that it is in Sete Rios that the extra trips are inserted in the schedule to deal with the additional morning peak demand from that point until Amoreiras.

In this section, I could have also looked at the stops where the bus bunching events dissipated most frequently, besides looking at the formations. However, the observations indicate that once a pair of buses gets bunched, they are more likely to continue bunched until the end of the route, i.e., dissipations are very uncommon and not representative as a pattern.

It is important to note that, even though the results on the BB detections are useful, the results on the headway deviations play a bigger role, since the headway control can lead to a more regular service and also because other regularity indicators can be derived from the headways statistics, which are more indicative of the passengers perspective of reliability. This will be further explored in section 4.3.

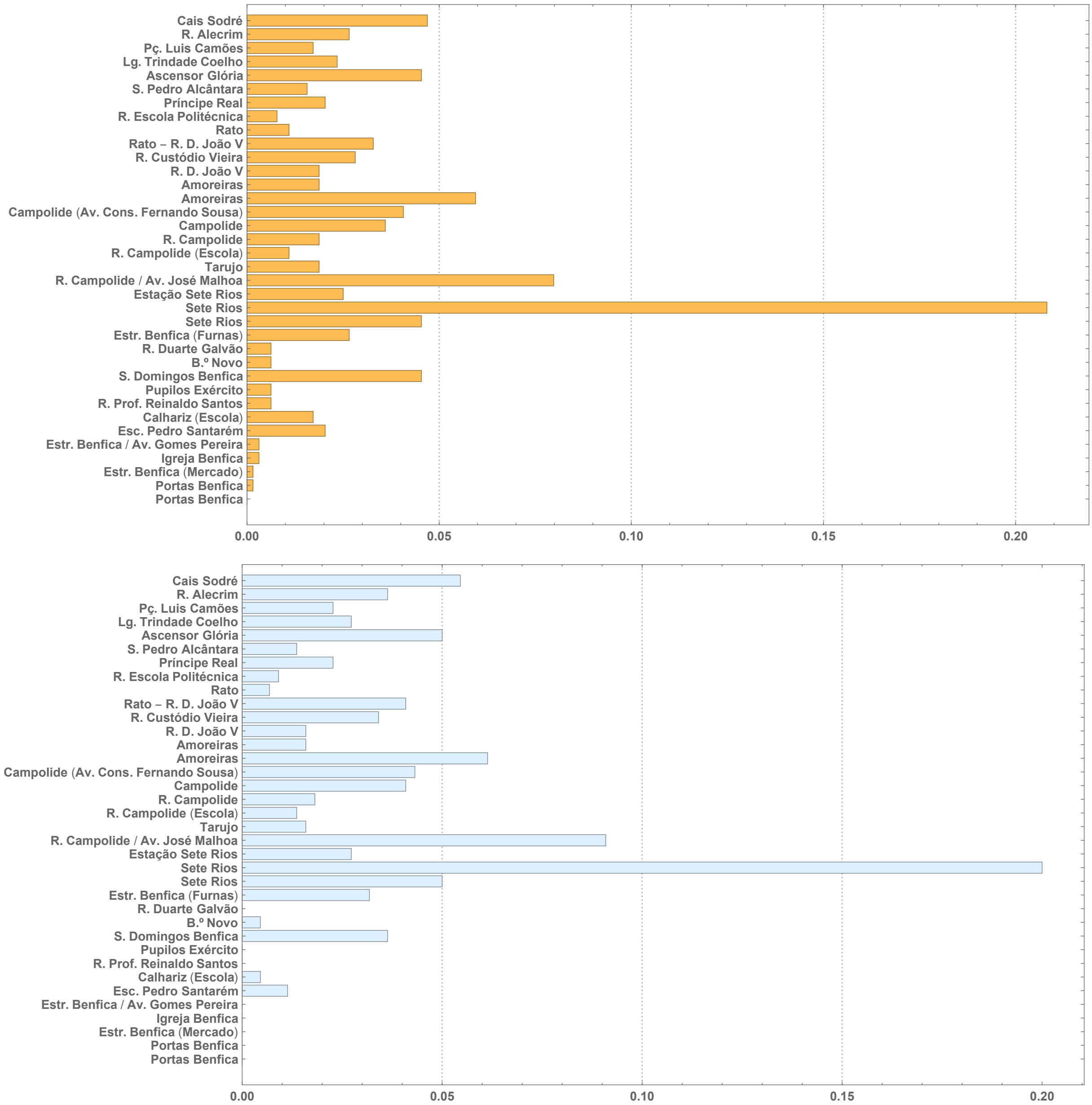


Figure 16 - Frequency of bus bunching formations at each stop in direction 2 (6h00-20h00) for the 2 minutes (orange) and 1 minute (blue) thresholds

4.2.5. The influence of the departure headway

In the previous section, we have seen the stops where BB is more likely to start in both directions. One of the conclusions 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 bus bunching detections for both directions for the complete period of analysis (between 6 a.m. and 8 p.m.). Consequently, I established a relation between the headway at the departure terminal between every two consecutive trips and the existence of bus bunching is any stop in the route for each of those trips.

Figure 17 shows the probabilities of downstream BB trips (considering the 2 minutes threshold) for different departure headway bins for direction 1 while figure 18 displays the same information for direction 2. For example, looking at direction 1, based on figure 17 and defining BB as the headway between consecutive buses being under 2 minutes, 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%. Similar interpretations apply to direction 2.

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 (Figliozzi et al., 2012).

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

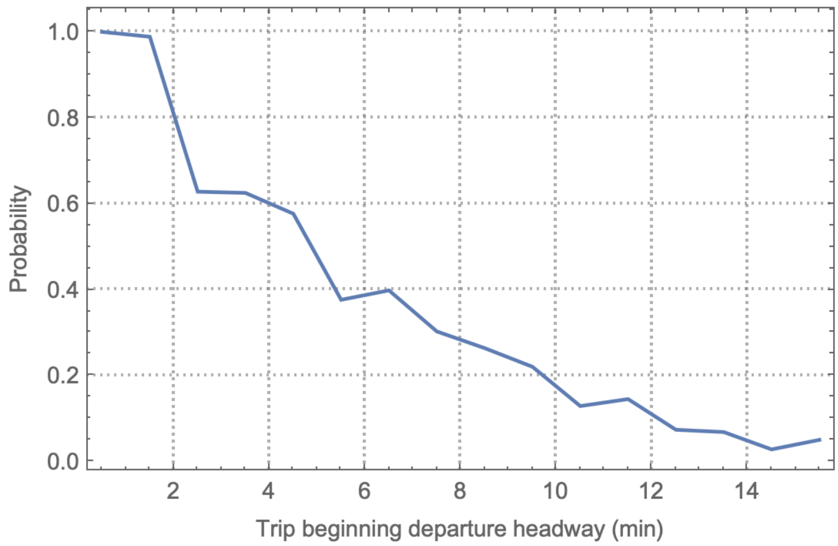


Figure 17 - Probability of downstream bunching with varying trip beginning departure headways for direction 1

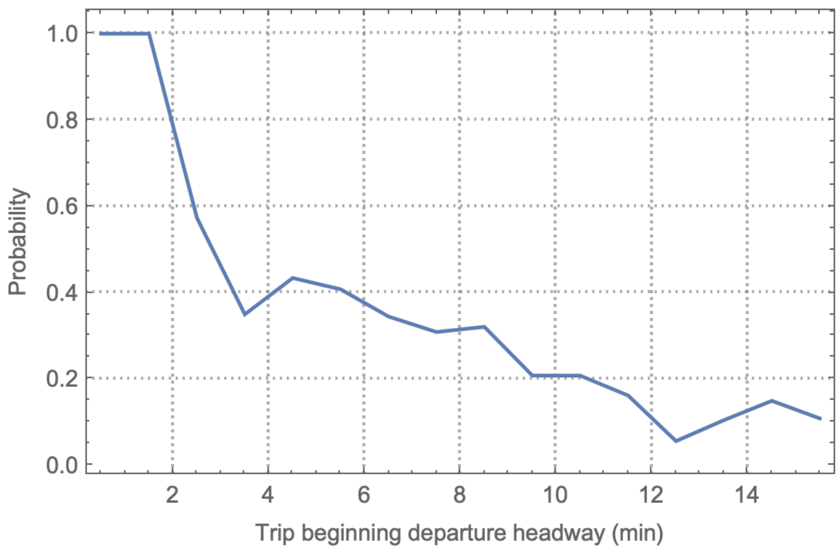


Figure 18 - Probability of downstream bunching with varying trip beginning departure headways for direction 2

4.2.6. The effect of the departure headway on the BB starting point

In the two previous sections, I have studied the stops where BB is more likely to begin in both directions and then the influence of the departure headway on the BB occurrences. In this section, I gathered these two types of data to assess the stops where BB is more likely to begin, depending on the departure headway. I expect that the trips starting with a narrower departure headway will have BB occurrences that start in the first stops of the route, since there is a tendency for small headways to become smaller quicker: it has already been seen that the BB appears as a positive feedback loop, in the assumption that passengers arrive uniformly to the stops (Cats, 2014).

In figures 19 and 20, I present the 3D representations of the proportion of trips where BB started in a specific stop, for a given departure headway. The x-axis represents the departure headway (in minutes), the y-axis represents the stop number in which BB started and the z-axis represents the contribution of each specific pair of departure headway and BB start in a specific stop, over the total amount of BB occurrences in that direction, where the total volume of the curve equals 1. The curves were obtained considering the operation for the complete daytime service (from 6h00 to 20h00).

In both figures, we can see that, just as expected, the trips where the departure headway is smaller are the trips where the BB occurrences start earlier in the route. Also, there is a bigger frequency of BB trips for departure headways that are smaller than 5 minutes, but also a considerable amount of trips where BB started for departure headways between 5 and 10 minutes, which is consistent with the results on section 4.2.5. Also, there are BB occurrences starting in all the stops of the route (33 stops for direction 1 and 36 for direction 2). In fact, for direction 2 (figure 20), the curve is wider, which is consistent with the fact that a significant number of BB trips start in the stops after Sete Rios, as seen in section 4.2.4. These 3D representations highlight once again the importance of tracking the departure headways, due to its impact on the performance in terms of the regularity of the service.

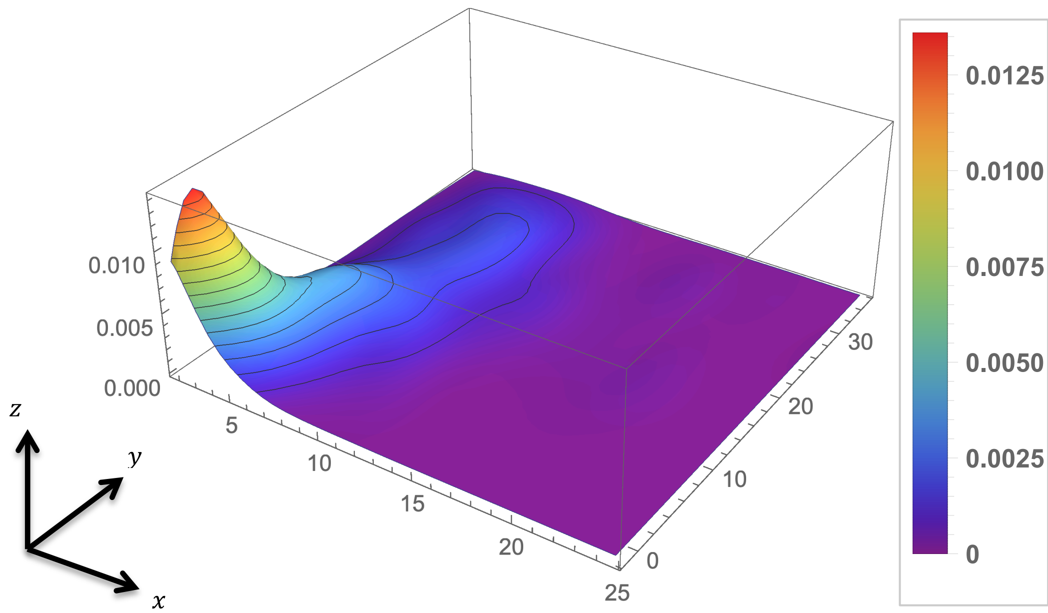


Figure 19 - 3D representation of the proportion of trips where BB started in a specific stop for a specific headway for direction 1. The x-axis represents the departure headway (in minutes), the y-axis represents the stop number in which BB started and the z-axis represents the contribution of each specific pair of departure headway and BB

start in a specific stop, over the total amount of BB occurrences in direction 1 (the total volume of the curve equals 1)

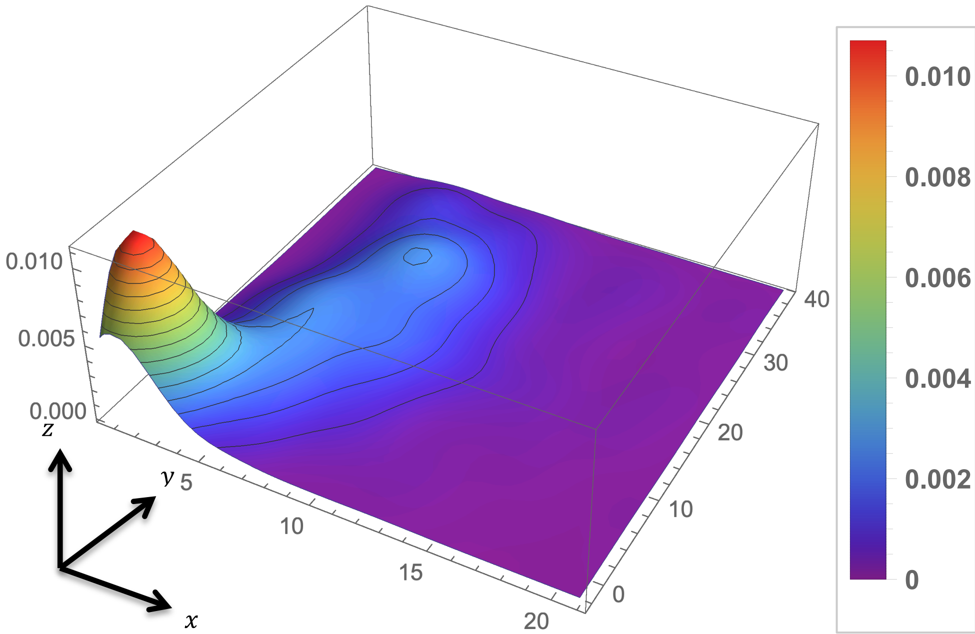


Figure 20 - 3D representation of the proportion of trips where BB started in a specific stop for a specific headway for direction 2. The x-axis represents the departure headway (in minutes), the y-axis represents the stop number in which BB started and the z-axis represents the contribution of each specific pair of departure headway and BB start in a specific stop, over the total amount of BB occurrences in direction 2 (the total volume of the curve equals 1)

4.3. Service regularity assessment

4.3.1. Headway distribution

As already mentioned, for frequent services, reliability needs to be interpreted in terms of regularity rather than punctuality. One of the categories of the measures of service regularity contemplates the indicators that are based on headway distribution and its relation to the planned headway.

Consequently, service regularity was primarily analyzed by calculating the distribution of all headways upon arrival to all stops. Headway distribution is a highly informative output as it allows observing the share of headways within each range of values. The planned headway is the focal value of the distribution and the variability it exercises reflects service regularity. The narrower the distribution the more regular the service is (Cats, 2013). The headway distribution provides therefore a useful evaluation tool that can be used to assess the success of the implementation of any corrective or preventive strategy.

Figures 21, 22 and 23 present the headway distributions for direction 2 (from Portas de Benfica to Cais do Sodré) for three different periods (6h00-20h00, 7h00-10h00 and 17h00-20h00). In each figure, 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. I started by computing the headway distribution for the complete days (6h00-20h00). We can see in figure 21 that the majority of the recorded headways is within the bins. However, the distribution is not very narrow, we can observe the existence of headways that are closer to 0 (0-2 and 2-4 minutes), as well as very large headways (bigger than 20 minutes). We can also see that overtaking took place for some pairs of buses since there are records of negative headways for consecutive buses, which is another consequence of BB.

