Disruption Management on High Frequency Lines: Measuring the effectiveness of recovery strategies

by

Alice Venancio Marques Serra

Master's Degree in Engineering
Ecole Polytechnique, 2014

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

JUNE 2016

Signature of author: Signature redacted
Department of Civil and Environmental Engineering
May 19, 2016

Certified by: Signature redacted
Nigel H.M. Wilson
Professor of Civil and Environmental Engineering
Thesis Supervisor

Certified by: Signature redacted
Haris N. Koutsopoulos
Professor of Civil and Environmental Engineering
Northeastern University
Thesis Supervisor

Accepted by: Signature redacted
Heidi Nepf
Donald and Martha Harleman Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Disruption Management on High Frequency Lines:
Measuring the effectiveness of recovery strategies

by

Alice Venancio Marques Serra

Submitted to the Department of Civil and Environmental Engineering
on May 19th, 2016 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in Transportation

Abstract

Incidents in urban rail systems are common. They vary in cause and severity, and can lead to minor or major disruptions on the network. Disruptions affect both the customers and the operating staff. To control and lessen the impact of an incident on the rail network, controllers implement corrective actions. Controllers rely mainly on experience, personal judgment and intuition to decide what recovery strategy to deploy for a given disruption. The recovery strategy deployed has a strong impact on both the service quality and the time to recovery. However, there is no systematic feedback loop to evaluate specific choices made in the control room. The scarcity of numerical data directly retracing operational actions makes ex poste aggregate analysis very difficult.

The objective of this research is to increase our knowledge about the impact of recovery strategies deployed on high frequency lines. A crucial step is to build a new dataset that accurately retraces controller's actions. The process is based on a comparison between observed train movements and scheduled train movements. Actual train movements can be obtained through various vehicle location databases. This research develops an efficient merger of several vehicle location databases to create a reliable and complete vehicle location dataset.

Building upon the reconstructed recovery strategy database, the research describes a framework to evaluate the effectiveness of recovery strategies. The framework includes a comparison between measures of recovery strategy characteristics and measures of recovery effectiveness. In particular, this methodology includes the definition of recovery effectiveness indices (REI) that take into account the impact of the disruption both for passengers and for the crew. For passengers, the research defines an integrated index based on a calculation of excess waiting time. For crew, the study focuses on lateness evaluated at crew relief points. The framework is applied to a case study based on the Piccadilly Line, a high-frequency line of the London Underground. In the context of the case study, the comparison of recovery strategies and effectiveness metrics suggests that an incremental implementation of cancelations compared to an aggressive cancelation strategy can have a positive overall impact on passengers.

Even though most of the research is applied to the Piccadilly Line, both the proposed framework and the conclusions of this thesis are transferable to other metro lines and systems.

Thesis Supervisor: Nigel H.M. Wilson

Title: Professor of Civil and Environmental Engineering

Thesis Supervisor: Haris N. Koutsopoulos

Title: Professor of Civil and Environmental Engineering Northeastern University

Acknowledgments

First and foremost, I am very grateful to my research advisors, professors Nigel Wilson and Haris Koutsopoulos. This research and thesis would have never been possible without their constant mentorship, guidance, support and suggestions. I would also like to thank others at MIT, faculty and staff, in particular Shannon, Mei-Chun and Kiley.

This research was generously supported by Transport for London. I would especially like to thank Chris Baitup, Charles Baker, David Watkins, Nigel Kelt, Sandra Weddell and Lauren Sager Weinstein for their help and for the valuable and productive meetings we had. I would also like to thank the entire staff at the Earls Court control center welcoming me so kindly during my time in London. Special thanks to Kevin Edwards for his constant availability, his generous help and support in the control center, and his friendliness. My experience at the London Underground would have not been the same without him.

Special thanks to my friends and classmates in transit lab, and more generally all the MIT students with whom I shared a class or project and who have made the past two years at MIT such an outstanding experience. Thank you to my family, friends and loved ones for providing me the support I needed during these two challenging years. Last but not least, I would like to thank Mike who made my experience at MIT and in Boston such a memorable one.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Incidents in urban rail systems are common. They vary widely in cause and severity, and can lead to minor or major disruptions on the network that affect both the customers and the operating staff. To control and lessen the impact of an incident on the rail network, controllers deploy a wide array of recovery strategies. On high frequency lines, the typical operational actions taken are train cancelation, rerouting or holding. Controllers rely mainly on experience, personal judgment and intuition to decide what recovery strategy to deploy for a given disruption. Similar incidents may be resolved with very different recovery strategies depending upon the controller in charge.

The recovery strategy deployed can have a strong impact on both the service quality and the time to recovery. However there is no systematic feedback loop to evaluate specific choices made in the control room. The scarcity of numerical data directly retracing operational actions such as train holding or rerouting makes ex poste aggregate analysis very difficult.

The objective of this research is to increase our knowledge about recovery strategies deployed on high frequency lines. A crucial step is to build a new dataset that accurately retraces controller's actions. The process is based on comparison between observed train movements and scheduled train movements. Actual train movements can be obtained through various vehicle location databases. This research develops an efficient merger of several vehicle location databases to create a reliable and complete vehicle location dataset. The result of this work will lead to a better understanding of current practices and should eventually lead to guidance for better management of future incidents.

The first section of this chapter provides background information on service control and high frequency lines. The second part presents the motivation for this research. The third part describes the Piccadilly Line which will be used as a case study throughout the thesis. Finally, the fourth part presents the general organization of this thesis.

1.1 Background

1.1.1 Rail transit disruptions

A disruption in an urban rail system can be defined as any unforeseen event that forces the system out of normal operations. The line is said to be disrupted if the observed train movements differ substantially from the scheduled movements. A disrupted rail system is generally characterized by accumulated delays and high variability in headways. The main cause of a disrupted rail system is the occurrence of an incident, such as a signal failure, a disabled train in a tunnel, or a passenger emergency. Other events, such as extreme overcrowding or internal staff problems can also lead to a disruption. In this research, an incident is defined as any event that leads to a disrupted state.

Babany (2015) identifies two distinct phases in any disruption as shown in Figure 1-1. The incident phase includes the period when the line is directly affected by the incident. An example could be a signal failure that forces a train on the line to stop. During this phase, controllers' main focuses is on safety and efficient communication. Specialized teams and engineers are contacted to resolve the incident in the field. The main actions taken by controllers are holding the trains closest to the incident. In addition, they often cancel trains due to the incident or to reduce the congestion caused by holding trains.

The recovery phase starts when the incident is fully resolved. The line is still disrupted but there are no longer any physical constraints imposed on train movements. In the case of a signal failure, all trains are now allowed to run at normal speed. However, the incident phase resulted in irregular headways and delays that have a direct impact on the observed train locations. The main goal during the recovery phase is to bring the trains back to their schedule. During this phase, controllers often reroute and renumber trains. They also re-introduce trains which were previously cancelled to get back to the scheduled frequency. The end of the recovery phase is a return to normal operations.

The disruption duration is defined as the time between the start of the incident and the end of the recovery phase. Time to recover varies depending on the characteristics of the line and the incident, as well as the type of actions taken by controllers. Further detail will be given in chapter 2.

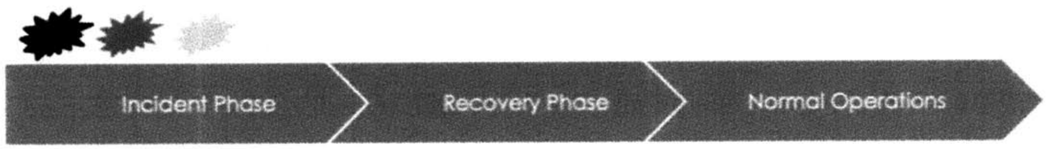


Figure 1-1: The phases of a disruption

1.1.2 Service control

Service control is the process deployed by controllers to bring a disrupted system back to normal operations. The process includes a variety of actions taken during the disruption. A recovery strategy is defined by the bundle of corrective actions taken by controllers to reduce the impact of a given incident. The recovery strategy is initiated during the incident phase, as controllers take corrective actions such as train cancelation to allow a smoother return to normal operations after the incident is resolved.

Carrel (2009) states that service control is a continuous process occurring throughout the day. Figure 1-2, adapted from Carrel (2009) and a service control manual of the RATP, the transit authority of Paris (Froloff, Rizzi, & Saporito, 1989), illustrates this process. The system starts at State 0 with normal operations. A first incident occurs, which leads to a disrupted State 1. If no action is taken, the system deteriorates to a State 2 that is characterized by higher delays and headway variability. To respond to State 1, the controllers take corrective actions A, that lead to a less disrupted line in State 2. If no additional actions are taken, the line can further deteriorate to State 3. Because the line is not back to normal operations, controllers deploy an additional set of corrective actions B that lead to State 3. State 3 could be characterized as a return to normal operations, but depending on the severity of Incident 1, it could also be characterized by residual disruptions. A second unforeseen event, Incident 2, occurs, forcing the controllers to initiate another set of corrective actions. Corrective actions A and B are seen as the recovery strategy in response to Incident 1. Corrective actions C are part of the recovery strategy in response to Incident 2. The whole process describes the service control strategy, and illustrates the dependency between various actions and the state of the system over time.

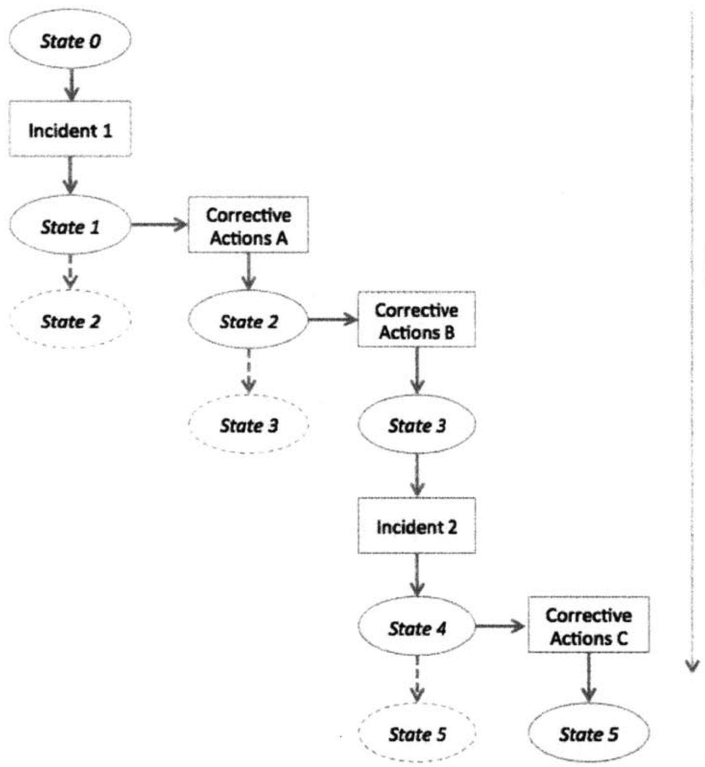


Figure 1-2: The continuous process of service recovery

1.1.3 High frequency lines and the timetable

The research focuses on disruptions on high frequency lines. Most examples in the thesis relate to rail systems and train operations. However, the general methodology and findings are also applicable to other modes, in particular bus rapid transit.

High frequency lines are defined as rapid transit lines with a frequent scheduled service. Observed headways on high frequency lines are typically under 15 minutes (at least 4 trains per hour and per direction). Most rapid rail transit systems are located in dense urban areas with high demand, such as Hong Kong (MTR), London (TfL), New York (MTA), Paris (RATP) and Boston (MBTA) which require frequent service on many lines.

Because of short headways, passengers do not plan their trip around a specific scheduled departure. The frequency and reliability of the service are high enough for the passenger to accept a short, random wait time. As we will expand on later, the average waiting time in case of random passenger arrivals is a function of both the mean and standard deviation of headway. Except in the case of fully automated systems, most high frequency lines are based on a timetable for train operations. Even though passengers do not use the timetable to plan their journeys, the operating staff depend on it. Each train is assigned a train number and the timetable consists of a list of *(location, time, train number)* triplets. Timetables are designed meticulously by the planning team to conform with various constraints including desired frequency, rolling stock characteristics, and crew work rules. Timetables are updated periodically to reflect changes in demand or operating performance on the line. Some latency is included in the timetable which allows for small amounts of lateness to be absorbed. However, larger incidents usually result in trains drifting away from their scheduled movements. Controllers' main focus during disruptions is to bring all the trains back to schedule to run smooth operations. This focus on lateness rather than headway will be discussed in detail throughout the thesis.

1.2 Motivation

As will be seen in more detail in chapter 2, in most systems there is variability in controller responses to incidents. Most agencies do not have a system to systematically record the corrective actions that were implemented. This makes it difficult, if not impossible, to reconstruct recovery strategies taken, or to evaluate them.

Addressing this question complements Babany's (2015) previous work on developing a tool to support controllers' choices during the recovery phase. Our research approach is based on analyzing existing strategies. This is made possible thanks to in-depth analysis of incident data. The results will help to put in perspective the recovery optimization tool proposed by Babany (2015).

1.3 Research Approach

The overall goal of this research is complex and can be broken down into a set of research objectives. The first milestone is to use different sources of Automated Vehicle Location (AVL) data to constitute a reliable database for further analysis. This includes an effective merging (implemented in R) between two incomplete, but complimentary, AVL databases : Netmis and CTFS. The second step is to infer the actions taken by controllers during both the incident and the recovery phases. Based on Carrel's (2009) methodology, the analysis will compare timetable values and observed train movements to retrace controllers' decisions. The final step is to use the recovery strategies database in parallel with reliability metrics calculated from AVL data to provide insight into current recovery strategy practices. Figure 1-3 illustrates the thesis workflow.

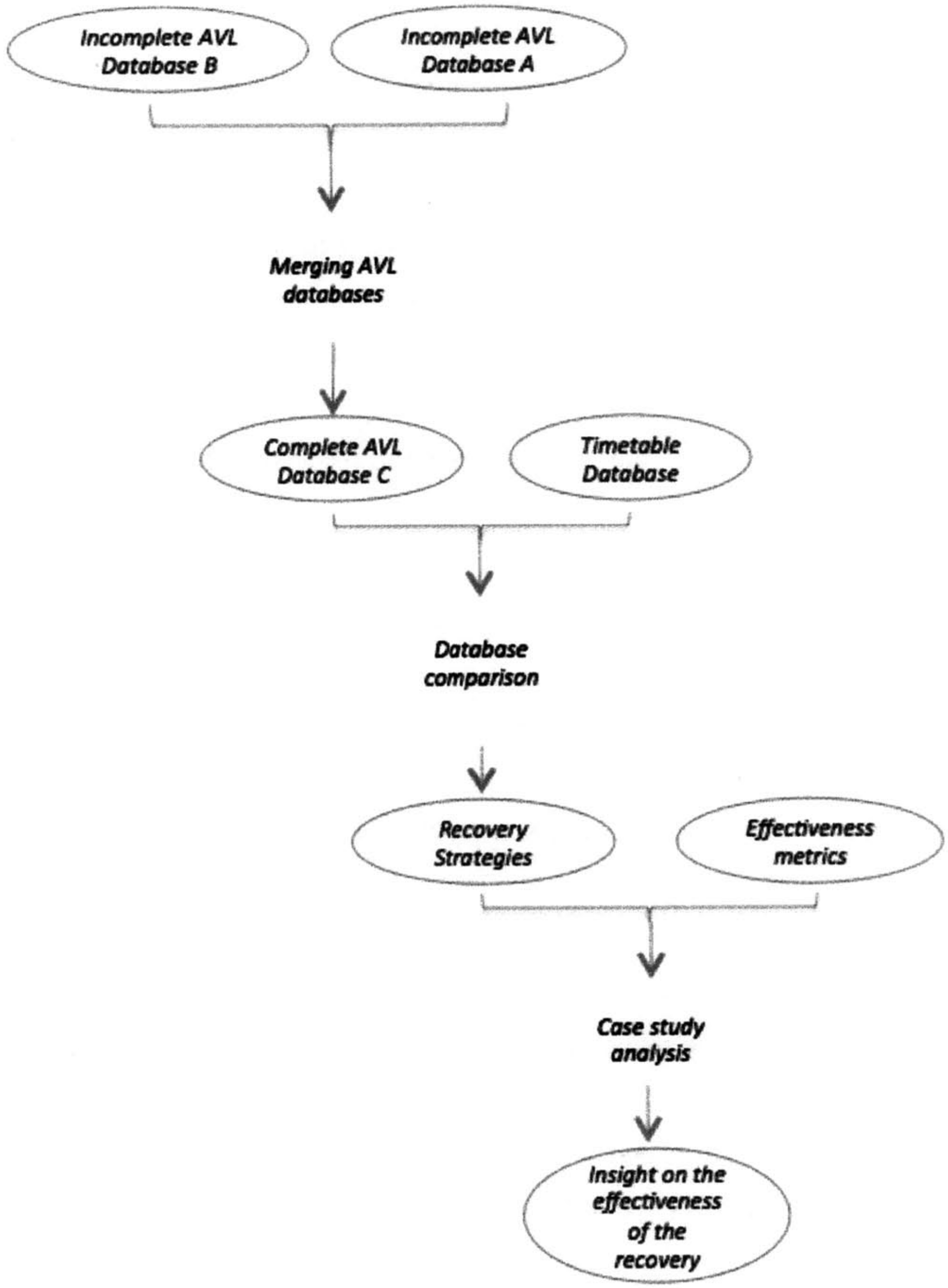


Figure 1-3: Research work flow

1.4 Literature review

Service recovery has been the subject of numerous past studies. Research concerning service control can be dated back to 1972. Most of the early work described corrective actions that are implemented without any knowledge on train locations throughout the line. In particular, Osuna and Newell (1972) developed a dynamic programing approach for optimal holding strategies at terminals. Osuna and Newell focused on understanding the impact of holding strategies for both the transit agency and the passengers. The holding strategy was computed to satisfy both the transit agency's goal which was to achieve the minimum possible running times (lower costs) and the passenger's goal which was to achieve maximum reliability.

The first elements of research concerning expressing and short-turning were conducted on the MBTA Green line. In particular, Deckoff (1990) studies the potential impacts of short-turning on passengers. He uses a simulation framework in which he generates data from a Monte Carlo simulation. He identifies several configurations where short turns provide a benefit to passengers. Even though his early research is an important first step in the field of disruption management, his results are specific to the characteristics of the MBTA Green line and cannot be easily applied to other high frequency lines. A few years later, Eberlein (1995) studied the usage of expressing, holding and deadheading for a linear rail line. She formulated two nonlinear integer programming models to minimize passenger wait time. She concluded that the effectiveness of the recovery strategies was mainly a function of the demand rather than the type of strategy.

With the increase in availability of real time information, more recent research provide comprehensive models to test holding and dispatching strategies on train or bus simulations. For example, Adamski & Turnau (1998) created a bus line simulation model to test a set of dispatching strategies, aiming mainly at good headway adherence. More recently, Sanchez Martinez (2015) developed an optimization model for holding-based control that accounts not only for the current state of the system, but also for expected changes in running times and demand.

While many research studies focus on the potential of simulations to provide insight on potential recovery strategies, only a very limited number of studies are fully based on non simulated data. Soeldner (1993) is one of the first researcher to have collected data on actual control decisions taken by local dispatchers. Because of the absence of numerical data available concerning corrective actions, which is still generally the case today, Soeldner collected data from manual logs for two days of Green Line operations. He classified the dispatcher's decisions as "good" or "bad" decisions by comparing them with model predictions. He concluded that a more structured control strategy which didn't rely solely on the individual judgment of dispatchers could increase the number of effective control decisions that were made. Carrel (2009) provided a great advance concerning the use of actual data for service recovery analysis. He develops a framework to infer various corrective actions based on non simulated train tracking data. His methodology is based on features of the data that are specific to the line of application, which could limit the potential for application to other lines. The general framework he develops will however be used in the context of this thesis to develop a more general and easily transferrable methodology.

1.5 Application to the Piccadilly Line

This section introduces the public transport system in London, with an emphasis on the Piccadilly Line which will be used as a case study and application throughout the thesis. The general research methodology, as well as the insights provided by this research, are applicable to other high frequency lines in London and throughout the world. Section 1.5.1 describes the London public transport system, in particular the transport agency Transport for London and the London Underground. Section 1.5.2 gives an overview of the Piccadilly Line. The Piccadilly Line is the high frequency line on which most of the data and field research of this thesis is based.

1.5.1 Transport for London and the London Underground

Greater London Region and Transport for London

Greater London is a region of England consisting of 32 London Boroughs and the City of London. With a population of 8.2 million in an area covering approximately 600 square miles (Transport for London, 2009), it is the most dense region of England. London has an extensive and well developed transport network, which includes both public and private services. Transport for London (TfL) (formerly "London Transport") is the local government body responsible for most aspects of the transport system in Greater London. The agency was created in 2000 as part of the Greater London Authority and is overseen by the Mayor of London. Transport for London is the public agency responsible for setting budget and fare policies for the London public transport system. TfL also oversees London's main road network and congestion charging scheme as well as the licensing, regulation, and inspection of London's taxi and private hire industries. In addition, it is responsible for the London Underground (LU), local buses, the London Overground, Tramlink, the Docklands Light Railway (DLR), the Barclays Bike Share network and more recently the Emirates Air Line (Transport for London, 2009). Except for the London Underground (LU) that is a subsidiary of TfL, operations are generally contracted out to the private sector. Long distance rail networks such as national rail lines are not the responsibility of TfL. They are overseen by the National Department for Transport (DfT). TfL is also responsible, jointly with DfT, for commissioning the construction of the new Crossrail line. The 2015-2016 annual budget of TfL was 11.5 billion pounds, covering some 23,000 employees (Transport for London, 2011)

London Underground

London Underground Ltd. (LU), which is a subsidiary of TfL, is the oldest transit system in the world with its first line opening in 1863. Today, LU is a rapid transit system that serves a large part of Greater London and extends into parts of Buckinghamshire and Hertfordshire as well as Essex. The network includes a total of 270 stations, 11 lines and approximately 250 miles of tracks. The annual ridership in the 2014/2015 fiscal year was 1.3

billion passengers (Transport for London, 2015). Fare collection has evolved from the paper format Travelcard ticket (1983) to the Oyster card (2003) which is a contactless (blue) smart card used to enter ("tap in") and exit ("tap out") the system. TfL is progressively contracting out the electronic ticketing business to the private sector. Since June 2014, contactless debit and credit bank card payments are accepted as a substitute for an Oyster card for the same fare.

Figure 1-4 shows the entire rail network overseen by TfL. The various colors of full lines represent the 11 different LU lines. The lighter colored lines represent the other systems (Crossrail, DLR, Cable Car, London Overground, London Tramlink).

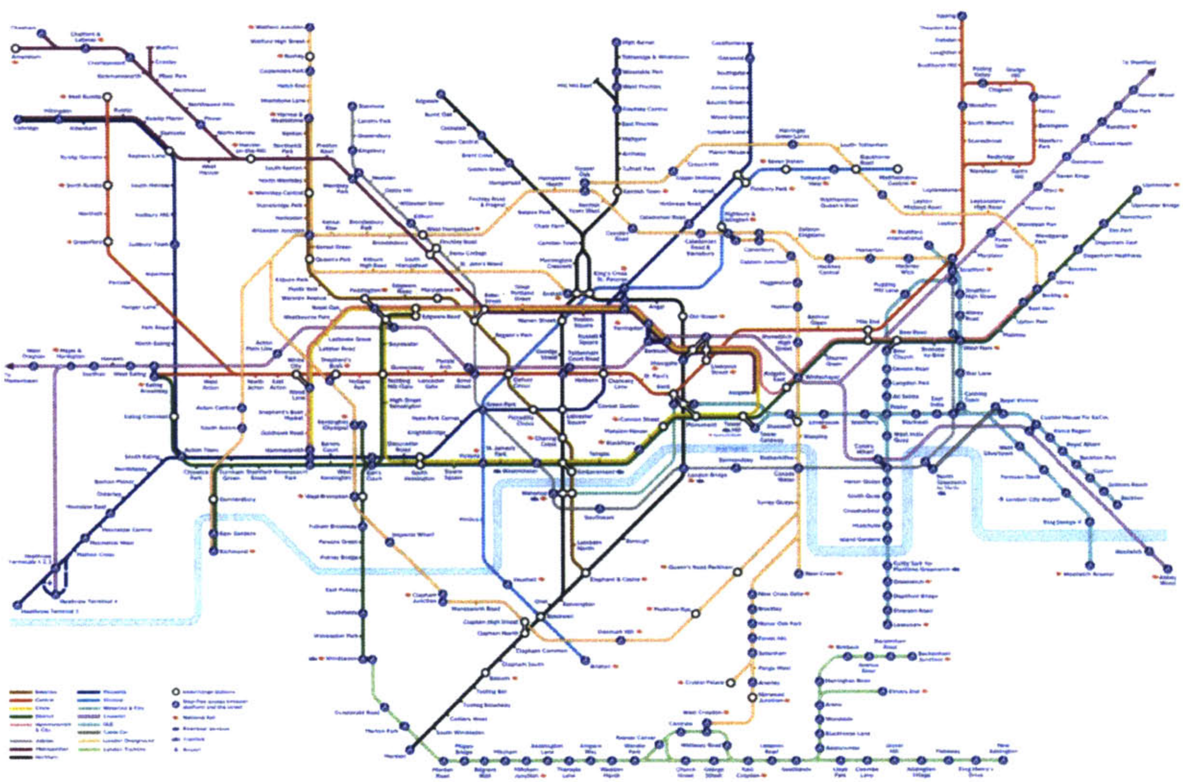


Figure 1-4: TfL Rail Network

The network is articulated around central London and extends through various branches and systems to the outer areas of London. It is a complex network, with many interchange stations. Some lines share portions of their route with other lines and many journeys can be completed using a variety of lines and routes. These redundancies make the whole network more robust. However interdependencies also increase the complexity of the network and the difficulty of planning and operations.

In the context of disruptions, the interdependencies are usually beneficial for passengers. In case of service suspension or extreme overcrowding on a given line, it is often possible for passengers to complete their journey by switching lines, even though this will inconvenience passengers, as will be discussed in detail later. However, it provides an opportunity for them to reach their destination without needing to change modes or to exit the public transport system. If interconnections and redundancies are numerous, the average journey time may be little changed using an alternate, second choice, route. The total negative effect on passengers of a disruption is therefore reduced due to redundancy in the network. On the planning side, shared tracks and interdependencies make operations more complex and add difficulty during recovery.

LU, and more broadly TfL as a whole, has a major economic impact on the London region. It provides a rapid transit system which connects millions of individuals to employment, shopping, entertainment, attractions and travel hubs such as airports or international train stations. The 21st century has seen a steady decline in total car trips in London, as shown in Figure 1-5. Coinciding with this decrease of car usage there has been a significant increase in public transport trips. As seen in Figure 1-6, the percentage of trips in London made by public transportation increased by almost 10% in the previous decade. The growing population of London as well as this shift from private to public transportation increases both the agencies' impact as well as the multiple challenges it faces. Many systems, in particular many London Underground lines, run at capacity during peak hours in Central London. Running at capacity both reduces flexibility in operations and sets higher expectations for reliability. At near capacity operations, incidents may lead to severe disruptions with larger negative impacts on passengers.

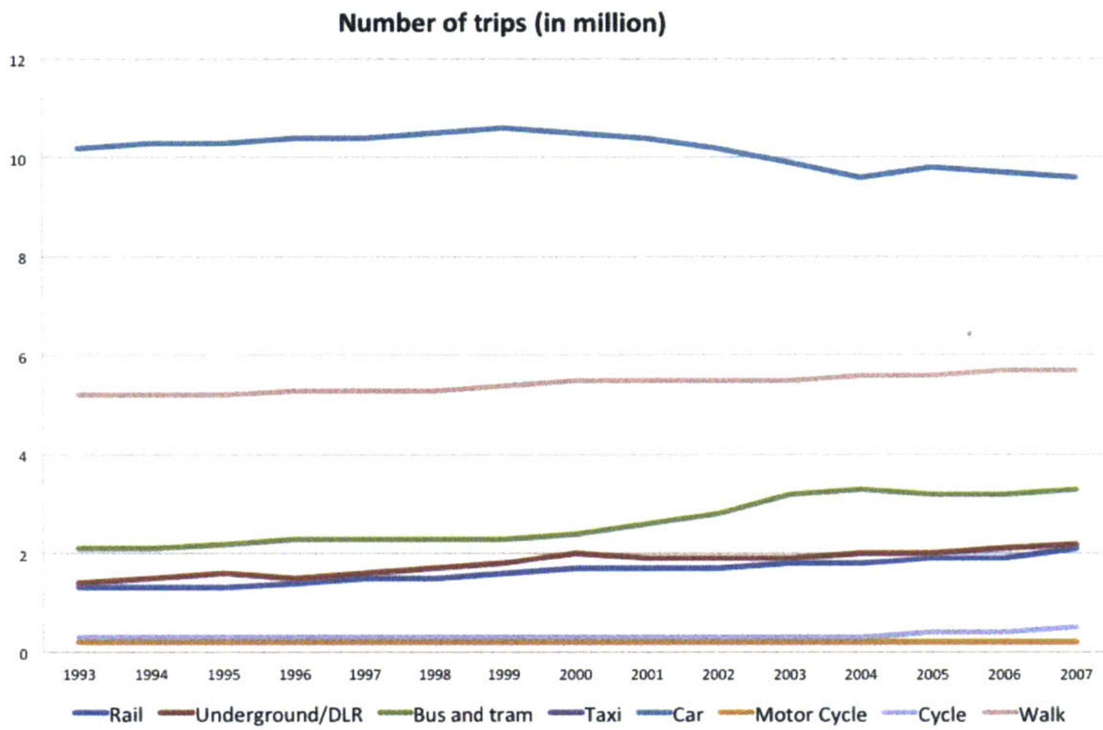


Figure 1-5: Average number of trips per weekday by mode

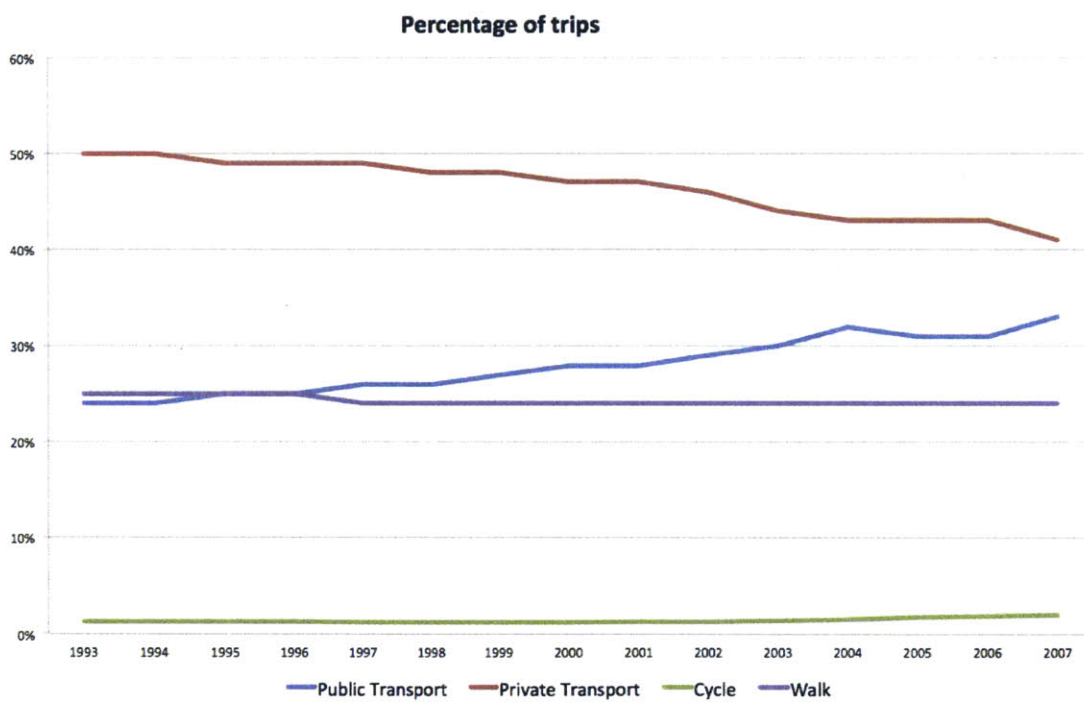


Figure 1-6: Percentage of trips in London by mode

1.5.2 The Piccadilly Line

In the context of the MIT-Transport for London partnership, the London Underground was chosen as the main Underground Line for analysis. More particularly, the Piccadilly Line is the line on which this thesis focuses. The research findings are also based on various field studies that complement the theoretical approach to disruption management. These field studies were mainly conducted in the Piccadilly Line control room at Earl's Court. Even though the thesis is based on Piccadilly Line data and observations, the general research framework and methodology are applicable to any other high frequency line in London. The metrics, concepts, and data used in the context of this research are applicable to many other lines throughout the world. Furthermore, the findings and insight from this research should be of value to many other transportation agencies beside TfL. This section presents the general characteristics of the Piccadilly Line.

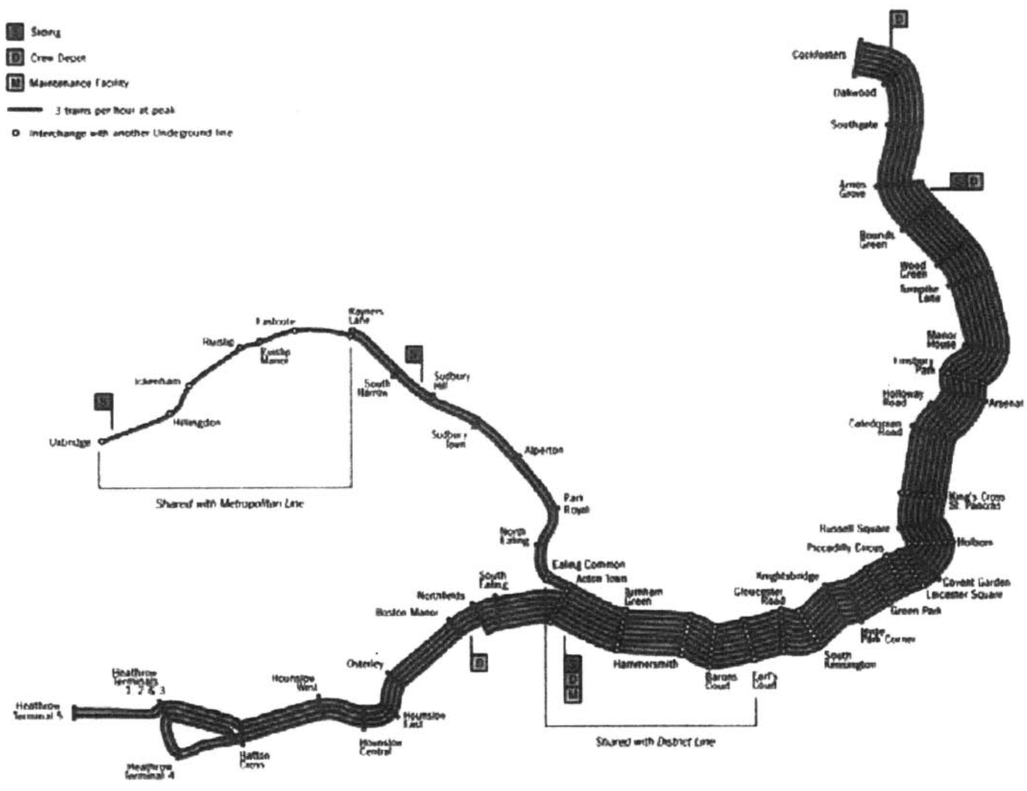


Figure 1-7: The Piccadilly Line, source Freemark (2013)

The Piccadilly Line is the 2^{nd} longest line of the London Underground with 71 km of tracks (Transport for London, 2014) and has 53 stations. The line is a crucial connector to major travel hubs, as it serves both the international rail station King's Cross and the terminals of Heathrow International Airport. It is also used by many tourists since it serves many famous London attractions. There are a more than 200 million journeys per year made on the Piccadilly Line.