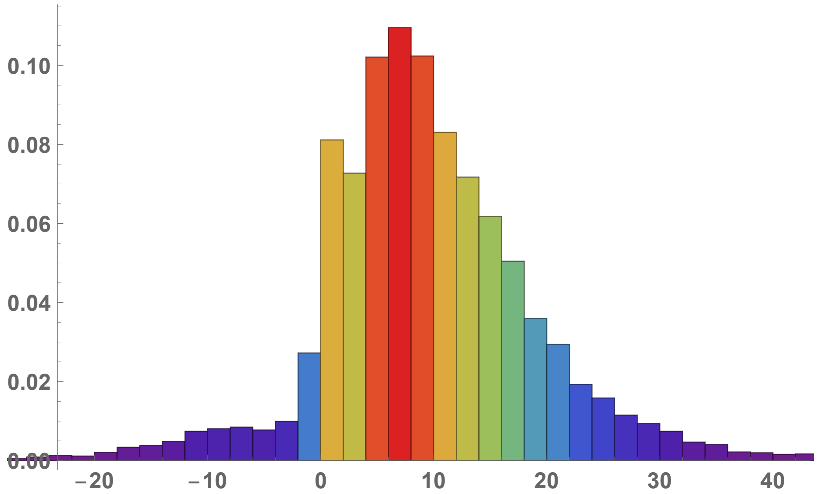


Figure 21 - Headway distribution for direction 2 for regular days of May of 2018 (6h00-20h00)

The platykurtic shape of the headway distribution outlines the poor performance in terms of regularity when the entire day is considered. However, since we are looking at a period of 14 hours throughout which the planned headway varies a lot, it is difficult to understand if the headways are close to the planned headway or not. For that reason, I analyzed the two peak periods, which have a planned headway that is close to 10 minutes, in order to improve the interpretation power.

In figure 22, we can see that the distribution of the headways during the morning peak period is narrower. More than 30% of the recorded headways are between 4-8 minutes and almost 20% of the headways are between 8-12 minutes. However, more than 20% of the recorded headways are between 0-4 minutes, which represents BB or close to BB situations. Besides, the remaining 30% of the headways are spread over negative values or above 12 minutes, which represents poor regularity.

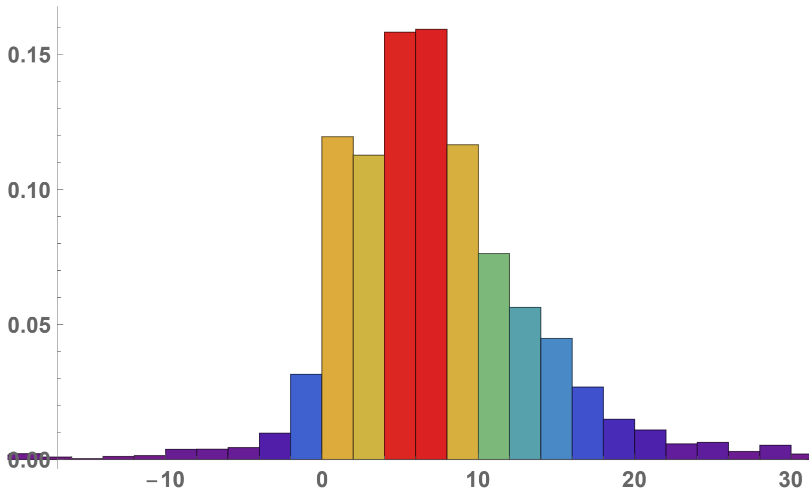


Figure 22 - Headway distribution for direction 2 for regular days of May of 2018 (7h00-10h00)

The performance in terms of regularity is much worse when the evening peak period is considered. In figure 23, we can see that the distribution of the headways is spread over a big range of values, being the most frequent bin 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 we compare this period to the morning peak period, and on the other hand, the share of regular headways decreased. It is true that according to the schedule presented in section 3.2.2, the headways are expected to be bigger than in the morning, but that should imply a horizontal translation of the distribution of about 2 minutes. Instead, we observe a wider range of headway values, with no clear focus on the planned headway (that should be around 10 minutes). This can be a sign that the irregularity of the operation in the morning period tends to propagate throughout the day until the evening period.

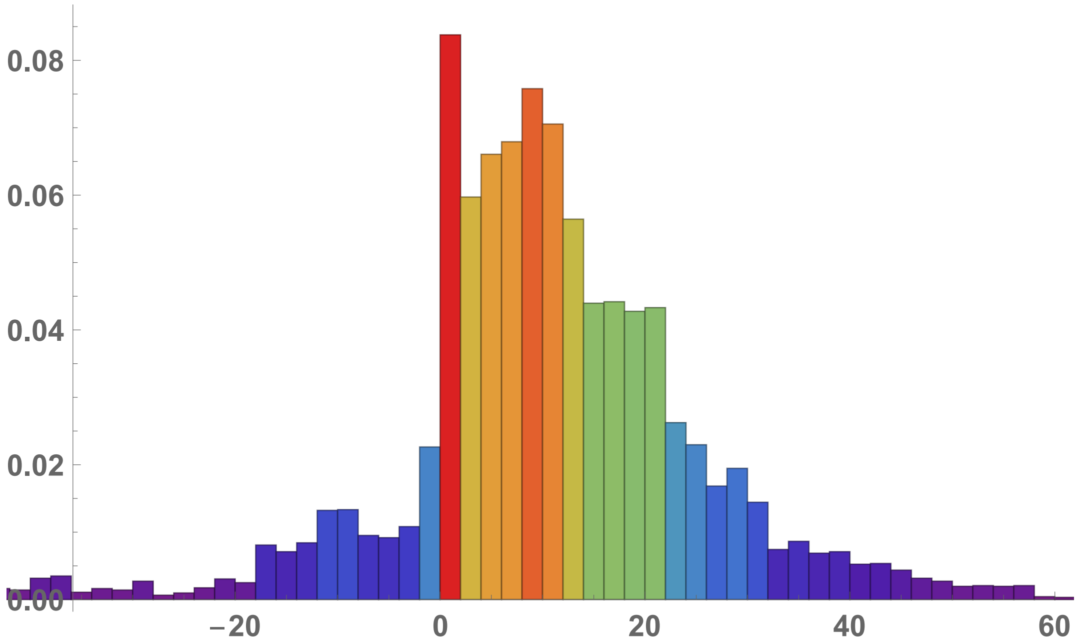


Figure 23 - Headway distribution for direction 2 for regular days of May of 2018 (17h00-20h00)

The headway distributions for direction 1 are presented in annex 8, and the analysis that follows can be similarly applied to that information.

4.3.2. Headway box-and-whisker plots

Besides looking at the distribution of the headways over the complete route for a specific direction, it is also useful to assess how the headways vary over the route. One of the ways to do so is to compute box-plots of the headways in each stop and to look at the dispersion of the values from stop to stop. As an example, the headway box-plots for the two peak periods for direction 1 are shown in figures 24 and 25. This visualization method allows for a quick inspection over the recorded headways. Each box-plot shows the range of the headways that were recorded in each stop (the orange boxes are limited by the upper and lower quartile values). The variability outside the upper and lower quartiles is seen through the vertical lines (whiskers) outside the boxes, which end up in the minimum and maximum value. Outliers are plotted in black and far outliers are plotted in grey.

The box-plots were plotted considering all the recorded headways (including the negative values), which explains the fact that the median value doesn't significantly increase along the line. The x-axis represents the stop numbers and the y-axis shows the headway values. Along the 33 stops, the variation around the median value increases (which is depicted by the size of the orange boxes). Likewise, the variation outside the upper and lower quartiles also increases (this can be seen in particular in the evening peak (figure 25), as the whiskers outside the boxes get longer along the stops. The median value in both morning peaks is around 8 minutes. The degree of dispersion in the data is clearly bigger than would be wanted, indicating poor headway adherence.

The operator can use these results to compare the data range of each box plot along the line with certain threshold lines by drawing two horizontal lines (a minimum and a maximum limit) to assess the proportion of headways that lie within that limits, as an indication of headway adherence.

In the context of low-frequency services, the box-and-whisker plots can be used to assess the schedule adherence, instead of the headway adherence, by studying the share of departures from stops that lie within a certain value around the scheduled departures. That can be done by plotting the schedule deviations (actual departure time minus scheduled departure time) for each stop and for a certain time period and by establishing a minimum and maximum limit for that deviation.

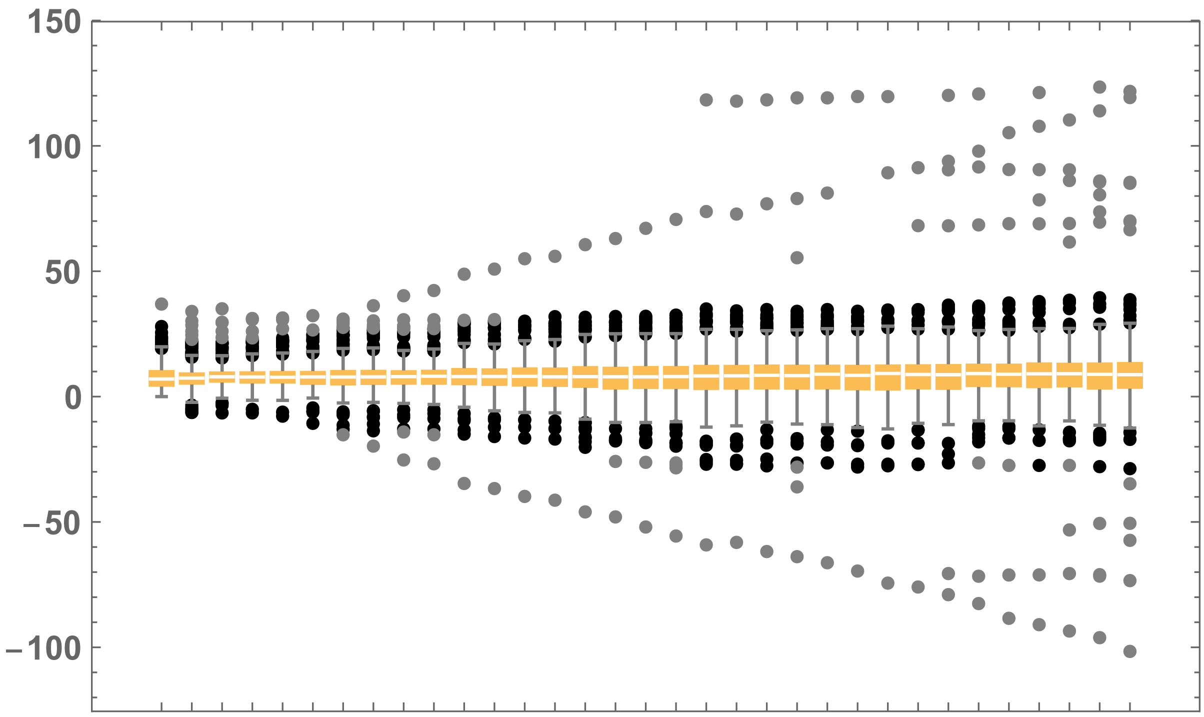


Figure 24 - Box-and-whisker plots for the headways in each stop (direction 1, 7h00-10h00)

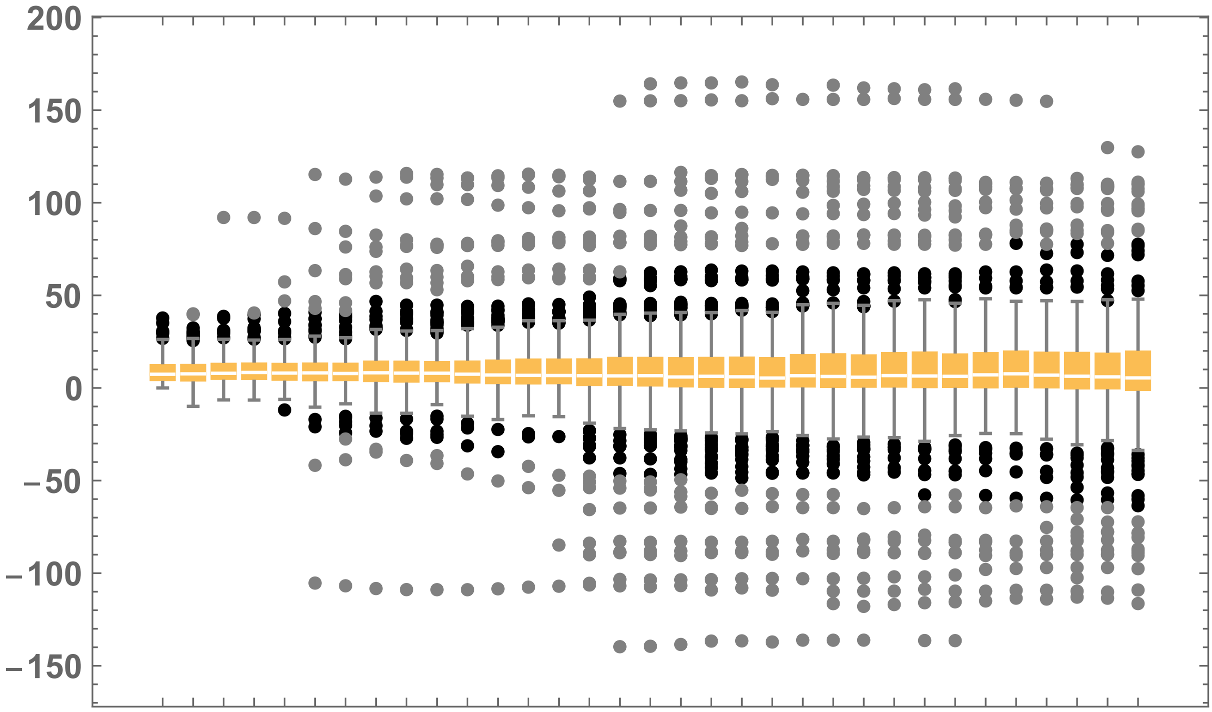


Figure 25 - Box-and-whisker plots for the headways in each stop (direction 1, 17h00-20h00)

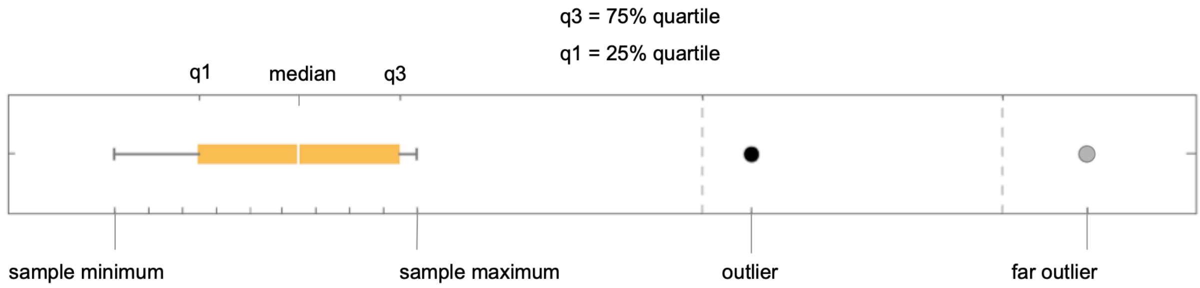


Figure 26 - Legend for figures 24 and 25

4.3.3. 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 $COV(h)$ represents the ratio between the standard deviation of the observed headways and the mean actual headway (section 2.2.4 for more details).

In this section, I will look at the variation of the $COV(h)$ amongst the analyzed days, for both directions, for the two peak periods of the day. To compute the $COV(h)$ for each day, I gathered the headways for every consecutive trip for every stop, for each day, considering the trips in each peak period.

According to (TCQSM 2nd Ed., 2003), six different levels of service (LOS) can be defined based on the values of $COV(h)$, as presented in table 2. The smaller the values of $COV(h)$, the more regular the service, and the less likely BB is to occur.

Figure 27 presents the extent to which regularity fluctuates from one day to the other in direction 1. From table 2, we can conclude that the levels of service in terms of regularity during the morning peak period range from E to F (the two worst levels, where BB is frequently expected), while during the evening peak period, the LOS is always F, the worst category in terms of regularity (TCQSM 2nd Ed., 2003). 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.

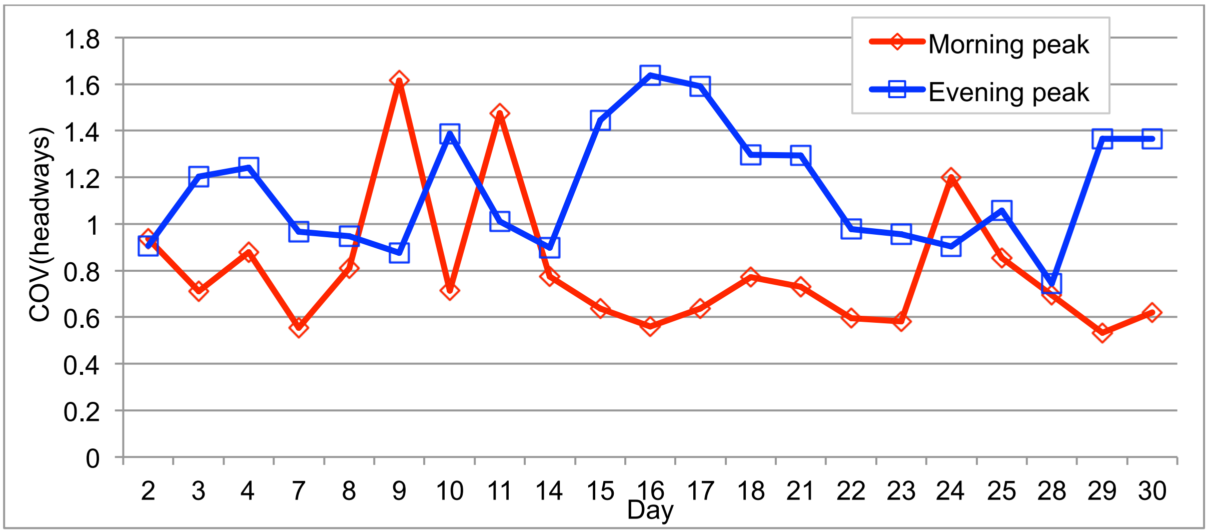


Figure 27 - Headway coefficient of variation (fluctuation throughout May for direction 1)

In figure 27, we can see that in general, the performance in terms of regularity is worse in the evening peak than in the morning peak. However, the same is not true for direction 2, as can be seen in figure 28. The fluctuations of regularity in direction 2 (figure 28) don't seem to have a relation with the fluctuations in direction 1 and, looking at direction 2, there is no clear relation between the changes in the morning period and in the evening period.

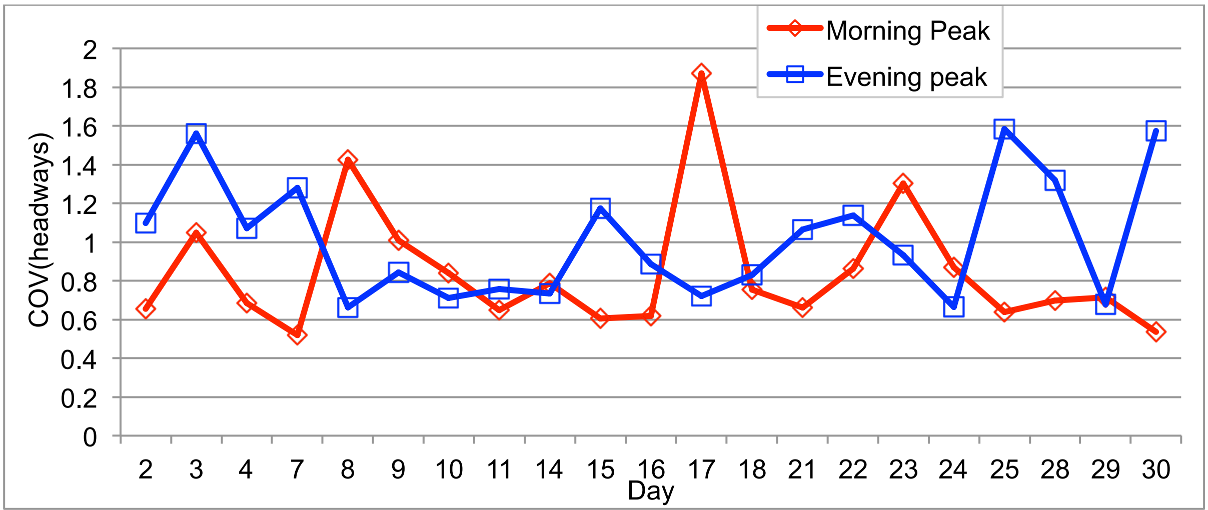


Figure 28 - Headway coefficient of variation (fluctuation throughout May for direction 2)

The performance in terms of regularity ranges from levels of service E to F. Therefore, the irregularity problems are common to both directions and periods of the day.

4.3.4. Headway coefficient of variation - fluctuation by stop

After looking at the variation of service regularity throughout the month of May, I also investigated the propagation of irregularity along the route, by computing the values of $COV(h)$ for each stop individually. To do so, I computed the average and standard deviation of the headways for every pair of consecutive buses passing in each stop over that time period.

In figure 29 we can see the variation of the $COV(h)$ along the line for direction 1 while in figure 30, the same information is shown for direction 2. Comparing the two figures, a pattern can be observed: the headway variability increases considerably along the line and buses arrive very irregularly in 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 (section 4.3.3).

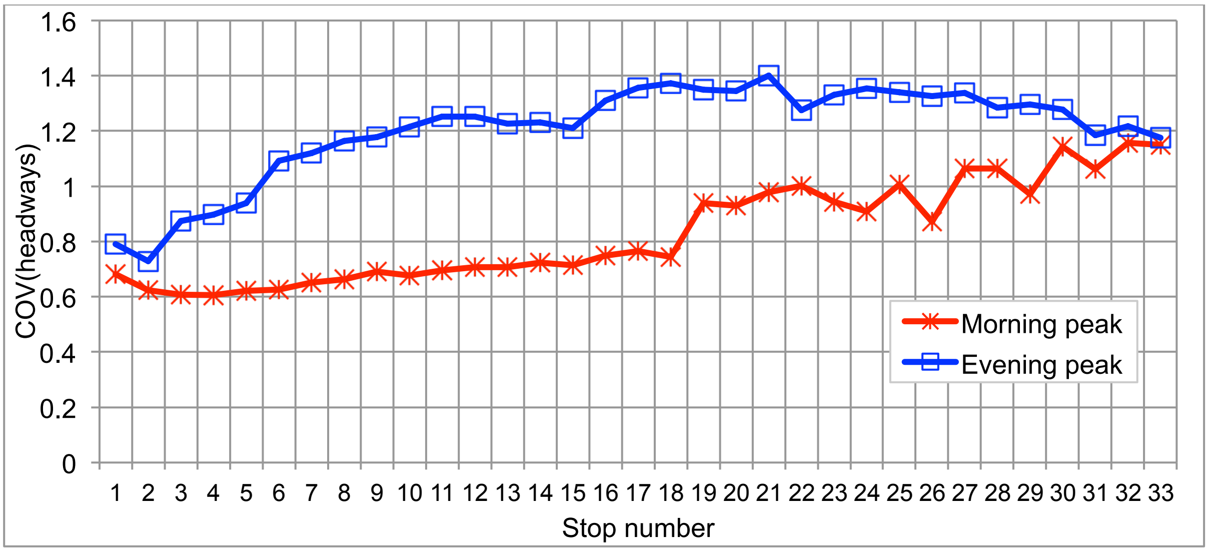


Figure 29 - Headway coefficient of variation (fluctuation along the line for direction 1)

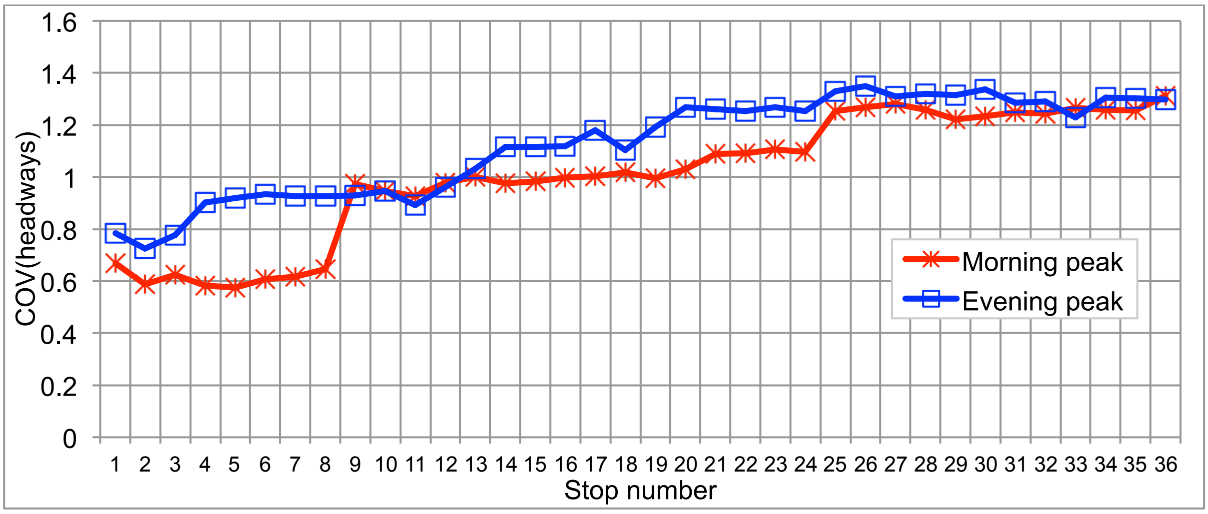


Figure 30 - Headway coefficient of variation (fluctuation along the line for direction 2)

Headway regularity in the last stop is important from the operator perspective as the operator can predict better when buses will become available. Even though the $COV(h)$ is a robust statistical

measure, it is not intuitive. Besides, in an ideal scenario, the mean headway in each stop would be similar from stop to stop. However, the mean headway in each stop also increases along the route (since only positive values are being considered). Thus, since the $COV(h)$ in each stop is normalized by the actual mean headway in that stop (which also varies), the dispersion of the headways may be further explored by looking at the standard deviation of the headways (which is presented in annex 9). There, it is possible to see that, for instance, for direction 1 and for the evening peak, the standard deviation on the first stop was 7.5 minutes while in the last stop the value was about 26 minutes (which represents a growth of about 250%). Thus, the dispersion of the headways around the average headway is much greater in the last stop than in the first stop. On the other hand, the $COV(h)$ varies from 0.8 to 1.2 (which represents a 50% growth), so by looking only at the $COV(h)$ we could have been led to conclude that the dispersion wasn't so considerable.
This indicates that the longer the distance, the more there's a possibility of a large headway delay and the more headways are expected to deviate from the scheduled headway.

4.3.5. Average passenger waiting time

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).
The excess waiting time (EWT) reflects the average extra time that passengers have to wait due to service reliability problems. If headways are perfectly even then $EWT = 0$. Hence, it is a more intuitive and informative regularity measure that takes the user perspective into consideration. The average EWT increases along the line as the service becomes more irregular (Cats, 2013). Hence, the extra waiting time depends on the stop at which one waits. However, in order to compute the EWT in each stop, detailed information is needed on the boarding rates on the stops and in this work, that information is not incorporated yet. That is left as a future development of the analysis method.
Nevertheless, I am still able to assess the effects of BB on the waiting times distribution through the computation of the average passenger waiting time ($E(waiting\ time)$), presented in section 2.2.4, which is expected to increase due to the irregularity (and in particular due to BB). Figures 31 and 32 show the results of this computation for directions 1 and 2, respectively.
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. In figure 32 it can be seen that the average passenger waiting times in the first stop for direction 2 and for the morning peak is close to that value. For the evening peak, however, the value almost doubles the scheduled headway.

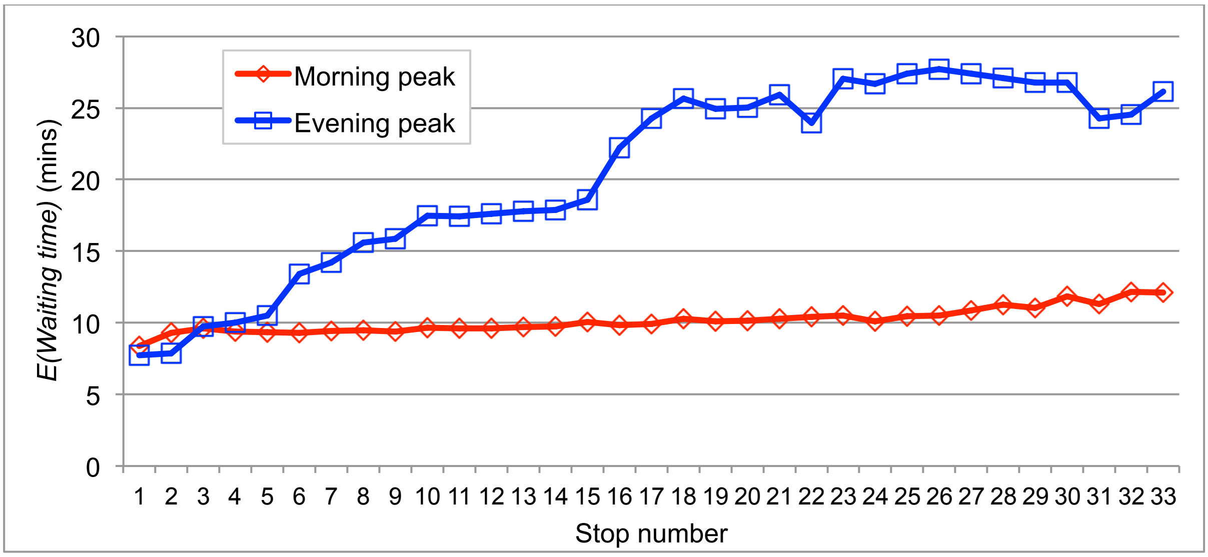


Figure 31 - Average passenger waiting time along the route (direction 1)

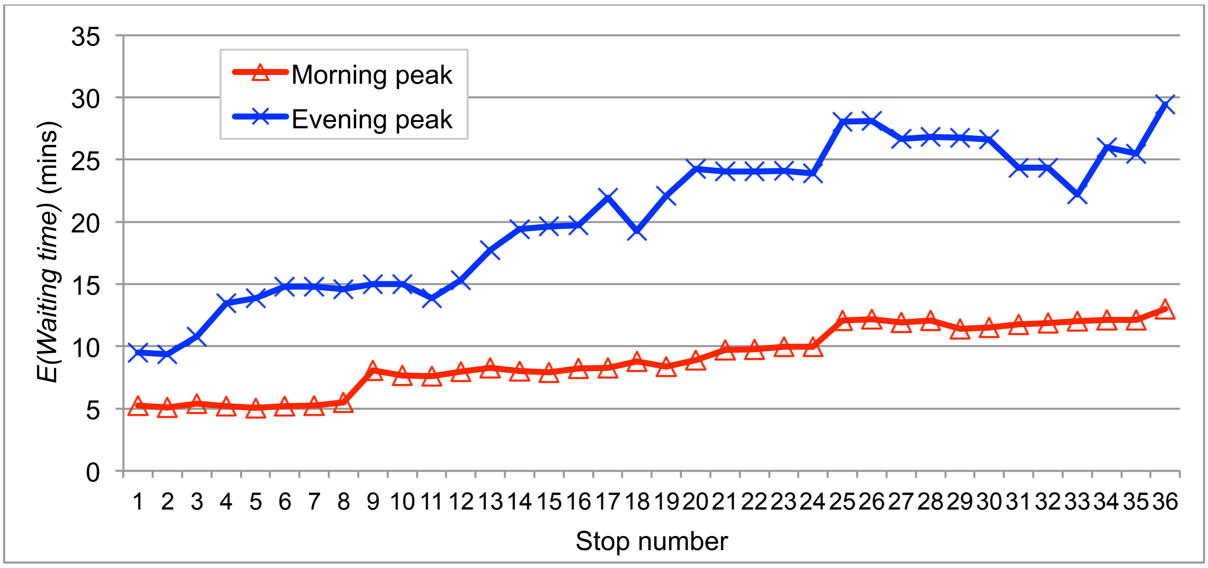


Figure 32 - Average passenger waiting time along the route (direction 2)

For direction 1 (figure 31), 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). In both figures, it can be seen that the propagation of the irregularity is felt more severely by passengers in the evening peak than in the morning peak, even though the $E(waiting\ time)$ deteriorates in both periods along the route. The peak period when the $E(waiting\ time)$ increased less along the route was the morning peak for direction 1, and even for that case, the $E(waiting\ time)$ in the last stop was about 12 minutes, which represent more than the double of the scheduled waiting time (there is a deviation of 7 minutes to the scheduled waiting time). In that same case, the $E(waiting\ time)$ in the stops in the middle of the route is about 10 minutes, which represents a deviation of 5 minutes to schedule. It may seem like not too much, but if I assume 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.