As seen in Figure 1-7 from Freemark (2013), the line runs from the Northeast to the Southwest, from Cockfosters to Uxbridge and Heathrow. The line splits into two branches at Acton Town. The portion between the stations Ealing Common and Earl's Court is shared with the District Line. The crew depots are at Arnos Grove and Acton Town, at each end of the line. The main points where trains can reverse directions are Arnos Grove, Acton Town, Uxbridge and South Harrow. The control room, which will be described in section 2.2.3, is at Earl's Court. The Piccadilly line is operated with 1973 tube rolling stock. The fleet is composed of a total of 86 trains, but only 79 are needed to run the scheduled service during the peak period. Table 1-1 describes the average train frequency scheduled by route.

Table 1-1: Train frequencies by route.

Portion of the line	**Scheduled frequency in trains per hour**	
	Off Peak	**On Peak**
Cockfosters-Heathrow T5	6	9
Cockfosters-Heathrow T4	6	9
Cockfosters-Uxbridge	3	4
Cockfosters-Rayners Lane	3	4
Arnos Grove-Northfields	3	4

The Piccadilly Line is of particular interest, as it presents many representative challenges to LU. It has seen a steady increase in ridership and yet relies on some of the oldest infrastructure and rolling stock. These characteristics make it a good line to study in the context of disruption management as incidents are frequent and impact a large number of passengers. The PICU project, which will renew parts of the signaling system and upgrade the control room, should reduce the occurrence of some types of incidents such as signal failures. However, there are no plans for a full update of the rolling stock and tracks in the next 10 years . These observations motivate our research. Given the capacity and age of the line, we will study alternate ways to lessen the overall impact of disruptions, both for passengers and for the operating staff.

1.6 Thesis Organization

Chapter 2 provides background information on service recovery applied to high frequency lines. First, the chapter defines types of disruptions and analyses incident causes. A precise understanding of disruptions is important to better grasp the utility and complexity of recovery strategies. The second part focuses on service recovery strategies. It describes the various corrective actions used to mitigate the effect of an incident. In addition, it explains the general framework used to measure the effectiveness of recovery strategies. The third part of Chapter 2 provides a detailed description of the Piccadilly Line including a description of the different roles of the operating staff as well as a discussion of communication and work flows during disruptions. It emphasizes the importance of the control room, but also describes the other units that play a role during service recovery. The Piccadilly Line is the high frequency line used as an example throughout this thesis and from which the data used originates. The findings and processes developed throughout this research are applicable to any high frequency line. However, some specifics of the algorithms deployed as well as some key assumptions are directly linked to the choice of the Piccadilly Line itself. It is therefore essential to present in detail the characteristics and challenges of this line.

Chapter 3 is focused on Automated Vehicle Location data. The chapter describes the characteristics and utility of AVL data. AVL data is crucial to track train movements and is the main type of data most public transport agencies rely on to measure service performance. Typical measures include headway variability or number of trains per hour. In the specific context of the research, AVL is a building block to infer controllers' actions through a fine comparison with the timetable. The chapter continues with an in depth description of the characteristics and limitations of the two distinct databases available for the Piccadilly Line. The CTFS database is limited in spatial and temporal resolution. The Netmis database is unreliable and many records are missing train numbers. A methodical comparison between scheduled train movements and observed train movements can reconstruct recovery strategies. The AVL that provides observed train movements should be as precise and reliable as possible to enable the reconstruction of recovery strategies. Missing train numbers in the AVL dataset could prevent the comparison with scheduled movements. Poor spatial resolution could limit the identification of short-turning strategies. Poor temporal resolution could limit the identification of holding strategies.

Chapter 4 presents the methodology and algorithm developed to resolve the AVL data quality issues. The chapter describes the fusion of several incomplete and unreliable databases to obtain a higher quality AVL dataset. The first part presents the methodology and general approach to the merging process. The knowledge acquired can be applied to multiple sets of complimentary but incomplete databases. The second part focuses on the AVL data and details the different steps of the merging algorithm. The third part of the chapter concentrates on the various parameters included in the merging process. The fourth part describes the implementation of the merging algorithm for the Piccadilly Line. The final part discusses results of the merging, the validation methodology and the choice of parameters based on a sensitivity analysis.

Chapter 5 concerns the inference of recovery strategies. The first part describes the general methodology applied to infer all the corrective actions implemented on the line. This methodology is based on a comparison between observed train movements and scheduled train movements. The algorithm developed both for the short-turn and cancelation inference is detailed in the second part of this chapter. A discussion on the choice of parameters for the inference is presented. The final part describes both the results and the limitations of the inference process. In particular, the inference is limited to a small subset of all the possible corrective actions that can be implemented on high frequency lines and further research could include other corrective actions such as expressing or holding.

Chapter 6 presents a methodology to evaluate the effectiveness of recovery strategies. The methodology is based on a comparison of days with similar incident conditions. The first part define recovery effectiveness indices for both passengers and crew. the protocol and defines recovery effectiveness indices. These indices provide a metric to evaluate the total impact of the disruption. The second part describes two similar incidents that occurred on the Piccadilly Line. These incidents are chosen as a case study to apply the proposed methodology. In particular, prior and posterior incidents during the chosen days are studied as these may have an effect on the assessment of recovery. Finally, the last part of the chapter presents the results of the case study.

Chapter 7 concludes with a summary of the main findings. It presents recommendations as well as the potential short and long term gains for Transport for London and other transportation agencies. Finally, it details the limitations of the research and the potential next steps.

Chapter 2: Disruptions and Recovery Strategies

This chapter provides background information on service recovery on high frequency lines. Section 2.1 characterizes incidents. Section 2.2 discusses the various impacts of disruptions on both passengers, crew and the travel agency. Section 2.3 discusses recovery strategies and how to measure their effectiveness. The final section presents in more detail the specifics of the Piccadilly Line control environment.

2.1 Incident characterization

All disruptions are initiated by the occurrence of an incident. Incidents are common on rail systems but they vary widely in cause, frequency and duration. This section presents different possible incident categorization schema and describes the large variability in incidents.

2.1.1 Variability in cause of incident

Most public transport agencies differentiate incidents by cause including customers, signals, staff errors, rolling stock, track, safety, fleet or external reasons such as extreme weather. The CuPID database records all incidents that occur on London Underground lines. The database is available for ex poste analysis but is not available in real time. Each incident on the London Underground is assigned a unique identification number. The CuPID data provides information on this incident, mainly the time and date of the incident, the location of the incident, the duration of the incident, the incident type as well as a more detailed description of the incident. Chapter 6 presents in more detail the limitations and information available through this database.

Based on CuPID data, Figure 2-1 illustrates the distribution of incidents by cause that occurred on the Piccadilly Line from October - December 2013. This illustrates the large variability in incident cause. Figure 2-2 references the total time of incident by cause. Even though incidents caused by staff are more frequent, the biggest portion of incident time is due to signals. As can be seen in Figure 2-2, 35% of the total incident time are due to signals. This observation is particular to the Piccadilly Line and can be explained by the aging infrastructure and rolling stock. The specific difficulties linked to the aging signaling system will be discussed in Chapter 3.

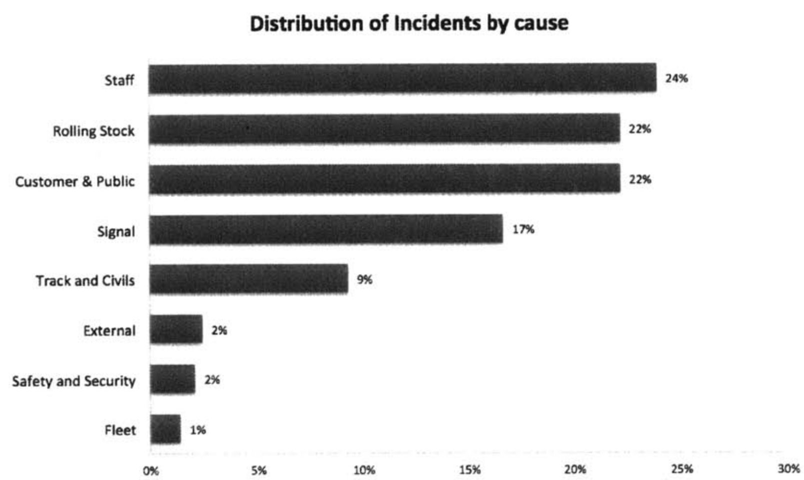


Figure 2-1: Distribution of incidents by cause on the Piccadilly Line

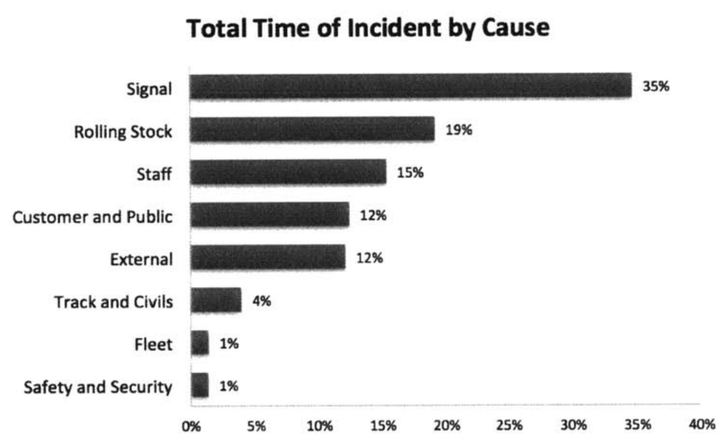


Figure 2-2: Total time of incident by cause on the Piccadilly Line

2.1.2 Variability in number of incidents

The variability in the number of incidents is also high. Figure 2-3 reports the number of signal failures on the Piccadilly Line over one month.

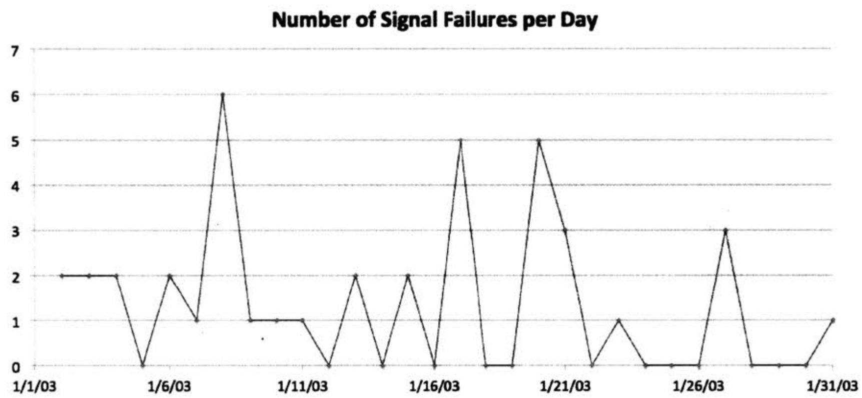
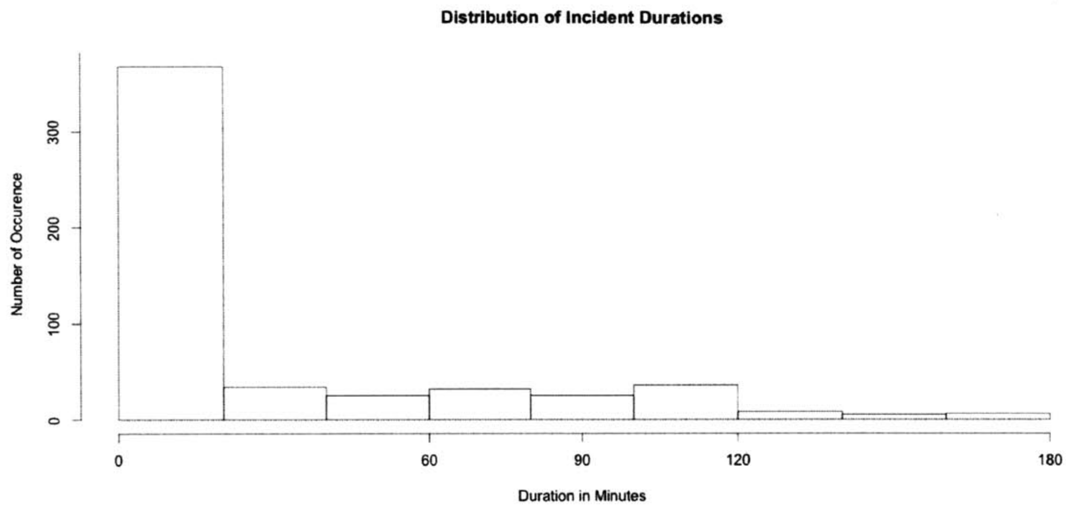


Figure 2-3: Number of signal failures

2.1.3 Variability in total time of incident

In addition to the variability in the cause of incidents and in their frequency, the total time of incidents (defined as the time from the start of the incident to the time of resolution) is also variable. Figure 2-4 represents the histogram and values of durations observed on the Piccadilly Line from October - December 2013.



(a) Histogram of Incident Durations

Min	1st Qu.	Median	Mean	3rd Qu.	Max	sd
1	4	8	30	43	177	41

(b) Distribution of Incident Duration

Figure 2-4: Incident duration

This high variability leads to complex decision processes in the control environment as corrective actions need to be implemented without a good idea of the likely incident duration and it's potential effect on the line.

2.2 Impact of disruptions

Disruptions that occur on rail systems have an impact both on passengers as well as on the operating staff. For example, drivers may often have to alter their scheduled shift because of an incident. Passengers' journey times may be increased and they may experience additional anxiety or inconvenience. Finally, the transport authority can also suffer severe negative impacts including financial loss and degradation of public image.

2.2.1 Impact on passengers:

- *Increased Waiting Time and Journey Time.*

The main negative impact on passengers resulting from a disruption is an increase in journey time. The incident usually results in disrupted service characterized by lower train frequency and uneven headways. Passengers located near the incident location or further away on the same line, as well as on substitute lines, are all affected because of the propagation of the disruption. Both longer journey times (because of potential reduced train speeds) as well as increase in platform waiting time have a direct negative impact on passengers.

- *Increased anxiety*

In addition to an increase in Journey Time, incidents result in increased anxiety for the passengers. Most common incidents require trains to be held for safety concerns. For example, if smoke is detected in a tunnel, the driver will bring the train to a stop and a specialized repair team will need to be sent into the tunnel. Signal failures are also a common cause of trains being held. Signals are located at various points along the tracks, and drivers must wait for a green signal to enter the next track segment. This simple system keeps a safe separation between trains. If the signal stays red, safety protocols require the driver to stop the train. Trains will also be brought to an abrupt stop in case of a passenger on the tracks. The engineering staff will turn the traction current off on the affected tracks for safety reasons.

These examples illustrate times when a train in service with passengers onboard is stopped. In some circumstances, the train may be stopped between stations. In cases of overcrowding, passengers may feel ill or faint. If the train must be held for an extended period between stations, the operating staff will assist passengers in alighting. A common strategy is to bring a second train up to the rear of the blocked train to allow passengers to pass through the second train to alight. In extreme cases, passengers will use the emergency paths along the tracks to reach the closest station.

In addition to the unknown duration of the hold, direct safety issues can lead to additional anxiety. Sudden stops caused by an emergency braking can harm passengers. The smell of smoke or loud unexpected noises can also cause additional anxiety. Rare events, such as doors opening while the train is moving, can have a large impact on passengers' experience. All these effects impact only the passengers directly affected by the incident. They are difficult to quantify precisely and they are likely to be only a small fraction of the total passengers on the network.

- *Increased inconvenience*

During disruptions, passengers throughout the disrupted line (or network) experience increased inconvenience. Disrupted service can result in higher levels of overall crowding. Tirachini (2013) explores the various effects of crowding on passenger wellbeing. The increased inconvenience due to crowding results from different factors. Crowding increases overall stress and feelings of tiredness, it reduces the productivity of passengers working on a train and can trigger a feeling of loss of privacy as well greater concern for safety. All these factors increase inconvenience.

The inconvenience can also be due to longer journey times, the increased possibility of arriving late at work or less time spent with friends or family. Even with an equivalent journey time, passengers needing to switch lines or modes may experience greater inconvenience. These passengers need to re-plan their routes, possibly walk further, wait for another train (or bus) or even switch from public to private transport. Inconvenience and anxiety can result in an increase in perceived journey time. An additional inconvenience in case of switching systems is an increased fare for the same journey. These issues will be further discussed in section 6.1.

2.2.2 Impact on crew :

- *Possible late relief or shorter meal break*

Disruptions generally result in longer shifts or shorter meal breaks for drivers. The driver of a late train may be forced to stay on duty for an extended period of time. The impacts of disruptions on drivers vary depending on the transport agency and the power of the driver unions. When driver's unions are powerful, the main priority of recovery strategies may be to ensure that drivers' shifts are respected. If a train is late, a driver may be replaced by a spare driver to respect his planned shift. In other systems where drivers' unions are weaker, respecting drivers' shifts and meal breaks may be less of a priority and drivers may often have to drive over schedule.

- *Increased complexity*

Incidents and disruptions result in increased overall complexity for drivers and staff. During normal operations, drivers are assigned to specific shifts and their main tasks are driving the train (advancing while respecting the signaling indications) and opening doors for passengers. In non-automated systems, the driver may also be responsible for announcing train destinations, next stations and connections to other lines. On the London Underground, some lines are automated and the main task for drivers is controlling the doors. However, when disruptions occur, drivers have many other tasks with the safety of passengers being most important. The driver needs to report any malfunctions or anomalies (smoke, passengers on tracks, etc). The driver must be able to communicate clearly with the rest of the operating staff. He needs to describe the incident as (s)he is experiencing it. In case of a signal failure (for example a signal being blocked at red), the driver may need to pass the signal at low speed if ordered to do so by operations control. The experienced complexity is very much linked to the overall lateness of the line. Staff do not have any additional tasks or uncertainty while working on a line running on schedule. On the contrary, the more the train are late the more the crew and staff will see changes in routine, scheduled operations and shifts.

As can be seen, both the complexity and the possible late reliefs, or changes in shift, are the direct result of lateness of trains. A measure of the observed lateness at crew reliefs is a good way to approximate the overall impact of lateness on crews. Section 6.1 will further elaborate on this point.

2.2.3 Impact for the transport authority

- *Financial loss :*

The first major impact is a financial loss for the transport authority. Transport authorities often reimburse passengers who experience severely disrupted journeys. On the London Underground specifically, passengers may request a total refund if their trip was delayed for more than 15 minutes (Transport for London, 2016). In addition to direct financial loss, the transportation agency loses passengers who switch to other modes due to disrupted services. Passenger counts at stations provides insight on this metric. By comparing the total number of station entries and exits between normal days and disrupted days, it is possible to infer how many passengers were lost by the agency. In some extreme disruption cases, entire stations or lines must be closed resulting in an increased financial loss for the transport authority. In addition to temporary switches, some customers may permanently switch modes, resulting in an additional financial loss for the agency.

- *Degraded public perception :*

High number of incidents and repeated disruptions often lead to a decline of popularity and trust from passengers. Reliability and safety are key for passengers using the service and disruptions can have a severe negative impact on their perception of travel. In dense networks where a large part of the population heavily relies on public transit, media extensively report and comment incidents and disruptions, feeding in the general degraded public perception. Social media such as twitter or Facebook that disseminate passengers' experience and dissatisfaction may increase this phenomenon. The degraded public perception may increase the number of passengers switching permanently from public to private modes.

Even though the impact of disruptions on the transport authority is important, the research will focus on the impact on both passengers and crew.

2.2.4 Metrics

Section 2.1.2 discusses the various impacts of disruptions. In particular, there is an important distinction between passenger impact and crew impact. Passenger impacts are related mainly to journey time metrics while crew impacts are related to lateness compared to the schedule. This research focuses on waiting time as the main measure of impact of a disruption. A possible next step for further research would be to include Automated Fare Collection data to incorporate metrics of observed journey times during disrupted days.

Assuming random arrival of passengers at stations, the average waiting time at a station is a function of the average headway and the headway variance, as given by Larson & Odoni (1981):

$$E(T) = \frac{E(H)}{2}(1 + cv(H)^2) \qquad (2-1)$$

E(T) = expected value of passenger waiting time,

E(H) = expected value of headway,

cv(H) = coefficient of variation of headway

In this case, only the headway value and regularity matter. Any measure of train lateness is irrelevant. However, the evolution of lateness is of great concern for controllers. As discussed in section 2.2.2, it is important to consider the impact of lateness on crews. Table 2-1 summarizes the main metrics that will be used throughout the thesis.

Table 2-1: Metrics to measure impact

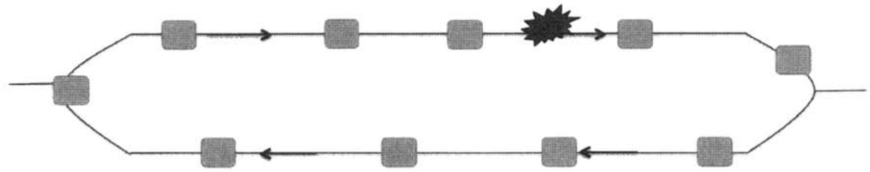
Proposed Metric	***Definition***	***Domain of Definition***	***Point of View***
Lateness	*Difference between observed train departure and scheduled train departure*	*Value associated by train number and station*	*Crew centric measure*
Frequency of trains	*Number of operating trains in both directions*	*Value associated to a portion of or an entire line or network*	*Passenger centric measure*
Evenness of headways	*Standard deviation of observed headways*	*Value associated to a portion of or an entire line or network*	*Passenger centric measure*

2.3 Service recoveries and metrics to measure their effectiveness

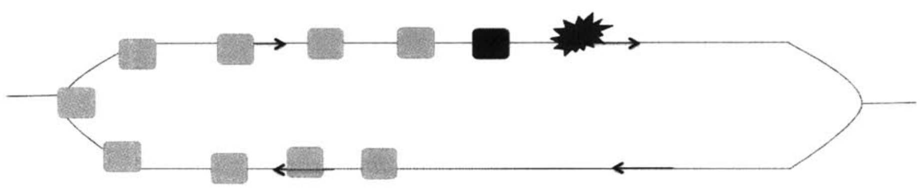
This section describes the corrective actions that are implemented by controllers to mitigate the impact of incidents. Section 2.3.1 describes the most common corrective actions that are used in rail systems. Section 2.3.2 presents the general characteristics of service recovery. Section 2.3.3 discusses recovery strategy effectiveness.

2.3.1 Description of corrective actions

Figure 2-5 (a) depicts schematically a network with trains represented in grey. A disruption is represented by the black star. With no corrective actions, the incident would lead to a highly disrupted service. Figure 2-5 (b) illustrates the possible state of the line with no control interventions. The eastern part of the line has extremely long headways and the western part of the line is bunched with trains blocked between stations. Controllers implement a variety of corrective actions to mitigate the effects of the incident. As explained in Chapter 1, the sequence of corrective actions is called the recovery strategy. This section details a few of the most common actions used on high frequency lines.



(a) Initial state of a line with disruption



(b) State of a line with disruption and no control

Figure 2-5: Schematic disruption

- *Holding*

Holding consists of purposefully stopping a train while in service. On some systems, automatic holding instructions are given to maintain, or increase, headway regularity. This type of holding occurs at stations and does not last longer than a few minutes. In the event of more severe disruptions, controllers generally give specific instructions to drivers concerning the location and the duration of a hold. Holding can cause immediate annoyance or anxiety for passengers in the concerned train. However, it is a useful strategy to reduce the possible bunching and keep a balanced distribution of trains on the line. The holding strategy has been studied and modeled in numerous previous research. Depending on the focus of the research, the holding strategies differ in general objectives (passengers or crew centric measures), underlying models used, and available information utilized for the optimization. Common objectives include schedule adherence (Adamski and Turnau, 1998), headway adherence (Rossetti and Turitto, 1998), headway regularity (Daganzo, 2009 and Bartholdi and Eisenstein, 2012), and general cost minimization (Delgado et al., 2009, Delgado et al., 2012, and S´aez et al., 2012).

- *Short tripping*

A train is short tripped when it changes direction before its scheduled destination. 66 Short tripping literature review A short trip can only be implemented at locations with reversing points. A reversing point allows a train to switch directions thanks to additional installed tracks. On the Piccadilly Line, there are a total of six reversing points. Short tripping impacts passengers whose final destination is no longer served. These passengers may have to change trains or wait longer and hence, have a longer journey time. Controllers implement short trips mainly to reduce lateness of a given train as illustrated in Figure 2-6. Train S goes through the reversing point. In the given example, the running time from A to B is 20 minutes and train S has accumulated a total of 25 minutes of lateness compared to schedule. A short trip for train S will result in the train arriving at point B only 5 minutes late.

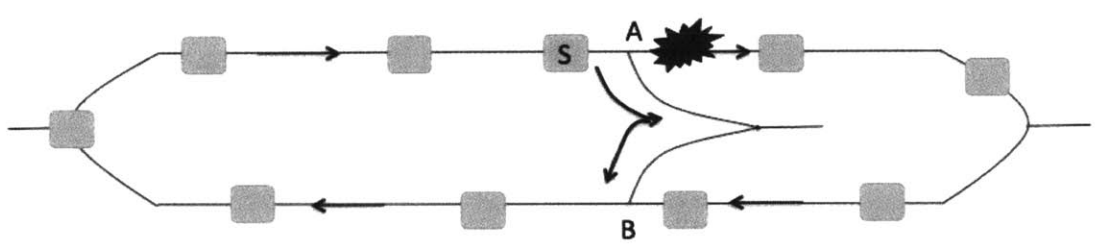


Figure 2-6: Short tripping

- *Expressing*

To bring a train back to schedule, another possible operation is expressing. 66 Short tripping literature review A This strategy is less common and is usually implemented in dense networks with closely spaced stations. A train is expressed when it does not stop at a given set of stations it was scheduled to stop at. For passengers planning to board or alight at the expressed stations, journey time will be increased and they will need to switch trains.

- *Train renumbering or train reformation*

A very common strategy is to implement train renumbering, also called reformation. This strategy consolidates lateness on a reduced number of trains. Train renumbering changes the train number of a train. Usually, controllers implement a series of reformations. Figure 2-7 illustrates the logbook with a reformation plan documented. Reformations bring all trains back to schedule except for one train that accumulates the delay of all other trains. This train is then expressed or short tripped to get back to schedule. Train renumbering does not directly impact passengers. The main impact is on crew and the operating staff. Drivers need to change the number of the train they are driving and a reformation will usually result in a change in their shift.

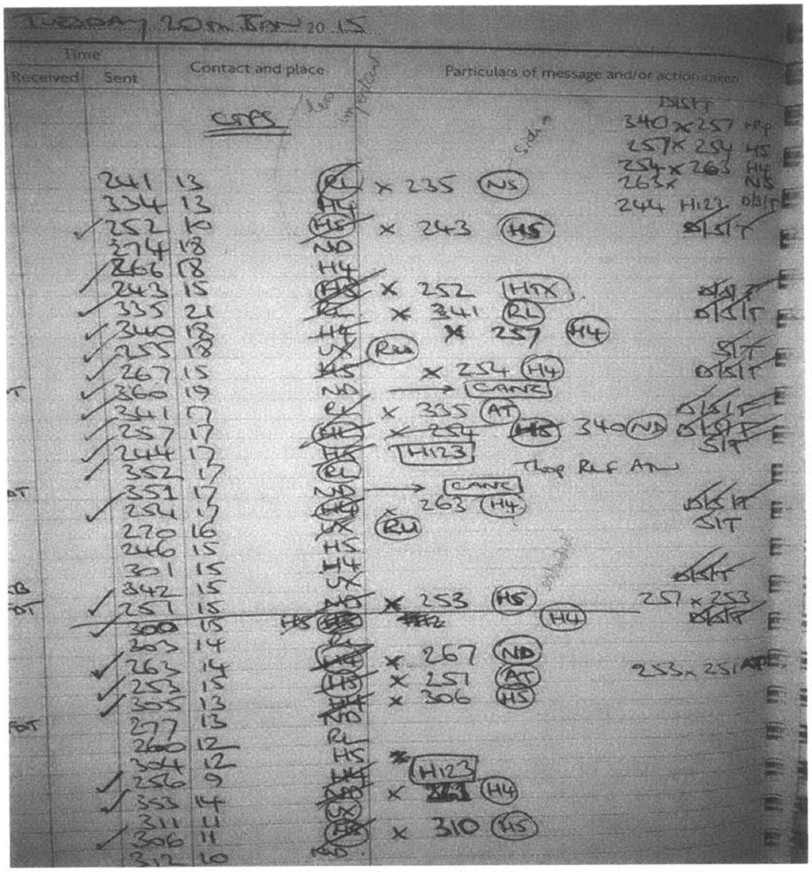


Figure 2-7: Logbook retracing a reformation plan

- Canceling

In addition to the strategies in place to reduce train lateness, controllers also cancel trains during disruptions. The main reasons for canceling trains is to reduce bunching and allow a smoother recovery. Canceling trains also creates more spare drivers that can be redeployed during the recovery. Cancelations have a direct negative impact on passengers as train frequency drops. Chapter 6 will discuss the possible benefits of aggressive or incremental cancelation policies.

2.3.2 Characteristics of service recovery strategies

As can be seen by the different actions just described, the main goal for the controllers and the operating staff is to bring the trains back to schedule. In some cases of severe disruptions, a return to scheduled operations may be very difficult to achieve. In those cases, it may better to concentrate on passenger metrics (train frequency, evenness of headways) rather than on lateness.

Even though there is a limited set of possible corrective actions that can be used during disruptions, there is a large variability in responses due to the experience, personal preferences and habits of controllers. For similar incident characteristics, two different controllers may implement very different recovery strategies. The lack of reliable information and the paper-based tools controllers work with also limit the scope of possible recovery strategies. For example, in the case of the Piccadilly Line, the controllers have a good understanding of train lateness but do not have access to real-time metrics on frequency or headway evenness. Both personal opinions and a lack of complete information can lead to sub-optimal choices of the corrective actions to implement.

Most agencies do not have any system to systematically record the corrective actions that were implemented. The research develops a framework to infer all the corrective actions that were deployed based on a comparison between train tracking data and schedule data.

2.3.3 Effectiveness of recovery strategies

As discussed previously, there are many ways to measure the performance of a line. This research focuses on a multi-criteria characterization of performance. Both passenger and operating metrics are incorporated in the evaluation. This section briefly describes the general concepts and framework that are used to measure the effectiveness of a recovery strategy.

One important concept is the time to recover. As seen in Figure 2-8, time to recover is defined as the time from the end of the incident to return to normal operations. Normal operations are usually described as scheduled operations, which relates to lateness measures. However, the passenger point of view, only train frequency and evenness of headway are important. Chapter 6 will include both passenger and operation metrics to take into account the different players in the system.

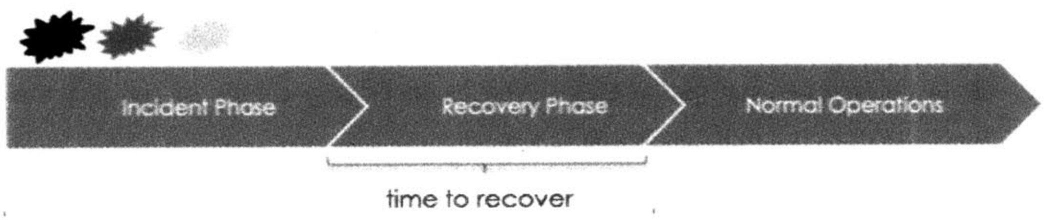


Figure 2-8: Time to recover

An effective recovery strategy is defined as a recovery strategy that mitigated a given incident's impact as much as possible. Because of the high variability in type, location and time of incidents as discussed in section 2.2.2, it is difficult to compare recovery strategies one to the other. Chapter 6 introduces a methodology to explore situations with similar incident characteristics. The protocol compares similar disruptions and studies the impact of different recovery strategies on metrics linked to reliability and time to recover.

2.4 Service Recovery on the Piccadilly Line

2.4.1 Organization of the control center

Figure 2-9 illustrates the organization of the Piccadilly Line control center. The control center is at Earl's Court on the southwest of the line. The control room combines four different roles. The signalers (I) are dispersed in a semi-circle facing the train tracking system represented by dashed lines. Each signaler is in charge of a portion of the line. On the right side of the room, an engineer is present to advise controllers during disruptions caused by signal failures or train malfunctions. The engineer is in charge of sending the appropriate repair staff to the field when needed. Behind the signalers and on a higher level (for visibility purposes), there are two controllers (IV) and the line information manager (III). Figure 2-10 is a picture of the control room.

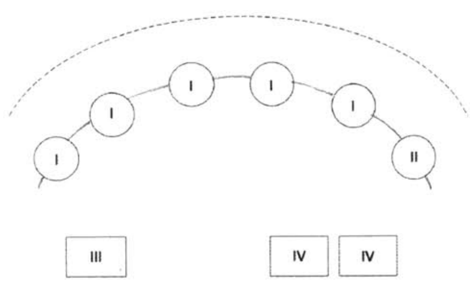


Figure 2-9: Display of Earl's Court Control Center

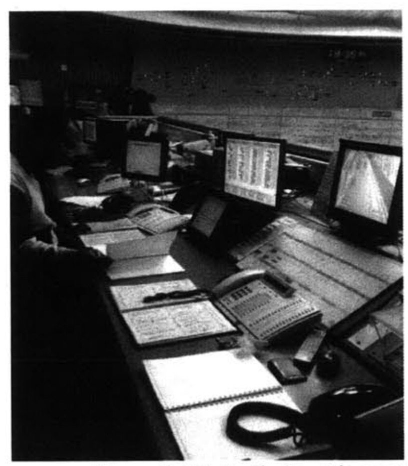


Figure 2-10: Earl's Court's controllers desk

The controllers are in charge of implementing all corrective actions when disruptions are observed. They work with a screen tracing train lateness. The screen indicates every value of train lateness at four given stations. A color code is used to indicate lateness. Green lines represent trains which are less than 5 minutes late. Orange lines indicate trains between 5 and 10 minutes late. Trains with more than 10 minutes of delay are displayed in red. Figure 2-11 is a screen shot of such a screen. This color coding may create a bias towards train lateness versus train frequency. During in-depth interviewing with controllers, the main goal that is mentioned is to return the lines to green.

Every corrective action is written in a manual logbook for personal records. These manual records are archived for several years but are accessible only in case of major fault or legal concerns. In most cases, the logbooks are used for controllers to transmit information from one shift to the other or for their own purpose to keep track of their operations. The service control manager does not sit in the control room but is also in the control center. His role is to manage the controllers and give feedback to the service manager who is usually a more senior employee. Service control managers are in charge of delivering daily summaries of operations. This system is effective to give direct feedback to operating staff in case of major faults. However, it does not allow more refined feedback on the type of recovery strategies deployed. The daily summary of operations does not precisely retrace the corrective actions that controllers implemented, as these are only recorded in paper logbooks. The service control manager therefore does not have access to enough information to give detailed feedback on the type of corrective actions that were implemented on a given day. Discussion with staff in place suggest that rebuilding such a database would be extremely helpful.

The line information manager communicates with the controllers and is the interface between the operation staff and the station staff as well as the public. Typically, the information manager is the person in charge of updating the publicly available dashboard to indicate the level of disruption to passengers. The display of the room enables every player to communicate easily with each other. To contact staff outside the control room, telephone is the main mode of communication. A new control room is in the process of being developed at South Kensington to replace Earl's Court. The general organization will stay the same. A major change will be the potential of rotating roles. As is done on other lines including the Central, new staff will be trained to operate interchangeably in the roles of signalers, controllers and line managers. This will allow each player to have a better understanding of the complexities of given roles and should ease communication between the groups.

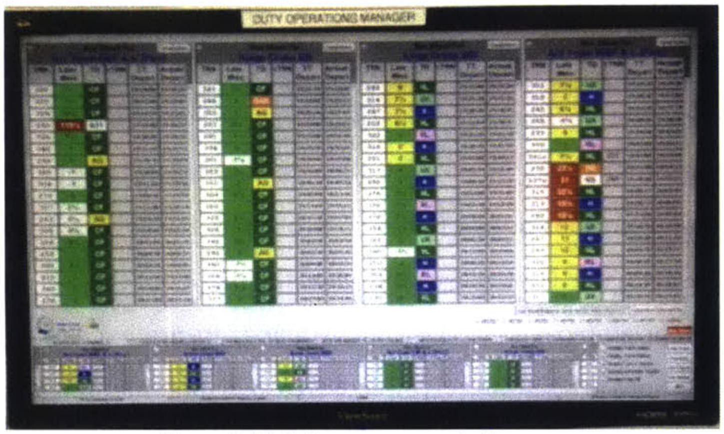


Figure 2-11: Lateness screen shot

2.4.2 Communication

Figure 2-12 illustrates the flow of information between the key players during disruptions. The control room is the main center of decision-making. However, in the case of the London Underground, train and staff management are separated. The duty manager trains, or DMT, is in charge of dispatching drivers to the trains. In case of a change of schedule, the controllers must first contact the crew depot to verify that their operations are feasible from the crew perspective. These calls are time consuming and communication issues can block a potentially successful recovery strategy. Controllers may propose a recovery strategy that includes a number of reformations. If the manager in the crew depot is not able to re-assign the drivers to the requested shifts, the controllers' optimal plan must be revisited. As described earlier, Babany (2015) proposes an optimization framework to automatically create an optimal recovery plan given crew constraints.

Besides the crew depot, both controllers and crew depot managers must communicate directly with the drivers in the field. Station staff may also be helpful to resolve some incidents. Since the 2012 Olympics, an incident response cell was created to deal with serious incidents. This cell is composed of specialized units and police to assist during a severe incident, notably for situations where legal issues are likely to arise (passenger fatality, severe injuries, ...).

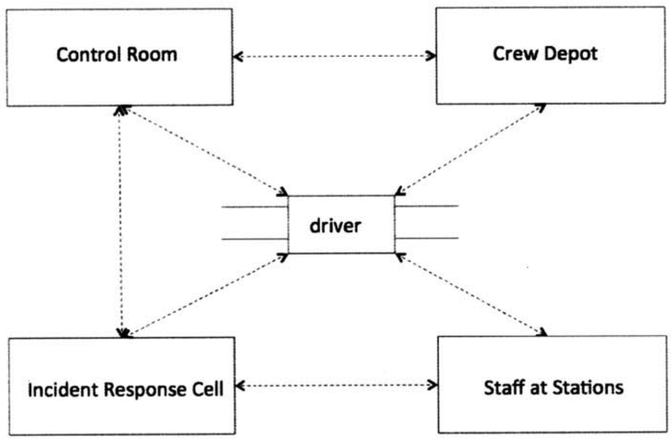


Figure 2-12: Communication during disruptions

Chapter 3: Track tracking data availability and reliability

Background information on service recovery on high frequency lines was provided in chapter 2. The emphasis of this chapter will be on Automated Vehicle Location (AVL) data which is the main source of data the research is based on. Section 3.1 describes the various data sources that can be used to qualify and quantify events linked to service recovery. Section 3.2 describes the various uses of AVL data, both during service recovery and for performance tracking. Section 3.3 and 3.4 describe the AVL data available on the Piccadilly Line where information concerning train location is provided through two independent systems: Netmis (Network Management Information System) and CTFS (Controlled Train Following System).

3.1 AVL in context

Section 3.1 introduces the range of data feeds available in the context of disruptions, giving their main characteristics and limitations. The second part of this section defines and describes the main characteristics of AVL.

3.1.1 Data generated during disruptions

This section lists some of the available data recorded during disruptions. As seen in Figure 3-1, some data is recorded automatically while other information is manually generated.

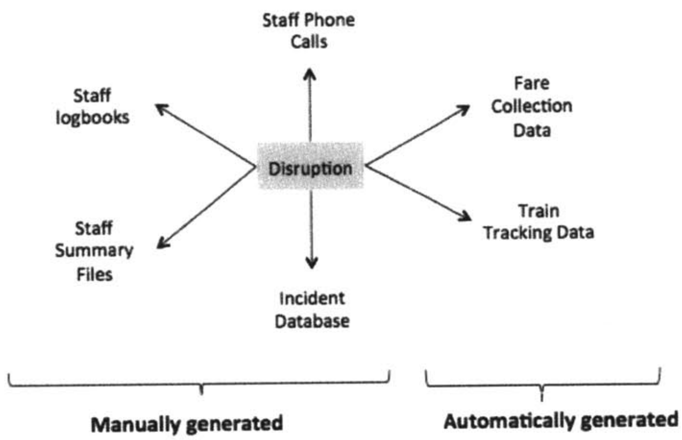


Figure 3-1: Data generated during disruptions

Most agencies rely on automatically generated data to monitor the system's performance, specifically fare collection data and train tracking data.

Fare collection data

Automated Fare Collection data, is recorded through the information system that is installed to collect fares from passengers. On the London Underground, all entries and exits are recorded through smartcard or a credit card taps, which provide a high level of spatial and temporal resolution for passenger data. The AFC data specific to the London Underground is referred to as Oyster data. Freemark (2008) uses Oyster data to measure the impact of disruptions on passengers. Even though Oyster data will not be the main source of data used in thi research, it is a reliable and valuable data source that provides an accurate description of demand. It is not available in real time. In other systems, fare transactions occur only at entry gates which limits the information available to entries per station. In such cases, it is possible to estimate the spatial and temporal demand using origin-destination inference (ODX) algorithms such as those developed by Gordon (2012).

Train tracking data

Train tracking systems provide data on train location. Most rail agencies have such systems since it is a key element in signaling and control systems. The most common train tracking systems rely on the signal system. Track are divided longitudinally into blocks. When a train crosses a block boundary, this event is detected by the track circuit and the time is recorded. Based on this data train trajectories can be developed. In addition to the signaling system, some agencies use radio signals as well as GPS to increase the reliability of train tracking. GPS is viable only at grade (or elevated) lines where direct communication between satellites and trains is feasible. Section 3.2 and 3.3 will discuss the characteristics and importance of this type of data in depth.

Staff phone calls

As described in Chapter 2, incidents require communication between many different players. In particular, controllers must call drivers and crew depot managers to inform them of recovery actions. Some of the communication occurs in the control room, as signalers and controllers are usually located in the same room. The control room of Piccadilly Line of the London Underground described in chapter 2 is an example of a control environment where signalers and controllers interact in the same space. However, most crucial information is communicated through phone calls which are typically recorded. The phone records are used mainly for legal or disciplinary purposes (homicide, operation error resulting in severe impact on passengers,etc). They can be accessed only through specific procedures which they were not feasible for this research. However, if the records were available, it would be an interesting additional source of data which could provide valuable information on both the decision making process as well as the recovery strategies themselves. A voice recognition system could be used to transcribe the messages and subsequently a machine-learning algorithm could be used to infer recovery strategies from these patterns.

Staff logbooks

Most transportation agencies require controllers to record all significant corrective actions in logbooks. The logbooks serve two distinct purposes. First, they are a valuable working tool for controllers who use them to organize their thoughts and to create the recovery strategies. They are also used to keep track of the current strategy being deployed (list of trains reformed, drivers contacted, trains canceled, etc). The logbooks are kept securely and can serve as additional evidence in case of legal or disciplinary issues. Because of the manual format of most logbooks as well as the variability of presentation depending on individual controller preferences, it is impractical to parse large amounts of data for aggregate analysis. Scrutiny of logbooks can however be useful for disaggregate analysis of particular incidents.

Staff summary files

Most operating centers produce a daily summaries of the main aspects of train operations and related events. These files provide a high level picture of the main events that occurred on a given day. On the London Underground, these files provide the number of canceled trains, information on train lateness, and what types of incidents (if any) occurred. The documents are usually used only on a daily basis by the service control manager and are not shared with upper management or the analytics team within TfL. For the Piccadilly Line, information on cancelations is approximate (number of trains canceled per three hours time intervals) and does not record the location of the cancelation. The information is recorded manually and is prone to errors. This limits the full understanding of the recovery strategies implemented. It can however serve as a good validation source for the recovery strategy inference that will be developed in Chapter 5. The information can be parsed easily as all documents are numerical and follow a similar format. Figure 3-2 is an example of such a daily report.

Daily Report

Time Printed: 15 Jan 2016 10:35:42

Piccadilly Line
Service Manager's Daily Review
For 24 hours commencing: 0200
Timetable In Use: WTT54 (PIC)
Monday 21 December 2015

Snapshot Cancellations and Station Closures

Time	Train Cancellations	Cancelled	Scheduled	Station Closures
06:00	Customer Action (10) : (+ 10 unlisted trains) Person under train Turnpike Lane	10	41	Turnpike Lane
09:00	Signal System Failure (25) : (+ 25 unlisted trains)	25	78	None
12:00	Signal System Failure (25) : (+ 25 unlisted trains) Acton Town failure	25	68	None
15:00	Signal System Failure (15) : 224, 225, 226, 300, 310, 316, 320, 323, 340, 342, 350, 351, 352, 354, 357 Due to the earlier failure in the Acton Town area.	15	68	None
18:00	Signal System Failure (2) : 226, 324 Due to the earlier failure in the Acton Town area.	2	77	None
21:00	Signal System Failure (1) : 322 Cancelled as part of service recovery which took place at Arnos Grove.	1	70	None
24:00	Full Service	0	56	None

Summary Of Late Running

Eastbound

Station	06:00	09:00	12:00	15:00	18:00	21:00	24:00
Arnos Grove	SEVERE	60	20	12	5	4	2

Figure 3-2: Daily report of the Piccadilly Line

Incident database

Incident databases record all incidents for a given line. These databases are generally generated manually by controllers or signalers, and are therefore prone to error. Important characteristics of incidents are type, duration, location, and time. As discussed in chapter 2, the CuPID system is the database used by the London Underground.

3.1.2 AVL features

As discussed in section 3.1.1, many different sources and formats of data are available in the context of disruption management. This research focuses on train tracking data to provide a reliable and easily reproducible framework for disruption analysis. Train tracking data is the most direct source of information that provides insight into train movements. It is available in most transportation agencies and the main features (timestamp, time number, station) are inherent in all systems. This makes it a valuable data type to base the research on as the framework developed can easily be applied to a large variety of transportation systems.

Depending on the structure providing AVL data (signaling system, radio signal, GPS, wifi etc), the precise fields included in the AVL data can change. In the context of this research, AVL data will be defined as data providing at least time-space information for a given train in the system. Generally, the train in service is represented by a unique train number that corresponds to the train number given in a schedule. AVL data sourced from train tracking systems provides information on train trajectories throughout the line. Other AVL data may be provided through train detection systems at stations, such as the Automated Vehicle Identification system used on the MBTA Green Line to detect when trains pass specific points on the line. More complex signaling systems or GPS tracking can provide additional information. In all cases, it is a key source of data to track train movements on a given rail system.

3.2 Uses of AVL data

AVL data is a key data source used by most transportation agencies to track overall performance. It is also the main source of data used for real-time control environments to inform the operating staff on train movements. Each of these functions is described below.

3.2.1 Use of AVL data for performance tracking

AVL is a main source of data that transportation agencies rely on to asses the performance of their system. Many different metrics can be calculated from AVL data. The main focus of most transportation agencies are metrics linked to train frequency, headway regularity, and lateness. These three metrics are directly linked to train movement data available through AVL.

Average waiting time is a key metric of system performance for the passenger side. As described in section 2.1.1, the average waiting time E(T) is a function of both the mean headway and it's coefficient of variation. Transportation agencies can therefore use AVL data to measure both the mean headway and it's coefficient of variation to extract the average waiting time.

$$E(T) = \frac{E(H)}{2}(1 + cv(H)^2) \qquad (3-1)$$

It is important to acknowledge that more sophisticated measures of a system's performance exist and include passenger data. Wood (2015) develops a framework for measuring transit reliability relying mainly on AFC data. The main metric is the reliability buffer time (RBT) that is defined as the difference between the Nth percentile and the median travel time over a given origin-destination pair and a time period. This metric is extremely useful to measure the reliability of a system and is more easily calculated with passenger data as input.

Aside from key metrics that are calculated at an aggregate level, transit agencies also use visual tools to track the performance of their systems. These tools provide a snapshot of the state of the line at a given time and are available both for real-time monitoring and ex-post analysis.

Figure 3-3 illustrates how the London Underground relies on graphs representing the scheduled headways (upper line) and the observed headways (lower line) at chosen stations. This tool is called "HQ clocks" and was inspired by the historic use of cards punched when a train passed a station. The tool enables easy tracking of headway regularity. Graphs representing headways can be used by the operating staff to give direct feedback to the drivers and controllers.

In addition, time-space diagrams are powerful tools used to illustrate train movements. The London Underground refers to them as waterfall diagrams. Figure 3-4 is an example of such a chart from Rahbee (2011). Many other transit agencies rely on similar time-space diagrams. A line represents the trajectory of a train, the slope of the line corresponding to its speed.

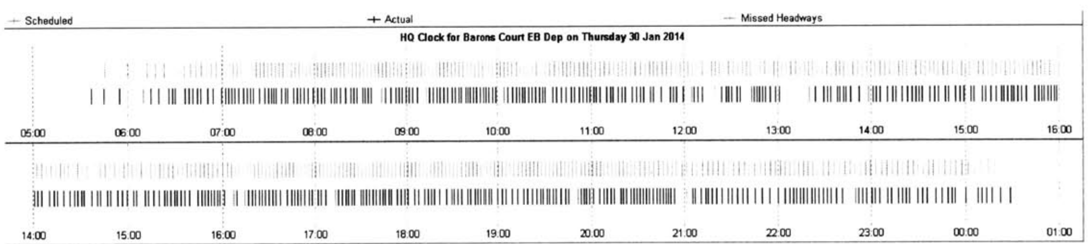


Figure 3-3: HQ Clocks

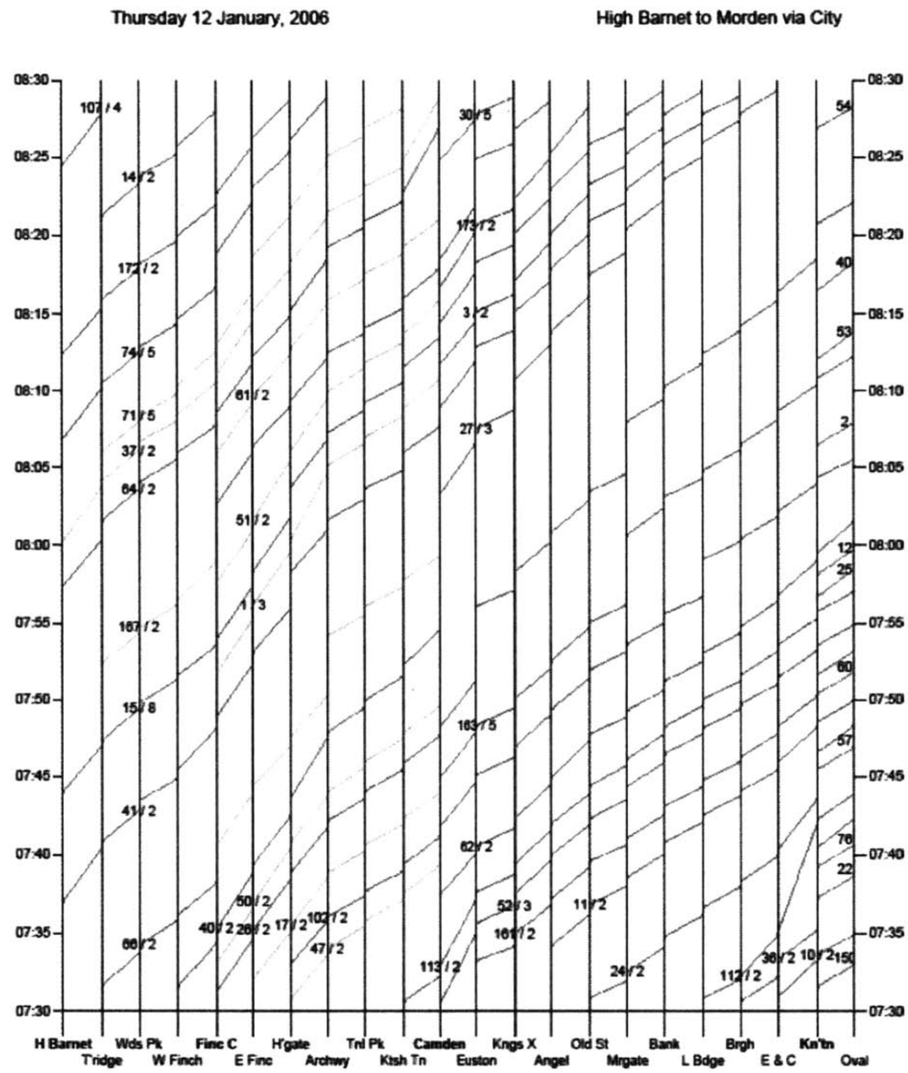


Figure 3-4: Waterfall diagram

3.2.2 Use of AVL data for service control

AVL data is constantly used during operations. Most control centers rely on AVL data to track trains in service. As an example, Figure 3-5 is a snapshot from the control center of Newcastle-upon-Tyne (Smith, 2011). Tracking the real-time position of trains enables controllers to react quickly in the case of anomalies and also to identify possible recovery strategies in the event of a disruption. Short trips can only be ended at reversing points. Similarly, AVL data can indicate bunching or large gaps, in which cases the controller may deploy holding strategies.

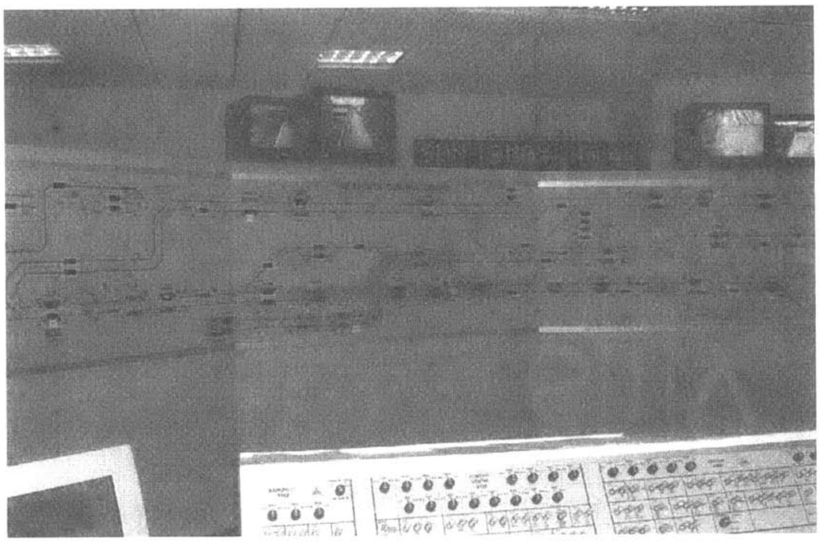


Figure 3-5: Newcastle upon Tyne control center

In addition to train tracking, it is also important for the operating staff to track train lateness. In the case of the Piccadilly Line, this information is available to controllers in real-time.

3.3 Netmis data

Both Netmis and CTFS systems provide data that contains train time and space information. According to our previous definition, Netmis and CTFS data are AVL data.

3.3.1 From TrackerNet to Netmis

Information concerning train location for every line of the London Underground is provided through a system called Netmis. The Netmis system is only the end system of a more complex flow of data represented in Figure 3-6. First, track circuits indicate the presence of a train on a given track circuit, or block. This information is sent to local computers that transfer it to the TrackerNet system. The TrackerNet system then aggregates the information received by the signaling system as well as information provided by radio signals sent from transmitters located at every station. The system correlates the radio signal with the track circuit information to retrace trajectories and infer the train number corresponding to a detected train movement. This information is sent to Netmis. Netmis processes TrackerNet information into a user friendly format with stations and time stamps (the movements between stations are not passed from TrackerNet to Netmis).

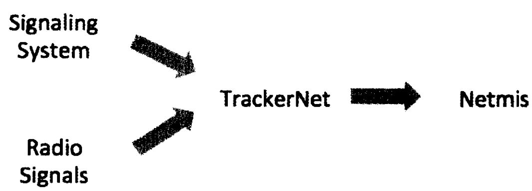


Figure 3-6: Flow of data from sensors to Netmis

The Signaling System

The signaling system on the Piccadilly Line is organized in six sections which operate with five different signaling systems. This is due to the history of contracts linked to the renewal of signaling infrastructure. This partitioning leads to lower reliability of train tracking data at the boundaries between different signaling systems. Figure 3-7 is a schematic illustration of the signaling partitions used by Trackernet based on discussions and maps provided by staff.

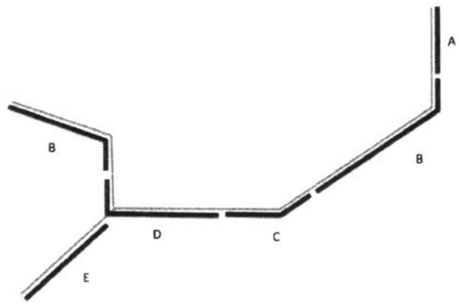
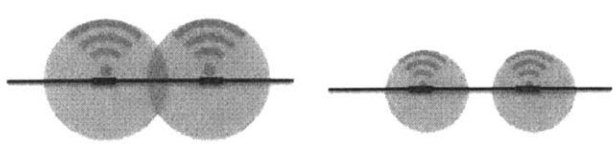


Figure 3-7: Sections of the signaling system

Radio Signals

In addition to the signaling system, TrackerNet relies on radio signal data to track train numbers and train destinations. The limited coverage of connect radio in certain tunnels of the Piccadilly line, also has a direct impact on data quality. Figure 3-8 (a) illustrates the theoretical coverage provided by connect radio. The line should be totally covered by connect radio in order for the system to effectively track a given train number. Humidity and aging infrastructure lead to situations represented in Figure 3-8 (b). The radio coverage is not continuous over the entire line. When the train enters a portion of the line that is not covered by radio, the computers lose track of the train. This results in gaps in the Trackernet data. The system relies on the continuity of information to infer train numbers. In cases of poor radio connect coverage, the system may assign an erroneous train number as the train number information will have been lost.



(a) Good coverage (b) Poor coverage

Figure 3-8: Radio connect coverage

3.3.2 Description of features

Netmis data is available on all London Underground lines and is the database on which the majority of the London Underground's internal performance metrics rely. It is retrieved in a user-friendly way through an Excel Macro via SQL queries. One month of AVL data for the Piccadilly Line corresponds to approximately 50 Mb of data. Table 3-1 summarizes the information included in Netmis database.

Table 3-1: Fields provided in Netmis data

TRNEVNT_ID:	*unique ID for a given data point*
TRAFFIC_DATE:	*Calendar date for the recorded event*
TIMESTAMP:	*time stamp (precise per second) for the recorded event*
TRACK_CIRCUIT_NAME:	*Reference of the track circuit corresponding to the location of the train*
SUTOR_CODE:	*Station code referred by three letters*
TRAIN_NUMBER:	*Train number (from 230 to 373), 0 if no train number associated with the event*
LINE_ID:	*7 for the Piccadilly Line*
TRAIN_DESTINATION:	*Code for the destination of the train*
TRAIN_IDENTIFICATION:	*Code referring to legs of trips made by a same train*
ACTUAL_DEPARTURE_TIME:	*Departure time of the train*
LEAD_CAR_NUMBER:	*Reference for the lead car of the train*
PASSENGER_LOAD_STATUS :	*NA for Piccadilly Line, but for other lines records the weight of the total train (proportional to the number of passengers riding)*

3.3.3 Time Space diagrams

Time-space diagrams are an effective way to visualize AVL data and this format will be used throughout the thesis to represent train trajectories and illustrate data reliability issues. These diagrams are similar to the waterfall diagrams presented in section 3.2.1 but only present one train per graph. This section describes the process through which these diagrams are created.

Raw Netmis data can be queried from TfL's data warehouse and extracted as a csv file. This data is exported to the data analysis software R for further treatment. Using R, the fields unrelated to date and timestamp, location, and train identification are discarded. The program filters the dataset by train number and date, as a time-space diagram is unique for every (train number, date) pair. The station codes are transformed into numerical codes. For the section of the line with no branches, consecutive stations on the line are given consecutive integers, from north to south. For the western branches, the program attributes integer numbers to the stations on the Heathrow branch and the same numbers shifted by a half unit for the Uxbridge branch. Thanks to this technique, the north-south trajectory is depicted and there are no large gaps on the diagram that correspond to different branches. Different shapes are used to represent the two branches. The triangular points correspond to the Uxbridge branch and the stars to the Heathrow branch.

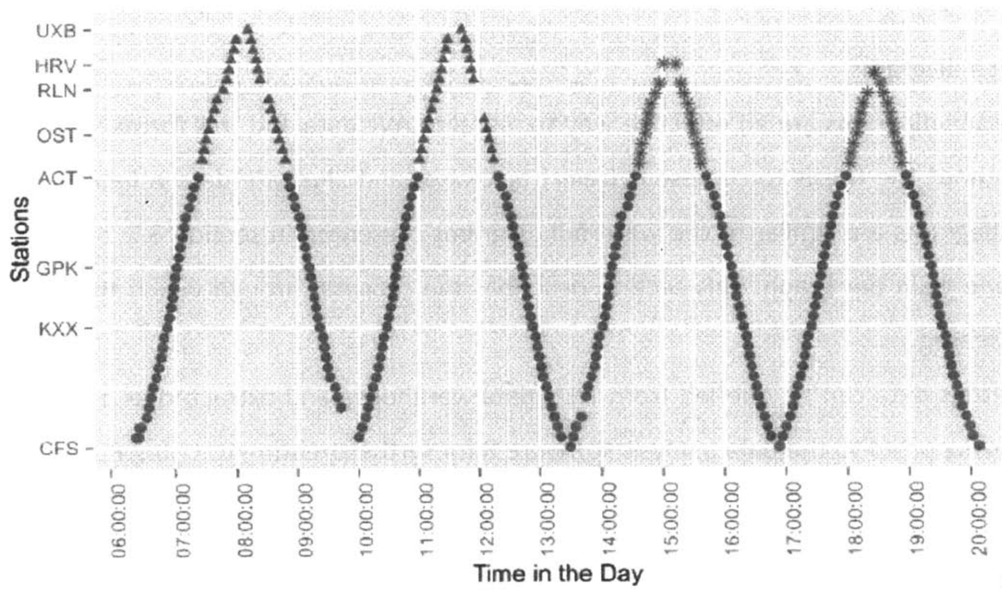


Figure 3-9: Sample Time Space diagram

Figure 3-9 illustrates such a time-space diagram, for train number 312 on 12 October 2013. We can observe that this train ran four round trips, twice to the Uxbridge branch in the morning and twice to the Heathrow branch in the afternoon. This diagram is also an effective way to observe any data discrepancies or quality issues in the data. A few gaps can be seen around 9:50 am near Cockfosters and around noon near Rayners Lane. These gaps where only a few data points are missing do not impede the general understanding of the train's trajectory. As will be seen later on, in some cases the gaps are much larger and result in large time windows where no reliable information on the train's trajectory is available.

3.3.4 Data quality

AVL data quality and reliability are key issues for train trajectory analysis. This section will describe the various data quality issues encountered with the Netmis data which can affect analysis. Table 3-2 illustrates the Netmis database imported into R after discarding the unwanted fields.

Missing Values

Table 3-2: Netmis data

TIMESTAMP	TRAIN_NUMBER	TRAIN_IDENTIFICATION	ACTUAL_DEPARTURE_TIME	SUTOR_CODE	TRAFFIC_DATE
	302	1013080	03:12:51	HOL	01/10/2013
	0	1013108	03:16:28	HMD	01/10/2013
	0	1013405	04:46:40	NFD	01/10/2013
04:47:40	0	1013405	04:47:25	SEL	01/10/2013
04:48:12	262	1013429	04:50:23	OST	01/10/2013
03:49:00	0	1013405			01/10/2013
03:49:30	0	1013405			01/10/2013
03:49:42	0	1013405			01/10/2013
04:50:44	0	1013405	04:51:11	ACT	01/10/2013
04:51:41	0	1013405		ACTS	01/10/2013
04:52:14	262	1013429	04:52:52	HNE	01/10/2013
	0	1013439	04:53:08	NFD	01/10/2013
04:54:00	0	1013439	04:53:53	SEL	01/10/2013
04:54:12	262	1013429	04:54:33	HNC	01/10/2013
03:55:19	0	1013439			01/10/2013
03:55:43	0	1013439			01/10/2013

As can be observed in Table 3-2, many fields have missing values. Table 3-3 summarizes the frequency of missing values in of the October 2013 Netmis database. The percentage of missing points are calculated in R by counting the total number of lines where a value in the corresponding field is missing, divided by the total number of rows in the database.

Table 3-3: Percentage of missing points

Field	Percentage of missing points
TIMESTAMP	1.3%
ACTUAL_DEPARTURE_TIME	16.0%
TRAIN_NUMBER	24.4%
SUTOR_CODE	15.0%

The field Timestamp rather than ACTUAL_DEPARTURE_TIME was used in this study as it more reliably available. The location as described by the SUTOR_CODE field is missing in 15% of cases, which is significant. The other key observation is the very high number of missing train number values. This research focuses on re-establishing the correct train number values in this database. This is made possible by the use of another database that has higher reliability in train tracking.

Another significant amelioration of the database would be to work on the missing information concerning locations. It has been observed that often missing location is correlated with missing train number. Approximately 25% of points that are missing location are also missing train numbers. This makes the reassignment more difficult as more information is missing for these points.

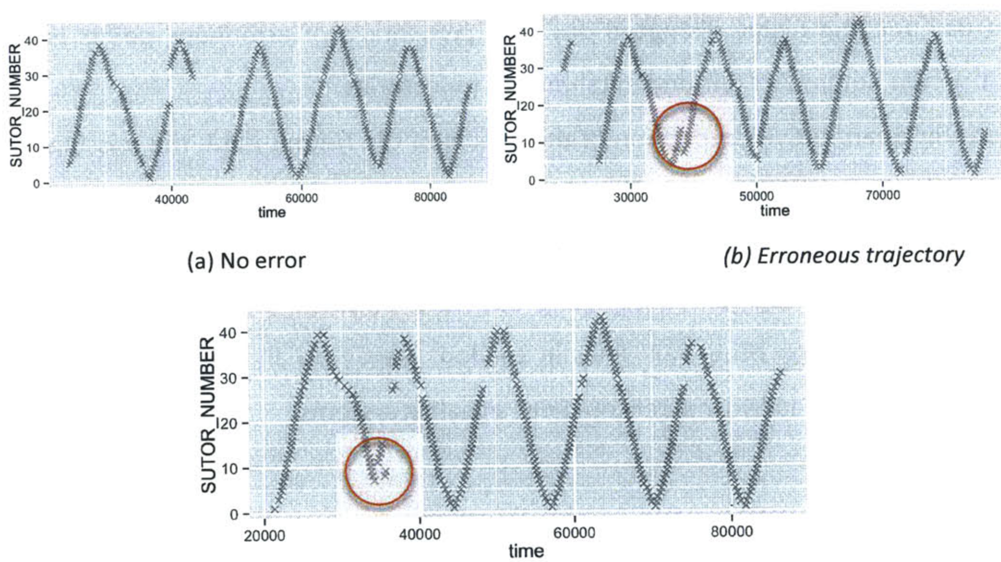
Erroneous data points

In addition to missing values, some data points are erroneous. To quantify the importance of these errors, a visual analysis was performed. Indeed, the difficulty to quantify correct and incorrect points algorithmically combined with the effective time-space diagram visualization tool strongly encourages a manual count rather than an algorithmic approach. It is very easy to use visual intuition (continuous trajectory) to detect erroneous legs or points on the diagram. It is much more complex to implement a code that will automatically distinguish an erroneous leg from a valid leg, or erroneous points. A more numerical and reproducible approach to the analysis of erroneous points would be beneficial to this study but is beyond the scope of our research.