4.3.6. Total travel time and commercial speed

In addition to the headways and waiting time distributions, it is also worthwhile tracking the evolution of the total travel times, and consequently the commercial speed. Moreover, the assessment of commercial speed is particularly important when the operators are interested in testing the implementation of corrective or preventive strategies, because the application of measures to improve service regularity may slow down the service (Cats, 2013), which is not desirable. To illustrate the capabilities of the developed script, I display in figures 33 and 34 the distribution of total travel times (from the first to the last stop) for direction 1 and 2, respectively, considering the complete period of analysis (7h00-20h00). The x-axis represents the total travel times in 5 minutes bins whereas the y-axis represents the frequency in which those travel times were seen amongst the dataset.
One should note that all the trips are considered; therefore the results include trips that only cover a part of the route (such as the trips between Sete Rios and Amoreiras in direction 2 or other trips where a mechanical breakdown took place, for instance).

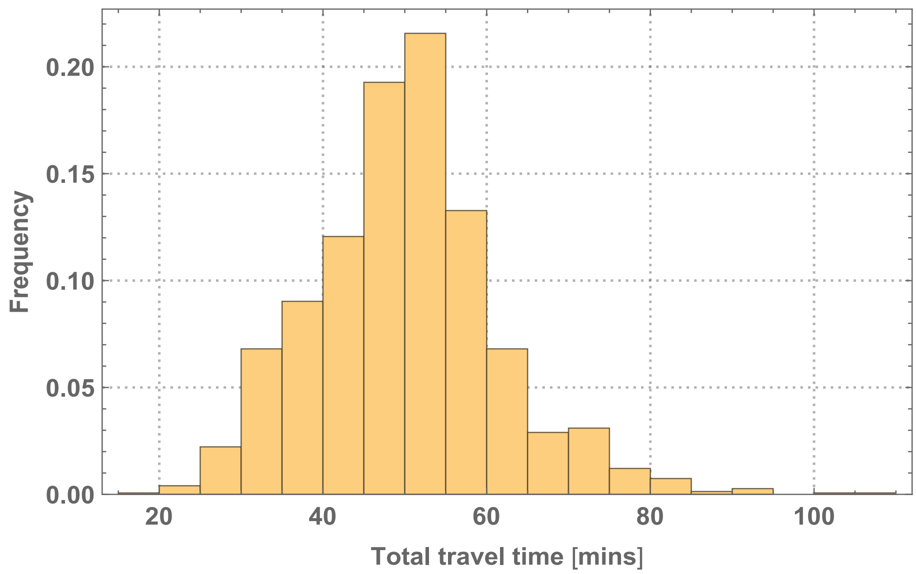


Figure 33 - Distribution of total travel times for direction 1 on line 758 (between 7h00 and 20h00)

The scheduled total travel time for the daytime service is 47 minutes for direction 1 and 53 minutes for direction 2, regardless of the period of the day. However, it is possible to conclude from figures 33 and 34 that the travel times vary considerably. For direction 1, more than 20% of the trips take between 55 to 65 minutes (which represents deviations of 8 to 18 minutes to the schedule) while, for direction 2, more than 20% of the trips take 60 to 70 minutes (which represents deviations of 7 to 17 minutes to the schedule). The same distributions can be obtained for specific times of the day.

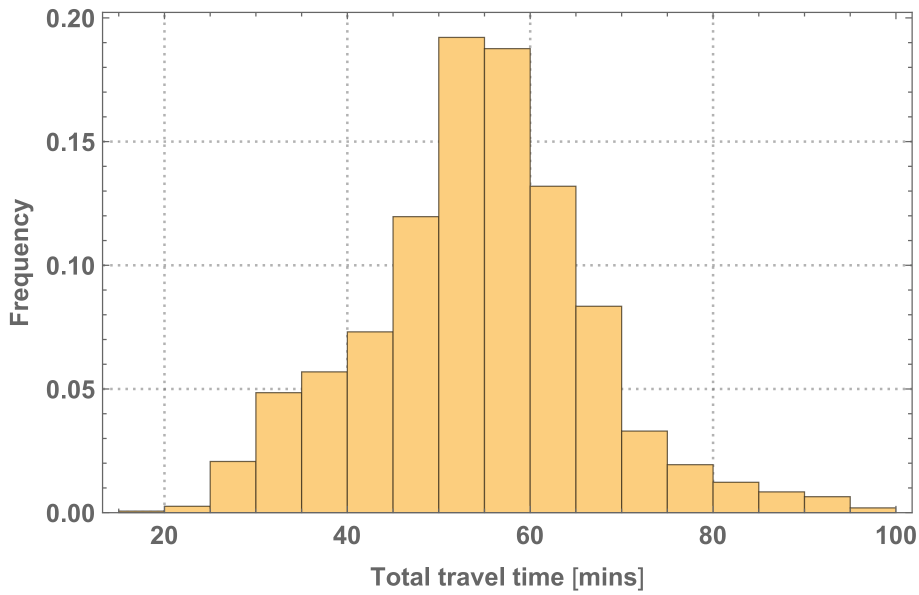


Figure 34 - Distribution of total travel times for direction 2 on line 758 (between 7h00 and 20h00)

In order to further evaluate the extension of these deviations, I computed the commercial speed for different periods of the day (tables 4 and 5). It is possible to see that the average commercial speed is under the average commercial speed of the complete network (which is 14 km/h, according to annex 2) for both directions. Direction 1 has a significant variation of the average commercial speed over the day while in direction 2, the average commercial speed remains close to 11 km/h, regardless of the period of the day that is considered. Also, the dispersion of the commercial speeds is wider in the evening peak periods than in the morning peak periods in both directions, which is consistent with the results from section 4.3.5).

Table 4 - Commercial speed [km/h] in direction 1 in line 758

	6h00 - 20h00	Morning peak	Evening peak
Average	12.4	12.7	10.7
Standard deviation	3.1	2.1	3.4

Table 5 - Commercial speed [km/h] in direction 2 in line 758

	6h00 - 20h00	Morning peak	Evening peak
Average	11.4	11.0	11.1
Standard deviation	3.1	3.2	3.9

5. Conclusions

5.1. Summary of findings

The ability to accurately generate and analyze operations performance measures is critical for transit agencies to ensure efficient transit operations and management (Figliozzi et al., 2012). This study used archived AVL data from Carris to develop an algorithm that identifies BB occurrences for each stop along both directions of bus line 758 during the month of May of 2018.

In section 4 of this dissertation, I analysed the results concerning the BB occurrences and looked at other key performance indicators to assess the regularity of the system. For the total of 4180 trips that were investigated in the regular weekdays through the GPS records, the script detected BB frequencies of 20 to 40%, in average, and there wasn't a specific weekday or direction in which BB detections were considerably more significant (considering the threshold of 2 minutes, close to a quarter of the planned headway, as indicated by TCQSM 2nd Ed., 2003). I looked at both the morning and the evening peak periods of demand and I saw that they were critical in terms of BB occurrences, which was also shown by the low performance in terms of regularity (LOS E-F regardless of the direction and period of the day).

Considering the propagation of irregularity, I saw that BB propagates along the route. As seen in (Cats, 2014), BB is known to function as a feedback loop in which a small deviation in the headways tends to aggravate further down the line, which was confirmed in section 4.2.3.

In terms of the stops where BB started more often, Sete Rios was the most problematic one and the section of the route between Sete Rios and Cais do Sodré was the most recurring in terms of BB formations, in particular when looking at direction 2 (from Portas de Benfica to Cais do Sodré).

Following the research of Figliozzi et al. (2012) that stated that most of the BB occurrences were due to deviations in the departure headway from the terminals, I investigated the influence of this variable in the BB detections and confirmed that the closer the buses depart from each other at the first stop, the more likely BB is to occur in the line. This was highlighted by looking at the stops were BB would start depending on the departure headway, where I concluded that BB was more likely to begin in the first stops of the route if the headways from the first stop were already irregular.

Moving on to a broader analysis of the regularity of the system, I saw that the shape of the distribution of the headways for the complete days (between 6h00 and 20h00) was very disperse and that around 25% of the recorded headways had extreme values (bigger than 20 minutes or smaller than 4 minutes). When looking at the peak periods, the evening had a worst performance than the morning period. I also confirmed that the variation of the headway values increased along the route, through the box-and-whisker plots of the recorded headways in every stop. Similarly, the headway coefficient of variation tends to increase along the line (the growth values are specified in section 4.3.4, depending on the considered period of the day). I also saw that, for instance, for the evening peak and considering direction 1 (from Cais do Sodré to Portas de Benfica), the standard deviation of the

headways increased almost 20 minutes along the route. However, the performance in the beginning of the route in terms of headway adherence was already poor (ranging from E to F, the two lowest levels according to TCQSM 2nd Ed., 2003).
Moreover, I computed the average passenger waiting times, which increased from 8 to 12 minutes along the route for the morning peak in direction 1 (and from 8 to 27 minutes in the evening peak). The propagation of irregularity was similar in direction 2.
At last, I looked at the total travel times of the investigated trips and concluded that the commercial speed in bus line 758 was below the average value given by Carris in its performance indicators. Carris indicates 14 km/h, whereas the computed commercial speed for the complete days in the trips made between 6h00 and 20h00 ranged from 11.4 to 12.4 km/h.

5.2. Key contributions

The work developed in this dissertation is of great value for Carris (and possibly for other transit operators) and its 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 that can be disaggregated for a single section of the route in a specific time-window;
- The development of a bus bunching detection model, based only on the AVL data: the model selects the files (vehicles) that need to be explored for the specific line and then orders the trips, comparing the headways between pairs of vehicles for flexible user-defined parameters (direction, time-window and detection threshold).

The visualization techniques presented in this dissertation help improve the generation and display of performance measures in the following ways:
- Well-developed visualized performance indicators are inherently easier to understand compared to tabular output, and attention can be drawn to specific stops or time periods of interest.
- The data visualization techniques demonstrated here provide thorough comprehension of how performance propagates over time and distance;
- The generation of the different visualization diagrams and graphs is automatic (the time-space diagrams, the several bus bunching statistics, the headway distribution, the box-and-whisker plots, the travel times and the average commercial speed). Only the headway coefficient of variation calculations and the average passenger waiting times were calculated only for the specific case study of bus line 758 and are still to be automated;
- The developed script can be easily applicable to other routes, by uploading the ID numbers of the vehicles that were affected to that bus line, as well as the information on the stops. The AVL data is in the same format for the different vehicles, however new issues can take place when cleaning the data for different routes due to route-specific particularities.

Besides the added value of the script developed to process the AVL data provided by Carris and generate results on the performance of the system, the value of the literature review made in section 2 cannot be disregarded. There, I saw how transit operators have assessed service reliability in the last years and that, in the context of high frequency services, regularity is more relevant than punctuality. In section 2.4.2, I investigated the different factors that have been reported to cause BB. Amongst these, I specifically evaluated the influence of the late (or irregular) departures from the terminal in section 4.2.5. Besides the causes, I reviewed the broad effects of BB in section 2.4.3. Amongst these, I computed the impact of BB in the average waiting times of the passengers, as well as the reduction of the commercial speed felt in line 758 in sections 4.3.5 and 4.3.6. I also studied the different preventive and corrective measures that have been followed by transit operators, the importance of real time operations management and the most recent modelling techniques that have been used to predict bus bunching, which are relevant for the application, refinement and extension of the work developed in this dissertation.

5.3. Recommendations and further work

There are several directions in which the work presented in this dissertation can be continued, that can be summarized as follows:

- The dataset - by including a larger dataset (including more days of operation or 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 - it can be used to generate passenger boarding and alighting patterns at each stop (even though Carris only collects the boarding information), and to generate passenger load profiles along the route. This would allow the measurement of the variations in passenger loads as a result of BB (the overcrowded leading bus is usually followed by an underutilized bus). Also, the capacity-utilization rate and seats-utilization rate can be computed from these, in order to assess the utilization of resources and the on-board conditions, respectively.
- Refinement of the BB detection model and extension of the regularity indicators with the dwell times at stops - currently, the precise evaluation of dwell times isn't possible since the GPS records are tracked every 30 seconds. The precise measurement of the dwell times would allow the drawing of statistics on the stops where buses spend more time standing still. Besides, the BB detection model would benefit from the precise arrival time stamp at each stop.
- Assessing the remaining causes of BB presented in section 2.4.2 (general traffic conditions and congestion; the presence of signalized intersections along the route; vehicle and driver availability; weather; driver behavior). This implies the existence of datasets containing these sources of unreliability and the development of a method to summarize causes of the identified BB incidents. The information on route configuration can be introduced to evaluate its impact on the BB detections.
- Developing a bus bunching prediction model in real-time that would recognize patterns in headway deviations and would indicate the probability of bus bunching 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.

- Testing the implementation of corrective actions (in particular holding and short-turning) based on RTI on service disruptions. Archived data may be used to analyse the impacts of these control strategy decisions, and to develop decision-support tools that enable supervisors to estimate the conditions under which these are most appropriate.
- 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). Between these stops, dedicated bus lanes were implemented (until the end of 2018), even though this section of the route doesn't seem to be the one where the route configuration influences more negatively the regularity of the service.

Apart from the directions in which the work of the dissertation can be extended, I also draw some recommendations on how Carris can take advantage of the work here presented in order to improve service reliability and therefore increase the ridership levels, namely by deploying preventive measures and assessing the differences in the BB detection results and in the key performance indicators before and after their implementation.

In chapter 4, I saw that most BB trips have origin between Cais do Sodré and Sete Rios, since this is the part of the line where the route configuration is more challenging: the narrow structure of the infrastructure closer to the more historical neighborhood (between Cais do Sodré and Rato), the fact that the bus operates almost always in mixed-traffic conditions and the presence of signalized intersections (also between Rato and Campolide) cause a variation in the cruising speed of the bus.

The implementation of dedicated or intermittent bus lanes should be useful for BB minimization, yet its implementation would be almost impossible in the more constrained parts of the route.

Traffic signal priority may also reduce running time variability by reducing delays at signalized intersections. Carris, in cooperation with the municipality of Lisbon, should implement conditional priority (providing a green phase to buses that are running late) in specific route sections such as:

- The intersection of Rua do Alecrim with Praça Luís de Camões;
- The traffic intersection in Largo de São Mamede;
- The intersection in Rato (by implementing green phase extension for approaching buses).

However, it should be noted that the impact of these changes in the green phases in the other directions should be accounted for in a traffic engineering study. It was also seen that Calhariz was a stop where buses had a dedicated lane, followed by a signalized intersection where buses were not given any priority (there were records of buses standing still for almost 2 minutes). A priority scheme should also be implemented here.

Furthermore, parking restrictions and the inspection on irregular parking in real time have to be seriously addressed: as for the current situation, it is common to find vehicles parked in illegal places (the scarce loading/unloading bays are often occupied and this leads to more illegal parking). Also, car drivers are not educated on the consequences of illegal parking on the vehicles' flow. In order to address this point, the focus should be on monitoring these events and applying more severe fines to drivers who disrespect the parking regulations. Illegal parking has a more negative impact on regularity in the areas of Príncipe Real (Rua da Escola Politécnica and Rua da Misericórdia) and in

Rua D.João V, between Rato and Amoreiras. Therefore, these sections should be the ones where entities such as EMEL should focus on.

Also, Carris must keep track of the deviations of the departure headways from Cais do Sodré and Portas de Benfica. Training the drivers and informing them about the importance of respecting the departure headways is essential, since the effects of BB occurrences in the first stops of the route will be broader, given the propagation of BB throughout a longer distance. Implementing incentives/penalties for drivers may help reducing the variability on drivers' work performance and its impact on the service.

The strategy on the number of spare vehicles and drivers is also important to improve regularity / reliability (in order to reduce gaps in service that also lead to bus bunching).

A more ambitious recommendation would be for Carris to test a new regularity-driven operation scheme in high frequency lines and assess its capability to mitigate bus bunching. 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 Cats, O. (2014). Headway control can be automated with computer-controlled signals and drivers can have continuous information on their performance with regards to regularity-oriented indicators. Consequently, drivers would depend less on the centralized control center, allowing for a quicker response to the dynamic changes in the transit service.