The manual analysis is performed on two days of data chosen to represent various states of the line. The first day (2 October 2013) is chosen as a day with few disruptions (no observed incidents reported on the daily summary) and a low number of reformations. The second day (10 October 2013) is chosen as a heavily disrupted day with a total of eight canceled trains and 55 reformations.

There are two main types of errors. The first is an invalid trajectory due to a leg that is visibly erroneous. This error is quantified as the percent of erroneous observed trajectories. A trajectory is defined as a back and forth movement of a train. This count is implemented by the number of observed trajectories and the number of erroneous trajectories.

The second type of error is erroneous points. These points are points that are visibly out of the scope of a feasible trajectory and exist due to noise in the data. The protocol to quantify this error is to (visually) count all the points that are considered erroneous and divide by the total number of observed points that have a train number (given directly by R). This method is easily applicable to a given day as only a few points per train are considered erroneous. To implement both of these counts, all the time-space diagrams corresponding to all the trains that have run for the given day are produced automatically through R and checked manually. Results are given in Figure 3-10. Figure 3-10 illustrates a) correct trajectories, b) erroneous legs and c) erroneous points observed. The time is in seconds after midnight.



(a) No error

(b) Erroneous trajectory

(c) Erroneous points

Figure 3-10: Time-Space Diagram showing data quality issues

Table 3-4: Results of the count of errors

Day:	**10/02/2013**	**10/10/2013**
Count of erroneous trajectories	5	4
Count of erroneous points	2	11
Percentage of erroneous trajectories	1.4%	1.2%
Percentage of single erroneous points	0.01%	0.05%
Percentage of total erroneous points (single points + erroneous trajectories)	0.76%	0.6%

Results show that there are a small number of erroneous data points both for the non- disrupted day and the highly disrupted day. The percentage of erroneous points is under 1% and it should not impact our understanding of trajectories.

Based on this quantitative description of the quality of the Netmis data, the missing train numbers on the Piccadilly Line is the largest limitation of the existing database. The issue of reliable train tracking is known by the operating staff and planners at TfL and is in part due to the aging and fragmented signaling structure of the Piccadilly Line. This limitation of AVL data exists on other rail lines in London, mainly on all the Overground Lines that, similarly to the Piccadilly Line, rely on older signaling systems.

3.4 CTFS Data

This section describes the second source of AVL data which is the Controlled Train Following System, or CTFS, data. It discusses both the data features and the complementarities between CTFS data and Netmis Data.

3.4.1 Description of features

CTFS data became available with the Piccadilly Line extension to Heathrow. The signaling system communicates with local computers that were installed as part of the extension. These computers transfer the information directly to the CTFS system. The data is sent in real-time to the Earl's Court control room to track train lateness, as discussed in Chapter 2. In addition, the data is saved and can be retrieved in raw binary format that is then processed by the IT department of TfL into a .txt format. This .txt file can be imported in R and transformed to a user-friendly .csv format resembling the Netmis format. One month of data is equivalent to 1.5 Mb. Table 3-5 (b) is a snapshot of the original text format. Table 3-5 (a) describes the data fields.

Table 3-5: CTFS fields

Field	Description
Station	Station at which the event is recorded
TRN	Train Number
rTRN	*: If the train was reformed, contains the initial number, otherwise NA*
TD	*Train destination*
Expected Time	Time of arrival at the station according to the schedule
Actual Time	Observed time of arrival at the station
Lateness (s)	Observed algebraic lateness in seconds

(a) Field Description

```
***** Sudbury Hill EB *****
TRN     TD       Expected Time    Actual Time      Lateness (s)     rTRN
-------------------------------------------------------------------------
274     A2S      06:53:30         06:52:48         -42              0
360     A2S      07:08:30         07:07:56         -34              0
275     A21      07:23:00         07:24:21          81              0
```

(b) CTFS snapshot

3.4.2 Data Quality

Missing stations

The spatial resolution of CTFS data is poor, as can be seen in Figure 3-11 where the available CTFS stations are represented in grey. The CTFS data available for this research covers only 33% of all stations and there is no available data toward the end of the Uxbridge and Cockfosters branches. This low spatial resolution on the branches limits in particular the detection of short trips, which is a key strategy used during disruptions.

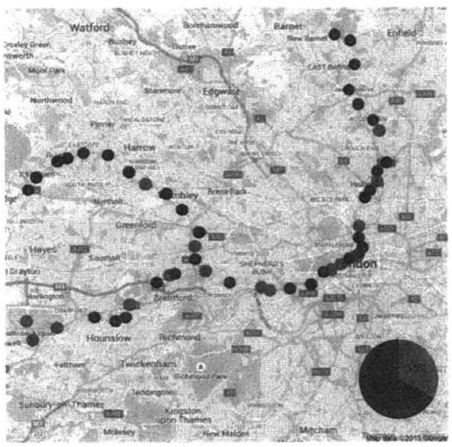


Figure 3-11: Stations available through CTFS data

Erroneous points

A similar methodology to the one presented in section 3.3.4 was implemented on the same days using CTFS data. Counting both single erroneous points and erroneous legs, the levels of error were less than 2% for both days. CTFS can therefore be considered a reliable source of data for tracking trains on the Piccadilly Line.

Chapter 4: Building a reliable train tracking database

A reliable and complete AVL dataset is crucial to accurately retracing train movements and operations. As described in chapter 3, the available AVL data is limited in terms of spatial and temporal resolution and incomplete, preventing the full representation of train movements. This chapter presents the methodology and algorithm used to build a more reliable AVL dataset, based on a combination of imperfect data sources. Section 4.1 describes the methodology, referred to as the *merging process*, which combines information from different AVL databases. The AVL databases used in this research are Netmis and CTFS which are both available on the Piccadilly Line, as described in chapter 3. However, this methodology could be applied to other databases that have similar characteristics. The merging algorithm is presented in section 4.2. This algorithm uses various parameters that are discussed in section 4.3. Section 4.4 outlines the application of the algorithm and section 4.5 presents the results and validation.

4.1 Methodology

4.1.1 Conceptual Framework

Different databases that record the same physical events can provide various types of data with different levels of accuracy and completeness. The merging method queries the different databases, matches data points that correspond to the same event, combines data for the event, and builds a new data set based on the combination.

For the Piccadilly Line, the Netmis and CTFS databases both track train movements but they originate from different information systems and therefore have different fields of interest and characteristics and are largely independent so that an error in one database may not exist in the other database. As described in chapter 3, Netmis has excellent spatiotemporal resolution but many records are incomplete. Specifically, about 25% of records do not contain a train number. The CTFS database on the other hand provides much weaker spatial coverage (only about 30% of all stations are covered) but is comprehensive in terms of train numbers. The merging process matches Netmis points with corresponding CTFS points for a given train trajectory and updates the train number in Netmis when the value is missing. Figure 4-1 represents the general workflow that was applied.

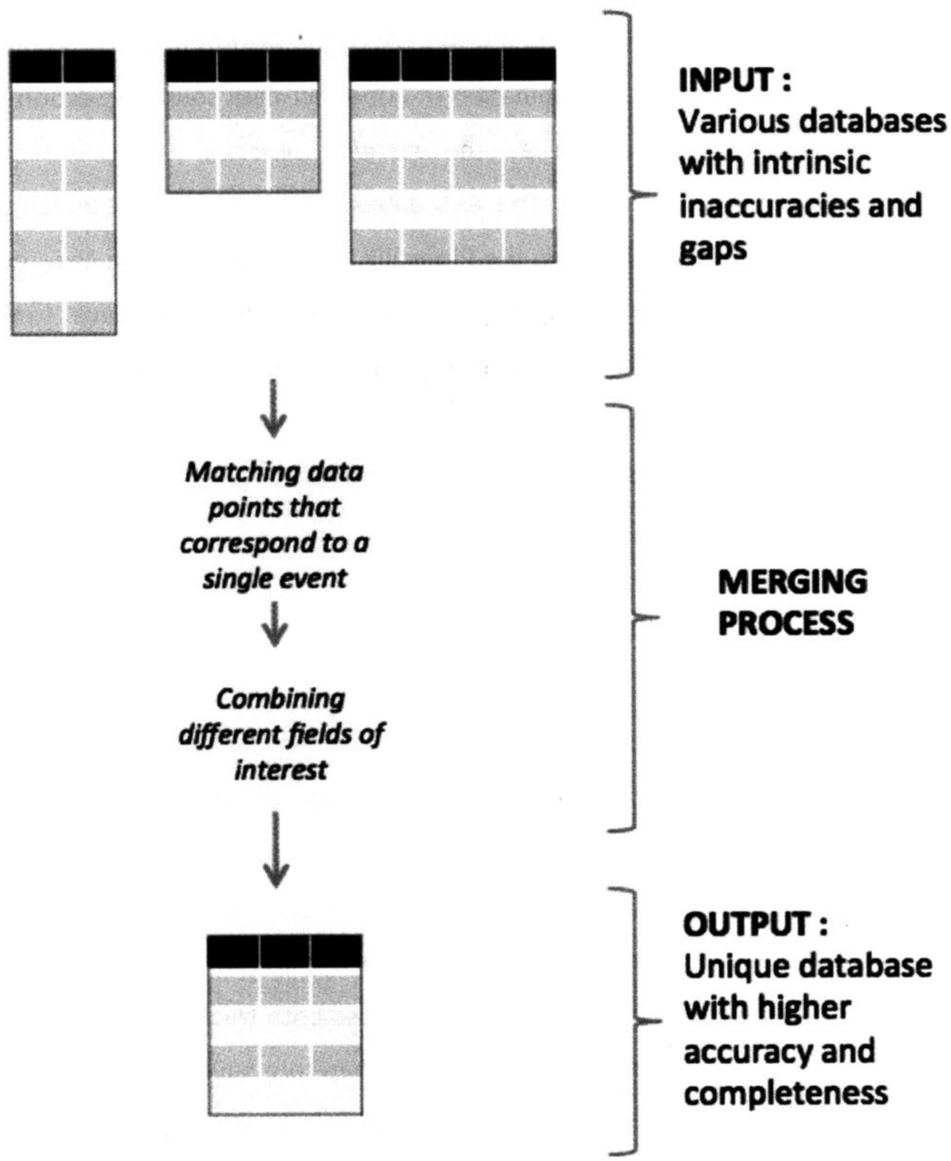


Figure 4-1: Fusion process

4.1.2 Database discrepancies

The matching of data points is based on fields that are common to both databases. In the case of train movements, the common fields are time and location. Because the databases originate from different physical information systems, there can be significant discrepancies in values for an identical field and the same event. For example, the timestamp of the Netmis and CTFS values are not strictly identical because the systems that provide data to the Netmis and CTFS databases have different characteristics (internal clock, underlying assumptions, etc). Furthermore the information system's occasional errors can lead to additional differences between values that should be identical. In the case of the Piccadilly Line, a signaling incident that leads to a momentary malfunction in the Trackernet system or a systematic delay of a few milliseconds on the connect radio in certain regions could both impact the accuracy of the timestamp value.

These possible discrepancies make the matching process more complex, as it must deal with both deterministic and stochastic sources. The matching process links data points that correspond to the same event with high probability, but there is a non zero chance of error. The sensitivity analysis and the test of the merging algorithm on a sample of data will be described in section 4.3 66 and will assess the accuracy of the matching process.

4.1.3 Preliminary study on the time difference

The matching of CTFS and Netmis data is based on comparing spatiotemporal values of data points. The value of time and location are indeed the unique and (generally) reliable common fields of the two databases. For a given station location, two points are assumed to correspond to the same physical event if they have similar timestamp values. A margin of error is applied to account for the discrepancies between databases noted above. A first important step is to understand the possible intrinsic differences between the CTFS and the Netmis data points. In the databases, both fields are named "ACTUAL_DEPARTURE_TIME" and should correspond to the moment when the train has left the reference station.

A case study on one day (the 16th of January 2014) with regular service is used to gain a better understanding of these systematic time differences. The approach is to calculate the time difference between CTFS and Netmis data points that are known to refer to the same event. The analysis uses the statistical programming language R. The data points with known time, location and train number are selected in both the Netmis and CTFS datasets. They are matched by location and train number. Because a given physical train will usually pass through the same station several times a day, there exist multiple data points for a given triplet of date, location, and train number. These triplets are ordered by time of day to ensure that they are matched correctly. Tables 4-1, 4-2 and 4-3 illustrate the matching process with a hypothetical example.

Table 4-1: Netmis database

Location	*Time*	*Train N°*
Station A	08:31:20	230
Station B	08:46:30	230
Station C	08:58:15	230
Station B	09:12:35	230
Station A	09:15:45	230
Station A	08:35:45	231
Station B	08:50:25	231
Station C	09:02:15	231
...	...	...

Table 4-2: CTFS database

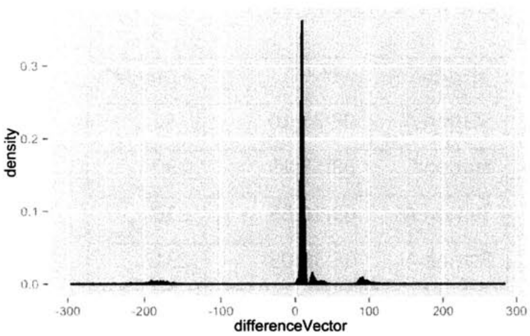
Location	*Time*	*Train N°*
Station A	08:33:10	230
Station C	08:58:15	230
Station A	09:16:55	230
Station A	08:36:05	231
Station C	09:03:05	231
...	...	...

Table 4-3: Simplified example of event matching

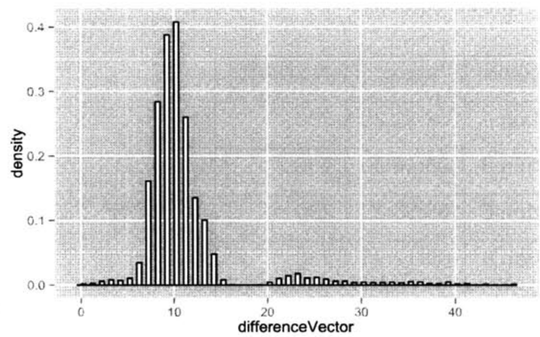
Location	*Netmis Time*	*CTFS Time*	*Train N°*
Station A	08:31:20	08:33:10	230
Station C	08:58:15	08:58:15	230
Station A	09:15:45	09:16:55	230
Station A	08:35:45	08:36:05	231
Station C	09:02:15	09:03:05	231

The difference between the CTFS and the Netmis timestamp for the same event is computed for a set of matched data points, over all stations and trains. The differenceVector is an array of the values of time differences in seconds between CTFS and Netmis for any matched data point.

$$differenceVector = CTFS\ Time - Netmis\ Time$$



(a): Error distribution



(b): Error distribution around zero

Range	Number of Points	Percentage
[-300,0)	298	5%
[0,50)	4808	89%
[50,100)	230	4%
[100,300)	83	2%

(c): Table of the global distribution

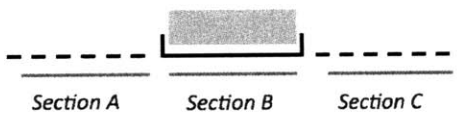
Range	Number of Points	Percentage
[0,5)	60	1%
[5,10)	2110	44%
[10,15)	2288	48%
[15,20)	21	0%
[20,25)	136	3%
[25,30)	87	2%
[30,35)	46	1%
[35,40)	44	1%

(d): Table of the near zero distribution

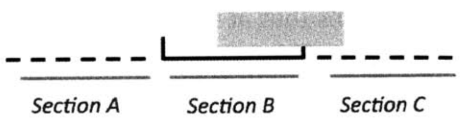
Figure 4-2: Distribution of the time difference

Figure 4-3 shows a systematic gap of approximately 10 seconds between CTFS and Netmis data points. This time difference exists due to intrinsic characteristics in the CTFS and Netmis information systems. The differences in underlying semantics explain this time gap. More precisely, different underlying semantics related to "DEPARTURE_TIME" (CTFS departure times are calculated according to Figure 4-3(c) and Netmis departure times correspond to Figure 4-3 (b)) are the main reason for the CTFS and Netmis discrepancies. The departure of the train from a station can be defined in different ways. A few seconds span between the moment the train starts moving and the moment the last car of the train exits the station. Figure 4-3 (a) shows a train at a given station, with the track sections that detect the train's presence. Figure 4-3 (b) illustrates a potential definition of departure time, when the first car of the train is detected on a track section beyond the platform.

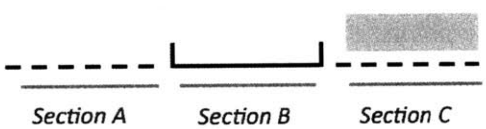
Figure 4-3 (c) is an alternate definition of departure time, when the train is no longer detected on any station track section.



(a) Train at station



(b) Head of train exits the station



(c) Last car exits the station

Figure 4-3: Departure time definition

The time difference τ is defined by the time between events in Figure 4-3 (b) and (c). With s the speed of a train leaving a station (approximately 33 k.h^{-1} or 33/3.6 m.s^{-1}), and L the length of the train (108 meters for the 6 car trains of the Piccadilly Line) we obtain an approximate value of time difference τ:

$$\tau = \frac{L}{s} = 12\,s \quad (4-1)$$

The time difference obtained via this simplified calculation is of similar magnitude to the systematic time gap between CTFS and Netmis. The matching process will incorporate this systematic time lapse to increase the chances of correctly matching two data points that are associated with the same event. Two other smaller group of points are detected around 23 and 90 seconds. These points correspond to erroneous CTFS time stamps that have an added time lag for a given event. These time points are rare and are not representative of the more usual 10 second time lag.

4.2 The merging algorithm

The goal of the matching algorithm is to fill in the correct train number in the Netmis dataset where initially the train number field in blank, thus it is applied only on the data points that do not have a train number. The algorithm has three steps that are illustrated in the workflow in Figure 4-4.

1) Identify on the Netmis data set legs of trips that relate to the same train
2) Match the CTFS data points that correspond to the given trip
3) Reassign the value of the train number to these data points. Figure 4-10 illustrates the various steps.

Filter Netmis points with no Train Number → ***Identify physical legs of trips*** → ***Compare data point values of the legs with CTFS*** → ***Match CTFS data points with Netmis legs*** → ***Assign CTFS train number to matching Netmis leg***

Figure 4-4: Workflow of the algorithm

4.2.1 Identifying legs of trips

Figure 4-5 illustrates the Netmis data that does not contain a value for the train number field on a particular day. Even though a pattern of superimposed trajectories similar to the ones presented in chapter 3 can be observed, it is difficult to group data points that correspond to a single train in all cases. Point-to-point processing, assigning each point to it's closest feasible neighbor, could potentially rebuild the trajectories of all trains, however with approximately 6000 data points to sort per day, this process would be extremely greedy in computational time and would not result in an efficient merging process.

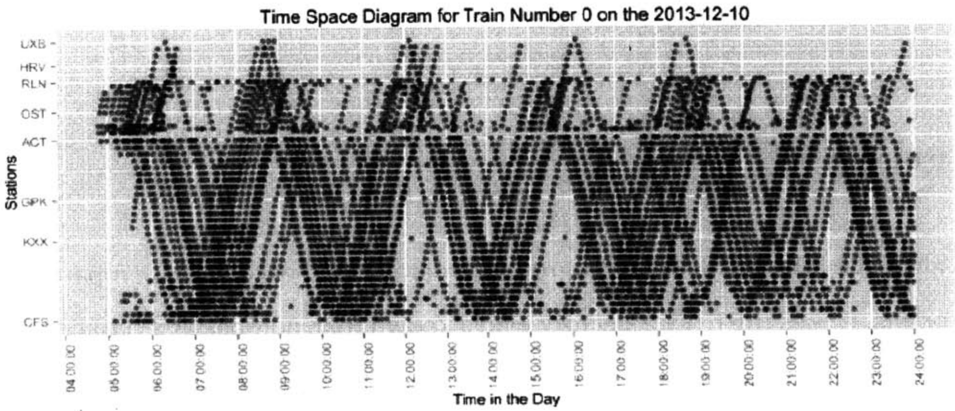


Figure 4-5: Time Space Diagram for train number 0

To reduce the computational time, the Train Identification field is used. This field, recorded directly through the signaling system, is linked to a car of a train. From the 6000 data points per day, there are approximately 300 unique train identification values for 86 unique train numbers. This leads to clusters of an average of 20 points that represent (parts of) trajectories. The time-space plot of points filtered by a given Train ID is presented in Figure 4-6.

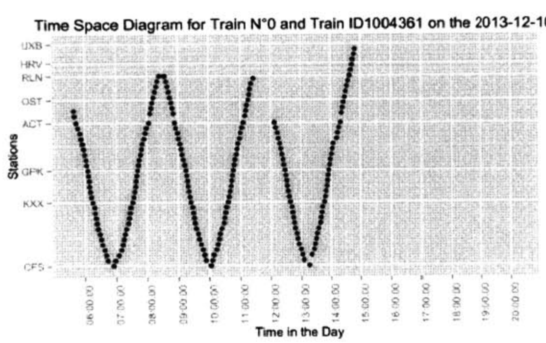
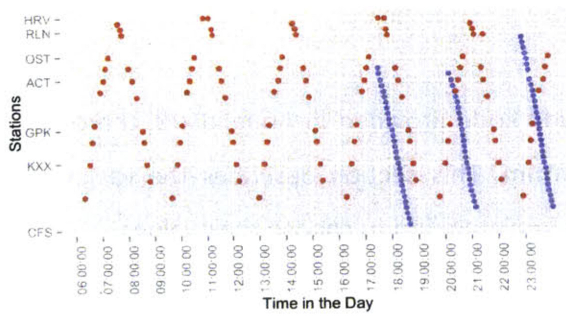


Figure 4-6: Time Space Diagram obtained after a filtering by Train Identification

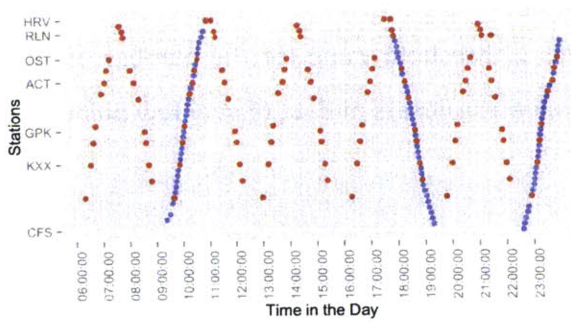
Time gaps in the data can be observed (for example between 11:45 and 12:00 during which the train movement is not recorded via the train tracking system. This could be due to internal processes in the Trackernet. The algorithm identifies these gaps and divides the data points into groups corresponding to continuous legs. These final groups are the entities that are considered as legs and that are compared to CTFS values to assign correct train numbers. Any given data point from this group will therefore be assigned the same train number.

4.2.2 Comparing the legs with CTFS values

Once the algorithm obtains the various groups of data points, each set is compared to the CTFS data points. It uses the CTFS trajectory with a known train number as a baseline, and compares it with the Netmis legs of trips. Figure 4-7 illustrates the concept. The blue points represent the Netmis legs and the red points represent the CTFS data points with the known train number. The algorithm looks at the intersection of the CTFS and Netmis data points by station, and computes the time difference between every point. Given the intrinsic difference in time measurement previously discussed, a certain margin of error is allowed. Points that are for the same station and close in time are considered to be matching. If the CTFS and Netmis data have a number of matching points above a threshold to be discussed in the parameters section, the Netmis leg is a match with the CTFS data.



(a) No match between CTFS and Netmis



(b) CTFS and Netmis match

Figure 4-7: CTFS and Netmis comparison

4.2.3 Reassigning the correct train number

When the algorithm concludes that a Netmis leg matches the CTFS data, the train number associated with the CTFS data is assigned to all the points of the Netmis leg. The process loops over all train identifications and all train numbers. The result is a database with the Netmis values and updated train numbers for the data points that were matched. The merging process can be seen as refining the data, as it uses one reference dataset to complete the other. The output is a dataset with high spatiotemporal resolution and more complete train numbers.

4.3 Setting the parameters

As seen in the pseudo code provided in Appendix B, there are six key parameters used in the merging algorithm. This section describes the various parameters and the sensitivity analysis that was implemented to choose their values.

4.3.1 Description of parameters

- *The number of points needed per unique train identification*

The algorithm applies a threshold value for the number of points obtained when filtering by train ID. This eliminates subsets of data of very few points that do not represent a trip.

- *The time gap between various legs obtained from the same train ID*

This parameter determines the maximum time gap authorized before grouping a given set of points into smaller groups of continuous points. As described previously, this parameter is useful to take account for reformations. The Train Identification is characteristic of a physical train but if a renumbering occurs, it is possible that various legs will have different train numbers.

- *The number of points obtained after dividing the data by continuous legs*

Similar to the first parameter, this parameter requires a minimum number of points from the groups that were created by dividing the data into continuous legs. If the leg is too short (typically only 2 or 3 points), the matching process will not be effective and it is possible that the wrong train number could be assigned to these points.

- *The time difference between a CTFS and Netmis point*

This parameter defines when a CTFS and a Netmis *(time,location)* pair is identical. As discussed previously, there exists an inherent time difference between the two databases. The parameter is the maximum that the absolute value of the difference between CTFS and Netmis time can have for two points to be viewed as identical.

- *The number of common points*

This parameter selects how many common points between CTFS and Netmis are required to conclude that a leg matches the CTFS data.

- *The ratio of common points to total leg parts*

This parameter is the minimum value of the ratio of the number of common points and the total number of points on the leg. If the leg contains a large number of points, a larger number of common points may be required to conclude that the two legs match. This ratio is low due to the low resolution of the CTFS data (a same leg is represented by many more points in Netmis compared to CTFS).

4.3.2 Choice of parameters and sensitivity analysis

The initial values of the various parameters were assigned using a heuristic approach. The merging process was applied on a few days with regular service used as case studies to develop the algorithm. For every matching leg, the algorithm automatically plots the given leg and the matching CTFS data. This facilitates the development of the identification of any major issues in the matching process and to apply a heuristic approach to determine an approximate value of parameters. In our case, the parameters are set as constants for the entire time window considered (three months).

A sensitivity analysis was then conducted to refine the values of certain parameters. For practical and computational reasons, the sensitivity analysis was implemented on single days with standard levels of service (no reported incidents and a low number of observed reformations). It could be beneficial to apply a similar methodology to a larger set of data. An other approach could be to vary the value of parameters for every day of the sample. This approach would result in an increase in merging accuracy but would simultaneously significantly increase the computational time.

The steps of the sensitivity analysis are as follows:

1) *Select the parameter to test as well as the range of values to be tested.*

2) *Apply the merging algorithm to a single day of data for a variety of values within the selected range. All other parameters are kept constant to test the effect of one parameter on the merging results.*

3) *The number of points matched for the various values of this parameter.*

This gives us access to the number of points matched for the chosen value of the parameter. The results of the sensitivity analysis suggest that there exists a threshold value for each parameter. In the example of the ratio value as seen in Figure 4-8, a ratio value below 0.6 results in a very high number of matched points. The constant high levels of matched points suggests that choosing a lower value of a threshold does not improve the matching rate. On the other hand, if the ratio value is set over 0.8, the threshold is too high and only a very small number of points will be matched. A similar methodology is applied to other parameters. However, in the context of the research, a choice based on visual judgment is applied. The results, as discussed in section 4.5, are satisfying for the level of precision required and do not necessitate more refined parameter choices.

Table 4-4 summarizes the parameters and the order of magnitude of the values used in the implementation.

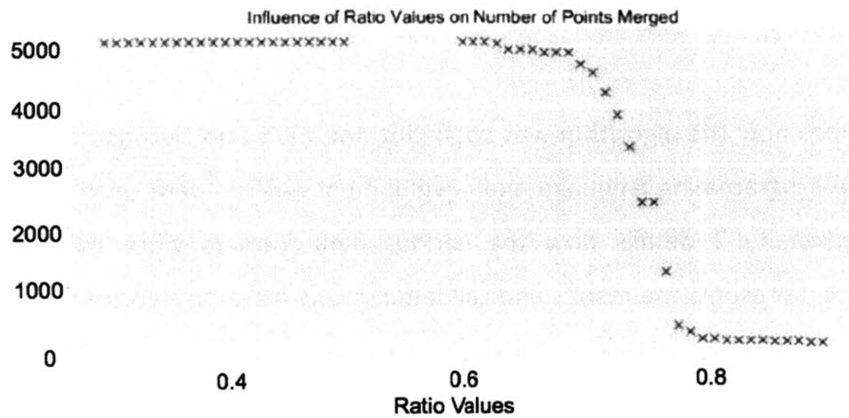


Figure 4-8: Influence of the threshold value on the number of points matched

Table 4-4: Parameter values

Parameter	*Threshold Value*
Number of points per Train ID	4 points
Time Gap between legs	30 minutes
Number of points after subsetting per time gap	4 points
Time difference between CTFS and Netmis	100 seconds
Common Points required	4 points
Ratio Value	0.7

The time difference between CTFS and Netmis in seconds is set to take into account the exceptional or recurring inaccuracies between CTFS and Netmis timestamps. The value should however not be higher than the headway as this could lead to data points being assigned to the number of the neighboring train.

4.4 Implementation

This section describes how the algorithm was applied to the CTFS and Netmis data. Section 4.4.1 presents the programming language and development environment used for the implementation. Section 4.4.2 details how the various data formats were made compatible, and section 4.4.3 presents the results and validation of the merging algorithm

4.4.1 RStudio

R was used as the primary programming language and statistical computing and graphics tool throughout the research. Rstudio, which is an open source integrated development environment compatible with R, was used as the user interface. R provides powerful statistical and graphical packages that were useful throughout the phases of data exploration as well as data analysis. In particular, most graphs were plotted through the popular package ggplot.

4.4.2 Formatting

An important part of the implementation consisted in formatting the various databases in compatible structures. The Netmis format was used as a template and the algorithm adapts the CTFS data to be compatible with the Netmis format. This choice was motivated by three factors. First of all, most of the employees in the planning department of Transport For London as well as in the London Underground rely on Netmis data. Our merging algorithm respects this format and can therefore easily be incorporated without having to adapt the existing analysis workflow that the London Underground uses. Second, Netmis data is available on all lines of the London Underground. Our merging methodology could be applied to other lines, it would be beneficial to have the output be compatible throughout lines. Third, the next step in the research workflow is to compare the merged database with timetable values. The schedule database is provided by Transport for London in a similar data format as the Netmis data is. Respecting the Netmis data format will therefore facilitate the comparison phase described in Chapter 5.

As presented in Chapter 3, the CTFS original data format is a text file representing all movements for a given day. The text file is a concatenation of numerous small tables containing the same columns presented in 3.4 66. and visible in Figure 4-9. These tables are linked to stations, introduced at the beginning of each new table by a line with stars and the spelled out name of the station. A first step was to write an algorithm that detects the star pattern and integrates the station name as a feature of the tables. The tables can then be merged into a single database that contains an additional "Station Name" column. The CTFS database was used per se for values concerning lateness and train reformations, two fields that are directly provided in this format. A second step was to format the CTFS database into a format similar to Netmis data. The column names were changed and the time stamp was converted into seconds after midnight. CTFS and Netmis do not use the same station name coding and an extra step was needed to transform the station column for further compatibility. Station numbers were used for simplicity.

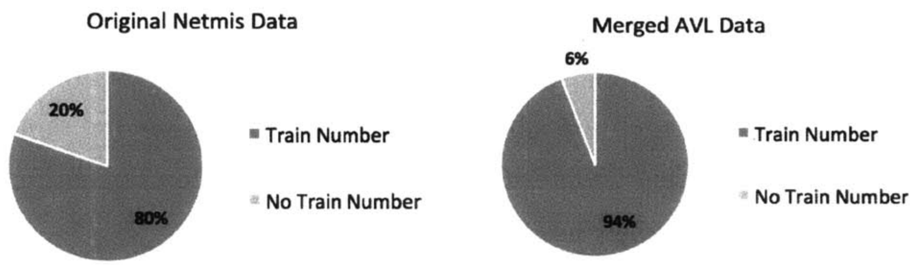
```
342     A2S     00:16:30        21:14:53        -97             0
232     A2S     00:31:30        21:31:20        -10             0
261     A21     00:46:30        21:45:27        -63             0
131     ATL     00:52:45        21:54:29         104            0
350     A21     00:59:00        21:59:09         9              0
--------------------------------------------------------------------

***** North Ealing WB *****
TRN     TD      Expected Time   Actual Time     Lateness (s)    rTRN
--------------------------------------------------------------------
232     RL      05:28:00        05:21:27        -393            0
234     RL      05:31:30        05:26:48        -282            0
130     RU      05:14:00        05:39:05         1505           0
243     UX      05:49:00        05:45:20        -220            0
```

Figure 4-9: Raw format of CTFS data

To process the databases efficiently during both the formatting and the merging process, the CTFS and Netmis databases are saved in lists, with one element of a list corresponding to one day of data. Using lists rather than two nested "for" loops resulted in a substantial gain in computational time.

4.4.3 Results

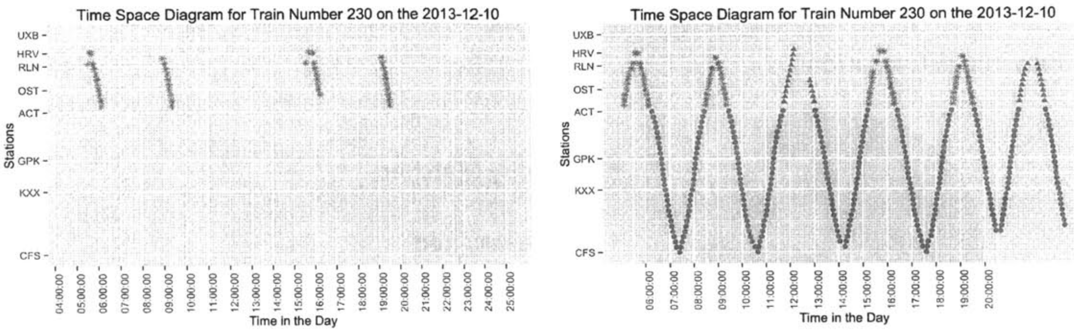
The merging leads to a significant improvement in the AVL completeness regarding train number values. One complete month of data (approximately 9*10^5 points) is used to measure the improvement of AVL quality. Figure 4-10 summarizes the proportion of rows from the data with no train numbers before and after the merging algorithm. In the original Netmis data, 20% of the data points lacked values for the train number field. In the merged data, only 6% of the data points lack values for the train number field.



(a) Original Data *(b) Merged Data*

Figure 4-10: Results of the AVL merging for on one month of data

This large improvement on AVL quality provides a better understanding of train movements. Figure 4-141shows the time space diagram of a given train using the original Netmis data in (a) and the merged AVL data in (b). Chapter 7 will discuss the multiple benefits of such an improvement in train number reliability, both for further research and direct use in transit agencies. In the case of this research, the merged database is useful to understand all the train movements that occurred during incidents, in particular to retrace the changes between the scheduled and observed movements. Chapter 5 will describe in more detail how the merged AVL database is used in parallel with the schedule data to infer the corrective actions implemented on the line. The next step is to validate the methodology on a test sample to verify that the inferred train numbers are correct.



(a) Original Netmis data *(b) Merged data*

Figure 4-11: Train trajectories

4.4.4 Validation methodology

The validation methodology tests the accuracy of our merging algorithm. It is implemented on a test sample of one month of data with available train numbers (approximately 9 x 10^5 points). The method assumes that the train number is not known and the merging algorithm described in section 4.2 is applied to the data. The last step is to compare the inferred train number with the actual train number. Given the relatively low rate of errors for existing values of train numbers discussed in Chapter 3, we assume that the original Netmis data provides reliable values for train numbers.

Results are presented in Figure 4-12. As can be seen, the merging appears to be effective as 83% of data points are correctly matched to their actual train number. A small amount (15%) of points form the sample test are not matched to any train number. Missing values for train numbers can introduce a bias in the inference of cancelations, as a trip may not appear complete in the AVL data but has actually been run. final percentage of missing values in the merged data is less than 10% of the data, which limits the potential bias in the cancelation inference. The comparison of a real count of cancelations and inferred count presented in Chapter 5 shows that this bias remains low. Missing values can bias the short turn inference, but only in the cases where the missing values correspond to the time window during which the train was short-turned. Finally, 2% of the number of points are matched to a wrong train number. It is important for this value to be very low as the wrong train number may result in a wrong inference of implemented corrective actions, both for cancelations and reformations.

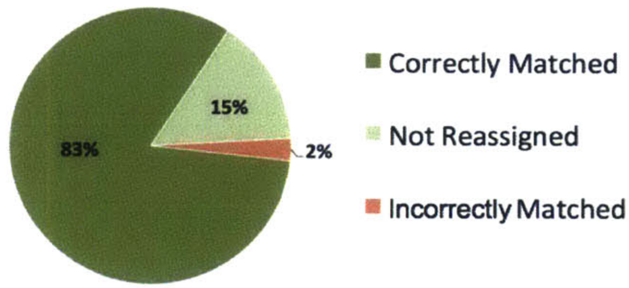


Figure 4-12: Results of the validation

Chapter 5: Inferring recovery strategies

Chapter 5 presents and discusses the recovery strategy inference. Section 5.1 describes the general methodology applied to infer corrective actions from train tracking and timetable information. In particular, it provides the pseudo-code for the inference algorithm and lists some key assumptions. Section 5.2 focuses on the implementation of the inference process on the Piccadilly Line. This section describes in more detail the tools used to implement the inference process and describes specific features of the Piccadilly Line relevant to this application. The final section presents the results of the implementation, at both the aggregate and disaggregate levels, and discusses the reliability of the results.

5.1 General Methodology

This section describes the general inference methodology. As discussed previously, most transportation agencies do not have numerical information documenting the corrective actions implemented during disruptions. The objective is to use the available numerical data, principally train tracking, to build a database that retraces these corrective actions. The database will provide a list of all corrective actions which will be useful to analyze recovery strategies at both an aggregate and disaggregate level.

5.1.1 Timetable and merged AVL comparison

The methodology is based on a comparison between the timetable and observed (inferred) train movements. The algorithm detects the differences between observed train movements and scheduled movements for a given train number. This comparison provides information on all the changes that were made to the planned service. As shown in Figure 5-1, the input to this process is the timetable data and the AVL data and the output is a database containing all corrective actions. The algorithm can be run on any time period where both timetable and AVL data are available.

In this research, the inference process is restricted to short-turns and cancelations. As described in Chapter 3, reformations are directly available through the CTFS database and are incorporated during the final stage of the reconstruction, as shown in Figure 5-1. Short-turns, cancelations, and reformations are the three main types of corrective action implemented on the line of interest. Given the low temporal resolution of the available AVL data, holding strategies are not inferred in this research, although this methodology could be applied in a similar way to take advantage of such information where it exists. In particular, the inference algorithm could detect unscheduled train delays in stations and attribute these to holding. A similar comparative methodology could also be applied to detect other corrective actions such as expressing. Expressing is particularly common on the New York metro system because of the existence of double tracks on several lines, and would be necessary to incorporate when applying this methodology to the MTA system.

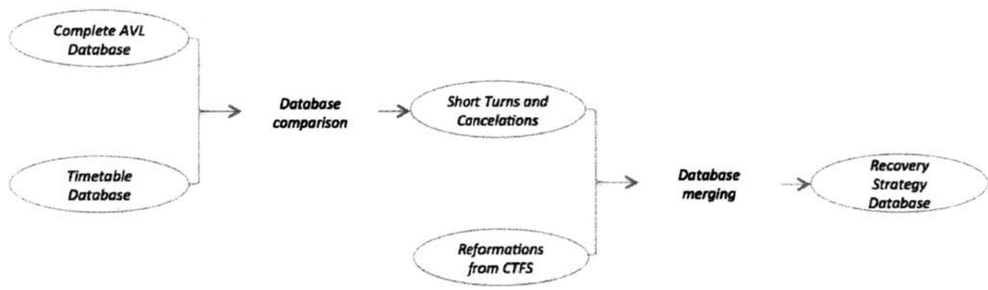


Figure 5-1: Inference process

5.1.2 Illustrative example

Figure 5-2 provides an illustration of the inference methodology. The green points on the time-space diagram represent the scheduled train movements for a given train on a given day. The blue points represent the observed trajectory for this train and day. This information is accessible respectively through timetable data and the merged AVL database presented in Chapter 4. In the example in Figure 5-2, there is a difference between the scheduled train movement and the actual train movement around 11:40 am. This train was scheduled to reverse at Uxbridge, the terminus of the northwest branch of the line but the blue points indicate that the train actually reversed a few stations earlier, at Ruislip. This is an example of an inferred short trip.

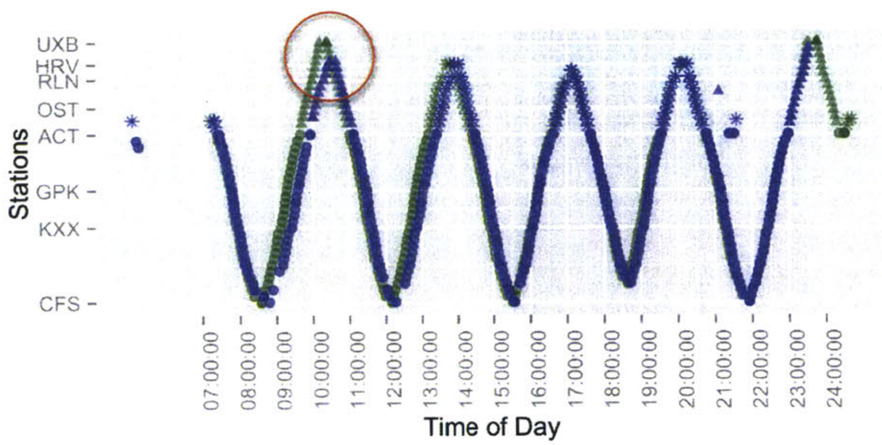


Figure 5-2: Short-turn inference example

Cancelations are detected in a similar manner. If there is a large time gap in the AVL data but not in the scheduled data, the inference is that the train was canceled for that time period. Section 5.1.2 presents the algorithm developed for the inference of both short-turns and cancelations. Table 5-1 gives an example of the output of the inference. The various fields are described below, with their significance varying across the type of action inferred.

Table 5-1: Recovery Strategy database

TRAIN_NU	Time1	Time2	Station1	Station2	Date	Type
300	26521	31435	34.25	33.25	12/7/13	Cancelation
303	26604	31848	39.25	39.25	12/7/13	Cancelation
342	26619	31821	36.25	35.25	12/7/13	Cancelation
302	26652	31986	37.25	36.25	12/7/13	Cancelation
310	26652	31476	32.25	33.25	12/7/13	Cancelation
301	26765	32068	39.25	37.25	12/7/13	Cancelation
306	26808	31866	36.25	38.25	12/7/13	Cancelation
305	27471	32028	39.25	39.25	12/7/13	Cancelation
307	28208	32998	39.25	39.25	12/7/13	Cancelation
340	32338	44711	4	7	12/7/13	Cancelation
237	32547	48104	7	2	12/7/13	Cancelation
314	32634	37494	4	7	12/7/13	Cancelation
254	34182	45220	30.25	30.25	12/7/13	Cancelation
303	34437	41556	28	34.75	12/7/13	Cancelation
357	34587	40231	4	5	12/7/13	Cancelation
317	35972	52132	4	6	12/7/13	Cancelation
313	36613	47333	4	6	12/7/13	Cancelation
256	38338	50168	1	2	12/7/13	Cancelation
301	38422	44239	1	4	12/7/13	Cancelation
315	38480	49867	4	6	12/7/13	Cancelation
243	38537	63172	10	2	12/7/13	Cancelation
307	40185	51458	1	4	12/7/13	Cancelation
245	27539	27780	36.75	1	12/7/13	Shortturn
311	27769	27750	30.25	36.75	12/7/13	Shortturn
275	28150	27930	36.75	1	12/7/13	Shortturn
255	29511	29550	36.75	1	12/7/13	Shortturn
352	31010	30480	6	1	12/7/13	Shortturn

For a cancelation:

- TRAIN_NUMBER : the train number affected by the cancelation
- Time1: the moment the train was canceled
- Time2: the time at which the train was brought back into service
- Sation1: station where the train cancellation was detected
- Station2: station where the train was detected as being put back into service

For a short-turn:

- TRAIN_NUMBER : the train number affected by the short-turn
- Time1: time at which the train was scheduled to reverse
- Time2: time at which the train was inferred to reverse
- Station1: station where the train was scheduled to reverse
- Station2: station where the train was inferred to reverse

The inference algorithm includes two independent parts. The first part detects all short-turns and the second part detects all cancelations. The function's input is a list of days on which to apply the inference process, and the output is a list of databases retracing all detected short turns (or cancelations) by day. An additional program is written to transform the two lists of databases into a unique database containing all days and all inferences (both cancelations and short turns). The last step is to incorporate reformations from the CTFS database.

5.2 Short-turn and cancelation inference

This section describes the algorithms to infer corrective actions. The general methodology is to compare all the reversals and large time gaps detected from the timetable with those detected from the AVL data. Section 5.2.1 presents the methodology used to infer the short-turns. Section 5.2.2 describes the algorithm developed for the cancelation inference. Section 5.2.3 presents the sensitivity analysis performed to choose the parameters for both short-turn and cancelation inference. Section 5.2.4 presents the CTFS data retracing reformations.

5.2.1 Short-turn inference

The algorithm is built as a function that can be applied to a list of days that correspond to the time window of interest. For this research, the inference was limited to a time period that corresponded to a single timetable, specifically from 1 October 2013 to 7 December 2013, the last day before a new timetable was implemented. This time period includes 68 days of operations but in practice, the inference was applied on 54 days because some days operated on different timetables and were discarded. Figure 5-3 shows a flow chart of the algorithm.

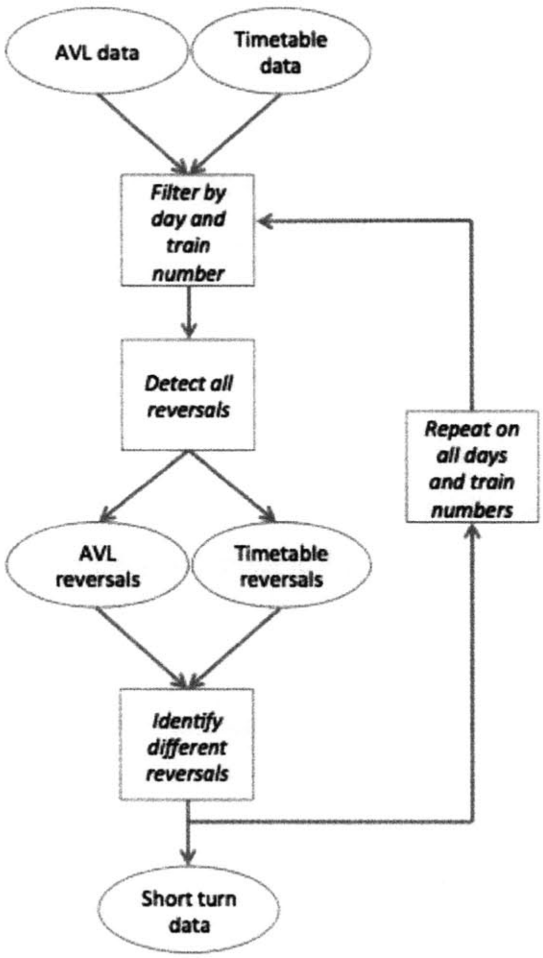


Figure 5-3: Flow chart of the short-turn algorithm

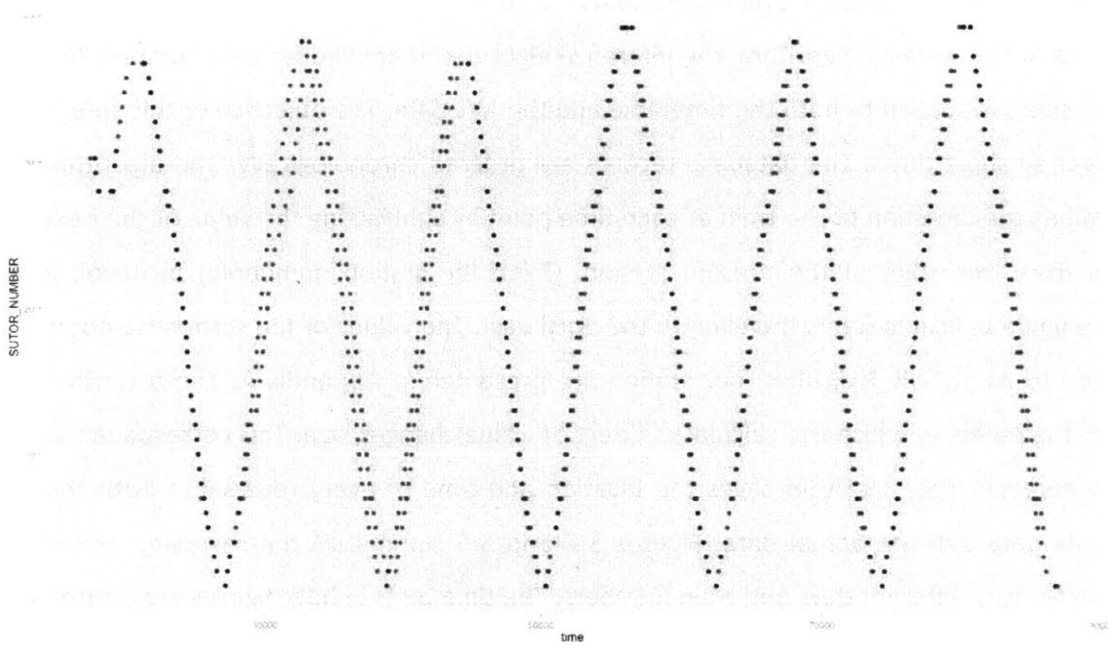
The main steps of the program are described in detail below.

a) Import timetable and AVL data and filter by train number

The first step is to import the AVL data and timetable data that corresponds to a day *d*. The AVL merger process outputs data in an identical format to the timetable, which allows the direct importation of the AVL merged database for use in this module.

b) Detect all reversals from the timetable and AVL data

As in the merging algorithm, the inference algorithm is applied by train number. The same process is applied to both the timetable and the AVL data. The objective of this step is to detect all short turns. The database records are ordered chronologically. The algorithm determines the direction of the train at each time point by subtracting the value of the next station from the value of the present station. Given the station numbering protocol, a positive value indicates trains traveling to the northeast. The values of the station numbers (referred to as SUTOR_NUMBER) per station are presented in Appendix A. The algorithm detects the points at which the calculated direction value changes sign. This corresponds to a train reversal. The algorithm saves the location and time of every reversal in both the timetable data and the actual data. Figures 5-4 and 5-5 summarize the reversing points inferred for two different days and train numbers. The time plots in both figures are plotted in seconds after midnight and SUTOR_NUMBER refers to the station numbers used throughout the research. Figure 5-4 corresponds to a train number and day for which there are no observed short-turns. Figure 5-5 corresponds to a day and train number with two inferred short-turns, at station number 37.25 around 9 am (32400 seconds after midnight) and at station 5 around 5 pm (61300 seconds after midnight).



(a) Time-Space Diagram

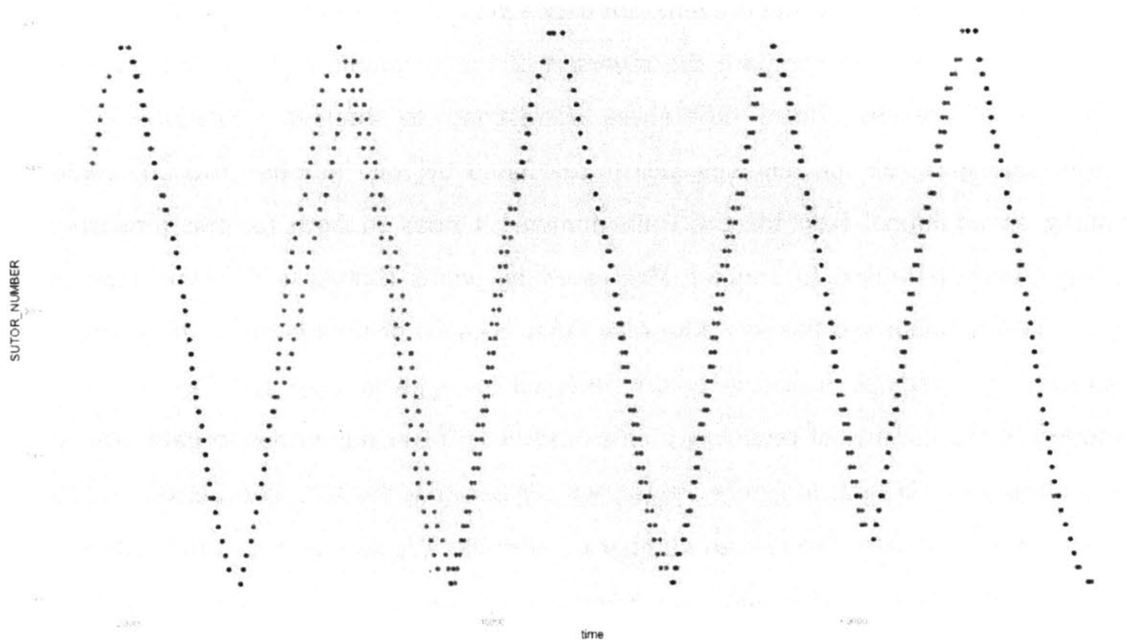
SUTOR_NUMBER	time
28.00	19279
36.75	21559
1.00	27154
38.25	33301
2.00	39465
36.75	44867
1.00	50392
39.25	56512
1.00	62541
38.25	68264
1.00	74328
39.25	80665

(b) Reversing points from AVL

SUTOR_NUMBER	time
28.00	19260
36.75	21540
1.00	27150
38.25	32790
1.00	38820
36.75	44820
1.00	50340
39.25	56670
1.00	62550
38.25	68100
1.00	74310
39.25	80670

(c) Reversing points from timetable

Figure 5-4: Reversing point detection A



(a) Time-Space Diagram

SUTOR_NUMBER	time
38.25	20055
1.00	25995
37.25	32441
1.00	37893
39.25	44082
1.00	50072
38.25	55731
5.00	61299
39.25	66823

(b) Reversing points from AVL

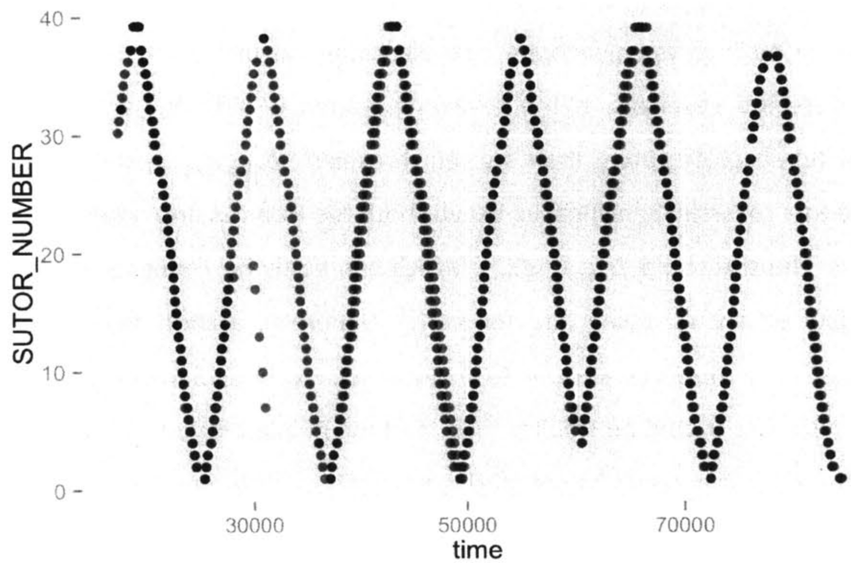
SUTOR_NUMBER	time
38.25	19620
1.00	25980
38.25	31590
1.00	37650
39.25	44070
1.00	49980
38.25	55620
4.00	61110
39.25	66870

(c) Reversing points from timetable

Figure 5-5: Reversing point detection B

c) Detect the reversing points that are different between the schedule and AVL data

The next step is to compare the reversals of the timetable and the AVL data to identify any differences. These differences correspond to inferred short-turns. The algorithm aggregates all the reversing points calculated by train number into one table containing an additional field for the train number. It does so both for the timetable reversing points (Database A) and the AVL reversing points (Database B). Time stamps between the timetable and the AVL data may differ because of train lateness. In addition the number of scheduled reversing points detected for a given train and day may not correspond to the number of reversing points observed. This is due either to data quality issues or train cancelations. In Figure 5-6 (a), we can see that the AVL data between 8:30 and 9:30 am is unreliable. The limited number of available AVL data points in that interval make it impossible to trace the train's movements. In this example, 14 reversing points are detected from the timetable data (Figure5-6 (b)) and only 11 reversing points are detected from the AVL data (Figure 5-6 (c)). It is not possible to merge databases row by row for the comparison as both databases do not have the same number of records. The next paragraph explains how the comparison between the two databases is implemented.



(a) Time-Space Diagram for Train 230

28.00	18160	230
39.25	20693	230
1.00	26578	230
28.00	30636	230
38.25	32239	230
39.25	57238	230
1.00	63260	230
38.25	68746	230
6.00	74482	230
36.75	79621	230
4.00	85327	230
38.25	18241	231
36.75	43700	231

(b) Reversing points from AVL

28.00	18120	230
39.25	20610	230
1.00	26580	230
38.25	32190	230
1.00	38160	230
42.75	44640	230
1.00	51000	230
39.25	57270	230
1.00	63210	230
38.25	68700	230
4.00	74160	230
36.75	79620	230
4.00	85320	230
19.00	87930	230
38.25	18120	231
1.00	24630	231

(c) Reversing points from timetable

Figure 5-6: Reversing point comparison

For every row k of Database B, the algorithm goes through each row l of Database A. If the train numbers of both rows correspond, the algorithm calculates the difference in time between the detected reversals. If the reversing station is different and the time difference is smaller (in absolute value) than a given threshold, $\Delta_{short\ turn}$, the algorithm records this instance as a reversal. Specifically, it saves both the location and times reported in Database A and B. The threshold $\Delta_{short\ turn}$, , which can easily be changed, takes into account the possibility of a train being late (or early). Generally, a short trip occurs to reduce train lateness, it is therefore crucial to include such a time parameter with an appropriate value. This value should be smaller than the time needed for a train to traverse the line, otherwise the algorithm could be comparing reversals on different trips on the line. Section 5.2.4 will discuss the value of the threshold used in the case of the Piccadilly Line.

The two nested loops result in a large computational time (o(k*l)). In practice, the number of reversing points is low and the total number of rows per day is generally less than 1000. This methodology may require higher computational time when applied to shorter lines with a large number of reversals.

d) Incorporate line constraints

Step c) results in a database listing all the inferred short trips. The last step is to refine the results by incorporating infrastructure constraints: short-turns can occur only at stations where cross-over tracks allow reversals. . In the case of the Piccadilly Line these stations are listed in the Table 5-2.

Table 5-2: Piccadilly Line Reversing points

Station	Station Number
Arnos Grove	4
Wood Green	6
Hatton Cross	36.25
Rayners Lane	36.75
Ruislip	39.75
Northfields	30.25
King's Cross	13

These constraints are used to partition the database obtained from step c) above. Other constraints, such as imposing fixed tuples of scheduled reversing location and observed reversing location, could be incorporated to increase the precision and reliability of the inference process. For example, if the train was scheduled to run to the terminus of the branch, it may not be possible to reverse in central London.

5.2.2 Cancelation inference

The cancelation inference uses a similar methodology to the short turn inference process: by identifying all the reversing points in both the timetable and the AVL data, the process compares all large observed time gaps. Similar steps can be used to describe the process. The flow chart of this algorithm is given in Figure 5-7.

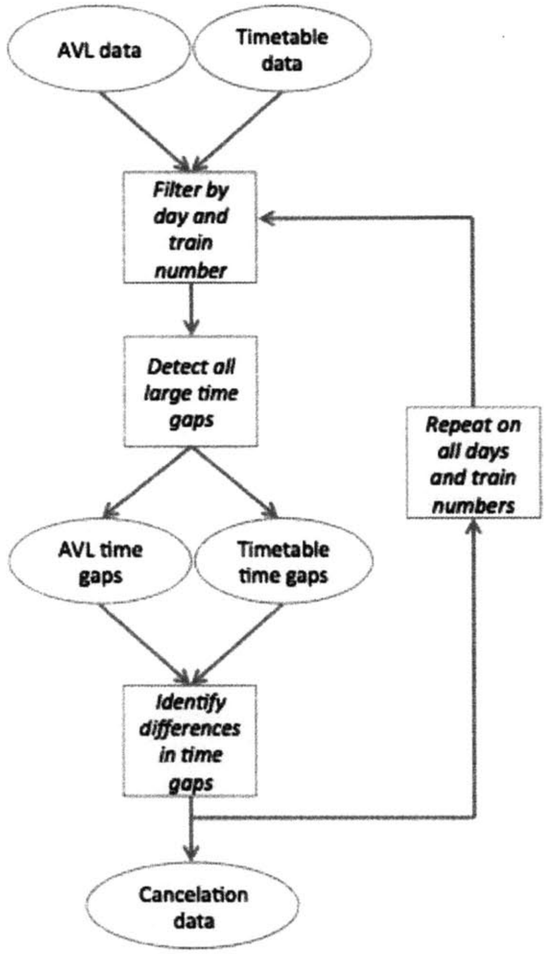


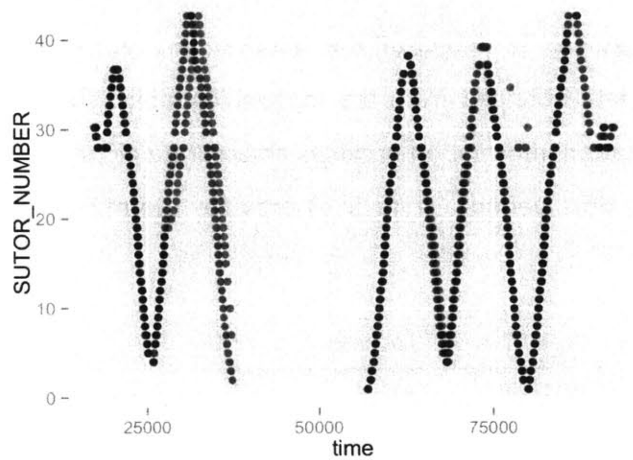
Figure 5-7: Flow chart of the cancelation algorithm

a) Import timetable and AVL data and filter by train number

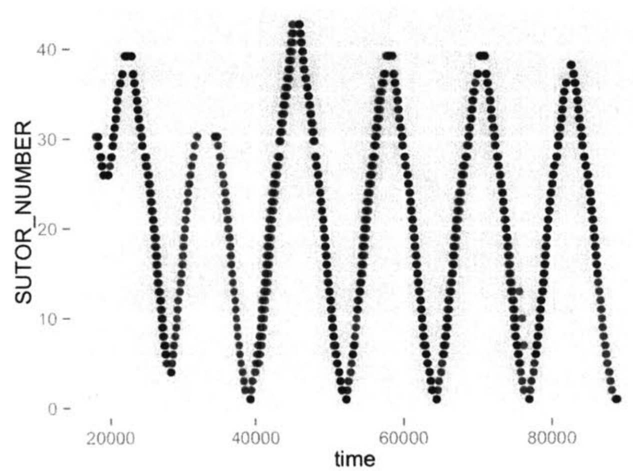
This step is identical to the first step in the short trip inference process.

b) and c) Detect all large time gaps and compare timetable and observed data

Rather than detecting all reversing points, this step detects all large time gaps. The process creates a database that records all these time gaps. The algorithm calculates the difference in time between two adjacent points, and if the time difference is larger than a given threshold, $\Delta_{cancelation}$, it saves the value. The threshold value is used to limit cases of cancelations being inferred as a result of missing data. Section 5.2.4 discusses the choice of this parameter. Figure 5-8 shows two cases where large time gaps are observed. In case (a) both the timetable and the observed data show large time gaps. In this case the process will not infer a cancelation. In case (b), the train was scheduled to run between 8:30 and 11 am but no data exists for train movement in this time interval. When large time gaps are observed in the AVL data but not in the timetable, the algorithm saves the time and location at the start and end of the time gap. These values correspond to the time and location of the cancelation and the time and location of the re-entry into service.



(a) No inferred cancelation



(b) Inferred cancelation

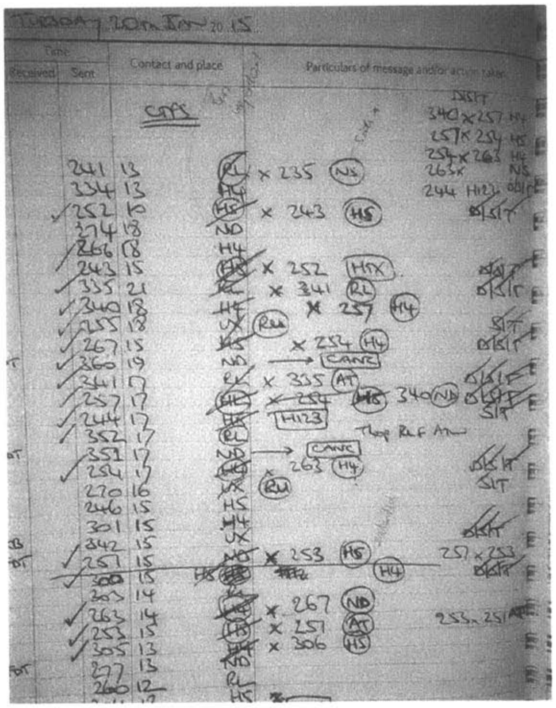
Figure 5-8: Inferring cancelations

5.2.3 Reformations

As described previously, the reformation information is directly available through the CTFS data. Figure 5-9 provides a snapshot of CTFS data with the last field corresponding to the original train number if the train has been reformed. The rTRN value is zero if the train has not been reformed. The algorithm detects all the rows with a non-zero rTRN value and incorporates that information in the corrective action database. A case study was conducted for a given day to verify that the information provided by this CTFS field is reliable. The methodology was to detect all the reformations recorded through the CTFS data and compare the data obtained from the manual logbooks. The case study showed that the retraced reformation information provides an accurate picture of the reformations implemented in a given time period. Figure 5-10 provides the manual logbook and data used.

```
***** Sudbury Hill EB *****
TRN     TD      Expected Time   Actual Time     Lateness (s)    rTRN
-------------------------------------------------------------------------
274     A2S     06:53:30        06:52:48        -42             0
360     A2S     07:08:30        07:07:56        -34             0
275     A21     07:23:00        07:24:21         81             0
```

Figure 5-9: Snapshot of CTFS data



(a) Manual Logbook

235	241	08:41:21
335	341	09:16:20
243	252	09:20:27
252	243	09:20:42
341	335	09:22:43
254	257	09:24:56
257	340	09:32:37
263	254	09:33:08
340	254	09:47:36
254	267	09:49:37
267	263	10:04:33
257	253	10:06:32
253	251	10:09:08
251	257	10:19:52
306	305	10:26:00

(b) Reformations from CTFS

Figure 5-10: Reformation comparison

For this time period, the logbook records 13 reformations while CTFS detects 15 reformations. From the 13 reformations noted in the logbook, only 1 reformation is not revealed in the CTFS data. Of the 15 reformations noted in CTFS, 3 are not shown in the logbook. These discrepancies could be due to an incomplete logbook or an error in the CTFS data. Assuming that the logbook accurately records all reformations, Table 5-3 is the matrix summarizing the differences between CTFS reformation and logbook reformations.

Table 5-3: Confusion matrix

	CTFS reformation	Not shown in CTFS reformation
Reformation on logbook	*12 (true positive or TP)*	*1 (false negative or FN)*
Not a reformation on logbook	*3 (false positive or FP)*	*NA*

We can calculate the sensitivity, TPR (true positive rate), and the precision, PPV (positive predictive value), defined in the literature as:

$$TPR = \frac{TP}{TP + FN} \qquad (5-1)$$

$$PPV = \frac{TP}{TP + FP} \qquad (5-2)$$

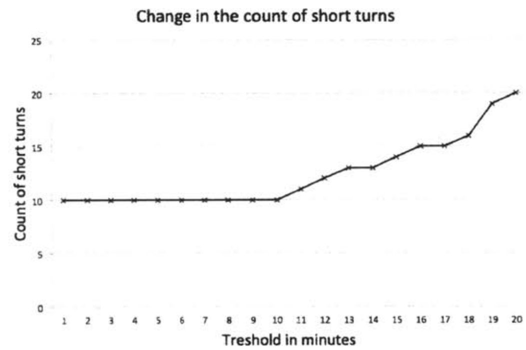
In our case, the sensitivity is 92% and the precision is 80%. Based on this example, the CTFS database has a tendency to overestimate the number of reformations (the CTFS count is 15% higher than the actual count) but it does not miss many of the reformations that were implemented. The calculated levels of sensitivity and precision are high enough to validate the usage of CTFS as reliable indicator of reformations.