A regularity-driven operation implies the removal of as many schedule constraints as possible from the operational routine (turnaround stops, regulation stops and driver relief stops). I saw that in line 758, Sete Rios was the driver relief point and at the same time, it was an important bottleneck for BB occurrences. Therefore, I recommend Carris to locate the exchange of drivers in the departure stations (Portas de Benfica or Cais do Sodré) instead of in Sete Rios, avoiding further deviations of the travel times between consecutive vehicles along the route.

Another way to reduce the total travel time variability would be to reduce the dwell time variability. Carris can do so by allowing boarding and alighting in all the entrances of the vehicle, rather than in the front one, only. Through smart card technologies and pre-boarding ticketing systems, boarding can also become faster. Furthermore, to cater to the needs of elderly people and wheelchair users, which require additional boarding times, the stops must be rebuilt so that they have aligned platforms with the bus entrances. All of these components reduce the amount of randomness that can be added to a bus trip, which results in less control needed to provide service regularity.

A shift to regularity-driven operations scheme should also be accompanied by consistent real-time fleet management strategies. Nonetheless, even if the allocation of buses is not made dynamically and in real-time, Carris must guarantee that the buses inserted in the line in Sete Rios (to cope with the demand peak between Sete Rios and Amoreiras), where the regularity already tends to deteriorate, don't deteriorate it even more. That can only be ensured if the drivers have RTI on the headways from the preceding vehicles and start their trips in a way that maximizes headway evenness between consecutive trips.

References

Abkowitz, M. 1980. The Impact of Service Reliability on Work Travel Behavior. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Abkowitz, M., H. Slavin, R. Waksman, L. Englisher, N. Wilson. 1978. Transit Service Reliability: Final Report. Technical Report MA-06-0049-78-1, Urban Mass Transportation Administration.

Abkowitz, M., A. Eiger, I. Engelstein. 1986. Optimal control of headway variation on transit routes. J. Adv. Transportation 20(1) 73–88.

Albright, E., Figliozzi, M. 2012. Analysis of the Impacts of Transit Signal Priority on Bus Bunching and Performance, Proceedings of the Conference on Advanced Systems for Public Transport (CASPT), Santiago, Chile.

Barnett, A., 1974. On controlling randomness in transit operations. Transportation Sci. 8 102–116.

Barnett, A.,1978. Control strategies for transport systems with nonlinear waiting costs. Transportation Sci. 12 (2) 119–136.

Barnett, A., D. J. Kleitman. 1978. On two-terminal control of a shuttle service. SIAM J. Appl. Math. 35(2) 229–234.

Bartholdi III, J.J., Eisenstein, D.D., 2012. A self-coördinating bus route to resist bus bunching. Transportation Research Part B 46 , p. 481–491

Bellei, G., Gkoumas, K., 2010. Transit vehicles' headway distribution and service irregularity. Publ. Transport 2 (4), 269–289.

Bin, Yu, Zhongzhen, Y & Baozhen, Y. (2006) Bus Arrival Time Prediction: Using Support Vector Machines, Journal of Intelligent Transportation Systems, 10:4, 151-158

Bing-Hong Wang, (2016) Influence of synchronized traffic light on the states of bus operating system, Physica A: Statistical Mechanics and its Applications, Volume 453, Pages 9-23

Cats, O., 2013 RETT3—Final Report, A Field Experiment for Improving Bus Service Regularity. Available at: 〈http://www.ctr.kth.se/research.php?research=rett3〉

Cats, O., 2014. Regularity-driven bus operation: Principles, implementation and business models, Transport Policy 36, pages 223–230

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.

Cats, O. Mach Rufi, Ferran & Koutsopoulos, Haris. (2014). Optimizing the number and location of time point stops. Public Transport. 6. 215-235.

Cats, O., Larijani, A.N., Burghout, W., Koutsopoulos, H.N., 2011. Impacts of holding control strategies on transit performance: a bus simulation model analysis. Transp. Res. Rec. 2216, 51–58

Cats, O. & Loutos, G. (2015). Real-Time Bus Arrival Information System – An Empirical Evaluation. Journal of Intelligent Transportation Systems Technology Planning and Operations. 10.1080/15472450.2015.1011638.

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

Chen, Q., Adida, E. & Lin, J. 2013. Implementation of an iterative headway-based bus holding strategy with real-time information. Public Transp (2013) 4: 165-186

Daganzo, C.F., 1997. Schedule instability and control. In: Fundamentals of Transportation and Transportation Operations. Elsevier, New York, NY, pp. 304–309.

Daganzo, C.F., 2009. A headway-based approach to eliminate bus bunching: systematic analysis and comparisons. Trans. Res. B 43, 913–921.

Daganzo, C.F., Pilachowski, J., 2011. Reducing bunching with bus-to-bus cooperation. Transportation Research Part B 45 (1), 267–277.

Delgado, F., Muñoz, J.C., Giesen, R., Cipriano, A., 2009. Real-time control of buses in a transit corridor based on vehicle holding and boarding limits. Transportation Research Record (2090), 59–67

Disney, J. (1999). Customer satisfaction and loyalty: The critical elements of service quality. Total Quality Management, 10(4e5), 491-497.

Eberlein, X. J. 1995. Real-time control strategies in transit operations: Models and analysis. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Eberlein, X.J., Wilson, N.H.M., Barnhart, C., Bernstein, D. 1998. "The Real-Time Deadheading Problem in Transit Operations Control," Transportation Research (Part B), Vol. 32, No. 2, pp. 77-100.

Eberlein, X. J., X.J., Wilson, N.H.M., Bernstein, D. 2001. The holding problem with real-time information available. Transportation Sci. 35(1) 1–18.

Eichler, M., Daganzo, C.F. 2006 Bus lanes with intermittent priority: Strategy formulae and an evaluation, Transportation Research Part B: Methodological, Volume 40, Issue 9, Pages 731-744.

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

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)

Fonzone, 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.

Furth, P.G., Hemily, B.J., Muller, T.H.J., Strathman, J.G. 2003. "Uses of Archived AVL-APC Data to Improve Transit Performance and Management: Review and Potential'. TCRP Web Document H-28. Transportation Research Board. Washington, D.C.

Furth, P. 1995. A headway control strategy for recovering from transit vehicle delays. Transportation Congress: Civil Engineers—Key to the World's Infrastructure, vol. 2. 2032–2039

Gershenson C, Pineda L.A., 2009, Why Does Public Transport Not Arrive on Time? The Pervasiveness of Equal Headway Instability. PLoS ONE 4(10).

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

Hammerle, M., Haynes, M., & McNeil, S. (2005). Use of Automatic Vehicle Location and Passenger Count Data to Evaluate Bus Operations. *Journal of the Transportation Research Record*, 27-34.

Hensher, D. A., & Houghton, E. (2002). Performance-based quality contracts for the bus Sector: Delivering social and commercial value for money. Transportation Research Part B: Methodological, 38(2), 123-146.

Hensher, D. A., & Stanley, J. (2003). Performance-based quality contracts in bus service provision. Transportation Research Part A: Policy and Practice, 37(6), 519-538.

Hernández, D., Muñoz, J.C., Giesen, R., Delgado, F. 2015. Analysis of real-time control strategies in a corridor with multiple bus services, Transportation Research Part B: Methodological, Volume 78, Pages 83-105

Hickman, Mark D., 2001, An Analytic Stochastic Model for the Transit Vehicle Holding Problem. Transportation Science 35 (3):215-237.

Hofmann, M. and O'Mahony, M. 2005. The impact of adverse weather conditions on urban bus performance measures, *Proceedings. IEEE Intelligent Transportation Systems*, Vienna, pp. 84-89

Ibarra-Rojas, O.J., Delgado, F., et al., 2015. Planning, operation, and control of bus transport systems: a literature review. Transportation Research Part B 77, 38–75.

INE, Instituto Nacional de Estatística, IP (2017), Statistical Yearbook of Portugal 2016

INE, Recenseamento Geral da População, 2001 and 2011

Kieu, Le Minh, Bhaskar, Ashish, & Chung, Edward (2012) Benefits and issues for bus travel time estimation and prediction. In *Australasian Trans- port Research Forum 2011 Proceedings*, Perth, WA.

Kimpel, T., Strathman, J.G., Callas, S. 2004. "Improving Scheduling Through Performance Monitoring Using A VL and APC Data". Portland, OR.

Leong, W., Goh, K., Hess, S., Murphy, P. 2016. Improving bus service reliability: The Singapore experience. Research in Transportation Economics, Volume 59, Pages 40-49

Liang, S., Zhao, S. Lu, C., Minghui Ma. 2016. A self-adaptive method to equalize headways: Numerical analysis and comparison, Transportation Research Part B: Methodological, Volume 87, Pages 33-43

Lin, J., Ruan, M., 2009. Probability-based bus headway regularity measure, IET Intelligent Transportation Systems, Vol. 3, Iss. 4, pp. 400–408

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

Moreira-Matias L., Gama J., Mendes-Moreira J., Freire de Sousa J. (2014) An Incremental Probabilistic Model to Predict Bus Bunching in Real-Time. In: Blockeel H., van Leeuwen M., Vinciotti V. (eds) Advances in Intelligent Data Analysis XIII. IDA 2014. Lecture Notes in Computer Science, vol 8819. Springer, Cham

Moreira-Matias, L., Ferreira, C., Gama J., Mendes-Moreira J., Freire de Sousa J.(2012). Bus bunching detection by mining sequences of headway deviations. Advances in Data Mining. Applications and Theoretical Aspects, LNCS, vol. 7377., Springer, pp. 77–91.

Muller, Th.H.J. and P.G. Furth. Conditional bus priority at signalized intersections: better service with less traffic disruption. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 1731, Transportation Research Board of the National Academies, Washington, D.C., 2000, pp.23-30.

Newell, G.F., 1974. Control of pairing of vehicles on a public transportation route, two vehicles, one control point. Transportation Science 8 (3), 248–264.

Newell, G.F., Potts, R.B.,1964. Maintaining a bus schedule. In Proc., Second Conference Australian Road Research Board, Melbourne, volume 2, pages 388–393.

O'Dell, S., N. H. M. Wilson. 1999. Optimal real-time control strategies for rail transit operations during disruptions. N. H. M. Wilson, ed. Computer-Aided Transit Scheduling: Lecture Notes in Economics and Mathematical Systems No. 471. Springer-Verlag, Berlin, Germany. 299–323.

Oort, N. van, 2011, *Service Reliability and Urban Public Transport Design,* T2011/2, April 2011, TRAIL Thesis Series, the Netherlands

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, The relative importance of public transport trip time attributes in route choice, *Proceedings PTRC*, London

Oort, Niels & H. M. Wilson, Nigel & Nes, Rob. (2010). Reliability Improvement in Short Headway Transit Services Schedule- and Headway-Based Holding Strategies. Transportation Research Record: Journal of the Transportation Research Board. 2143. 67-76.

Osuna, E.E., Newell, G.F., 1972. Operational strategies for an idealized public transportation system. Trans. Sci. 6 (1), 52–72.

PAMUS-AML (2016), Plano de Ação de Mobilidade Urbana Sustentável da Área Metropolitana de Lisboa, VOL I - Relatório,

Pilachowski, J.M., 2009. An Approach to Reducing Bus Bunching. Dissertations & Theses - Gradworks.

Redman, L., Friman, M., Gärling, T., Hartig, T., 2013. Quality attributes of public transport that attracts car users: a research review. Trans. Policy 25, 119–127.

Schmöcker, J., Sun, W., Fonzone, A. Liu, R. 2016, Bus bunching along a corridor served by two lines, Transportation Research Part B: Methodological, Volume 93, Part A, Pages 300-317

Shen, S., N. H. M. Wilson. 2000. An optimal integrated real-time disruption control model for rail transit systems. Preprints from 8th Internat. Conf. on Computer-Aided Scheduling of Public Transport.

Shi-Gong Ge, Zhong-Jun Ding, Rui Jiang, Qin Shi, Reinhart Kühne, Jiancheng Long, Jian-Xun Ding, Wardman, M., Public transport values of time, Transport Policy 11 (2004) 363–377

Strathman, J., Dueker, K.J., Kimpel, T., 1999. Automated bus dispatching operational control and service reliability baseline analysis. Transp. Res. Rec. 1666, 28–36.

TCQSM 2nd Ed. (2003). Transit Capacity and Quality of Service Manual (TCQSM), Second Edition, Transportation Research Board, TCRP Report 100, Washington, DC.

Transport for London, 2012. London Buses—Quality of Service Indicators.

Transport for London. (2014). Travel in London: Report 7 [online] Available from: https://www.tfl.gov.uk/cdn/static/cms/documents/travel-in-london-report-7.pdf (Accessed March 12th 2018).

Transport for London. (2015). London's bus contracting and tendering process [online] Available from: http://content.tfl.gov.uk/uploads/forms/lbsl-tendering-and- contracting.pdf (Accessed March 12th 2018).

Turnquist M.A., Blume, S. W. 1980. Evaluating potential effectiveness of headway control strategies for transit systems. Transportation Res. Record 746 25–29.

Turnquist M.A. (1981) Strategies for improving reliability of bus transit service. Transp Res Rec 818:7–13

Viegas, J., Lu, B., 2004. The intermittent bus lane signals setting within an area. Transportation Research, Part C 12, 453–469.

Wilson, N.H.M., Nelson, D., Palmere, A., Grayson, T.H., Cedequist, C. 1992 "Service-Quality Monitoring for High-Frequency Transit Lines". Transportation Research Record 1349. Pages 3-11.

Xuan, Y., Argote, J., Daganzo, C.F. 2011. Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis, Transportation Research Part B: Methodological, Volume 45, Issue 10, Pages 1831-1845

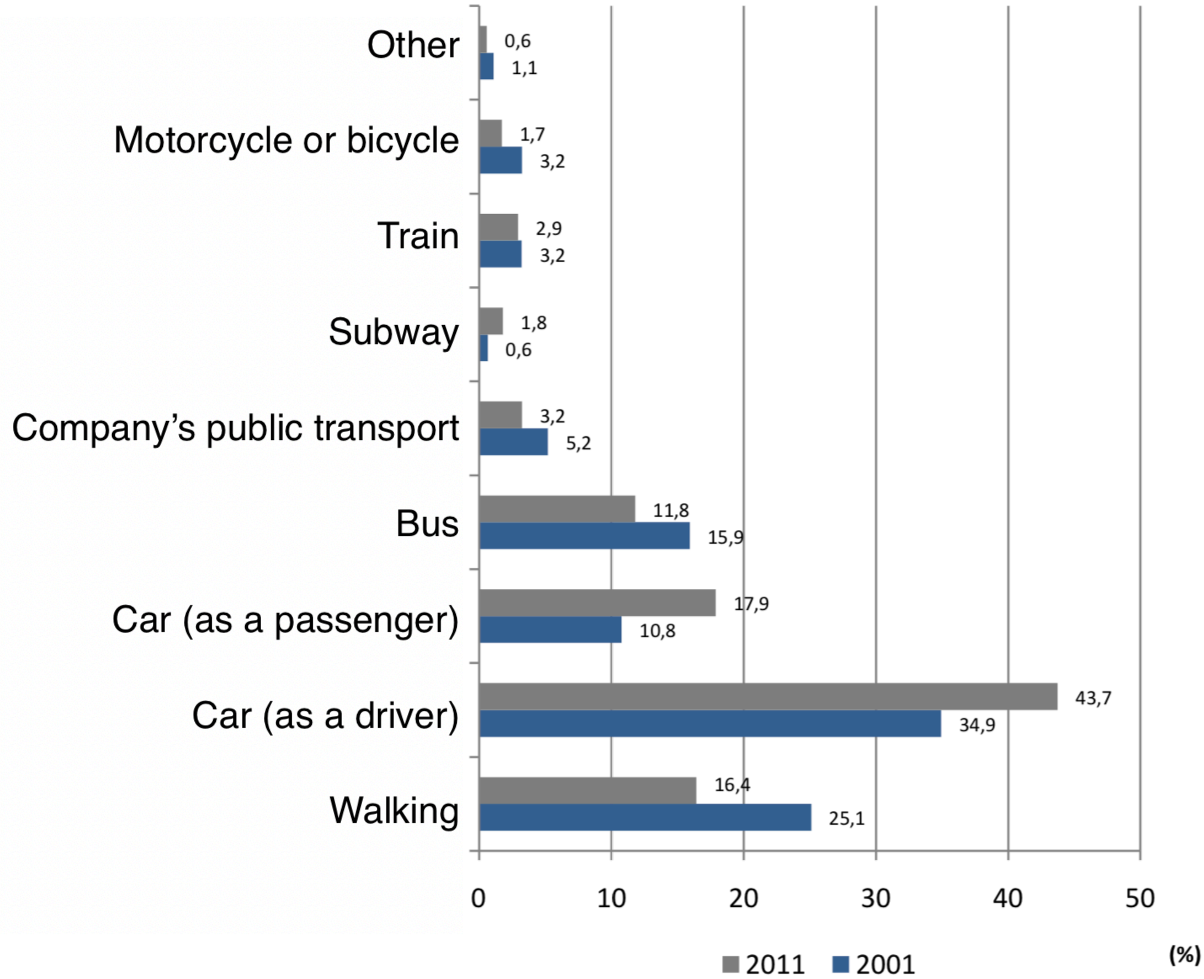
Zhao, J., Dessouky, M., Bukkapatnam, S., 2006. Optimal slack time for schedule-based transit operations. Transportation Science 40 (4), 529–539