5.2.4 Choice of parameters

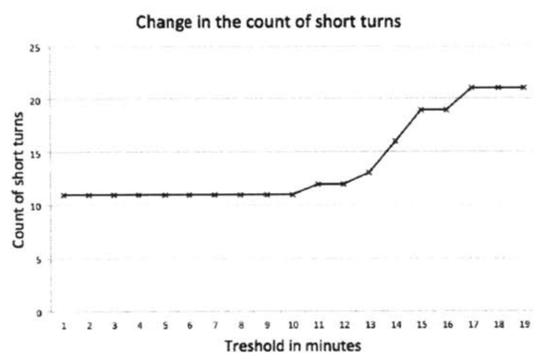
As described in sections 5.2.2 and 5.2.3, both inference algorithms rely on the value of parameters, $\Delta_{short\ turn}$ in the case of short turn inference, and $\Delta_{cancelation}$ in the case of cancelation inference. A sensitivity analysis is used to choose the value of these parameters. The general methodology is similar to the sensitivity analysis implemented in the merging algorithm. The algorithm is run with a variety of values for the parameters, and we observe the output of the algorithm for each value. In the case of the inference algorithms, we focus on the number of short trips and the number of cancelations inferred.

a) Short-turn parameter

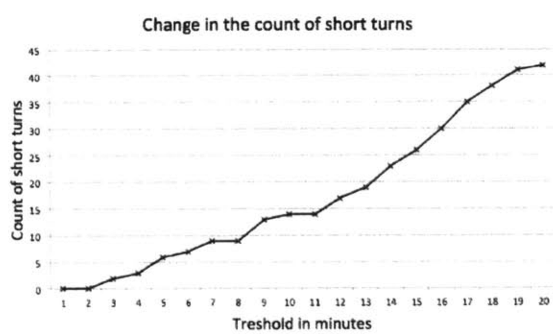
$\Delta_{short\ turn}$ is used as a threshold to compare the times of reversals detected in the timetable and in the AVL data. The sensitivity analysis is implemented on three different days.



(a) 1st November



(b) 4th November



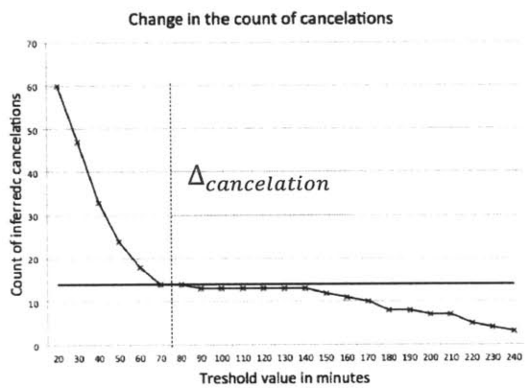
(c) 6th November

Figure 5-11: Sensitivity analysis for short-turn time threshold

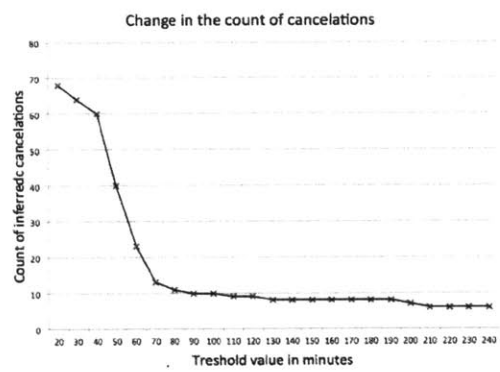
Figure 5-11 shows that for cases (a) and (b) there seems to exist a given threshold above which the algorithm overestimates the number of short turns. The value of 10 minutes is used as a threshold for our analysis, but it is understood that this choice increases the uncertainty in the inference count. In particular, the choice of 10 minutes seems more arbitrary given the evolution of the counts as seen in Figure 5-11 (c). A next step would be to validate this threshold by comparing the number of inferred short turns with the actual number of short turns. A one-day case study during which the researcher in the field manually records all the short turns that are implemented by the controllers and then compares this value with the inferred value would be a good way to validate the inference. In some AVL systems, the data retraces the scheduled train destination and the actual train destination. If such fields are proven to be reliable, it would be interesting to use the inference methodology to compare the inference results with the short turns detected directly through destination changes.

b) Cancelation parameter

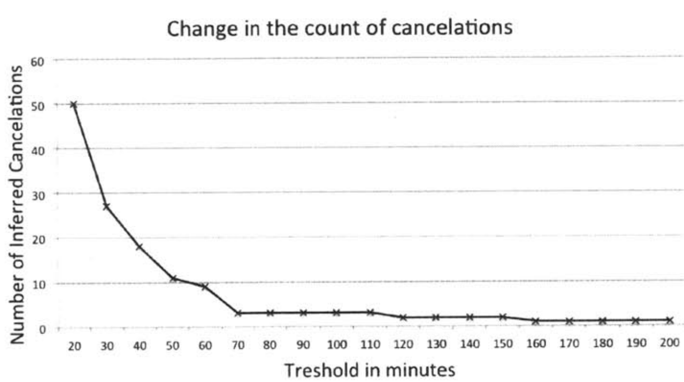
The parameter $\Delta_{cancelation}$ is used as a threshold to detect large time gaps. It is important to use such a threshold as small observed time gaps may be due to data gaps rather than train cancelations. The sensitivity analysis was run on three days with various levels of service. Case (a) in Figure 5-12 corresponds to a high level of disruptions, case (b) to a medium level of disruption and case (c) to normal operations. The analysis looks at the evolution of the number of inferred cancelations per day given a certain value of $\Delta_{cancelation}$. As expected, too small a threshold value results in a large overestimate of the number of cancelations. In the three cases, there seems to be a threshold value beyond which the number of cancelations is less sensitive to changes in parameters. In some cases ((a) and (b)), the count seems to be underestimated for high values of $\Delta_{cancelation}$. The value that corresponds to the beginning of a plateau for the counts is chosen as our threshold. The daily summary for 6 November (case (a)) allows us to compare the inferred cancelations and the reported cancelation (horizontal line at 14). This confirms that the plateau value provides a realistic estimate of the actual number of cancelations. The value of 75 minutes is chosen as the threshold.



(a) 6th November



(b) 12th November



(c) 30th November

Figure 5-12: Sensitivity analysis for cancelation time gap threshold

5.3 Results and limitations of the inference process

Sections 5.1 and 5.2 described the inference algorithm and its implementation. This section provides some examples of how this recovery strategy database can be used to provide insight into operations, both on an aggregate and disaggregate level. The last part of this section discusses the limitations and assumptions made during the inference process and their implications on the reliability of the results. Further analysis based on the information provided through the recovery strategy database will be presented in Chapter 6.

5.3.1 Count of actions per day

The recovery strategy database can be used to better understand operations and the corrective actions that were implemented in response to a given incident. This numerical information could be useful for service control managers to better understand the corrective actions that were implemented. It would allow them to have a more precise and direct record of the number of cancelations, short trips and reformations, rather than relying on oral reports and manual logbooks. Daily disaggregate analysis that focuses on the evolution of counts of corrective actions over the day will be presented in Chapter 6.

Information on the type, number, location, and time of corrective actions could be incorporated in daily reviews as well as the general performance proxy used within the London Underground. In particular, having access to information concerning short trips and cancelations provides a better understanding of the total impact of the disruption and recovery on passengers.

Other illustrations based on the recovery strategy database may be useful for further analysis. Figure 5-13 illustrates the evolution of the daily number of cancelations, short-turns and reformations over the time window of interest. Even though there seems to be some correlation between the counts of different actions, some instances show a high number of a particular type of corrective action and a lower number of another type. This suggests that the recovery strategies implemented vary across days. This type of diagram can be used by the operating staff as it is also useful to detect days that experienced larger or smaller numbers of interventions.

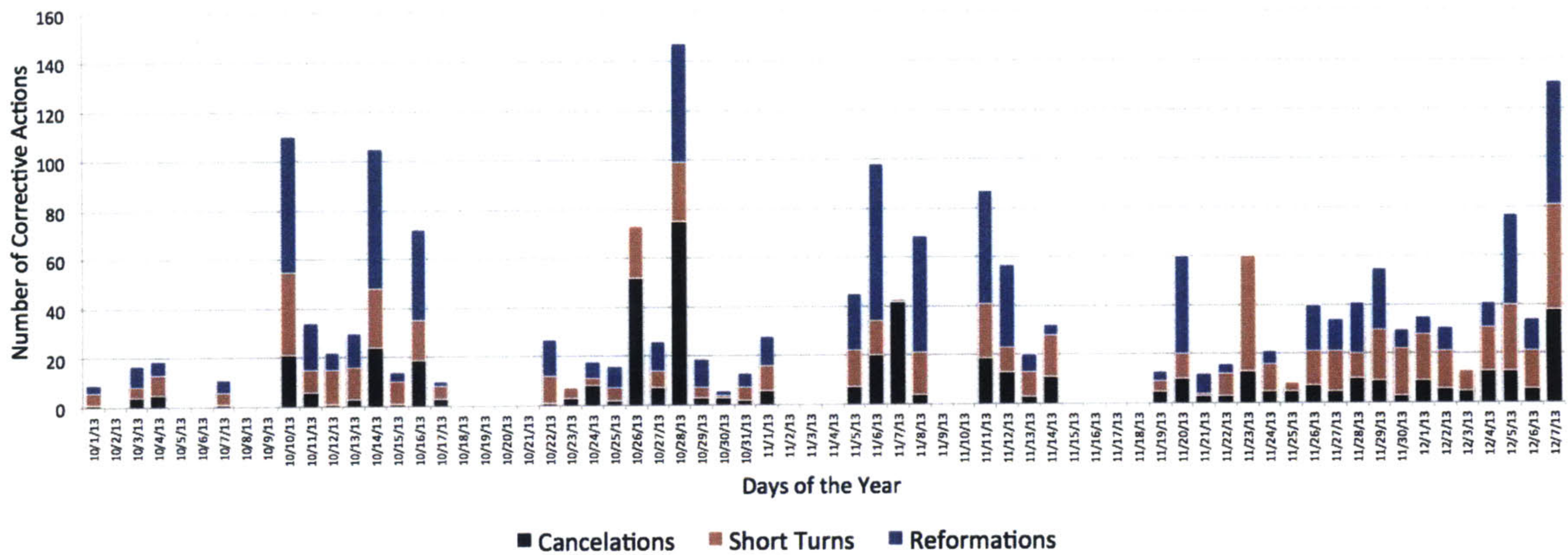


Figure 5-13: Corrective actions by day.

5.3.2 Short trips

The recovery strategy database can also be used to illustrate the locations and times of all short trips implemented during a given time period. Figure 5-14 illustrates the short trips implemented in October 2013. The red points indicate the scheduled reversing points and the blue points indicate the inferred reversing points. This graph is a user-friendly way to observe all short-turns in a single format. The majority of short trips are operated between 7:00 am and 11:00 pm and there is a higher density of short trips on the northern branch of the line.

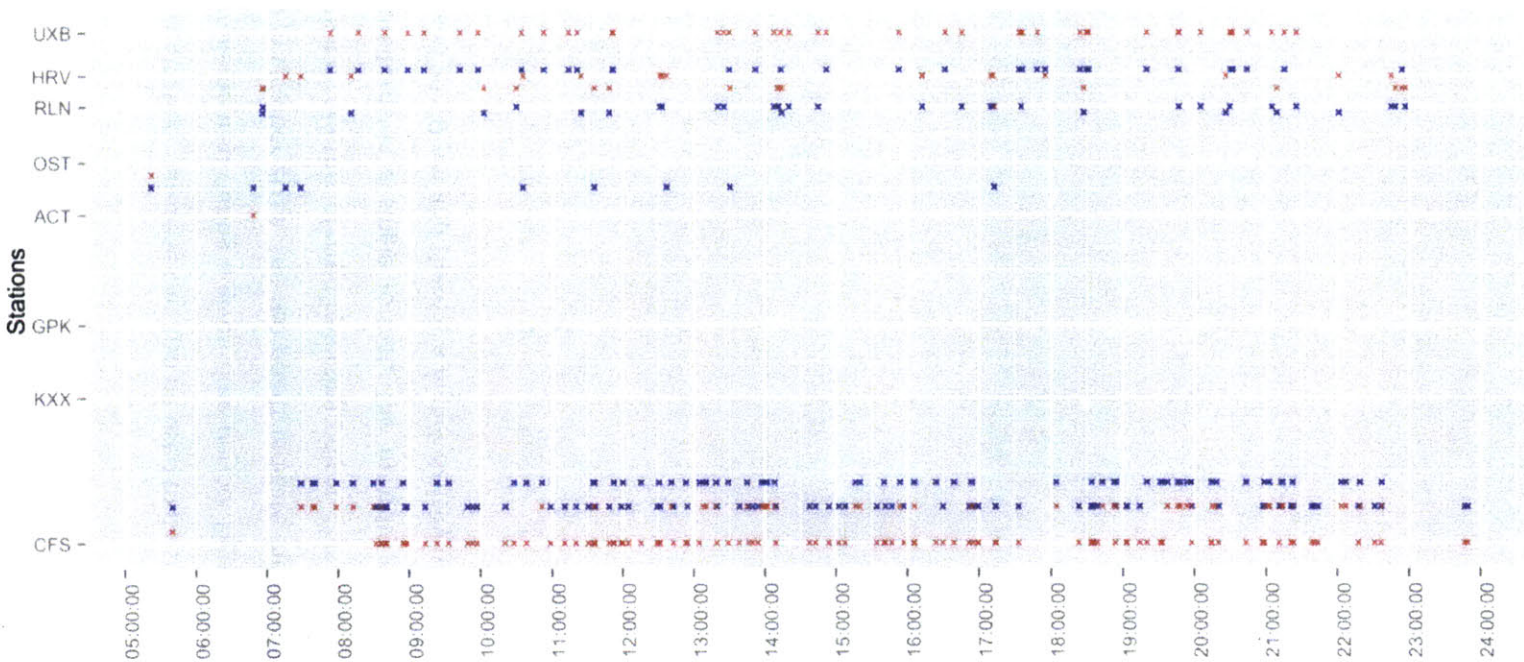
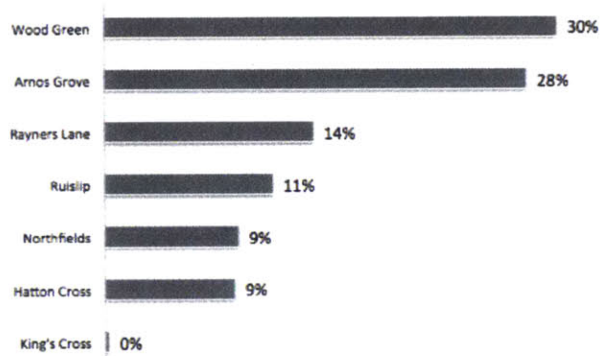
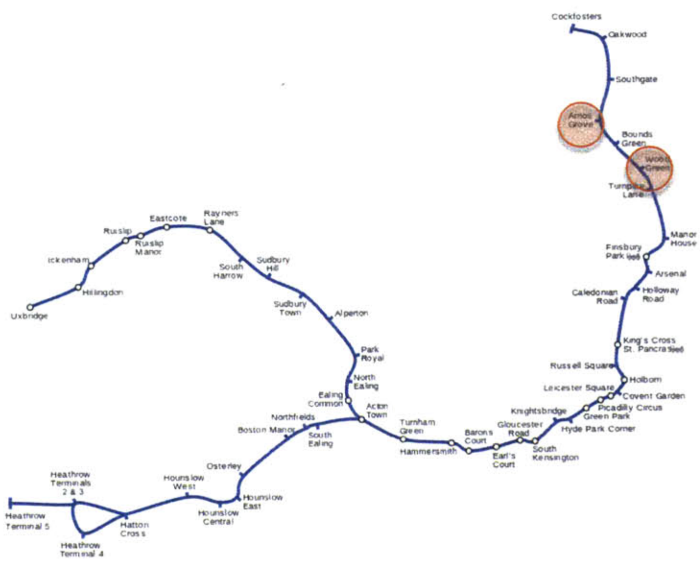


Figure 5-14: Short Trips for October 2013

Figure 5-15 (a) illustrates the percentage of short trips by station of occurrence. As can be seen in Figure 5-15 (b), the two largest reversing locations are on the northwest branch of the line. This information can be shared with local controllers and service control managers, as well as employees in charge of infrastructure. It could also be useful to observe how this distribution and the total number of short-turns may change in response to new timetables or changes in infrastructure (new crew depot, new train depot, etc).



(a) Most common reversing points

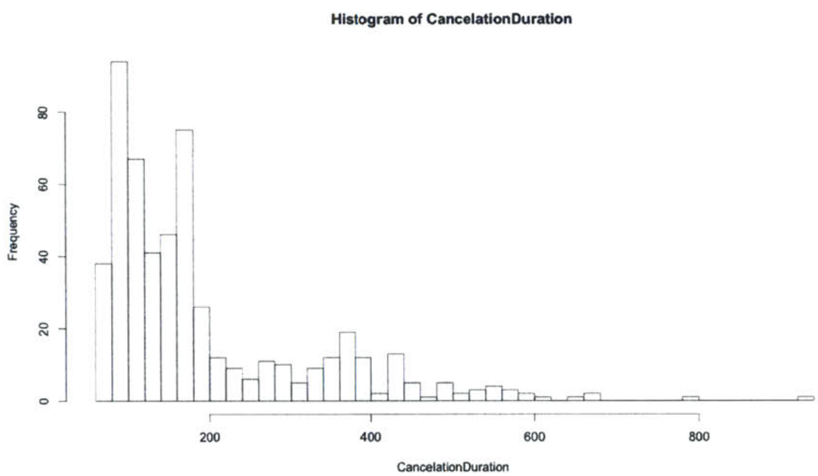


(b) Location of these points

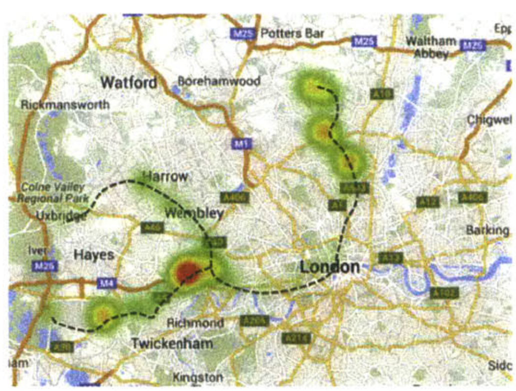
Figure 5-15: Most used reversing points

5.3.3 Cancelations

The corrective action database can also provide insight into cancelations. As can be seen in Figure 5-16 (a), in the case of the Piccadilly Line, it is not surprising that the main location of cancelations is Arnos Grove, where a train depot is located. This information is useful to confirm the accuracy of the inference. The average duration of a cancelation (plotted in minutes in Figure 5-16 (b)) is 3 hours and 15 minutes, with a large group of cancelations shorter than 3 hours and a smaller group of cancelations that appear to be much longer (around 7 hours).



(a) Distribution of the duration of cancelations



(b) Heat map of the location of cancelations

Figure 5-16: Cancelation characterization

5.3.4 Assumptions and limitations

The method of comparing the timetable and AVL data is applicable to a large variety of high frequency lines that operate with a schedule and with human controllers. In this section the assumptions as well as the limitations of this methodology are discussed. The process can only function with reliable data for both scheduled and observed train movements. In addition, all discrepancies between scheduled and observed movements are inferred to be corrective actions that are part of a recovery strategy. In some cases, trains are canceled or short turned directly because of an incident, for example a train that is damaged will be sent to the depot. The actions affecting these trains are for safety and are not part of the recovery strategy as described in Chapter 1. These actions are assumed to be rare in comparison with similar actions implemented as part of a recovery strategy. Further analysis and research on differences between actions taken to resolve an incident and actions taken to respond to the line disruption would be useful.

The main limitations of this methodology are related to train tracking reliability. The quality of this data can indeed impact the precision and reliability of the inference results. The AVL data used needs to present a high enough spatial resolution to correctly detect all reversing points to infer short trips. In addition, the inference of cancelations may be biased to over-estimating the number of cancelations because of train tracking data gaps. As described in Chapter 3, significant incidents such as signal failures may result in large gaps in data for a given train. In that case, the train has in fact run in service but this is not identified through the train tracking data. The inference algorithm will detect a gap and erroneously conclude that a cancelation occurred. The use of a threshold of an hour and fifteen minutes for the time gaps detected is a first step towards dealing with this issue (usually data gaps due to missing data are shorter than an hour), but this limitation must be kept in mind throughout the analysis of the data.

Chapter 6: Evaluating the effectiveness of recovery strategies

Chapters 4 and 5 described the process of AVL data fusion and the methodology used to infer recovery strategies based on comparison between observed and scheduled train movements. Chapter 6 illustrates how the inferred information on recovery strategies can provide insight and recommendations on current practices. This chapter presents a methodology and metrics that can be used to describe and assess the effectiveness of a recovery strategy. It focuses on a case study of the Piccadilly Line to illustrate the knowledge gained from the recovery strategy database.

6.1 General Framework

The general approach is to compare similar incidents to assess how different recovery strategies lead to different recovery characteristics. The protocol described here is based on an analysis of historical data. This methodology complements previous research on disruption management by Babany (2015) that is based on a simulated optimization platform.

From an initial condition characterized by incident features and the state of the network, controllers implement a recovery strategy that leads to the recovery of the line. The proposed methodology selects days with similar initial conditions and compares how different recovery strategies lead to a differences in the way the line recovers. It is important to select days with similar incident characteristics. Indeed, a similar recovery strategy may result in different recovery characteristics if the incident characteristics are different. The general methodology is illustrated in Figure 6-1.

Identify equivalent incident conditions → *Identify the implemented recovery strategies* → *Calculate the metrics related to recovery* → *Use recovery characteristics to compare recovery strategies*

Figure 6-1: Methodology to evaluate the effectiveness of a recovery strategy

In particular, it is important to define two different recovery effectiveness indices that quantify respectively the effectiveness of the recovery for passengers and the effectiveness of the recovery for the crew. This approach complements the approach developed by Babany (2015). Babany develops an optimization tool based on an objective function that includes both passenger service costs, schedule adherence, and a measure of the complexity of the proposed recovery schedule. He defines the passenger cost as a measure of passenger waiting time. He neglects the possible irregularity of headways and uses half the average headway as a measure of waiting time. Babany defines schedule adherence as the observed lateness during crew swaps. Indeed, if all the crew swaps are on time, the line is operating on schedule which corresponds to normal operations. This description of crew impact will be incorporated in this research. A main difference between Babany's work and the current research framework is linked to the context in which the metrics to assess passenger and crew impact are used. Babany defines the passenger and crew cost in the context of a simulated optimization tool that predicts an optimal recovery strategy given a set of fixed constraints. This research focuses on metrics for the ex poste analysis of actual recovery data. Building on Babany's previous work, section 6.1.1 and section 6.1.2 introduce metrics to evaluate the impact of disruptions on passengers and on crew.

6.1.1 Recovery Effectiveness Index for passengers

Excess Waiting Time:

As described in Chapter 2, disruptions have various impacts on passengers. A comprehensive measure of passenger impact should include average and regularity of headways, demand levels to take into account the number of passengers impacted by the disruption, measures of in-train times and measures of additional transfers. In particular, additional short-turns or cancelations may require passengers to make more transfers than planned and could increase their total travel time. In the context of this thesis, the analysis focuses on metrics related to platform waiting time as this is assumed to have the highest impact on passengers.

This section defines the metric R that can be used to represent the excess waiting time. We calculate the average and variance of headway, and compute the total wait time after weighting the expected wait time by passenger counts. To incorporate the variability in passenger counts, the line is divided into N sections. All values are calculated by ten-minute time intervals (the function WT is therefore defined as one value per ten-minute interval). $E[T]_{i,t}$ corresponds to the average observed waiting time per section and per time interval t. $E[H]_{I,t}$ is the average headway observed on section i during the time interval t. $cv(H)_{I,t}$ is the coefficient of variation of headways observed on section i during the time interval t and $D_{i,t}$ is the percentage of demand observed on section i during the time interval t. Finally, $E^0[H]_{i,t}$ corresponds to the scheduled headway on section i during the time interval t. Subtracting half this value (equivalent to the scheduled waiting time) from the observed waiting time gives an approximation of the excess waiting time. The excess waiting time is defined per direction since both the headway average and headway variance are defined per direction. The measure defined in (6-1) and (6-2) complements Babany's definition of passenger cost as it represents the excess waiting time rather than the average waiting time and it includes headway irregularity measures. This metric does not capture the effect of denied boarding for passengers.

$$E[T]_{i,t} = \frac{E[H]_{i,t}(1 + cv(H)_{i,t}{}^2)}{2} \qquad (6-1)$$

$$WT_t = \sum_{i=1}^{N} (E[T]_{i,t} - 0.5 * E^0[H]_{i,t}) * D_i \qquad (6-2)$$

Integrated Passenger Index

The excess waiting time as defined in (6-2) is a good metric to measure the impact of disruptions on passengers. Typically, the excess waiting time will sharply increase after the beginning of the incident. It will reach a maximum value and then gradually decrease until the line is back to schedule and the excess waiting time is null. Different recovery strategies may lead to different values of maximum excess waiting time as well as different times to recover. It is important to incorporate both the time taken to reach normal operations and the observed values of excess waiting time. Figure 6-2 represents schematic examples of two different recoveries. Part (a) depicts an aggressive recovery strategy leading to a short time to recover. Part (b) depicts a more incremental recovery strategy with a longer time to recover but smaller maximum values of waiting time. We define a single integrated recovery effectiveness index (REI) that captures the impact of maximum values but also the total time to recover. It corresponds more precisely to the shaded area shown in Figure 6-2.

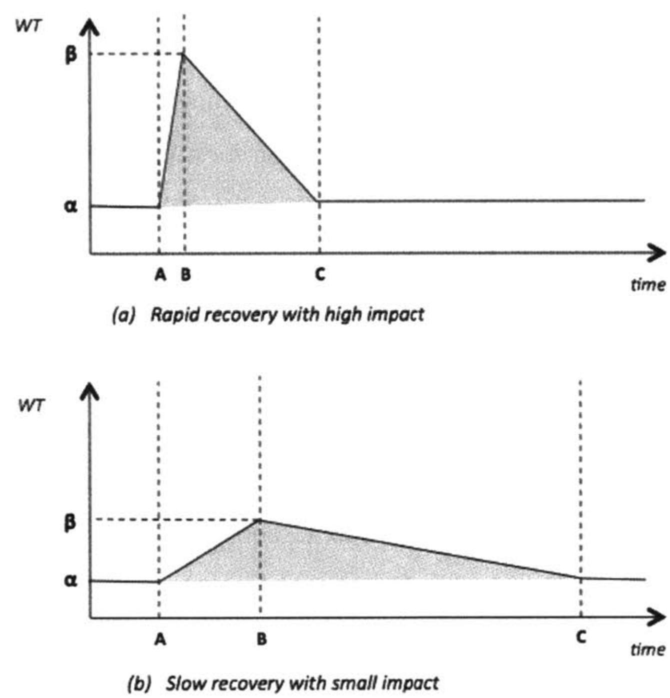


Figure 6-2: Different types of recoveries

In the case of the theoretical example presented in Figure 6-2, REI can be defined as:

$$REI = \frac{(\beta - \alpha) * (C - A)}{2} \qquad (6-3)$$

As the general form of excess waiting time is not linear let alone triangular, REI is more broadly defined in the form:

$$REI = \int_A^C (WT(t) - WT(\alpha))dt \qquad (6-4)$$

As it is based on the waiting time, REI is specific to a direction. It can be calculated both for the overall line as well as a specific section of the line.

Calculating the average headway on the Piccadilly Line:

The average headway E[H] can be calculated directly from the refined AVL data. The direction of the train is inferred with the same methodology used for the short-turn inference. The data is then sorted by direction, station and time. The headway is estimated as the difference of two consecutive time stamps at a given location for a given direction. The next step is to calculate the average observed headway by direction. The analysis uses ten-minute time intervals, but this value can be changed depending on the desired level of granularity. For each direction and line section, all data points are grouped into ten-minute intervals. The analysis computes the average for each group of data. This results in the average headway value per ten-minute interval for each direction and section. In the case of the Piccadilly Line, the line can be divided into four zones, A to D, as shown in Figure 6-3, with the list of all the stations in each zone given in Appendix C. Zone B includes Central London and has a higher density of stations and a higher frequency of service.

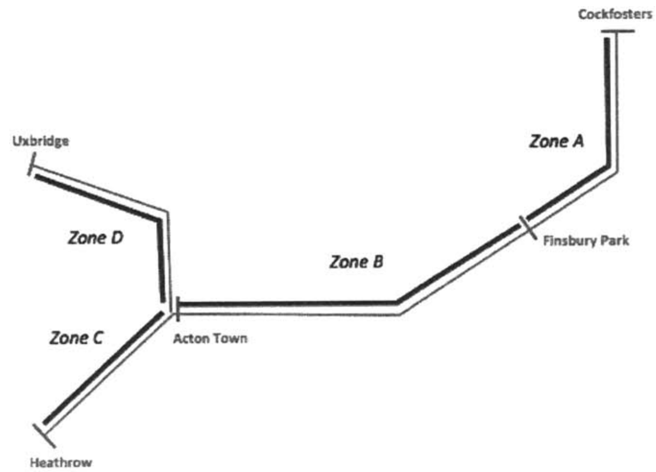
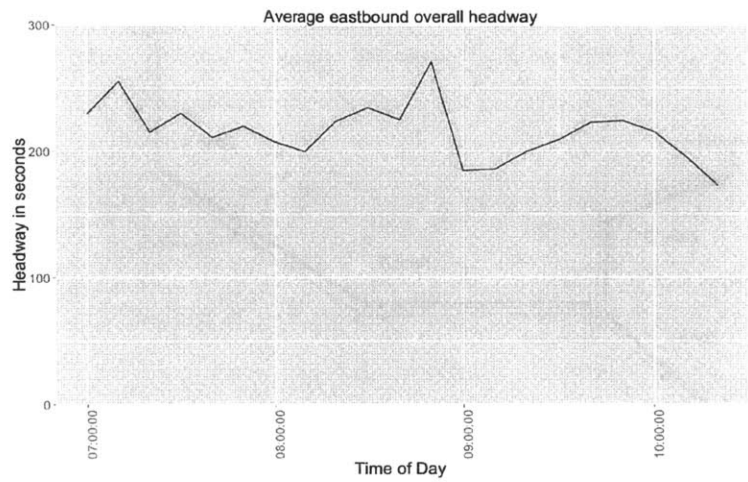
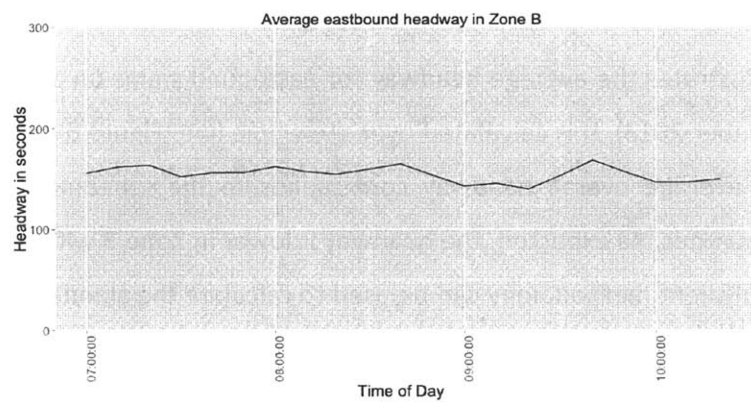


Figure 6-3: Geographic segmentation of the Piccadilly Line

Figure 6-4 illustrates the average headway for eastbound trains on a normal day of operations (3 December 2013). It is calculated over the whole line (Figure 6-4 (a)) as well as a for Zone B. The average overall headway corresponds to the scheduled headway of approximately 3.3 minutes. As expected, the headway is lower in Zone B with an average of 2.5 minutes. An equivalent methodology can be used to calculate the standard deviation of the headway per direction, section and 10 minute time interval.



(a) Overall



(b) Zone B

Figure 6-4: Average observed southbound headway

6.1.2 Recovery Effectiveness Index for the crew

Metrics related to headway values are not of importance for the crew. A different recovery effectiveness index should be defined to measure the impact of the disruption on the crew, in particular on the drivers. To accurately measure the impact of recovery on the crew, it is important to understand the operations planning process.

Crew scheduling:

In his work on evaluating the robustness of crew schedules under disrupted services, Ravichandran (2013) provides a detailed description of operations planning. The operations plan is described as a set of plans which fully describe the utilization of the transit agency's resources. A key constituent of the operations plan includes the timetable, which describes all the scheduled train movements. Another important element is the crew schedule that describes all the driver shifts to cover the scheduled train movements. The crew schedule is designed to operate under normal operations. A single driver is assigned to a specific duty that corresponds to part of a trajectory from the timetable. Generally, a small amount of slack time is incorporated to allow for small levels of train lateness. However, in the case of larger train lateness, drivers may experience late reliefs and a change in their scheduled duties. Most transit agencies rely on the availability of spare drivers to provide relief to drivers assigned to late trains. Ravichandran discusses the benefit of spare drivers but also the difficulty of utilizing this resource effectively. In the context of this framework, the availability of spare drivers is not incorporated in the final measure of the impact of disruptions.

Recovery Effectiveness Indices for crew:

Given the crew scheduling process, the best metric to assess the impact of a disruption on the crew should be linked to schedule adherence at crew relief locations. Indeed, if the lateness of a train is recovered before the end of a drivers' shift, the driver will not suffer any negative impacts. A first metric that can be used is the total number of late trains observed at crew relief locations. Late trains are defined by trains arriving at the crew depot with a lateness above a certain threshold. The REI linked to the count of late trains corresponds to the sum of all trains that arrived late between the start of the incident and the return to normal operations. The return to normal operations is defined as the moment when all trains are back on schedule.

In addition to the count of late trains, it can be useful to study the cumulated value of lateness at crew reliefs for all late trains. This takes in account the fact that a driver scheduled on a train that is 30 minutes late at the relief point is more impacted than a driver scheduled on a train that is only 12 minutes late at the relief point. The REI linked to the cumulated lateness of trains corresponds to the sum of the observed lateness of trains that are out of schedule. The sum is implemented from the start of the incident to the return to normal operations. With L(T) the value of lateness for train T and (T_1...T_N) the list of all trains that arrived late at the crew depot between the beginning of the incident and the end of recovery, we can define:

$$REI_{crew} = \sum_{i=1}^{N} L(T_i) \qquad (6-5)$$

One limitation of using the lateness at crew relief locations is that it overestimates the impact on drivers. Indeed, it is possible that a driver passes a crew relief station without a driver swap. In that case, even though the train is late, if it recovers the lateness before the next actual relief point, the driver will not be impacted. For more precision, the analysis should incorporate the actual scheduled relief points to count late trains at scheduled reliefs.

Even though the median lateness observed throughout the line does not accurately represent the impact of the disruption on drivers, it is a good metric to understand how the line gradually recovers from a disruption. In particular, it can be assumed that a return to low levels of median lateness corresponds to a return to normal operations. This leads us to define two important metrics of time. The time of disruption as described in Figure 6-5 corresponds to the time from the beginning of the incident to the end of the recovery. It does not take into account the length of the incident. The time to recover is defined as the time of the recovery phase.

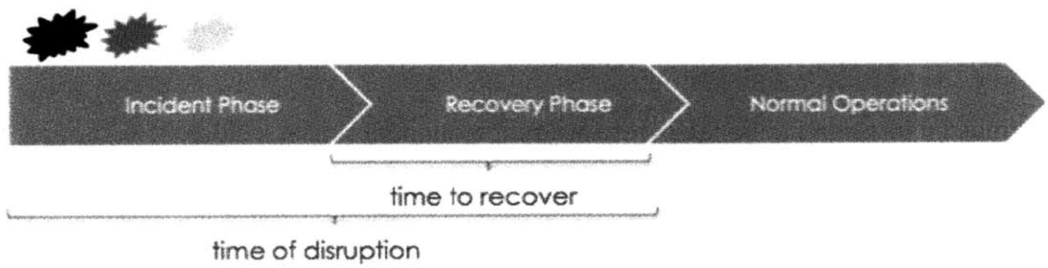


Figure 6-5: Time of disruption

Finally, the number of canceled trains can have an impact on the crew as it results in a change in drivers' shifts and additional maneuvers. TIn the case of the Piccadilly Line, on field interviews suggest the majority of cancelations are implemented at crew depots. This suggests that they do not have a large negative impact on the crew as drivers shifts are usually shortened as a result. The number of cancelations is therefore not included as a measure of crew impact in this methodology.

Application to the Piccadilly Line

In the case of the Piccadilly Line, the lateness value is not directly available from Netmis data but it is accessible through CTFS data as described in Chapter 3. CTFS data contains a field that directly indicates lateness in minutes. This field is used by controllers in real time to identify trains that are significantly behind schedule. As presented in Chapter 3, one limitation of this database is it's low spatio-temporal resolution. In particular, data is available for only one of the two crew relief points of the Piccadilly Line (Acton Town and not Arnos Grove)

6.1.3 Identification of similar incidents

Section 6.1.1 focused on defining recovery effectiveness indices both for passenger and crew impact. These indices are important to integrate to the comparison methodology as they provide metrics to assess the performance of various recovery strategies. This section discusses the selection of similar incidents to apply our comparison methodology on.

Various factors need to be taken into account to identify similar incidents. First of all, it is desirable to select disruptions that occurred in similar locations and during the same time period. The recovery strategy implemented due to a signal failure occurring off peak outside central London can be difficult to compare with one in response to a similar failure that occurred in the peak in central London. Lower train frequencies and higher passenger loads may result in a larger accumulation of overall lateness and a longer time to recover in the latter case. This difference can bias the comparison of the line recovery strategies and effectiveness. In addition, during the peaks in central London, high levels of crowding and smaller headways can raise safety concerns. As an example, no overcrowded train should be held in tunnels due to the increased risk of passenger illness. This can constrain the possible recovery strategies. Depending on the desired level of precision, comparing incidents that occurred in the peak (or off peak) in central zones (or outside the center) might suffice. If more data is available, it would be beneficial to compare incidents that occurred in the same hour and at the same station. If available, the track direction affected by the disruption should also be incorporated in the analysis.

In addition to the time and location of the disruption, it is important to identify similar incident characteristics. The duration as well as the type of incident both have a direct impact on the recovery. In the case of the London Underground, incidents are classified by type as described in Chapter 2. This classification will be incorporated in the choice of case study days. Section 6.2.1 provides more insight into the selection process used for the Piccadilly Line case.

The uncertainty on the total duration of the incident may have a large impact on controller decisions. As seen in Chapter 2, this uncertainty is characteristic of most incident conditions. It is important to take into account the fact that uncertainty may be a driving reason for differences in recovery strategies. In particular, on field interviews suggest that controllers dealing with an incident they are confident will last long are more likely to implement a high number of cancelations in a short period of time. On the contrary, if the controller has very little information on the incident, he may prefer implementing cancelations incrementally. One way of reducing the effect of uncertainty in our methodology is to select incident types that are as equivalent as possible. In particular, the analysis should focus on incident types that have similar distributions in duration. As will be discussed in Chapter 7, the implementation of the methodology on a large sample size may reduce the impact of uncertainty on our conclusions.

Finally, once similar incidents have been identified, the last step is to determine the time window for analysis. It is important to be aware of any previous and following incidents as these may also have an impact on the recovery. The study should focus on a time window during which only the incident of interest occurred.

6.1.4 Identification and comparison of recovery strategies

An important step in the comparison framework is to identify the recovery strategies that were implemented. The reconstructed recovery strategy database described in Chapter 5 is the source of data used for this step. In the case of the Piccadilly Line, the corrective actions that are identified are cancelations, reformations and short-turns. The case study focuses on the evolution of the number of corrective actions that were implemented during the chosen time window. A further step in the analysis would be to study the duration of cancelations as well as the location of short turns. This information is available in the inference database described in Chapter 5.

Figure 6-6 provides an example of the count of corrective actions. It represents the recovery strategy implemented in response to a customer incident that occurred around 13:40 and lasted approximately 20 minutes at Covent Garden on the 4 October. Once the recovery strategy is identified, the study compares the metrics related to the line's recovery with the metrics linked to recovery strategies. This comparison can lead to insight on the impact of specific corrective actions on the recovery of the line.

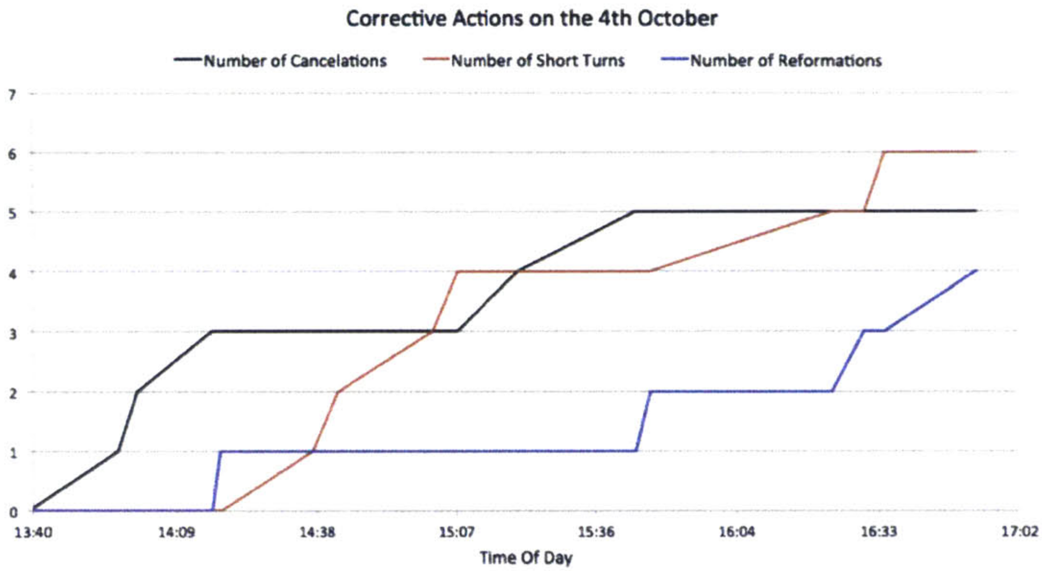


Figure 6-6: Recovery Strategy on the 4th October

6.2 Case Study Presentation

This section describes the case study days selected for the case study analysis. Section 6.2.1 focuses on the methodology implemented to choose the case study days. Section 6.2.2 describes the incidents of the case study and section 6.2.3 details the previous and posterior incidents of the chosen days. This last step is important to identify recovery strategies deployed in response to a specific incident.

6.2.1 Choice of days

The available data for the recovery strategy database covers nine weeks from the 1 October 2013 to the 7 December 2013. As described previously, the case study days must be chosen based on similar incidents, locations, times and characteristics. A good metric to provide insight into these three factors are Lost Customer Hours (LCH) calculated by the London Underground and provided throughout the CuPID database.

Lost Customer Hours are obtained through a modeling tool developed internally by LU that uses the cause, duration, time and location of the incident. It is a metric that reflects the direct impact of the incident on customers' journey times, but does not consider the recovery strategy deployed. This is therefore an appropriate metric to characterize the incident itself. Figure 6-7 illustrates the value of LCH for incidents from October - December 2013. It represents every value per observed incident. As can be seen, the value of LCH varies widely over this period. The average per incident is 917 and the standard deviation of all the values per incident is 4108.

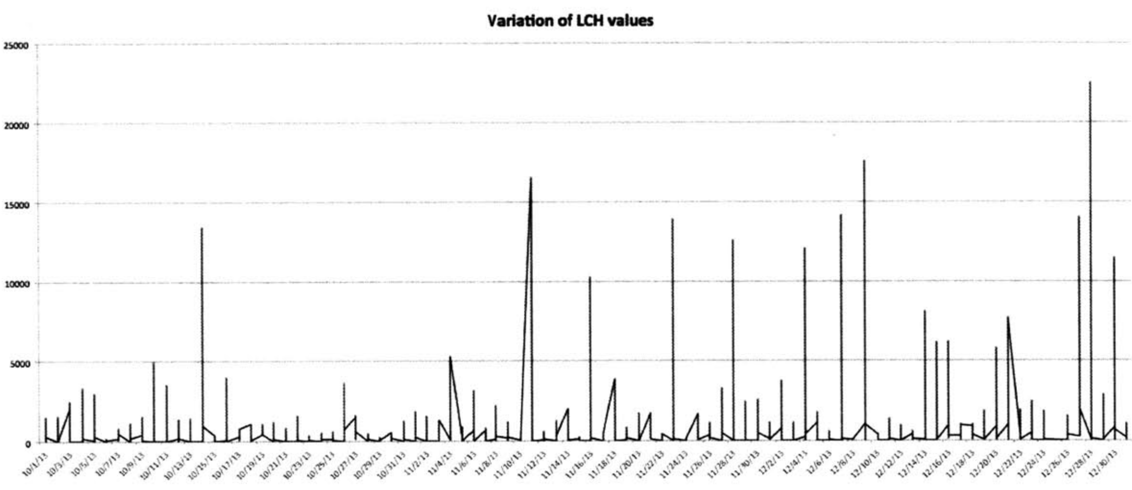


Figure 6-7: Variability in LCH value

Signal failures have a direct impact on data quality. In particular, a signal failure results in a higher number of missing data points as the train tracking information used in this analysis comes from the signaling system. The analysis therefore focuses on customer incidents. Customer incidents are relatively homogeneous and do not impact train movement data collection. As can be seen in Figure 6-8, customer incidents occur throughout the line.

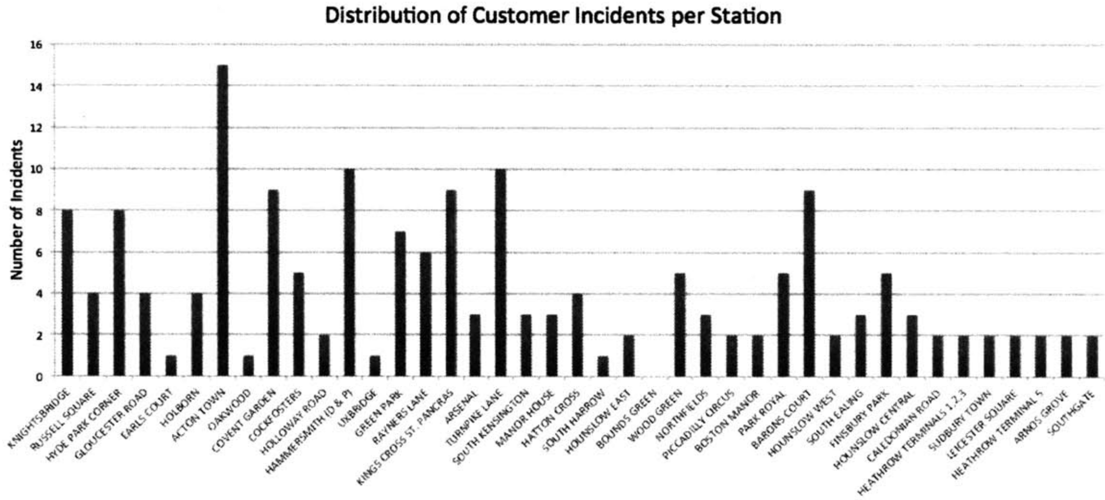


Figure 6-8: Location of customer incidents

The data is sorted by location and time to help identify similar incidents that occurred at the same station. Table 6-1 provides a snapshot of the table used. Due to a limited time period as well as the need for similar time, location and incident characteristics, the number of possible case studies is constrained.

Table 6-1: CuPID data snapshot

Date	Time	Train N	Location	Direction ID	Initial Delay (mins)	LCHs	Cause Category
10/10/13	18:23	241	HAMMERSMITH (D & P)	WB	3	250.61	Customers & Public
11/22/13	8:09	306	ACTON TOWN	EB	2	267.51	Customers & Public
11/8/13	14:08	313	SOUTH EALING	WB	7	300.95	Customers & Public
10/12/13	22:13	350	ARSENAL	WB	5	315.76	Customers & Public
11/11/13	9:17	251	KINGS CROSS ST. PANCRAS	WB	3	365	Customers & Public
11/12/13	18:43	334	ACTON TOWN	WB	4	395.58	Customers & Public
10/2/13	13:55	306	EARLS COURT	WB	6	413	Customers & Public
11/20/13	8:50	314	ACTON TOWN	EB	3	440.37	Customers & Public
11/22/13	22:37	253	MANOR HOUSE	WB	7	444.44	Customers & Public
11/30/13	15:22	360	HYDE PARK CORNER	EB	3	504.48	Customers & Public
12/21/13	13:18		COVENT GARDEN		0	506.54	Customers & Public
12/11/13	17:28	351	KNIGHTSBRIDGE	WB	4	516.05	Customers & Public
10/21/13	18:46	342	HYDE PARK CORNER	EB	4	570.24	Customers & Public
10/25/13	8:45	353	GREEN PARK	EB	4	607.72	Customers & Public
10/27/13	18:20	303	PARK ROYAL	WB	20	611.7	Customers & Public
10/1/13	8:50	313	KNIGHTSBRIDGE	EB	4	626.83	Customers & Public
11/27/13	8:23	304	GREEN PARK	EB	4	694.75	Customers & Public
12/2/13	8:11	310	NORTHFIELDS	EB	5	760.98	Customers & Public
11/7/13	14:14	316	HAMMERSMITH (D & P)	WB	9	836.68	Customers & Public
12/18/13	8:32	276	KINGS CROSS ST. PANCRAS	WB	4	923.94	Customers & Public
10/14/13	18:37	230	MANOR HOUSE	WB	6	965.22	Customers & Public
10/1/13	10:10	312	HYDE PARK CORNER	WB	11	1207.95	Customers & Public
10/16/13	8:48	310	HYDE PARK CORNER	EB	6	1405.62	Customers & Public
10/13/13	18:30	307	SOUTH KENSINGTON	WB	12	1433.27	Customers & Public
10/10/13	8:20	303	HAMMERSMITH (D & P)	EB	5	1667.48	Customers & Public
10/16/13	8:50	254	GLOUCESTER ROAD	EB	6	1822.13	Customers & Public
11/6/13	**10:44**	**257**	**HOUNSLOW WEST**	**EB**	**44**	**3194.27**	**Customers & Public**
12/7/13	**7:24**	**300**	**HOUNSLOW EAST**	**EB**	**84**	**3331.17**	**Customers & Public**
10/4/13	13:51	275	COVENT GARDEN	WB	16	3338.05	Customers & Public
10/11/13	8:29	276	KINGS CROSS ST. PANCRAS	WB	8	3537.79	Customers & Public
10/10/13	8:00	342	GREEN PARK	EB	11	4983.63	Customers & Public
11/16/13	16:25	304	GREEN PARK	EB	16	10294.92	Customers & Public
10/14/13	15:38	342	RUSSELL SQUARE	EB	56	13476.14	Customers & Public

The case study focuses on a comparison between the customer incidents on the 6 November and 7 December. They are characterized by similar, high levels of Lost Customer Hours. Furthermore, they occurred at the same station (Hounslow) and in the same direction (Eastbound). Finally, both incidents occurred off peak, around 7 am on a Saturday for 7 December and around 11 am on a Wednesday for 6 December.

6.2.2 Description of the chosen incidents

This case study is of particular interest because of the strong similarities between the incidents themselves. Both incidents occurred at Hounslow East, which is located on the southwest branch of the Piccadilly Line as shown in Figure 6-9.

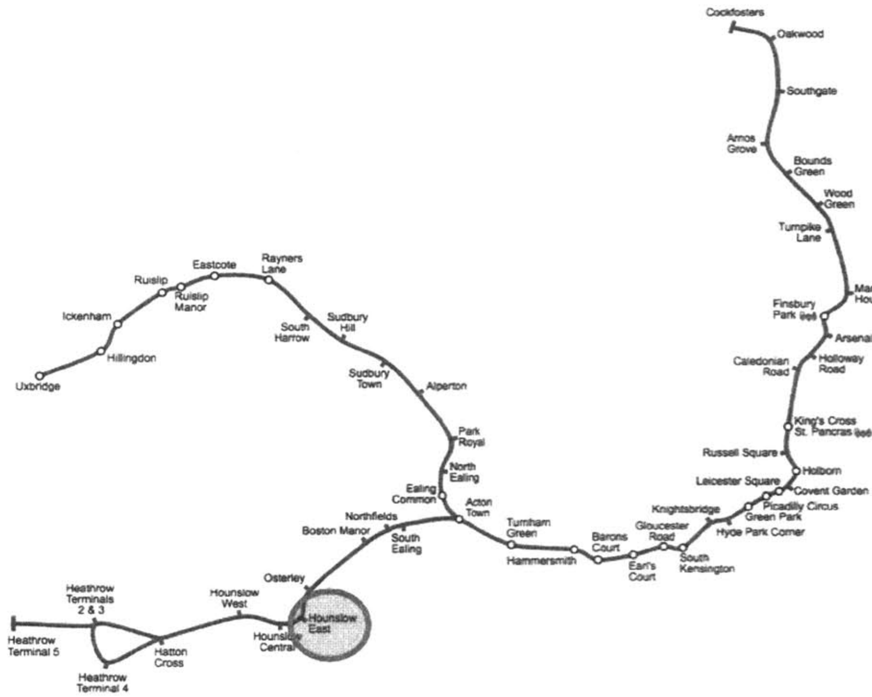


Figure 6-9: Case Study incident location

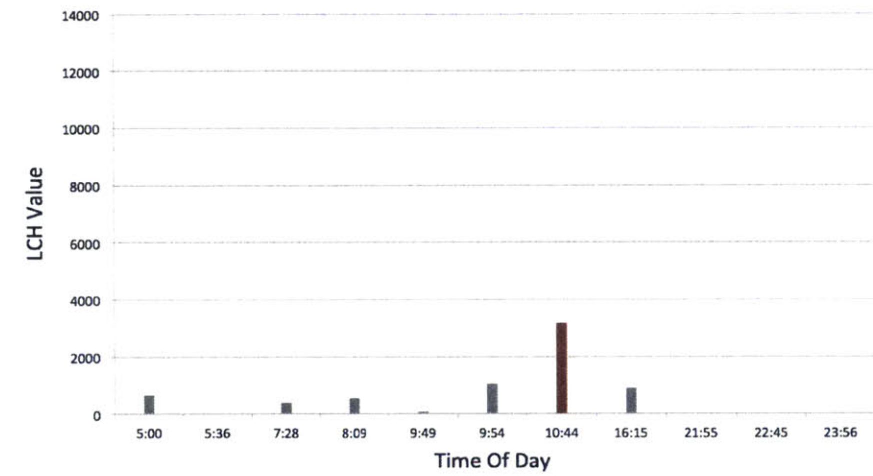
The CuPID data provides additional information that confirms that both incidents are similar. The two incidents are characterized by the Customers & Public – Illness/Accident cause. Appendix A provides details on the reported incidents, as reported by the controller on duty at the time of the incident. On 6 November, even though the incident is reported in CuPID as occurring at 10:44, the detailed description of the incident suggests it occurred at 10:13. According to the CuPID data, a customer fell on the tracks at 10:13 while a train was approaching the station. The traction current was turned off, and the station was evacuated and closed. The Piccadilly Line was suspended between Northfields and Heathrow for the emergency team to intervene. After the victim was removed from the tracks, the traction current was turned on again at 10:50. According to the CuPID data, the service was resumed in both directions at 10:58, after a 45 minute delay.

The information available for the incident on 7 December does not provide a precise start time for the incident. We will therefore assume that the start time is the one indicated in the CuPID time value (07:24), even though this information may not be accurate. The incident is also characterized as a person under a train. Similarly to the 6 November, the traction current was turned off to let the emergency group intervene. The line was suspended between Northfields and Heathrow. An important difference to note is the duration of the incident. According to the controller's description, service resumed at 08:48, 84 minutes after the initial incident. In this case the victim died a few minutes after the incident, which may account for its duration.

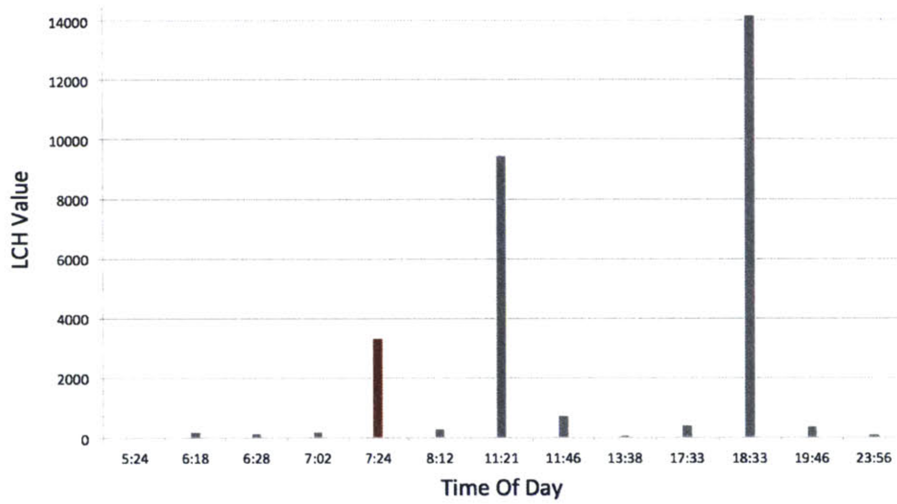
6.2.3 Prior and posterior incidents

The next step was to study the occurrence of other incidents on the chosen days. Incidents that occurred both before and after the selected incident may have an impact on the line recovery and affect the findings. Figure 6-10 illustrates the LCH value of all the incidents observed for the chosen days.

On 6 November, the CuPID data describes the occurrence of various smaller incidents prior to, and after, the major customer incident at Hounslow East. Even though these incidents have lower LCH values, they may have had a negative impact on the recovery of the line. On 7December, two signal failures occurred, the first at Holborn at 11:21 followed by one at Arnos Grove at 18:33. The high LCH values seen in Figure 6-9 and the detailed information provided in the incident descriptions indicate that these signal failures had a significant impact on the overall performance of the line that day. These additional incidents could bias our understanding of the line recovery. Therefore, to understand precisely the effect of a given recovery strategy, the case study limits the analysis to a time window that corresponds to the single incident. In our case, a four hour time window is chosen in which the only major incident observed on both days is the customer incident at Hounslow East. The time window includes time before the reported start of the incident to account for the uncertainty on the exact time of the incident. Section 6.3.1 shows how the reported CuPID time of incident does not seem to correspond to the actual incident start time.



(a) 6 November



(b) 7 December

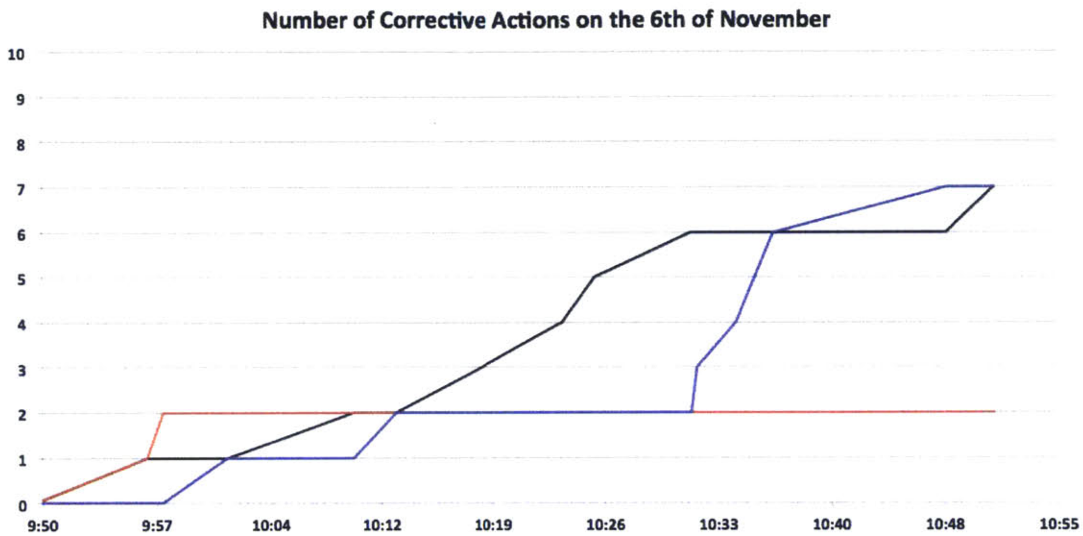
Figure 6-10: LCH Values

6.3 Case Study Results

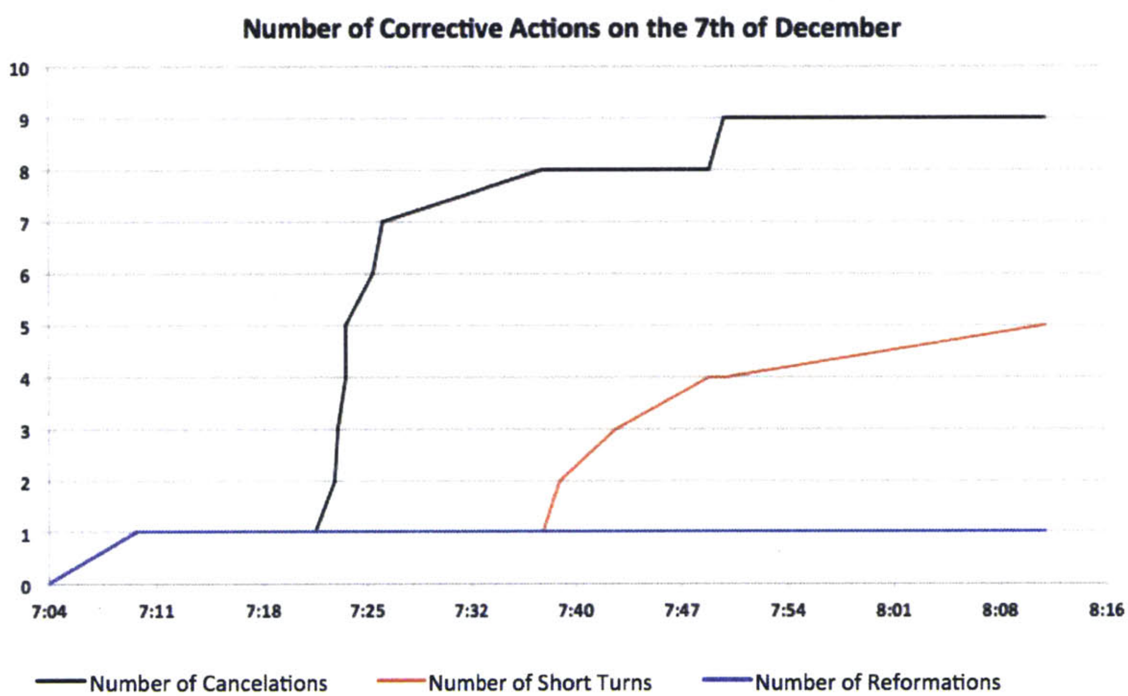
This section presents the results and conclusions of the case study. The case study is an application of the methodology described in Section 6.1. It is presented to illustrate the methodology and is not intended to be a comprehensive analysis of the effectiveness of recovery strategies on the Piccadilly Line.

6.3.1 Recovery Strategy

As described in Figure 6-1, the first step is to analyze the recovery strategies implemented in the time window of interest. Figure 6-11 provides insight into the recovery strategy implemented during the hour after the customer incident. The cumulative number of inferred cancelations, short-turns and reformations are represented in black, red and blue.



(a) 6 November



(b) 7 November

Figure 6-11: Recovery Strategy

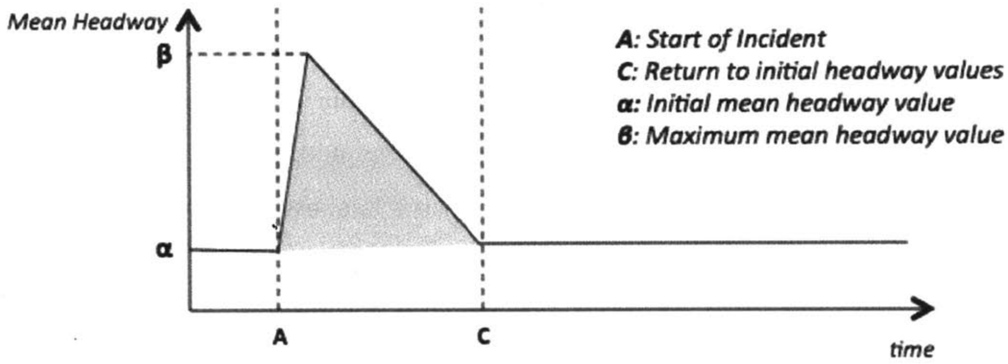
The main difference observed is the speed at which cancelations are implemented on the line. Within the first 20 minutes of the 7 December incident controllers canceled 8 trains, compared to 2 on the 6 November. This aggressive use of cancelations certainly lead to higher headways but perhaps a quicker recovery. The analysis of reliability metrics will provide more insight on the impact of this strategy. The first hour of recovery on the 7 December is characterized by a higher number of short turns and a lower number of reformations compared to the 6 November. The study of the lateness at crew reliefs will provide insight on the impact of this lower number of reformations.

6.3.2 Passenger Recovery Index

The second step is to analyze different performance metrics associated with the chosen time window. This section focuses on passenger REI.

- *Simplified REI:*

The general form of REI was defined in section 6.1. The time window for this case study corresponds to off peak operations with stable levels of demand and stable train frequencies. We therefore assume that the levels of demand as well as the scheduled values of average headway and headway regularity are constant. Additionally, the case study assumes that the headways are regular, and limits the numbers of sections defined in equation (6-2) to 1. Finally, the case study uses the simplified formula defined in equation 6-1, assuming that the REI can be approximated by the area under the triangle as defined in Figure 6-12. The period corresponding to the time between the start of the incident and the return to normal headway values will be referred as the passenger time of disruption. The difference between the maximum average headway and the initial average headway will be called excess average headway.



(a) Rapid recovery with high impact

Figure 6-12: Simplified REI

$$REI = \frac{(\beta - \alpha) * (C - A)}{2} \qquad (6-2)$$

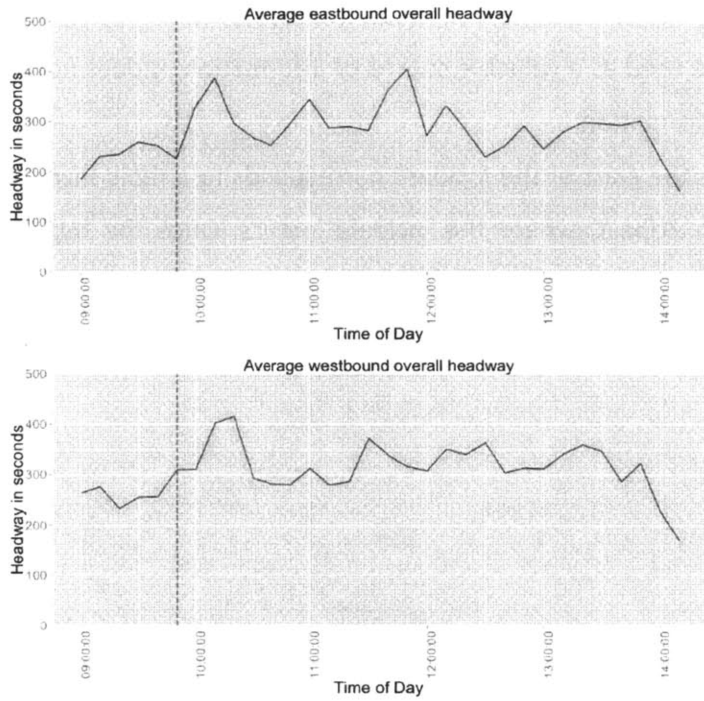
$$REI = \frac{ExcessAverageHeadway * PassengerDisruptionTime}{2} \qquad (6-6)$$

This formulation is a very simplified application of the theoretical REI defined earlier. Further research should incorporate the more comprehensive definition of REI for a fine tuning of results. It is however a good first order representation of the impact on passengers as it captures both the maximum observed values of average headway and the total time during which headway values are observed. The next step is to calculate these maximum values of average headway and total time.