Zhao, J., Bukkapatnam, S., Dessouky, M., 2003. Distributed architecture for real-time coordination of bus holding in transit networks. IEEE Transactions on Intelligent Transportation Systems 4 (1), 43–51

Zimmermann, L., Kraus, W., Koehler, L.A., Camponogara, 2016. E. Holding Control of Bus Bunching without Explicit Service Headways, IFAC-PapersOnLine, Volume 49, Issue 3,Pages 209-214

Annexes

Annex 1 - Percentage of transport modes used in the commuting trips in 2001 and 2011



Source: INE, Recenseamento Geral da População, 2001 and 2011

Annex 2 - Performance indicators for Carris operation between 2004 and 2016

Total public space (P.S.)	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Offer													
Vehicles x kms P.S.(millions)	42,2	40,1	39,7	40,6	41	41,6	41,4	38,2	34,5	32,8	31,8	30,7	28,8
Seats x kms P.S. (millions)	3.856	3.605	3.534	3.621	3.717	3.749	3.728	3.432	3.125	2.967	2.885	2.199	2.042
Demand													
Passengers P.S. (milions)	256,6	240,8	234,9	236,4	234,4	240,4	184,4*	183,7*	155,7*	149,7*	144,4*	144,8*	140,6*
Passengers x kms P.S. (milions)	850,5	796,3	775,8	768,9	758,3	786,1	613,6*	614,8*	510,9*	467,3*	448,2*	453,8*	440,4*
Quality													
Commercial speed (km/h)	14,5	14,5	14,4	14,3	14,4	14,3	14,4	14,4	14,4	14,3	14,2	14,2	14
Passengers-kms P.S. / seats-kms P.S.	22,1	22,1	22	21,2	20,4	21	16,5*	17,9*	16,3*	15,7*	15,5*	20,6*	21,6

*Values calculated based on a new methodology (with passenger information validations), starting in 2010

Source: www.carris.pt (visited on the 22nd of January of 2018)

Annex 3 - On-time performance levels

(Source: TCQSM 2nd Ed. (2003))

In the context of low-frequency routes, (TCQSM 2nd Ed. 2003) proposes levels of service for on-time performance, with "on-time" performance defined as 0 to 5 minutes late applied to either arrivals or departures. This means that both early arrivals (or departures) and arrivals (or departures) that happen 5 minutes after the scheduled are considered off-time.

Level of Service (LOS)	On-Time Percentage [%]
A	95 - 100
B	90 - 94,9
C	85 - 89,9
D	80 - 84,4
E	75 - 79,9
F	< 75

Annex 4 - Worst results for non-compliance rate (%) amongst the bus lines in the first trimester of 2018

Bus Line	January	February	March	Trimester
758	5,8	7,1	8,7	7,2
797	5,0	6,0	5,7	5,5
716	5,5	4,8	5,6	5,3
773	4,2	4,5	6,2	5,0
726	3,1	3,2	8,2	4,8
750	4,1	4,2	4,7	4,3
768	3,7	4,4	4,6	4,2
706	3,3	3,4	5,4	4,0
729	3,8	3,6	4,3	3,9
751	2,7	3,4	5,2	3,8
728	2,5	3,3	5,1	3,6
730	3,1	2,6	4,9	3,5
754	3,6	3,8	3,0	3,5
713	2,5	3,2	4,5	3,4
734	3,4	3,9	3,0	3,4
736	3,1	2,8	4,1	3,3
781	2,6	3,0	4,3	3,3
774	1,9	3,5	4,0	3,2
783	2,9	2,8	3,6	3,1
767	3,4	3,0	2,8	3,1
714	2,0	2,3	4,8	3,0

(the non-compliance rate compares the lost kms with the scheduled offered kms for each line)

Annex 5 - Bus bunching model script (main functions)

```
CoordinateConverter[lat_, long_] :=
 {lat / 60 000, -Mod[((4 294 967 295 - long) + 1), 4 294 967 295] / 60 000} // N
                 operação do módulo                                           v

Dist[{lat1_, lon1_}, {lat2_, lon2_}] :=
  Module[{R, latDistance, lonDistance, a, c, distance},
  módulo de código
   R = 6371; (* Radius of the earth *)
   latDistance = lat2 - lat1;
   lonDistance = lon2 - lon1;
   a = Sin[(latDistance / 2) Degree] * Sin[(latDistance / 2) Degree] +
       seno                    grau       seno                    grau
     Cos[lat1 Degree] * Cos[lat2 Degree] * Sin[(lonDistance / 2) Degree] *
     cosseno  grau      cosseno  grau      seno                    grau
      Sin[(lonDistance / 2) Degree];
      seno                  grau
   c = 2 ArcTan[Sqrt[a / (1 - a)]];
         arco ta⋯ raiz quadrada
   distance = R * c * 1000;
   Sqrt[distance ^ 2]
   raiz quadrada
  ];
```

```
RemoveLeiturasRepetidasErradas[w_, distRepReg_:20] := Module[{i, j, discard},
                                                      módulo de código
discard = {};
i = 1;
While[ i ≤ Length[w],
repete até q··· comprimento
For[j = 1, j + i ≤ Length[w] && w[[i, 1]] == w[[i + j, 1]], j++];
para cada                comprimento
If[j > 1,
se
Do[
repete
      If[Dist[CoordinateConverter @@ w[[i + k - 1, {2, 3}]],
      se
          CoordinateConverter @@ w[[i + k, {2, 3}]]] > distRepReg &&
        w[[i + k, 4]] ≠ "STOP_VOYAGE_REQ",
AppendTo[discard, i + k]], {k, 1, j - 1}]
adiciona a
];
i += j;
];
w[[Complement[Range[Length[w]], discard]]]
  conjunto co···  interv··· comprimento
];
```

```
LongParseTXT[path_, horaIn_, horaFin_, distRepReg_] := Module[{res, ConverteHoras},
                                                        módulo de código
(* Funcao que importa os ficheiros AVL txt, para o intervalo de hora pretendido *)
res = Quiet[Check[Import[path], -1]];
      silenc… verifica importa
If[res ≠ -1,
se
res = StringSplit[#, "|"] & /@ StringSplit[res, "\n"];
      subdivide cadeia de caracteres   subdivide cadeia de caracteres
res = Select[res, Length[#] ≥ 19 &];
      selecione   comprimento
(* Descartar linhas do AVL que não têm registos de posição,
   nao têm todas as colunas *)

res = res[[ ;; , {1, 15, 16, 2, 8, 9, 10, 11, 7}]];
(* ficar com as colunas da
    1hora, 15lat, 16long, 2tipoderegisto, 8chapa, 9viagem, 10sentido,
   11variacaosentido, 7carreira *)
   res = Map[Function[y, {StringSplit[y[[1]]][[2]], StringSplit[y[[2]]][[2]],
             apl… função            subdivide cadeia de caracteres subdivide cadeia de caracteres
    StringSplit[y[[3]]][[2]], StringSplit[y[[4]]][[2]], StringSplit[y[[5]]][[2]],
    subdivide cadeia de caracteres subdivide cadeia de caracteres subdivide cadeia de caracteres
       StringSplit[y[[6]]][[2]],
       subdivide cadeia de caracteres
    StringSplit[y[[7]]][[2]], StringSplit[y[[8]]][[2]], StringSplit[y[[9]]][[2]]}], res];
    subdivide cadeia de caracteres subdivide cadeia de caracteres subdivide cadeia de caracteres
(* extrair a informação que interessa de cada uma das colunas,
   ou seja ficar so com os numeros e descartar siglas *)
ConverteHoras[h_] :=
    Function[w, FromDigits[w[[1]]] * 3600 + FromDigits[w[[2]]] * 60 + FromDigits[w[[3]]]] @
    função      define o número por uma seqüê… define o número por uma se… define o número por uma seqü
     StringSplit[h, ":"];
     subdivide cadeia de caracteres
res =
    Map[Function[w, If[FromDigits[w[[2]]] == 0, Nothing,
    apl… função     se define o número por uma se… nada
       {ConverteHoras[w[[1]]], FromDigits[w[[2]]], FromDigits[w[[3]]]} ~ Join ~ w[[4 ;;]]]],
                               define o número por u… define o número por uma… junta
     res];
(* Mapear função que converte um triplo de strings <hora,lat,
   lon> em um triplo com o números respectivos, eliminando os registos lat=0 *)
res = If[Length[Flatten[res]] > 0,
      se compr… achatar
Select[res, (ConverteHoras[horaIn] ≤ #[[1]] ≤ ConverteHoras[horaFin]) &]
selecione
     (* Data das x h às y h *)
,
{}
];
```

```
posSTOP[data_, chapa_] :=
  Flatten[Position[data⟦ ;; , 4⟧,
  achatar  posição
    If[chapa == 21, "START_VOYAGE_REQ" | "STOP_VOYAGE_REQ", "STOP_VOYAGE_REQ"]]];
    se
```

```
SplitPerStop[data_, posSTOP_] :=
 (* SplitPerStop - função que vai dividir a informaçao de cada radio
   para todo o dia entre mudanças de sentido
      ou seja, vai pegar na informaçao do dia todo e dividir a informacao
  quando encontra uma posSTOP
         como já visto,
 posSTOP fica com as posições (linhas correspondentes) das paragens para
  inversão sentido   *)
 If[Length[posSTOP] == 0,
 se comprimento
  {},
  {data⟦ ;; posSTOP⟦1⟧⟧}~Join~If[
                       junta  se
    (* ⟦;;posSTOP⟦1⟧ é todas as linhas até encontrar a primeira posSTOP *)
    Length[posSTOP] == 1,
    comprimento
    {},
    Table[data⟦posSTOP⟦i⟧ + 1 ;; posSTOP⟦i + 1⟧ (* MUDAR: juntei -1 aqui*) - 1⟧,
    tabela
     {i, Length[posSTOP] - 1}]] ~Join~ {data⟦posSTOP⟦-1⟧ + 1 ;;⟧}
         comprimento                junta
 ]
(* se tiver 1 posSTOP → divido em duas listas. a Table vai da posSTOP i ao i+
  1 e junta elementos quando aparece uma posSTOP nova,
criando assim uma divisao dos elementos entre cada posSTOP *)
```

```
CheckSentido[w_] := Module[{counts},
                     módulo de código
(* CheckSentido - funçao que vai servir para perceber em que sentido
     é que a informação que já está dividida entre cada paragem de percurso está.
    ou seja, se não houver registos (MissingQ vê se falta um elemento)
     sentido 1 e maioria contagens 2 →
    entao sentido 2
     (as contagens sentido 0 vao ficar agrupadas a um dos dois sentidos,
    é irrelevante pq isso é quando o autocarro está parado ou ainda não
      está dentro do percurso)  *)
   counts = CountsBy[w, #[[7]] &];
                contagens por
   If[MissingQ[counts["1"]] && MissingQ[counts["2"]],
   se  dados ausentes?                dados ausentes?
    3,
    If[MissingQ[counts["1"]],
    se  dados ausentes?
If[! MissingQ[counts["0"]],
se     dados ausentes?
      2 + Boole[counts["0"] > counts["2"]],
          função booleana
2],
     If[MissingQ[counts["2"]],
     se  dados ausentes?
      1,
      1 + Boole[counts["1"] < counts["2"]]
          função booleana
(* 1 + true = 1+1 = sentido 2
   1 + false = 1+0 = sebntido 1 *)
     ]
    ]
   ]
  ];
```

```
CheckCarreira[w_, numCarreira_: 758] := Module[{counts},
                                       módulo de código

   counts = CountsBy[w, #[[9]] &];
               contagens por
    (*O número de vezes que numCarreira 758 aparece tem de ser mais de
     metade das vezes *)
If[MissingQ[#], 0, #] &@counts[ToString[numCarreira]]
se  dados ausentes?                converte em cadeia de caracteres
     (*+If[MissingQ[#],0,#]&@counts[ToString[0]]*) > Length[w] / 2
                                                      comprimento
  ];
```

```
AllBusFromDaySentidoCarreira[Radios_, horaIn_, horaFin_, sentido_, anchor_,
  distFun_, plotOnOff_, dia_] :=
Module[{subRadios, aux, cores, plots, posSTOPs, datas, SplitPerStops,
módulo de código
   dadosGroupBySentido, vals, path, chapa, legs, radios, filter},
(*path=SystemDialogInput["Directory",
     WindowTitle→"Selecione a Pasta com os Ficheiros"];*)
path = "/Users/joaoalmeida/Desktop/LogsSAEMaio/20180" <>
    If[dia == 31, "601", "5" <> If[dia < 9, "0", ""] <> ToString[dia + 1]] <> "/";
    se                      se                      converte em cadeia de caracteres
Print[path];
apresenta o resultado
radios = Radios⟦dia⟧;
If[path == Null, Print["Erro: Pasta errada!"]; Abort[]];
se        expr···  apresenta o resultado         interrompe
datas = Table[LongParseTXT[path <> "UM " <> ToString[d] <>
        tabela                                  converte em cadeia de caracteres
".txt", horaIn, horaFin, 20], {d, radios}];
filter = Complement[Range[Length[radios]],
         conjunto co···  interv··· comprimento
    Flatten[Position[datas, u_ /; u == -1]]];
    achatar  posição
radios = radios⟦filter⟧;
datas = datas⟦filter⟧;
chapa = Table[If[MemberQ[ToExpression[el⟦ ;; , 5⟧], 21], 21, 0], {el, datas}];
        tabela se  está incl··· converte em expressão
posSTOPs = Table[posSTOP[datas⟦i⟧, chapa⟦i⟧], {i, Length[datas]}];
           tabela                                  comprimento
  SplitPerStops =
   Table[If[! CheckCarreira[#] || CheckSentido[#] == 3, Nothing, #] & /@
   tabela  se                                          nada
       SplitPerStop[datas⟦i⟧, posSTOPs⟦i⟧], {i, Length[datas]}] ;;
                                                  comprimento
dadosGroupBySentido =
    Map[Function[y, Select[y, (CheckSentido[#] == sentido && CheckCarreira[#]) &]],
    apl·· função      selecione
     SplitPerStops];
cores = Table[RGBColor[Random[], Random[], Random[]],
        tabela cores do sistema RGB
    {j, Length[dadosGroupBySentido]}];
        comprimento
dadosGroupBySentido = Select[dadosGroupBySentido, Length[#] > 0 &];
                      selecione                  comprimento


vals = Table[(distFun[datas⟦j⟧, anchor]), {j, Length[datas]}];
       tabela                                  comprimento
If[plotOnOff,
se
aux = Table[
      tabela
ListLinePlot[Table[(# / {3600, 1}) & /@ distFun[dadosGroupBySentido⟦j, i⟧,
gráfico de linha··· tabela
                 anchor], {i, Length[dadosGroupBySentido⟦j⟧]}],
                           comprimento
                 PlotStyle → {cores⟦j⟧}, InterpolationOrder → 1,
                 estilo do gráfico        ordem de interpolação
      PlotLegends → {("Radio " <> ToString[dadosGroupBySentido⟦j, 1, 1, -1⟧])},
      legenda do gráfico         converte em cadeia de caracteres
      GridLines → Automatic, FrameLabel → {"Time [hours]", "Distance [meters]"},
      grade de linhas automático  legenda do quadro
      Frame → True, GridLinesStyle → LightGray]
      quadro  verd···  estilo de linhas de g··· cinza claro
            , {j, Length[dadosGroupBySentido]}];
                  comprimento
If[Length[aux] > 0,
```

```
If[plotOnOff,
se
aux = Table[
      tabela
ListLinePlot[Table[(# / {3600, 1}) & /@ distFun[dadosGroupBySentido⟦j, i⟧,
gráfico de linha··· tabela
                    anchor], {i, Length[dadosGroupBySentido⟦j⟧]}],
                                 comprimento
                    PlotStyle → {cores⟦j⟧}, InterpolationOrder → 1,
                    estilo do gráfico            ordem de interpolação
        PlotLegends → {("Radio " <> ToString[dadosGroupBySentido⟦j, 1, 1, -1⟧])},
        legenda do gráfico           converte em cadeia de caracteres
        GridLines → Automatic, FrameLabel → {"Time [hours]", "Distance [meters]"},
        grade de linhas automático   legenda do quadro
        Frame → True, GridLinesStyle → LightGray]
        quadro  verd···  estilo de linhas de g···  cinza claro
              , {j, Length[dadosGroupBySentido]}];
                    comprimento
If[Length[aux] > 0,
se comprimento


Print[Show[aux, PlotRange → All]];
apres··· mostra  intervalo do g··· tudo
(*Export["~/Desktop/ex3.pdf",Show[aux,PlotRange→All]];*)


];
Print[Show[Table[ListLinePlot[vals⟦j⟧, PlotStyle → {cores⟦j⟧},
apres··· mos··· tabela  gráfico de linha de uma lista ···  estilo do gráfico
          InterpolationOrder → 1, PlotRange → Full,
          ordem de interpolação     intervalo do g··· completo
          PlotLegends → {"Bus " <> ToString[radios⟦j⟧]}, PlotTheme -> "Detailed"],
          legenda do gráfico        converte em cadeia de caract···  tema do gráfico
         {j, Length[datas]}] ~ Join ~
             comprimento          junta
        Table[ListPlot[vals⟦j, posSTOPs⟦j⟧⟧,
        tabela  gráfico de uma lista de valores
PlotStyle → {Darker[Red, Random[]]}, PlotMarkers → Style["×", 16]],
estilo do gráfico  mais e···  vermelho     marcadores do ···  estilo
         {j, Length[datas]}], PlotRange → All]];
             comprimento          intervalo do g··· tudo
];
{(* Primeiro elemento: lista de dados separados por rádio e separados
        por viagem *)
Map[{#⟦1⟧} ~ Join ~ CoordinateConverter[#⟦2⟧, #⟦3⟧] ~ Join ~ (#⟦4 ;;⟧) &,
aplica-se ao p··· junta                                     junta
     dadosGroupBySentido, {3}],
(* Segundo elemento: lista com as respectivas distâncias dos pontos à âncora *)
Table[distFun[dadosGroupBySentido⟦j, i⟧, anchor], {j, Length[dadosGroupBySentido]},
tabela                                                comprimento
     {i, Length[dadosGroupBySentido⟦j⟧]}],
         comprimento
radios
} (* Devolve par < dados por sentido,
  distancias à âncora > numa correspondência 1 para 1 *)
]
(* A função responde com estes dois elementos *)
```

```
DistByTimeToCoord[data_, pontoIn_] :=
  {First /@ data, (Dist[CoordinateConverter[pontoIn[[1]], pontoIn[[2]]],
   primeiro
CoordinateConverter[#[[2]], #[[3]]]] & /@ data)}^T;
(* DistByTimeCoord - função que calcula a distancia ao ponto inicial (garagem),
mantendo o registo desse tempo. vai
dar pares de <tempo,distancia ao ponto inicial> que depois são usados no graf   *)
DistByTimeToCoordProg[data_, pontoIn_] := {First /@ data, Accumulate[
                                            primeiro        acumula
(* DistByTimeCoordProg - função que calcula a distancia ao ponto
       inicial (garagem) para o primeiro registo e depois
calcula a distancia em relaçao aos registos anteriores,
     para os outros registos. vai dar pares < tempo, distancia acumulada >   *)
Table[
tabela
If[i == 1,
se
Dist[CoordinateConverter[pontoIn[[1]], pontoIn[[2]]],
        CoordinateConverter[data[[i, 2]], data[[i, 3]]]],
Dist[CoordinateConverter[data[[i - 1, 2]], data[[i - 1, 3]]],
        CoordinateConverter[data[[i, 2]], data[[i, 3]]]]
], {i, Length[data]}]]}^T;
      comprimento
```

```
PosMinDistToStop[viagem_, stopCoord_, distTol_] := Module[{aux, pos},
                                                    módulo de código
   (* Rotina que recebe os dados de uma viagem,
   as coordenadas de uma paragem e a distância mínima que se considera
    para que se considerar que houve um registo na paragem,
   e devolve a posição correspondende ao ponto mais próximo da viagem à paragem,
   se este for inferior à distância mínima disTol ou
    -1 caso esse ponto não exista *)
   (* RESOLVE APENAS PARA 1 PARAGEM. Vai ver os dados de cada viagem e
    vai ver qual o ponto nessa viagem (nos registos de 30 em 30s) que
    se aproxima mais dessa paragem,
   para que se tenha um tempo aproximado de cada viagem na paragem.*)
   aux = Table[Dist[viagem[[i, {2, 3}]], stopCoord], {i, Length[viagem]}];
         tabela                                          comprimento
   pos = First[Ordering[aux, 1]];
         prime··· ordenação
   If[aux[[pos]] ≤ distTol,
   se
    pos,
    -1
   ]
  ];
LocateStops[data_, dataStops_, radio_, minDist_] :=
 Table[PosMinDistToStop[data[[r, viagem]], dataStops[[it, {2, 3}]], minDist],
 tabela
  {r, 1, Length[data]}, {viagem, 1, Length[data[[r]]]}, {it, Length[dataStops]}]
        comprimento             comprimento                 comprimento
```

```
DetetaBunching[w_List, tempoMax_, NumMinRegis_:0, NumMinBunch_:0] :=
 Module[{ExisteBunching, r, aux},
 módulo de código
  (* recebe uma lista w, que é uma lista de listas. lista para 1 dia e
   1 sentido com todos os rádios. em cada rádio,
  temos várias listas (viagens). em cada viagem temos o tempo de passagem
   nas paragens. *)
  (* recebe tempoMax, que é o tempo definido como critério bunching,
  em segundos *)

  r = {};
  ExisteBunching[timeDiffs_, t_, r1_, r2_, v1_, v2_] :=
   If[Length[Flatten[Select[timeDiffs, IntegerQ[#] && Abs[#] < t &]]] > NumMinBunch &&
   se  compr…  achatar  selecione          número inteiro?  valor absoluto
     Length[Flatten[Select[timeDiffs, IntegerQ[#] &]]] > NumMinRegis
     compr…  achatar  selecione          número inteiro?
    , {timeDiffs, {{r1, v1}, {r2, v2}}}, {}];
  (* ExisteBunching é uma função auxiliar que recebe todos aqueles
   elementos. vai ver que timeDiffs são inteiras
   (para não ver os 0.0 quando não há registo nas paragens) e quando
   existir pelo menos 1 valor absoluto das timeDiffs inferior ao limite
   definido t na lista,
  então guarda uma lista com {timeDiffs e os respetivos pares radio,
    viagem entre os quais houve bunching} *)

  Do[
  repete
   aux = ExisteBunching[Map[First, w[[r1, v1]]] - Map[First, w[[r2, v2]]], tempoMax,
                        apl… primeiro              apl… primeiro
     r1, r2, v1, v2];
   If[Length[aux] > 0, AppendTo[r, aux]];
   se  comprimento         adiciona a

   ,
   {r1, 1, Length[w]},
          comprimento
   {r2, r1 + 1, Length[w]}, {v1, 1, Length[w[[r1]]]}, {v2, 1, Length[w[[r2]]]}];
                comprimento              comprimento                  comprimento
  r
  (* compara cada par rádio,viagem com os outros, sem repetir comparações.
    no final, r devolve todas as {timeDiffs com os pares {{r1,v1},{r2,v2}} *)
 ]
```