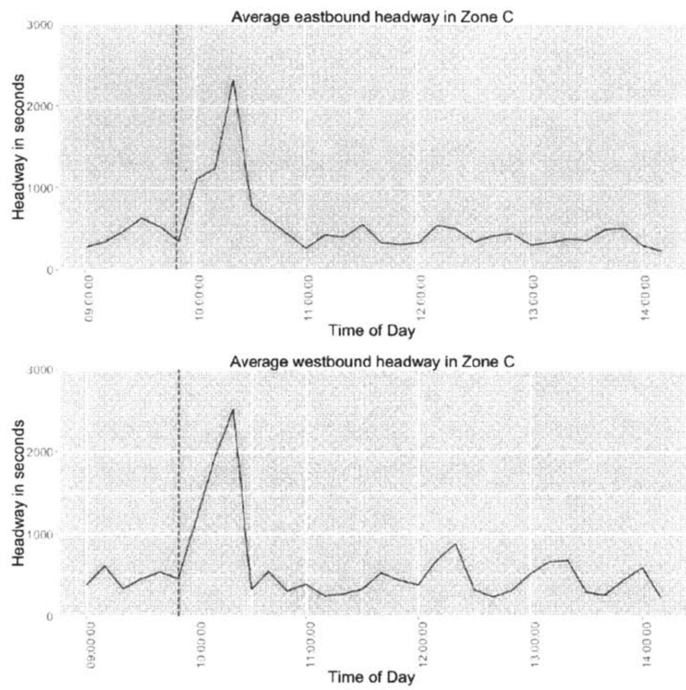
- *Average Headway :*

The passenger REI is calculated both for the overall line and for Zone C where the incident occurred. Figures 6-13 and 6-14 illustrate the average headway in both direction on the 6 November and 7 December. The dashed line is the start of the incident phase. In both cases, this time is different from the reported start time of the incident. On the 6 November, the incident starts around 9:50 am, rather than 10:13 as stated in CuPID. Similarly, the dashed line for the 7 December is at 7:05 am, 19 minutes before the reported incident start time of 7:24 am. From these observations, we can conclude that even though the CuPID data can be used as a valuable source of information, it does not always provide completely reliable information.

For both days, the start of the incident corresponds to a rapid increase in headway values, particularly in Zone C where the incident occurs. Once the headway reaches a maximum value, there is a decrease of average headway. Depending on the days and zones, the decrease in headway is more or less sharp, corresponding to a quicker or slower recovery.

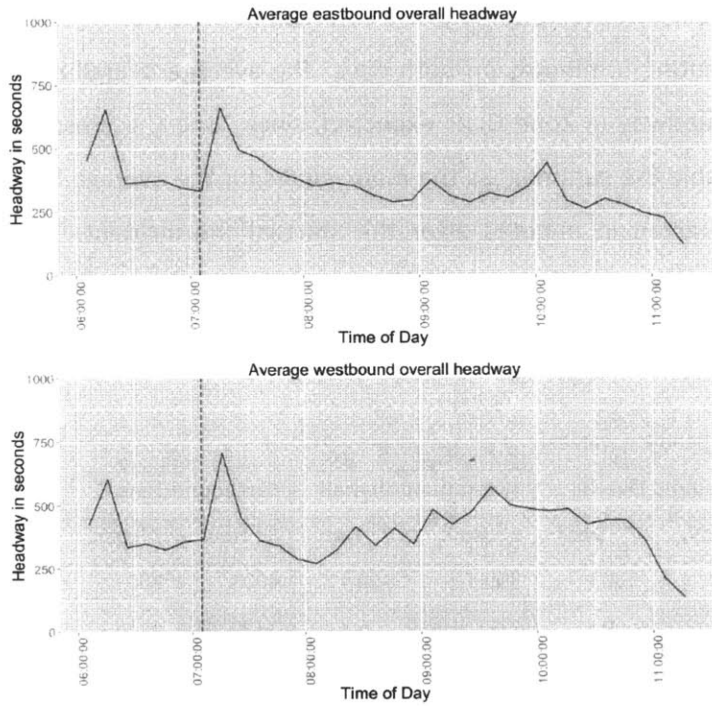


(a) Overall Headway

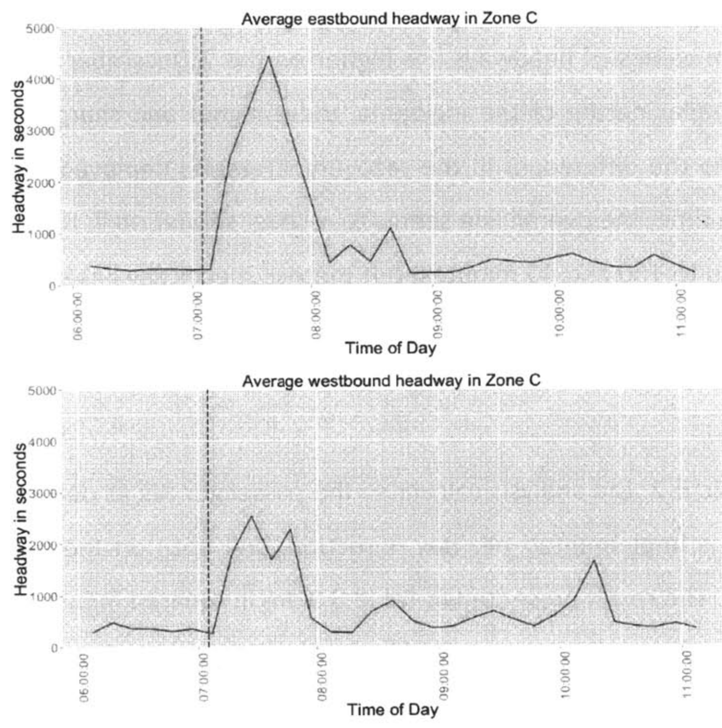


(b) Headway in Zone C

Figure 6-13: Average Headway for the 6 November



(a) Overall Headway



(b) Headway in Zone C

Figure 6-14: Average Headway for the 7 December

The initial values of observed headways correspond to values slightly above normal operations, approximating 5 minutes on both days. The average overall headway is lower than the observed headway in Zone C, as expected, since Zone C corresponds to a single branch of the line. Table 6-2 summarizes the main values for the average headway on both days. The times are given in minutes after the start of the incident. The shaded cells correspond to the smallest values, both for the maximum observed value as well as the passenger disruption time.

Table 6-2: Comparison of headway values

	Eastbound Overall		**Westbound Overall**		**Eastbound Zone C**		**Westbound Zone C**	
	6th Nov.	**7th Dec.**	**6th Nov.**	**7th Dec.**	**6th Nov.**	**7th Dec.**	**6th Nov.**	**7th Dec.**
α	230 s	330 s	300 s	360 s	400 s	300 s	500 s	380 s
β	390 s	660 s	410 s	705 s	2300 s	4400 s	2500 s	2500 s
Passenger disruption time	160 min	85 min	40 min	30 min	60 min	102 min	40 min	60 min

The maximum values of headways are higher on the 7 December and are reached sooner. Given the similar nature of the incidents, these higher and sharper peaks can be assumed to be due to the differences in the recovery strategies deployed. Concerning the passenger disruption time, the overall line seems to recover sooner on 7 December (85 and 30 minutes compared to 160 and 40 minutes) but metrics specifically linked to Zone C have a shorter recovery time on 6 November.

- *REI calculation:*

Based on Table 6-2, the analysis calculates the passenger REI as defined in equation 6-5. The calculation is implemented per day, direction, and both on the overall line and specifically to zone C. The result presented in Table 6-3 are in squared minutes.

Table 6-3: Comparison of passenger REI

	Eastbound Overall		Westbound Overall		Eastbound Zone C		WestBound Zone C	
	6^{th} Nov.	7^{th} Dec.	6^{th} Nov.	7^{th} Dec.	6^{th} Nov.	7^{th} Dec.	6^{th} Nov.	7^{th} Dec.
$REI_{passenger}$	213	233	37	86	950	3485	667	1060

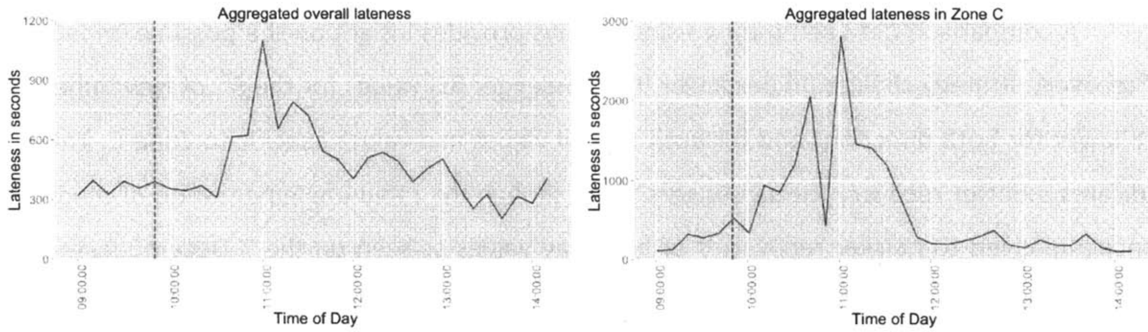
- *Conclusion for passengers:*

A comparison between the two similar days provides insight on the possible impact of recovery strategy choices. In particular, the passenger REI values for the 6^{th} of November are lower for both directions and for both overall values and values specific to Zone C. This indicates a better recovery for passengers. Even though the rapid implementation of many cancelations led to quicker recoveries of headway values as seen on the 7 December, the overall negative impact on passengers was higher. These results complement Babany's findings (2015). The application of his optimization tool to a simulated case suggests indeed that a higher number of cancelations has a positive impact on total time of passenger disruption. This can be confirmed in this case study. However, to complement Babany's (2015) work, it is important to consider the total impact on passengers rather than only the time taken to recover usual headway or waiting time values. Indeed, when taking into account the overall integrated impact on passengers, a more incremental implementation of cancelations has a better output on passengers.

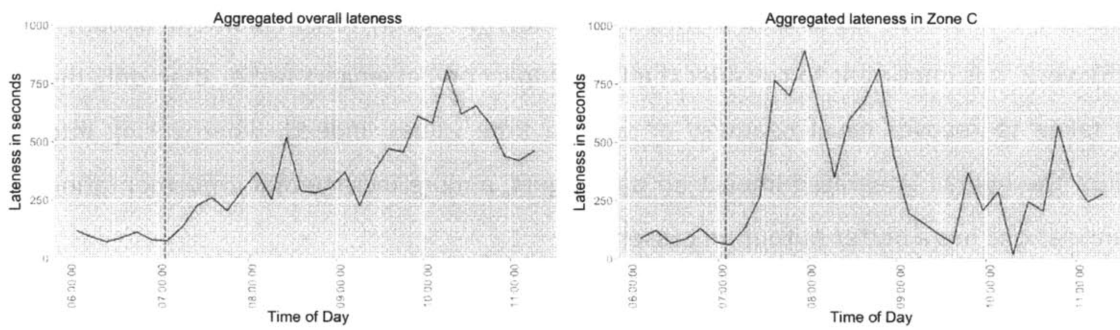
6.3.3 Crew Recovery Index

-Time of disruption and time to recover:

The first step is to look at median values of lateness. This provides a good approximation of the time of disruption and the time to recover.



(b) 6 November



(b) 7 December

Figure 6-14: Lateness evolution

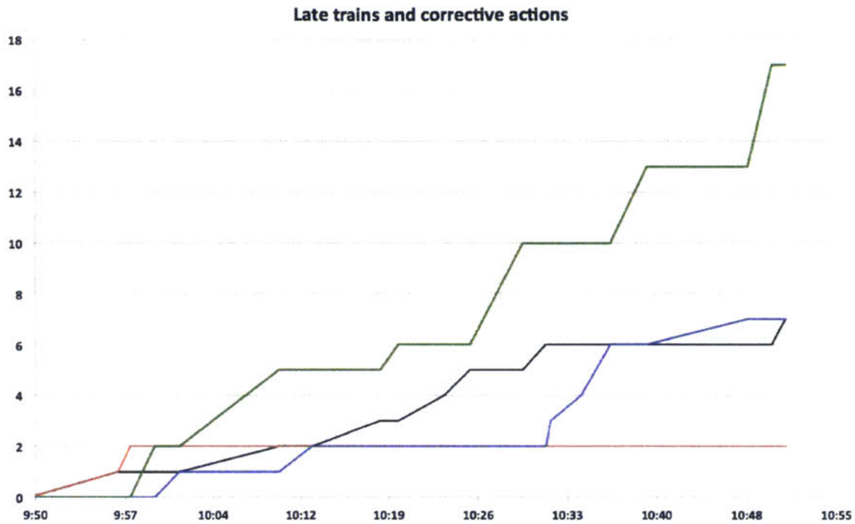
Table 6-4: Time of disruption and time to recover

	Overall		**Zone C**	
	6th November	**7th December**	**6th November**	**7th December**
Time of Disruption	200 min	NA	115 min	150 min
Time to Recover	155 min	NA	70 min	65 min

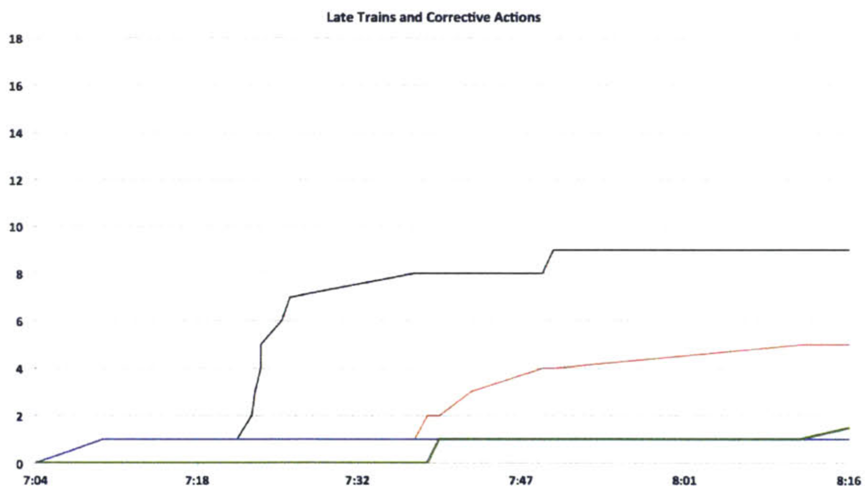
As seen in Figure 6-14, the 6 November has higher maximum values of lateness, both on the overall line and in Zone C. The plot of the median overall lateness on the 7 December shows that the line does not seem to recover the lateness from the incident within the defined time window. In Zone C, even though the time of disruption is shorter on the 6 November, when taking into account the total length of the incident, the 7 December performs better with a lower time to recover. This quick recovery time can be linked with the aggressive cancelation technique.

- *Number of late trains*

Let us observe the evolution of the cumulative count of late trains within the first hour after the incident. A train is considered late if it arrives at the crew relief more than 10 minutes after the scheduled arrival. This analysis is provided in Figure 6-15 and illustrates how the incremental cancelation technique leads to a higher number of late trains compared to a more aggressive approach.



(a) 6 November



(c)

(b) 7 December

—Number of Cancelations —Number of Late Trains —Number of Short Turns —Number of Reformations

Figure 6-15: Impact of corrective actions on the number of late trains

We can calculate a recovery effectiveness index as defined in 6.1. The recovery effectiveness index sums the total number of observed late trains during the disruption (from the start of the incident to the end of the recovery). Table 6-5 summarizes the values in number of trains.

Table 6-5: REI values for crew

	6 November	**7 December**
$REI_{passenger}$	17	13

The 6 November is characteristic of a higher impact on crew. This is due to the large number of cancelations implemented by the controllers within the first hour of the incident.

6.3.4 Conclusion

This case study shows the insight gained from using the methodology presented here to identify corrective actions and evaluate their potential. The analysis illustrated the various steps of the comparison methodology and applied simplified versions of the REI. Comparing the recovery strategy characteristics with various REI measures provide important insight on the impacts of corrective actions on passengers and crew. In particular, the main difference between both days lied in an aggressive implementation of cancelations on 7 December. This analysis of the rate of cancelations complements Babany's research that focused only on the total number of canceled trains.

Even though this aggressive strategy lead to a quicker recovery in headway values, the overall impact on passengers as measured by $REI_{passenger}$ was worse compared to the more incremental approach implemented on the 6 November. Concerning lateness, the aggressive cancelation strategy of the 7 December resulted in a smaller impact on crew in Zone C. However, the overall lateness was not recovered in the time window of analysis. Based on this case study, an implementation of incremental cancelations is recommended to mitigate as best as possible the impact on passengers. This conclusion complements Babany's findings concerning cancelations. Based on a case study, Babany concludes that the implementation of cancelations immediately after the incident leads to better overall recovery. As we have seen, the quick implementation of cancelations does lead to a smaller impact on crew and a quicker recovery to schedule. However, the aggressive implementation of cancelations has an overall negative impact on passengers. This case study is presented to illustrate the methodology developed in section 6.1. In particular, the next step for this research is to implement a similar comparison methodology to a larger number of incident occurrences to provide statistically significant conclusions and recommendations.

Chapter 7: Conclusion

7.1 Summary of the main findings

The first part of the thesis examined the benefit of the merging of several incomplete but complementary databases. The second part of the thesis developed a methodology to infer corrective actions based on a comparison between observed train movements and scheduled train movements. Thanks to the reliable reconstructed AVL database, the comparison successfully inferred the various corrective actions. The methodology applied on the London Underground was limited to short turns and cancelations. A similar framework can be developed to best represent additional corrective actions on other metro systems. The inference methodology provided a numerical database with key information on recovery strategies. This advance provides deeper insight in disruption management and is a useful input for the last part of the thesis.

The last part of the thesis develops a framework to measure the effectiveness of recovery strategies. Thanks to the inference step, precise information concerning recovery strategies is available for any given incident. The analysis defines recovery effectiveness indexes that capture the total impact of the disruption both for passengers and for crew. The methodology compares the recovery strategies implemented and the recovery effectiveness indexes for similar incident conditions. A case study is developed to illustrate the proposed framework. Based on the case study conclusions, this research suggests that an incremental implementation of cancelations leads to a smaller overall impact on both passengers and crew.

7.2 Direct applications and recommendations for the London Underground

This section details the recommendations and potential applications of the research for the London Underground. First of all, the thesis developed a methodology to increase the quality of AVL data. For the Piccadilly Line, the use of a second complementary database helped reduce the percentage of missing train numbers from 23% to 4%. The thesis processed the merging on three months of available data. This available refined data could be used by the London Underground for further analysis. The methodology could be industrialized by the London Underground to significantly reduce issues concerning missing train numbers. This could be useful both to improve the reliability of performance metrics calculated by the London Underground that are based on AVL data, as well as provide a better near to real-time tracking of trains for operations.

The corrective action database can also be of direct use for the London Underground. The available database retraces 9 weeks of data. It provides aggregate information on the location and frequency of the corrective actions as presented in Chapter 5. This information can provide support for infrastructure decisions such as the construction of a new crew depot. It can also be used to assess the impact of a new timetable. The database can also be used in a disaggregate way as illustrated in the case study. Information on the corrective actions that were implemented could be integrated in the daily reviews to encourage a constructive feedback loop between managers and controllers. On the longer term, this comparison methodology could be integrated in a near to real time control tool. It could help controllers keep track of the recovery strategy and simplify communication between drivers, signalers, controllers and managers. It would also be particularly useful during shifts swaps or when several controllers work in parallel.

The methodology developed to measure the effectiveness of recovery strategies provides key insights to disruption management. In particular, it emphasis the difference between crew impacts and passenger impacts. Control service tends to focus primarily on lateness measures, which implies neglecting the impact of disruptions on passengers. Integrating both the passenger and the crew recovery effectiveness indexes to the London Underground analysis tools could provide insight on the true impact of a recovery strategy on both crew and passengers. In particular, the case study illustrated how an aggressive cancelation strategy resulted in a smaller time to recover lateness locally but a larger impact on passengers and crew globally. Other measures such as providing real time information on headway values and regularity could encourage controllers to incorporate passenger centric measures of disruption. In the longer term, a larger scale implementation of the described methodology could lead to statistically significant conclusions and recommendations for best practices.

7.3 Limitations and next steps

- *Limitations concerning the recovery effectiveness index:*

The REI defined for passenger and crew impact are a first step towards an integrated understanding of the impact of recovery. On the passenger side, the proposed index proposed reflects mainly a measure of excess waiting time as a function of average headway average and regularity. It takes into account demand to measure the total impact on passengers. However, other passenger impacts of disruptions described in chapter 2 are not reflected in this measure. For example, disruptions lead in additional transfers or extended in train time. Automated Fare Collecting data could provide a significant improvement for the measure of the recovery effectiveness index for passengers. Further research could address the need to add other negative impacts of disruptions on passengers that are less easily measurable such as added anxiety could be approximated.

On the crew side, the REI is defined as a function of median lateness and total number of late trains. This metric captures the overall impact of lateness on crews and counts in the actual number of affected drivers thanks to the number of late trains. However, it must be acknowledged that train lateness has a variable impact on crew depending of location. In particular, high values of lateness observed at crew relief points are likely to have the highest negative impact on crew. Furthermore, a measure of the number of available spare drivers could be integrated as on site interviews suggested that spare drivers could considerably lessen the negative impact of lateness on crew by providing additional reliefs to respect driver's shifts.

- *Limitations on the developed methodology:*

The methodology developed to measure the effectiveness of recovery strategies also has limitations First of all, each incident is unique and it is difficult to find strictly identical incident conditions for the comparison. For the case study, even though the incident descriptions are very similar and the incidents occur at equivalent locations and time, the total duration of the incident was reported to be 30 minutes longer on the second day. Limited available reliable information on the start time and end time of the incident may limit our understanding of the precise impact of the length of the incident on recovery effectiveness. In particular, the uncertainty concerning the total length of the incident should be acknowledged as a factor that may impact decisions of controllers. The implementation of the comparison methodology on a large number of incidents could reduce the impact of these discrepancies and limitations. Given the low number of similar incidents, the methodology would need to be implemented on a large time window of several years to lead to statistically significant results and recommendations.

-*Further research on decision making:*

Future research could study more precisely the process of controllers' decision making. In particular, it would be interesting to study the impact of personal opinion and biases in controller choices. The databases retracing corrective actions could be used in this context. Provided with information on controller shifts, the analysis could link a controller with the set of actions he implemented. The analysis could thencompare characteristics (such as the number of trains canceled within the first minutes or the total number of reformations between several drivers) over the course of a significant number of incidents. This would provide more insight on the variability of recovery strategies due to personal biases and choices.

Bibliography

Babany, M. (2015) Decision Support for Disruption Management On High Frequency Transit Lines *MSc thesis.* Massachusetts Institute of Technology

Carrel, A. (2009) Diagnosis and Assessment of Operations Control Interventions: Framework and Applications to a High Frequency Metro Line *MSc thesis*. Massachusetts Institute of Technology

Larson, R. C., & Odoni, A. R. (1981). Urban Operations Research. Englewood Cliffs, N.J.: Prentice Hall.

Office for National Statistics (2012). 2011 Census - Population and Household Estimates for England and Wales, March 2011. url:
webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171778_270487.pdf
Retrieved 11th November 2015

Tirachini, A., Hensher, D.A., & Rose, J.M. (2013). Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. Transportation Research Part A: Policy and Practice, Vol. 53, pp. 36-52.

Transport for London. (2016)
url: https://tfl.gov.uk/fares-and-payments/replacements-and-refunds/refunds-for-delays?intcmp=16681
Retrieved 13th December 2015

Transport for London. (2009) Travel in London Report
url: http://content.tfl.gov.uk/Travel-in-London-report-1.pdf
Retrieved 7th December 2015

Transport for London. (2011) Work force report
url: http://content.tfl.gov.uk/tfl-workforcereport.pdf
Retrieved 14th December 2015

Froloff, E., Rizzi, M., & Saporito, A. (1989). Bases et pratiques de la régulation. RATP, Direction du Réseau Routier RC/MSE.

Transport for London. (2015) TFL Annual Report and Statement of Accounts 2014/15
url: http://content.tfl.gov.uk/annual-report-2014-15.pdf
Retrieved 3rd October 2015

Transport for London. (2014) Piccadilly line facts url:
http://web.archive.org/web/20140210035233/http://www.tfl.gov.uk/corporate/modesoftransport/londonunderground/keyfacts/13175.aspx
Retrieved 10th September 2015

Freemark Y. (2013) "Assessing Journey Time Impacts of Disruptions on London's Piccadilly Line". *MSc thesis.* Massachusetts Institute of Technology

Trevor Smith, 2011 url:
https://roundtreessidings.wordpress.com/control/
Retrieved 18th November 2015

Gordon, J. (2012) Intermodal Passenger Flows on London's Public Transport Network: Automated Inference of Full Passenger Journeys Using Fare-Transaction and Vehicle-Loading Data. *MSc thesis*. Massachusetts Institute of Technology

Daniel Alden Wood (2015) A Framework for Measuring Passenger-Experienced Transit Reliability Using Automated Data. *MSc thesis*. Massachusetts Institute of Technology

Rahbee A.B. (2011) Service Reliability At London Underground: A Forensic Approach. Independent Publication

Osuna, E., & Newell, G. (1972). Control strategies for an idealized public transportation system.Transportation Science, 6, 57-72.Deckoff (1990)

Eberlein, X. J. (1995). Real-Time Control Strategies in Transit Operations: Models and Analysis. PhD Thesis. Massachusetts Institute of Technology

Sanchez Martinez G.E. (2015) Real-Time Operations Planning and Control of High-Frequency Transit PhD Thesis. Massachusetts Institute of Technology

Adamski, A., & Turnau, A. (1998). Simulation support tool for real-time dispatching control in public transport. Transportation Research, 32A, 73-87.

Rossetti, M. D. and Turitto, T. (1998) Comparing Static and Dynamic Threshold Based Control Strategies, Transportation Research Part A, Vol. 32, No. 8, pp. 607-620.

Daganzo, C.F., (2009) A headway-based approach to eliminate bus bunching. Transportation Research Part B 43(10) 913-921.

Bartholdi, J.J. and Eisenstein, D.D. (2012). A self-coordinating bus route to resist bus bunching. Transportation Research Part B: Methodological, In Print.

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

Soeldner, D. (1993). A Comparison of Control Options on the MBTA Green Line. MSc Thesis. Massachusetts Institute of Technology

Harshavardhan Ravichandran 2013 Evaluating the robustness of crew schedules for rail transit systems". MSc thesis. Massachusetts Institute of Technology

Appendix A: Station Names and Codes

SUTOR_NAME	SUTOR_CODE	SUTOR_NUMBER
Cockfosters	CFS	1
Oakwood	OAK	2
Southgate	SGT	3
Arnos Grove	AGR	4
Bounds Green	BGR	5
Wood Green	WGN	6
Turnpike Lane	TPL	7
Manor House	MNR	8
Finsbury Park	FPK	9
Arsenal	ARL	10
Holloway Road	HRD	11
Caledonian Road	CRD	12
King's Cross St. Pancras	KXX	13
Russell Square	RSQ	14
Holborn	HOL	15
Covent Garden	COV	16
Leicester Square	LSQ	17
Piccadilly Circus	PIC	18
Green Park	GPK	19
Hyde Park Corner	HPC	20
Knightsbridge	KNB	21
South Kensington	SKN	22
Gloucester Road	GRD	23
Earl's Court	ECT	24
Barons Court	BCT	25
Hammersmith (District & Picc)	HMD	26
Turnham Green	TGR	27
Acton Town	ACT	28
South Ealing	SEL	29.25
Ealing Common	ECM	29.75
Northfields	NFD	30.25
North Ealing	NEL	30.75
Boston Manor	BOS	31.25
Park Royal	PRY	31.75
Osterley	OST	32.25
Alperton	ALP	32.75
Hounslow East	HNE	33.25
Sudbury Town	STN	33.75
Hounslow Central	HNC	34.25
Sudbury Hill	SHL	34.75
Hounslow West	HNW	35.25
South Harrow	SHR	35.75
Hatton Cross	HTX	36.25
Rayners Lane	RLN	36.75
Heathrow Terminals 123	HRC	37.25
Eastcote	ETE	37.75
Heathrow Terminal 4	HRF	38.25
Ruislip Manor	RUM	38.75
Heathrow Terminal 5	HRV	39.25
Ruislip	RUI	39.75
Ickenham	ICK	40.75

Appendix B: Pseudo-Code of the Merging Algorithm

d indexes the day, from the 2013-10-01 to 2013-12-31

RawDataCTFS : AVL data frame from CTFS for the given day *d*, that contains correct train numbers

RawDataNetmis : AVL data frame from Netmis for the given day *d*, with correct and zero train numbers

```
for d in all the dates

  TrainsNetmis <- vector of all the unique Train Numbers of RawDataCTFS
  NetmisZero <- subset of RawDataNetmis with Train Number = 0
  TrainIdentification <- vector of all the unique values of Train Identifications of
  NetmisZero

  for i in TrainsNetmis

    ExtractCTFS <- subset of rawdataCTFS_d with Train Number = i

    for k in TrainIdentification

      ExtractNetmis <- subset of NetmisZero with Train Identification = k
      ExtractNetmis$difference <- difference of time between two rows of ExtractNetmis

      Cuts <- vector of indexes of ExtractNetmis where difference is > TimeGap

      for c in Cuts

        ExtractNetmisCut <- subset of ExtractNetmis where time is > time[c] and <= time[c+1]

        if number of rows of ExtractNetmisCut is > NumberRows

          Intermediate <- merge ExtractNetmisCut and ExtractCTFS by Train Station
          ListStations <- vector of all the unique Train Stations of intermediate

          PointsNetmis <- number of rows of the merge between ExtractNetmisCut and ListStations

          Intermediate2 <- subset of Intermediate with difference between Netmis
            and CTFS < TimeDifference

          CommonPoints <- number of rows of intermediate2

          Ratio <- CommonPoints/PointsNetmis

          if Ratio > RatioValue and CommonPoints > CommonPointsValue

            update the values in Netmis of these rows to the Train Number i

            save in a separate file the plot of the superposition of CTFS and this subpart of Netmis data

          end if
        end if
      end for
    end for
  end for
```

Appendix C: Zone segmentation

SUTOR_NAME	Zone
Cockfosters	A
Oakwood	A
Southgate	A
Arnos Grove	A
Bounds Green	A
Wood Green	A
Turnpike Lane	A
Manor House	A
Finsbury Park	A
Arsenal	B
Holloway Road	B
Caledonian Road	B
King's Cross St. Pancras	B
Russell Square	B
Holborn	B
Covent Garden	B
Leicester Square	B
Piccadilly Circus	B
Green Park	B
Hyde Park Corner	B
Knightsbridge	B
South Kensington	B
Gloucester Road	B
Earl's Court	B
Barons Court	B
Hammersmith (District & Picc)	B
Turnham Green	B
SUTOR_NAME	Zone
Acton Town	B
South Ealing	C
Ealing Common	D
Northfields	C
North Ealing	D
Boston Manor	C
Park Royal	D
Osterley	C
Alperton	D
Hounslow East	C
Sudbury Town	D
Hounslow Central	C
Sudbury Hill	D
Hounslow West	C
South Harrow	D
Hatton Cross	C
Rayners Lane	D
Heathrow Terminals 123	C
Eastcote	D
Heathrow Terminal 4	C
Ruislip Manor	D
Heathrow Terminal 5	C
Ruislip	D
Ickenham	D
Hillingdon	D
Uxbridge	D

Appendix D: Detailed CuPID incident information

- **6th of November**

****PRELIM EIRF 596794*** Person Under Train. 1013 SS Hounslow Central alerted via emergency operation of PHP that female customer had apparently suffered a fit while at the PTI on the eastbound platform as T.257 was on the approach. Female customer had fallen on to the track and subsequently T.257 had travelled over her location. SS informed Service Controller and requested traction current be turned off. Station evacuated and closed. Piccadilly Line suspended Northfields to Heathrow in both directions. Trains berthed in platforms and reversed as appropriate. Incident Channel 27 in use. SOO Gold Control. Piccadilly Line SM Silver control. 1030 LAS, LFB and BTP on site. Customer conscious underneath the train, approx. 10m from the front cab. SS and T.Op laid SCDs front and rear. SS and emergency services accessed track and to assess casualty. 1035 ERU on site. 1040 DRM, TOSMs and relieving T.Op on site. Casualty moved to platform. Conscious and first assessments suggest injuries to the right pelvis. LAS continuing to treat. 1047 NIRM on site. 1050 All staff and equipment clear of the track. Request for traction current to be turned on made. 1052 Incident train departs Hounslow West to Northfields Depot with relieving T.Op and TOSM on board. Track search commenced, personal items of casualty found and ERU access track to gather them. Casualty moved from platform area to ambulance by LAS. NIRM becomes Silver Control. 1058 Service resumes in both directions. Hounslow West station reopened. Customer suffered fit while near edge of platform which caused her to fall on to the track as T.257 approached. Situation very well controlled by SS on site which aided swift resolution of incident.*

- **7th of December 2013**

DTSM EIRF 602658: Piccadilly Line service suspended between Northfields and Heathrow EB & WB due to a person under a train at Hounslow East eastbound. When spoken to by the undersigned the Train Operator of eastbound train 300 was able to briefly explain that he had approached the station at normal line speed, as he came around the left hand bend into the platform he could see a number of people waving. Initially he though they may have been messing about on their way home after a night out, however he then saw a body laying on the track ahead. The train was stopped by an emergency brake application and came to a stop approximately three cars into the platform and the Service Controller then alerted by train radio. At the same time a customer on the eastbound platform used a help point to alert the Station Supervisor to the same situation. Traction current was discharged Hatton Cross - Hounslow East EB and LAS & LFB arrived on site shortly after. The LFB requested that traction current also be discharged Hounslow East - Hatton Cross westbound to allow safe access to the site. BTP, ERU & NIRM arrived on site and incident talk group 27 established to maintain communications. Initially the Station Supervisor took responsibility as Bronze Control and this was later handed over to the DRM with the NIRM assuming Silver Control. The Train Operator of train 300 was supported throughout the incident by a spare Train Operator from Acton Town. He was interviewed by the BTP and later returned to Arnos Grove by special taxi where he was offered any additional support he may require. Eastbound train 342 was stalled approximately ten metres from the platform at Hounslow West, an emergency detrainment was authorised and implemented by the Station Supervisor and additional staff. This was completed at 0818 hours with approximately 120 customers being walked to the station and no reported injuries or complaints. Eastbound train 302 was stalled on the eastbound approach to Hatton Cross at signal WW2, the Hounslow West Station Supervisor was authorised to open section switches 844 & 844A to provide a single end feed allowing train 302 to berth into the platform. Section switches 844 & 844A were authorised to be left open and were subsequently closed by the night DRM. BTP officers 1963 & 4296 viewed the station CCTV (camera 23) and declared that the incident was not deemed suspicious and stated that services were free resume following an all clear. It appears the person was asleep on a bench, woke up and then was walking around a family on the platfrom before falling on the track. The person that had fallen onto the track received medical attention under the train and was able to be removed alive from under the train but subsequently declared deceased on the platform. Traction current sections Hatton Cross - Hounslow EB & WB were recharged to allow the incident train to move. At 0848 hours, the all clear was given, and services resumed with trains being reformed, reversed and cancelled as required.