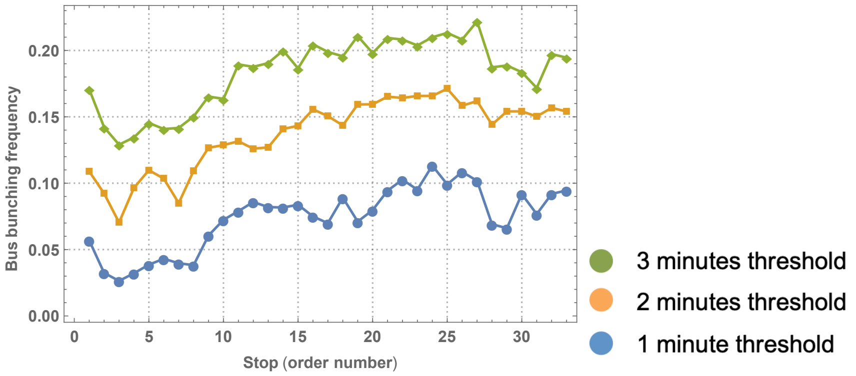
```
SortByFirstRegister[w_] :=
  Module[{nZeros, wZeros, wNonZeros, first, i, in, fim, m, tempo},
  módulo de código
   For[nZeros = 0, ! IntegerQ[w[[nZeros + 1, 1, 1]]] && nZeros < Length[w], nZeros ++];
   para cada                  número inteiro?                                  comprimento
   If[nZeros > 0,
   se
    wZeros = w[[ ;; nZeros]];
    wNonZeros = w[[nZeros + 1 ;;]];


    While[Length[wZeros] > 0,
    repete·· comprimento
     first = wZeros[[1]];
     For[i = 1, ! IntegerQ[first[[i, 1]]], i ++];
     para cada    número inteiro?
     tempo = first[[i, 1]];
     in = 1;
     fim = Length[wNonZeros];
           comprimento
     While[fim > in,
     repete até que não retorne um valor verdadeiro
      m = IntegerPart[(fim + in) / 2];
          parte inteira
      If[wNonZeros[[m, i, 1]] > tempo,
      se
       fim = m - 1,
       If[wNonZeros[[m, i, 1]] < tempo,
       se
         in = m + 1,
         fim = in = m
        ];
      ];
     ];
     wNonZeros = Insert[wNonZeros, first, in];
                 insere
     wZeros = Rest[wZeros];
              todos exceto o primeiro
    ];
    wNonZeros
    ,
    w
   ]
  ];
```

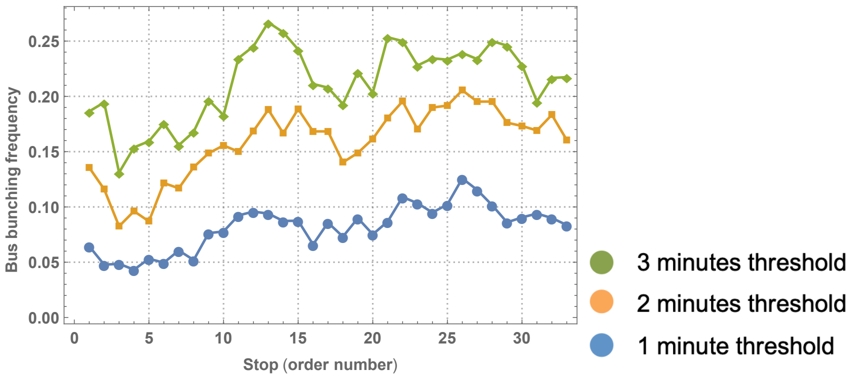
```
FindFirstBunchingStop[bunchings_, paragens758_, valsDataOrdStops_,
   paragens758Sentido_, time_] := Module[{b1, b2, c1, bunchFirstStop},
                                  módulo de código
   {b1, b2} = {(First /@ #), Last /@ #} &@bunchings;
               primeiro      último
   c1 = Table[If[Abs[b1⟦i, j⟧] < time, j, Nothing], {i, Length[b1]},
        tabela se valor absoluto        nada           comprimento
     {j, Length[b1⟦i⟧]}];
         comprimento
   bunchFirstStop =
    Table[paragens758Sentido⟦
    tabela
      x⟦First[Ordering[First /@ valsDataOrdStops⟦1, 1, x⟧, 1]]⟧⟧, {x, c1}];
        prime··· ordenação primeiro
   Table[{x, Select[paragens758, IntegerPart[#⟦4⟧] == x &]⟦1, -4⟧},
   tabela    selecione            parte inteira
    {x, bunchFirstStop}]
  ];
```

Annex 6 - Bus bunching along the route in the peak periods

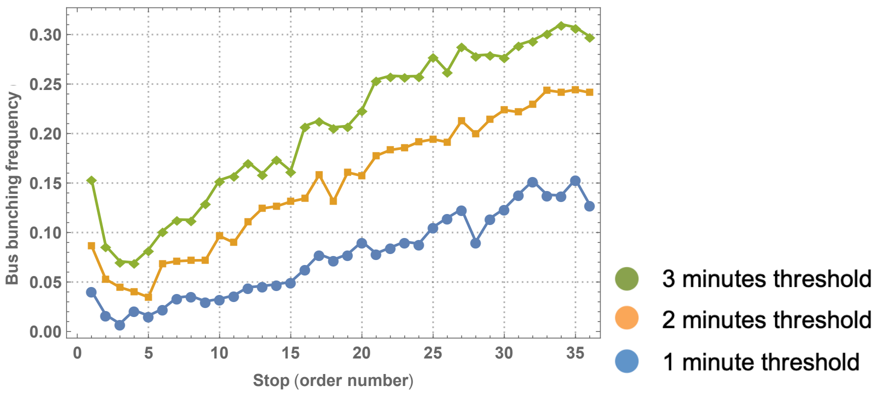
Direction 1, morning peak:



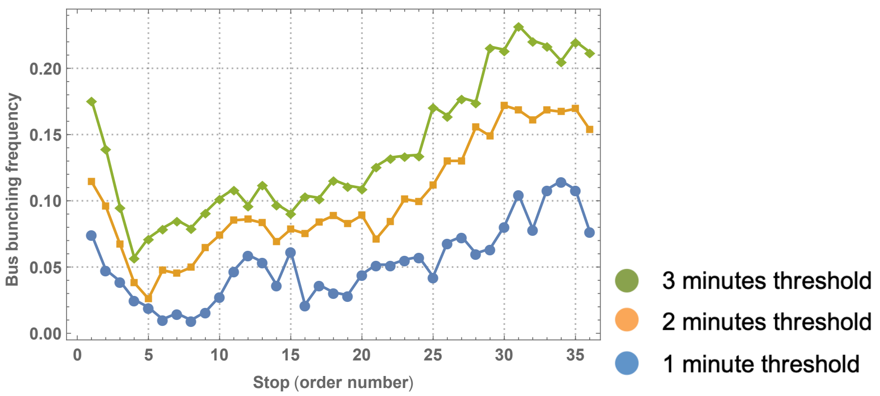
Direction 1, evening peak:



Direction 2, morning peak:



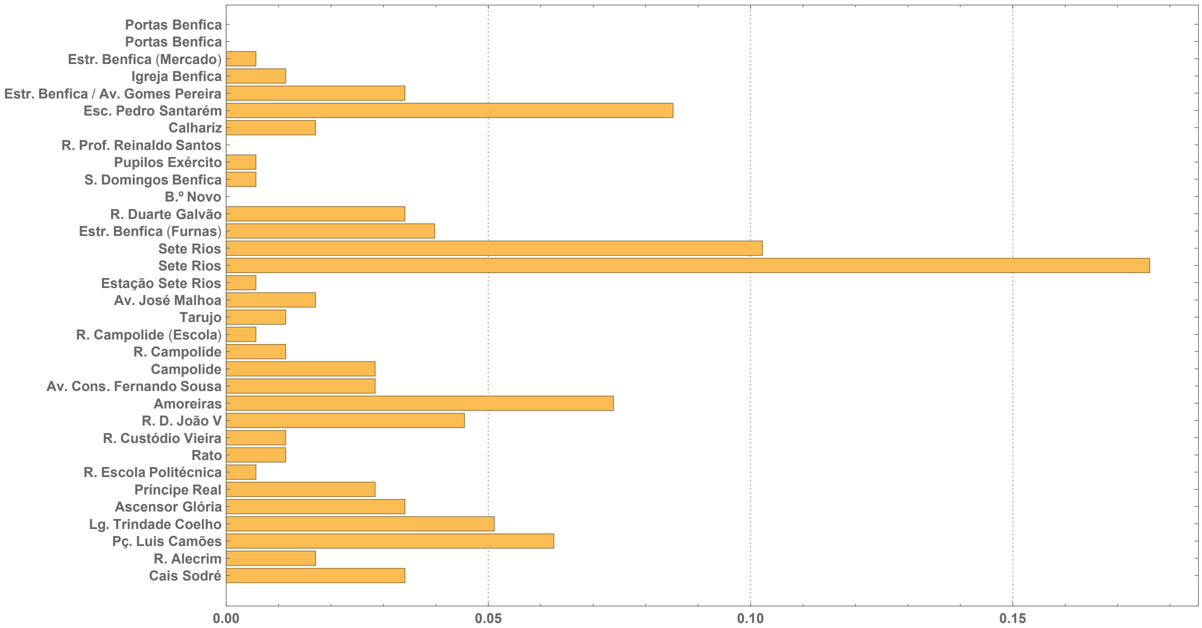
Direction 2, evening peak:



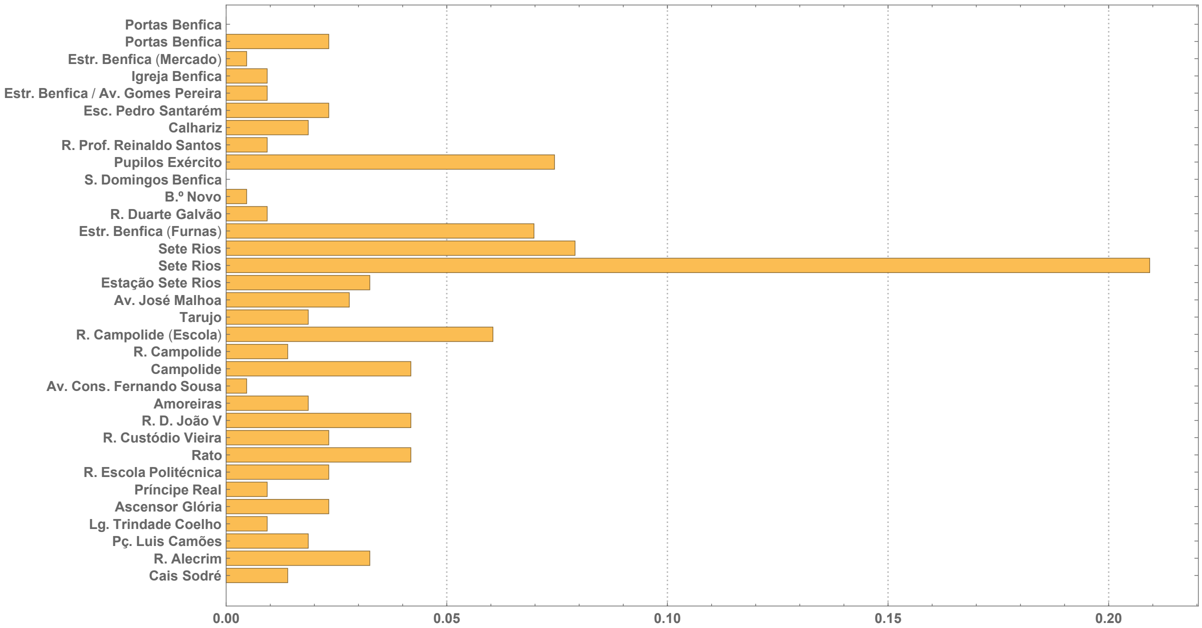
Annex 7 - Bus bunching starts in the peak periods

In these charts, the y-axis represents the ordered stops (from Cais do Sodré to Portas de Benfica) while the x-axis represents the frequencies in which the BB started in that specific stop

Direction 1, morning peak (threshold of 2 minutes)

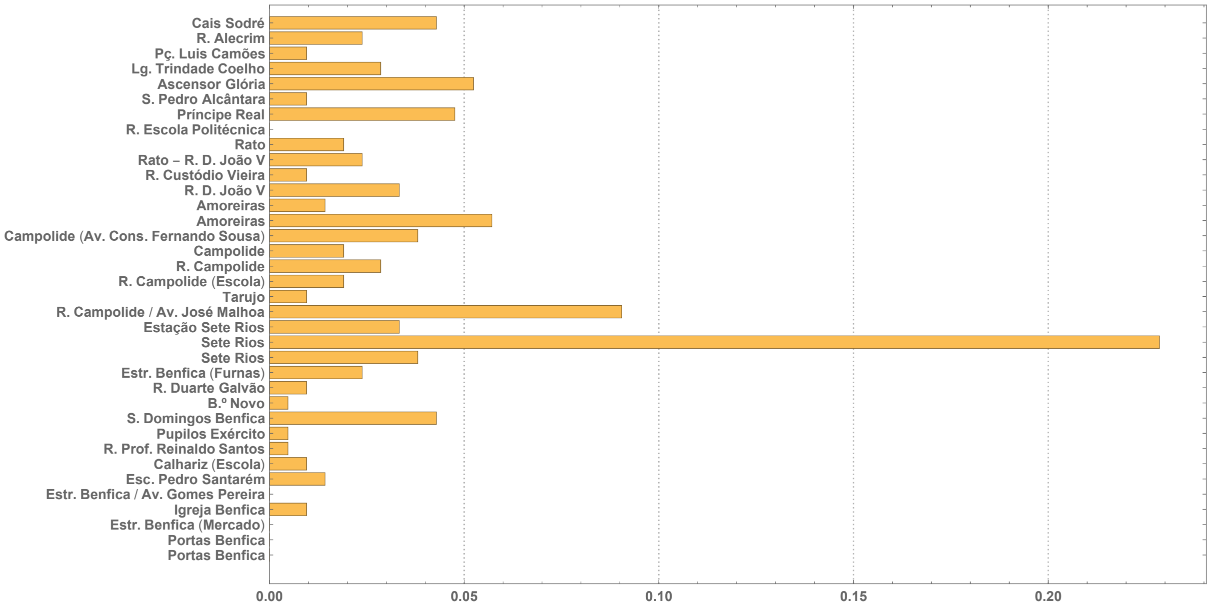


Direction 1, evening peak (threshold of 2 minutes)

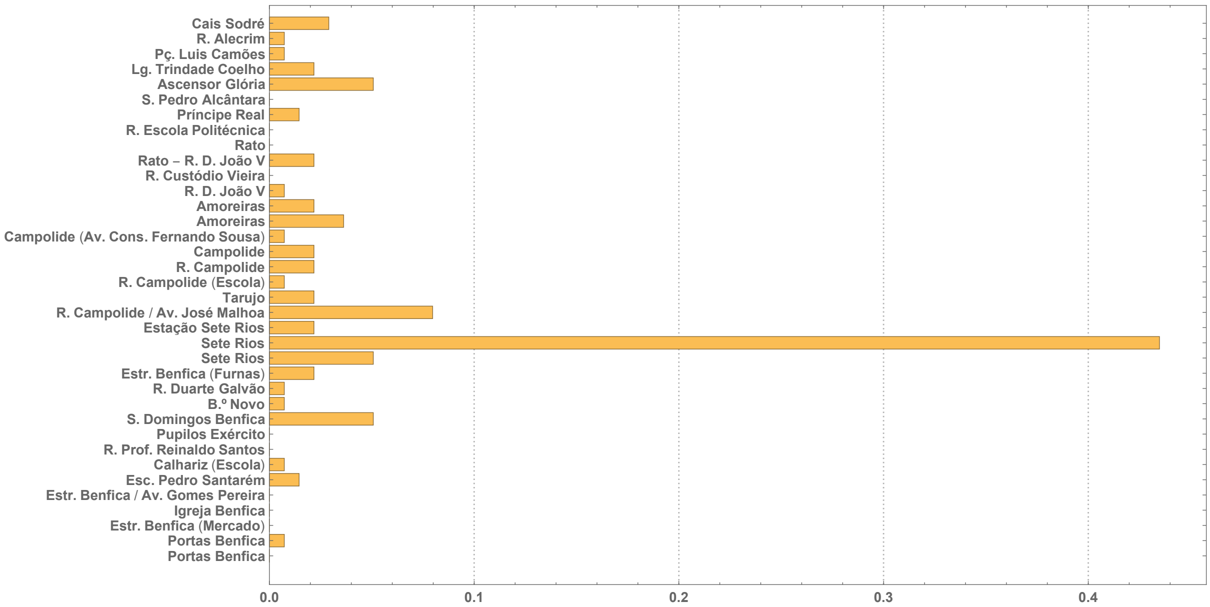


In these charts, the y-axis represents the ordered stops (from Portas de Benfica to Cais do Sodré) while the x-axis represents the frequencies in which the BB started in that specific stop

Direction 2, morning peak (threshold of 2 minutes)

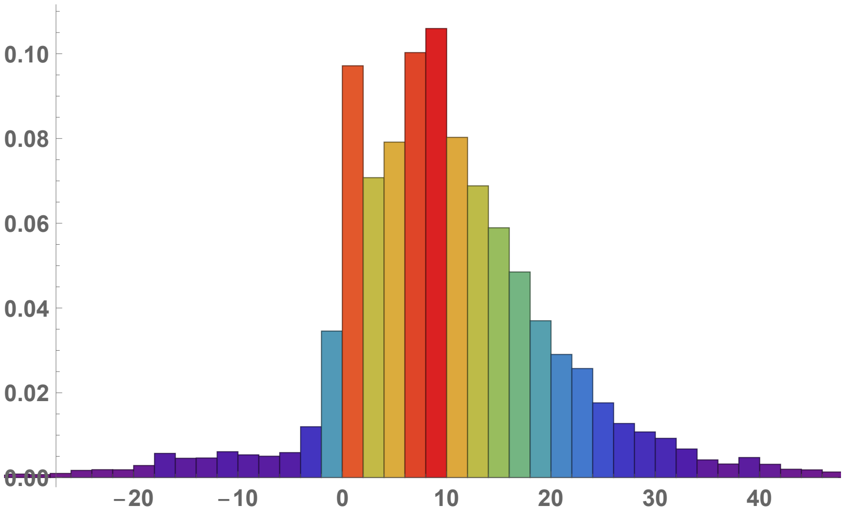


Direction 2, evening peak (threshold of 2 minutes)

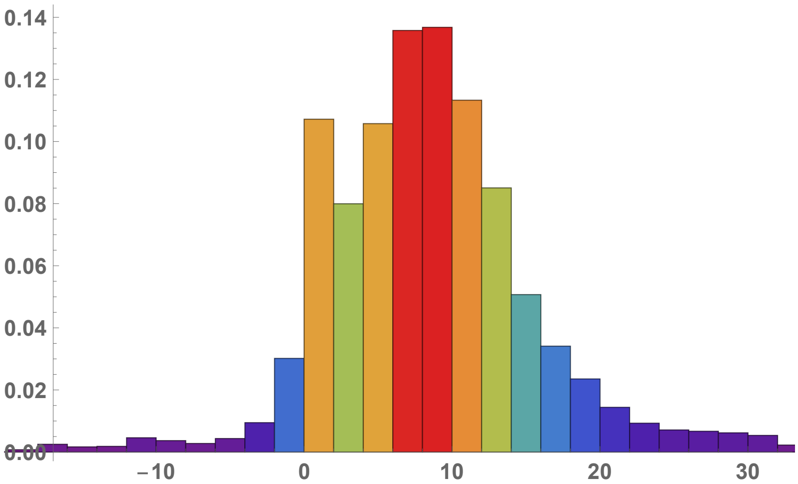


Annex 8 - Headway distribution for direction 1 for regular days of May of 2018

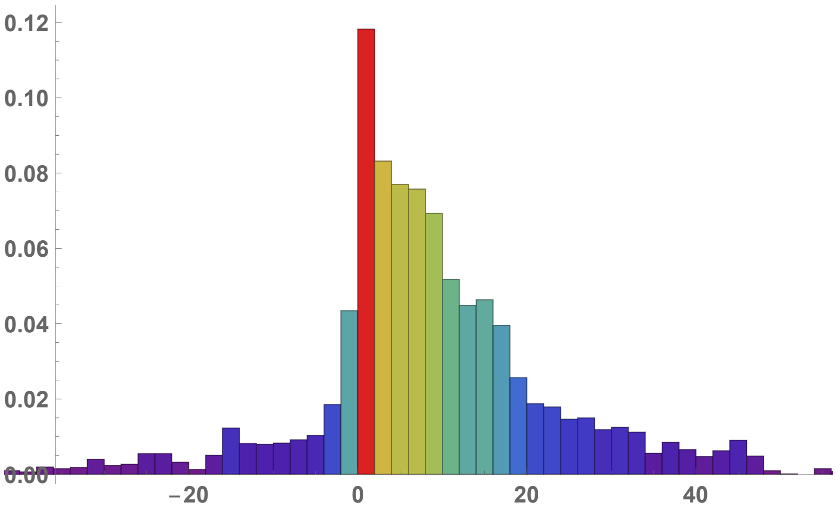
6h00-20h00



Morning peak - 7h00-10h00

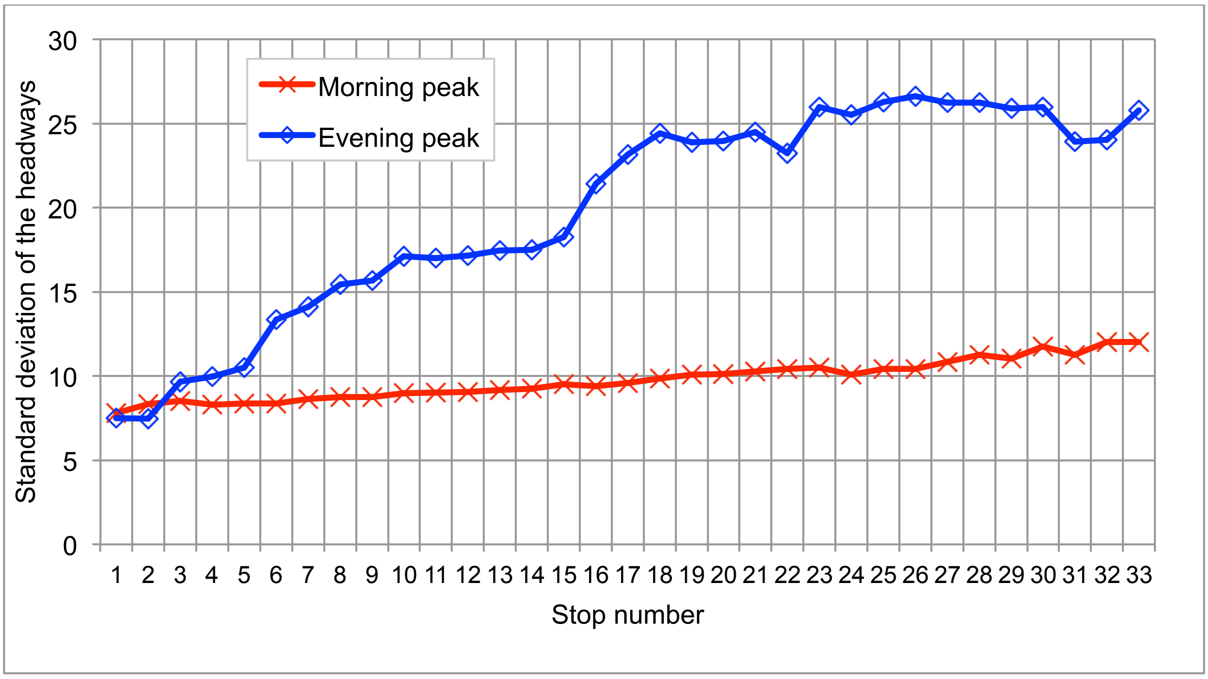


Evening peak - 17h00-20h00



Annex 9 - Standard deviation of the headways (fluctuation along the lines)

Direction 1:



Direction 2: