Diagnosis and Assessment of Operations Control Interventions: Framework and Applications to a High Frequency Metro Line

by

André Carrel

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

September 2009

Author.......................... ....
Department of Civil and Environmental Engineering
June 12, 2009

Certified by...................... ....
Nigel H.M. Wilson
Professor of Civil and Environmental Engineering
Thesis Advisor

Certified by.................. ....
Rabi G. Mishalani
Visiting Associate Professor of Civil and Environmental Engineering
Thesis Advisor

Certified by.......... ....
John P. Attanucci
Research Associate, Department of Civil and Environmental Engineering
Thesis Advisor

Accepted by..
Daniele Veneziano
Chairman, Departmental Committee for Graduate Students

Diagnosis and Assessment of Operations Control Interventions: Framework and Applications to a High Frequency Metro Line

by

André Carrel

Submitted to the Department of Civil and Environmental Engineering
on June 12, 2009 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in Transportation

ABSTRACT

Service control, the task of implementing the timetable in daily operations on a metro line, plays a key role in service delivery, as it determines the quality of the service as provided to passengers. This thesis proposes a framework for the study of rail service control which builds on the integration and analysis of data from multiple sources and on background knowledge about service control. The framework takes into account the decision environment in which service control takes place and acknowledges that the reliability of the system depends on many factors which are endogenous to it, aspects previously not recognized in a comprehensive manner by researchers and practitioners alike. This research makes use of automatically generated operational and passenger data, which are increasingly available and accessible to transit agencies and allow for addressing questions in service control from multiple perspectives. As a result, this study takes a distinctly different approach than previous research, which has mostly focused on individual service control strategies and relied heavily on modeling and on simplifying assumptions about the objectives and constraints of service control on a metro line.

The developed framework consists of four main elements. First, the controller's decision environment is integrated and described based on an extended visit of the author to a control center. Second, an algorithm for reconstructing train operations from signaling data and identifying service control interventions is presented. Third, a measure for assessing the impact of the interventions on operations is introduced. The fourth and final element is a set of passenger travel time and reliability measures.

The framework is applied to the Central line, a high-frequency line of the London Underground where the control center observations were also made. Three common service control strategies are assessed in terms of their impact on operations and on passengers, and the influence of timetable variables on the frequency of service control interventions is investigated. From observations at the control center, it is found that aside from the objective of maintaining adequate levels of service from an operations perspective and minimizing the impact of schedule deviations on passengers, considerations relating to crew and rolling stock management, safety and infrastructure capacity have a major influence on service control decisions. Given the uncertain environment in which service control operates, a strong preference among controllers for manageable and robust solutions is observed. In the analysis of common control strategies, it is found that in the absence of official policies on the response to certain types of problems on the line and in the presence of the multitude of factors mentioned above, service controllers have developed rules of thumb which may not always be optimal from the passengers' perspective. Furthermore, the fundamental tradeoff is highlighted between the availability of spare resources in form of drivers, trains and infrastructure capacity and the need for service control interventions. Regarding the influence of timetable variables, it is found that an increase in scheduled service frequency and in running times on an otherwise unchanged line operating close to its capacity caused significant increases in numbers of service control interventions, mostly due to a higher rolling stock requirement and reduced operational flexibility.

Recommendations are made with regards to service control policies, the structure of responsibilities among operational staff, the design of the timetable and the design of the operations control system. Although the results provided by the applications are specific to the Central line, they demonstrate how the elements of the framework can be implemented in a practical setting, and many of the conclusions of this thesis are transferable to other metro lines and systems. Finally, future research in passenger behavior and crew management in the presence of service control interventions is proposed.

Thesis advisor: Nigel H.M. Wilson
Title: Professor of Civil and Environmental Engineering

Thesis advisor: Rabi G. Mishalani
Title: Visiting Associate Professor of Civil and Environmental Engineering

Thesis advisor: John P. Attanucci
Title: Research Associate, Department of Civil and Environmental Engineering

Acknowledgements

I am very grateful to my research advisors, Prof. Nigel Wilson, Prof. Mishalani and John Attanucci for providing me the opportunity to work on this topic and for their guidance and insightful feedback over the past two years. Thank you also to Dr. George Kocur for the interesting and thought-provoking meetings we had at the beginning of my research.

I would also like to express my gratitude to Adam Rahbee, who introduced me to this field of study and continued supporting and encouraging me throughout my second year at MIT.

This research was generously supported by Transport for London. I would especially like to thank Charles Horsey, Chris Baitup, Sandra Weddell and Nick Whelan for their help and for the valuable and productive meetings we had. Many thanks also to Mike Donnithorne and the entire staff at the Wood Lane control center for letting me be part of their team over the course of two weeks.

I am very indebted to Elizabeth Paul for all her efforts and contributions to this research, and for bearing with me for this entire thesis. Special thanks to my friends and classmates from the MST program who have made the past two years at MIT such an outstanding experience, especially to my roommates David Uniman and Harvey Scorcia.

And last but not least, many thanks to Ginny Siggia for all the administrative support she has provided throughout the past two years.

Table of Contents

List of Figures

List of Tables

1 Introduction

This thesis presents research into service control with a focus on a high-frequency metro. It highlights the importance of studying service control both from an academic and practical point of view and shows what value a transit agency can draw from institutionalizing the analysis of service control interventions. Specifically, it describes the decision environment of rail service controllers and shows how automatically collected signaling data can be used to reconstruct operations, including interventions by service controllers. Furthermore, it presents measures with which a transit agency can assess the impact of control decisions and demonstrates how an agency can analyze those decisions from multiple perspectives.

The approach of this research differs from previous research efforts on this topic, which have been strongly focused on mathematical modeling and have suffered from the simplifications which had to be made in order to make the models tractable, thus limiting their applicability. The description of the service controller's decision environment is intended to point out those gaps and to show what would need to be done in order to fill them. The chapters on diagnosing and assessing service control interventions then present an alternative to using models for analyzing and improving service control.

1.1 Motivation

Service control is, broadly speaking, the task of operating the schedule under everyday conditions and modifying it in real-time to deal with the inherent variability in dwell times and running times which are experienced by vehicles on the line as well as unforeseen events which disrupt the service. These sources of service unreliability affect passengers negatively and are of major concern to most transit agencies. However, although unreliability can be observed in the variability of passenger travel times and in passenger complaints, its root causes are clearly on the supply side and cannot be effectively addressed without building a good understanding of operational processes on the transit line of interest. Building such an understanding can only be achieved by studying daily operations on the system, of which service control is an important component. This holds especially true on a high-frequency, high-demand metro line where non-ideal ("disrupted") conditions exist for a large fraction of total operating time.

Despite recent advances in vehicle and signaling technology and improved control systems which are at the disposal of service controllers, the field of service control remains heavily reliant on human judgment and on informal, undocumented practices. It is often poorly understood by outsiders even within the transit agency; in the course of the author's work, service control was described more than once by transit agency staff as a "black art". Owing to the complexity of the field, there is little published on service control, and in many transit agencies managers and members of planning staff are often not well informed of the decisions and problems faced by service controllers as they try to implement the schedules on a daily basis. This has several implications: On one hand, management decisions and agency policies aimed at improving service control, despite the best of intentions, may simply be unrealistic or inapplicable in the real-world context. On the other hand, in the absence of an understanding of the role of service control, planners may have difficulty interpreting performance metrics correctly and verifying whether assumptions and models used in the scheduling process were in fact correct.

This research is a step towards remedying such problems by attempting to deliver a complete description of the rail service controller's decision environment and by demonstrating the practical value of studying service control. Ultimately, this research can enhance the feedback loop to rail transit managers with the help of several concrete application examples. Furthermore, it shows that this can be done with the help of data which are readily available within the transit agency; the only cost consists of the effort needed to extract, process, integrate and analyze the necessary data.

1.2 The service delivery process

One can imagine a transit service as a business process, as shown in Figure 1-1. The overall service policies such as span of service, frequency and routing are determined at the management levels of the transit agency. Service policy decisions are usually based on expected or actual demand, network connectivity considerations, financial constraints and political considerations. These policies are then used by the planning department, which is responsible for developing an operations plan. The operations plan, which will be described in more detail in section 2.1, is the detailed plan which describes the utilization of the agency's resources – rolling stock, personnel and infrastructure – in order to meet the service policies. The most important component of the operations plan is the timetable, which features all train movements reflecting where and when transportation service should be provided to customers. The last piece within the transit agency comes together at the operational level, where the operations plan is implemented. It includes all front-line staff (train operators, doormen) as well as vehicle and infrastructure maintenance divisions, engineers and operational support personnel. Service control[1], which is an essential component at the operational level, oversees and coordinates the implementation of the operations plan and modifies it in order to

---

[1] The term "service control" is synonymous with "operations control" and "service management".

cope with unforeseen events and short-term infeasibilities: It is the centralized, real-time control of schedule-based operations. Its centralized nature, in which controllers are provided with information on the state of the entire system, sets it apart from local dispatching techniques, where a supervisor is positioned at a terminal or a station along the line. The service delivery process results in the daily operations which are provided to passengers. It is important to note that the passenger experience is heavily influenced by service control. That is, in daily operations, passengers do not experience how the service was planned to be operated, they experience the actual operations.

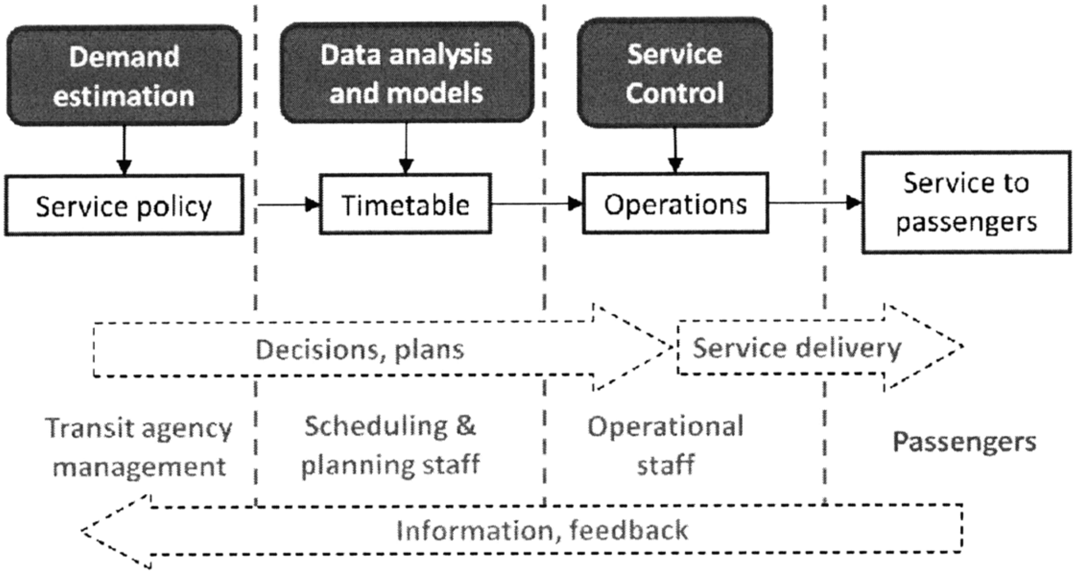


Figure 1-1: Transit service delivery as a business process

As is shown in Figure 1-1, decisions flow from left to right in the diagram; the high-level, strategic decisions are taken at the management level and the planning level then decides how to implement them in practice; the operational level, located at the end of the decision chain, should ideally be able to carry out the resulting plans directly but in reality often needs to fine-tune them. In order to make realistic and informed decisions, the higher levels rely on information about daily operations and the performance of the system, which is provided by the operational level to the planning level and then on to the management level. In addition, some feedback is also given by passengers in the form of complaints or responses to surveys.

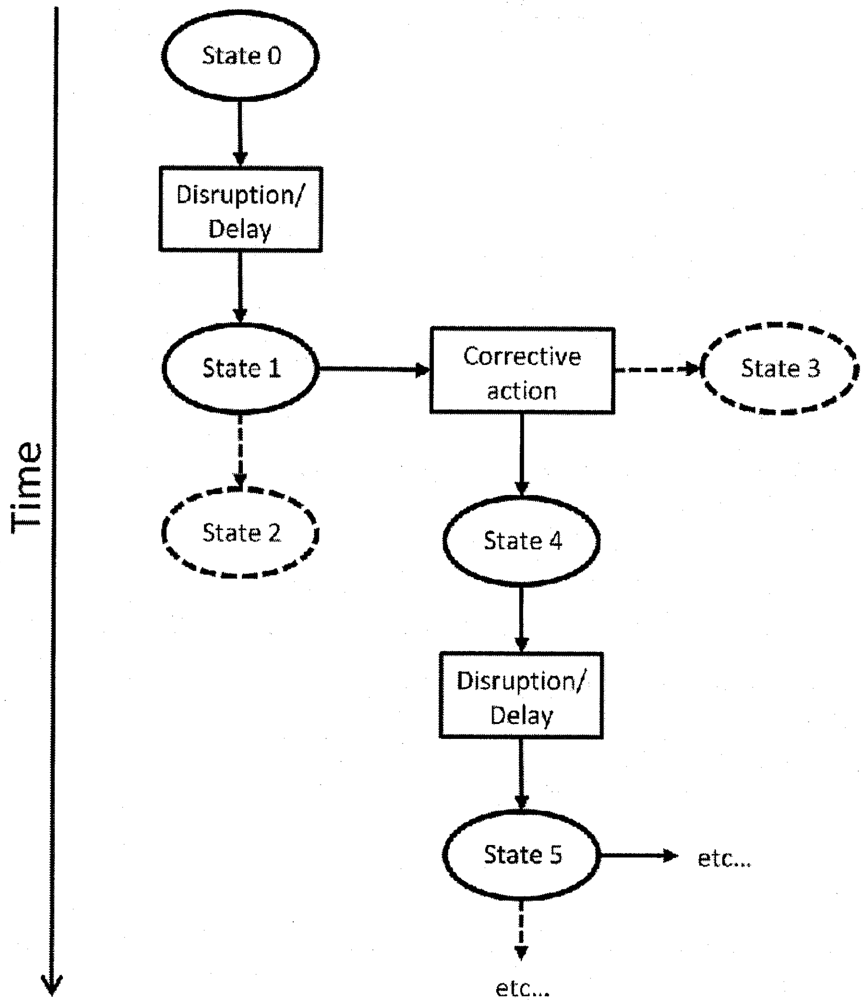


Figure 1-2: An idealized representation of the service control process (Adapted from Froloff, Rizzi & Saporito, 1989)

As one can see, service control is not only a crucial element of the service delivery process, but it also directly governs the interaction between the supply of transportation capacity and demand for it on a daily basis. As a matter of fact, the service control center (or centers, depending on the degree of centralization) of a metro system can be described as the network's communication, decision-making and coordination node(s). The continuous nature of this process is illustrated by Figure 1-2, which shows an idealized representation of the service control process, adapted from a service control manual written for the transit authority of Paris, the RATP (Froloff, Rizzi, & Saporito, 1989). The transit line (or system) starts at a state of optimal operations, denoted here as state 0. Due to unforeseen events (also known as disruptions) or delays, it then moves to a disrupted state, denoted as state 1. If nothing were done, operations would further deteriorate and the line would drift towards state 2. To prevent this and to restore good operations, service controllers take corrective actions towards a target state, state 3. These actions extend over time and the system only slowly moves towards its target state via other states, such as state 4. During that process, further unforeseen events may occur, prompting controllers to intervene again. In Figure 1-2, the line only improves to state 4 before a further disruption or delay occurs, causing it to deteriorate to state 5. As the timeline to the left of the figure suggests, this is a continuous process over time which takes place throughout the day, from the beginning of service until the end.

1.3 Why study service control?

It can safely be said that service control is one of the most poorly understood aspects of rapid transit. In a certain sense, it is a "black box" to managers and planners alike, despite its crucial role in service delivery. Therefore, the flow of information from the operational to the planning and management levels of a transit service, as shown in Figure 1-1, is often not as strong as it should ideally be. This and a number of other reasons outlined in what follows justify studying and better understanding service control:

- As Rahbee (2001) points out, service control techniques and objectives are typically passed down by word-of-mouth, and they vary across lines, let alone agencies. In many transit agencies, there is not much formal training for service controllers. Yet, given the level of influence that service control has over the functioning of the line, it would be desirable for transit agencies to move towards more formal training, and ultimately more unified service control policies.
- Generally, transit agencies track the performance of their lines with performance metrics, high-level measures of service quality and customer satisfaction. With the help of these metrics, deviations of daily operations from the service plan can be observed at an aggregate level (unless the metrics are inappropriate), but only the study of service control will help build an understanding of the nature of and the reason for these deviations. As Rahbee (2001) notes, it may also point towards problematic performance metrics or agency policies. Such research should rely on an objective and preferably automated identification of service control interventions (for example by using a program such as the one developed in this research.) Having to rely on manual logs filled out by service controllers is often problematic since such logs will generally be filled in only when controllers have time. During the most stressful moments in control centers, for instance during major disruptions, controllers are

likely to be consumed by their service management tasks, to the detriment of the reliability of the manual logs. Yet, the response to such disruptions would be of great interest for an analyst to study. Aside from that, a further problem with manual logs is that controllers may record poor service management decisions incompletely or not at all.

- As previously pointed out, much value can be added to investment decisions by taking into account the needs of service control. A detailed analysis of the usage of infrastructure providing operational flexibility (e.g., crossovers and reversing tracks) can show where upgrades or the deployment of spare trains, drivers or mechanics would be most sensible. In addition, an analysis of the disruption patterns on a line and of controller's responses to disruptions where no such infrastructure is available can give indications of where the construction of additional reversing capacity would have the largest positive impact on service stability.
- When upgrades are made to operations control systems, a thorough assessment of the needs of controllers can provide valuable benefits during operation. Many of the systems are only customizable before installation and are very difficult to upgrade or amend at a later stage. The experience of the author has shown that controllers are not always able to clearly formulate their procedures and needs; thus there is a need for an analyst to observe daily operations in the control room and analyze operational data to gain a good understanding of those needs.

1.4 Objectives

This research has three objectives:

1. To help researchers, practitioners and analysts who have not been closely involved with service control understand the decision environment of service controllers with its objectives and constraints. The author's intention is that this can improve cross-divisional communications within an agency by defining a common knowledge basis and by giving planning and management staff an understanding of what causes service controllers to take the dispatching decisions which are observed in daily operations. Furthermore, it can help researchers move towards more realistic models which serve as the basis for decision making, operations planning and strategic planning.
2. To demonstrate how archived data from train signaling systems can be used to reconstruct operations and service control decisions on a high-frequency metro line, thus allowing the

analyst an objective, unbiased view of how the schedule is implemented on a daily basis and what changes are made by service controllers. This information can be used to verify assumptions made during the scheduling process, validate models and identify operational problems which were previously unknown. Furthermore, a set of measures is proposed to determine the impact of service control decisions on passengers and on the train service, thus allowing the comparison between different strategies for regulating service.

3. To illustrate the value of studying operations and service control by way of four applications on the London Underground Central line, aimed at showing how the operations data and impact measures can be tied together in order to evaluate decisions taken by service controllers in a real-world context.

1.5 Research approach

This research aims to be of value to any researcher or analyst wishing to study service control on a high-frequency metro line. However, the details of how service control functions are highly dependent on the characteristics of the transit agency and even the line in question, so it is difficult to conduct research on service control that will lead to universally applicable findings which are at the same time relevant in practice. This thesis uses one line on the London Underground network, the Central line, to conduct the research. The findings of this research are directly relevant to the London Underground, both with respect to the Central line and with respect to its other lines since all lines are managed in a similar way. In addition, throughout the thesis, discussions are provided on how the findings from the Central line can be generalized to other metro systems. As the general ideas behind the methods used would remain the same for any metro line, and only the specific methodology to implement them would vary.

The objectives outlined in section 1.4 were achieved through the analysis of data available from the London Underground and an extended visit by the author to the Central line control center. Observations made during the latter served as a basis for objective 1. For objective 2, the author developed an application for processing operations data available for the Central line. With experience from reconstructing operations and knowledge gained during the visit to the Central line control center, a set of cases were identified and selected for detailed study to achieve objective 3. Based on those cases and on the available data sources, several impacts of service control decisions were defined and measured and the requirements for each measure were detailed. They can be used for any analysis of service control interventions, provided the data are available, and therefore they are not restricted to the applications which are presented as part of objective 3. Thus, the applications are intended to serve as examples of how such research can be conducted on other metro lines.

Although this section already refers to the London Underground and the Central line, the following chapter is general in nature and will not yet focus on the Central line. The London Underground context will be discussed in more detail in chapter 3.

1.6 Organization of this thesis

Chapter 2 gives a detailed introduction to service control, establishing a common terminology, introducing the operations plan, which constitutes the main input into the service control process, and showing how in daily operations, train service can deviate from the plan. It then presents the set of tools which service controllers can use to correct those deviations. Furthermore, it reviews past research and literature on service control and points out the gaps which this research is intended to fill. Chapter 3 then introduces the London Underground and specifically the Central line. Chapter 4 is dedicated to a detailed description of the objectives, constraints and other decision factors in service control which are derived from the author's observations during an extended visit to the control center of the London Underground Central line, and the end of the section reflects on how these observations can be generalized to other metro lines and systems. The thesis then moves to the specific application examples on the Central line. Chapter 5 introduces the data used for studying service control on the Central line and presents the algorithm which was developed to assemble the different data sources into one coherent dataset and to identify the interventions taken. Chapter 6 then shows the set of measures which were found to capture the impacts of service control decisions on the train service and on passengers. Chapter 7 shows how the data and the impact measures were used to evaluate three common control strategies used on the Central line and also points out how the workload of service controllers can be influenced by scheduling parameters by comparing data from before and after a timetable change on the Central line. Finally, chapter 8 summarizes the findings and suggests possible future research directions.

2 Service control: Background, motivation and research framework

This chapter presents a brief introduction to the service control process, to its function and to the techniques used by controllers to perform their job. Before conducting any detailed analysis of the service control process, one needs to have an understanding of the inputs which go into it. Sections 2.1 and 2.2 describe the plans which form the basis of service control and the types of deviations controllers face in daily operations. That is followed, in section 2.3, by an introduction to the techniques which service controllers use to correct those deviations. Section 2.4 reflects on the relation between service control and planning, and section 2.5 presents some of the previous research and literature on the subject. Finally, section 2.6 presents the research framework which will be developed in detail in the following chapters. Although this thesis focuses on rail services, it will become evident to the reader that there are many parallels between bus and rail service control and that many of the findings can also be applied to bus service control.

2.1 The operations plan

The operations plan is, generally speaking, the set of plans which fully describe the utilization of transit agency resources in daily operations. As Moore (2002) describes it, the operations plan represents the ideal operational procedures, crafted in advance. It is designed to meet the service policy requirements set forth by management while complying with crew work rules, vehicle management and infrastructure capacity and maintenance requirements. Typically the operations plan is built around the service plan. While the service plan is focused on customer services, that is, all train movements in passenger service, the operations plan includes everything needed to produce the service plan, including a working timetable (including train movements which are not in passenger service), a crew schedule, a vehicle assignment plan and a crew roster. The crew schedule assigns driver shifts to cover all parts of the working timetable, and the crew roster then links individual employees to those shifts. The vehicle roster assigns individual trains to vehicle blocks (i.e., sets of linked trips), which are embedded in the working timetable. Unlike the working timetable and crew schedule, which are created in advance by the agency's scheduling staff, vehicle assignments are often created at shorter intervals as a function of rolling stock maintenance needs. The crew roster development process is agency-specific and depends heavily on agreements between the agency and unions in terms of work rules.

The published timetable is the most important component of the operations plan. Given its inputs and its function, it essentially represents a plan of how to achieve the best possible customer service subject to constraints on agency resources. In a sense, it is a promise by the transit agency to the customer. However, the timetable is also very important for asset management and strategic planning on a metro system. Rolling stock procurement, line upgrade and investment plans are always based on a hypothetical timetable representing how the transit agency envisions service in the future, and the future timetable is then an input into a simulation model which is used to identify needs for targeted investments for infrastructure capacity improvement.

Before describing service control in more detail, this section establishes a common terminology for describing schedules. The terms timetable and schedule will be used synonymously throughout this thesis. The operations planning process is described in detail by Ceder (2007). It consists of four steps: network route design, timetable development, vehicle scheduling and crew scheduling. In an existing metro system, network design will not be of great importance, leaving the three other steps to be repeated whenever a new timetable is needed.

The primary scheduling parameter for high-frequency rail services is the service frequency per line section, which is a function of passenger demand, maximum and minimum headway constraints and infrastructure characteristics. It is usually expressed in trains per hour (tph). On lines where different vehicle types are used or train lengths can be altered, train capacity is an additional degree of freedom. If that is held constant, i.e., all trains have the same capacity, the number of trains per hour is directly proportional to the capacity provided, which is based on demand and on passenger loading standards or policies.

The timetable development process builds on the above mentioned factors and service frequency policies. Furthermore, it requires running times, dwell times, layover times and capacity limitations as input. Historically, these variables were primarily derived from models and assumptions since operational data were difficult to collect and often involved significant uncertainties. However, as many transit agencies have installed digital signaling and train control systems in recent years, the availability, accessibility and quality of operational data have greatly improved and the roles of data analysis and modeling in the operations planning process are shifting. Where possible, the analysis of operational data can provide information on many of the variables mentioned above without a direct need for models. The increased availability of data can also assist in the development and calibration of better predictive models to determine the effect of future changes.

The result of the timetable development is a set of individual end-to-end train trips on the line which constitute the timetable for passenger service. The individual trips are then linked together in a chaining process often referred to as "vehicle blocking", where a block is the sequence of revenue and non-revenue trips for an individual train, including train recovery times at terminals to account for variability in running times and dwell times (Ceder, 2007). The last step in the operations planning process is crew scheduling, which aims to define crew duty pieces such that all vehicle movements have a driver assigned to them, at minimal overall crew cost. Aside from the scheduled crews, a certain number of spare operators is usually allocated to crew depots to cover for absences and unforeseen needs. The crew and vehicle rostering processes will not be described in detail here, as they do not have a strong influence on the design of the working timetable and the crew schedule.

For the sake of continuity, although the case of the London Underground (LU) will be introduced in chapter 3, the commonly used LU terminology will be defined here:

- A (physical) train, consisting of a number of permanently coupled cars, is referred to as a *train unit*.
- A vehicle block is identified by a ***train set number***. The set number is assigned to a train unit at its pull-out in the morning and is (ideally) retained by it throughout the day (Allen, 1981). There is no continuity in set numbers overnight. The assignment of train units to set numbers (vehicle rostering) is based on where a train unit needs to be stored in the evening, i.e., a train which is scheduled for maintenance at a certain depot will be assigned to a set number which ends at that depot in the evening.
- A set number is a collection of *train trips*, where a train trip is defined as the trip made by the train from one terminal to another. By convention, the trip number starts with 1 in the morning and increments by 1 at each terminal. Therefore, *every trip scheduled in the working timetable is uniquely identified by a set number/trip number combination.*
- Since all LU trains have only one train operator, the terms *crew*, *driver* and *train operator* are used synonymously in this thesis.

2.2 Deviations from the operations plan

In daily operations, the train service can deviate from the service plan due to unforeseen constraints and events (disruptions). Due to these events, certain parts of the operations plan, such as scheduled vehicle and crew movements, become infeasible. Service control can be described as the work of modifying the operations plan in real-time to deal with the aforementioned unforeseen constraints and disruptions. It is a process which is both proactive and reactive. Some constraints are known before they become immediately relevant for service delivery; for instance, the availability of rolling stock at the beginning of service can be limited due to maintenance requirements or defective trains, or certain track sections might be unavailable due to engineering work. In that case, service controllers can plan ahead and modify the schedule in order to allocate the remaining resources optimally and avoid conflicts. However, more commonly, service control must deal with disruptions and train delays as they occur, causing deviations from the schedule. Section 2.2.1 examines the common types of disruptions and schedule deviations in more detail. From the service controller's point of view, these events immediately affect service quality or make the schedule infeasible, thus requiring controllers to reschedule the service or to dispatch vehicles in real-time in order to maintain the service at the best possible level, subject to the momentary constraints, and to eventually restore it to schedule.

The real-time control of a transit line functions much like a control loop, in which the two elements are the service control center and the operations on the transit line (Horsey, 2009). The service controllers constantly monitor the state of the system and compare indicators for the level of service to the service plan and to other service quality objectives. Deviations will cause controllers to perform a corrective intervention (a list of possible interventions will be given in section 2.3). The choice of intervention is informed by the service controller's knowledge of the system, the momentary constraints, the target state (which is often the timetable) and a projection of the effect of the intervention on the system.

It should be clear that a deviation from the service plan is at the same time a deviation from the operations plan, which, as discussed, also includes train movements not providing service to passengers. A deviation from the service plan is very likely to cause controllers to perform service control interventions (in order to maintain good service to passengers), whereas other deviations from the operations plan (e.g., non-revenue vehicle movements) may not have an impact on the service and may therefore not be of concern to controllers.

2.2.1 Disruptions and schedule deviations

A disruption (or incident) is defined as a single, unforeseen event which causes one (or more) trains to be unable to complete their trips as scheduled. A disruption has a beginning and an end in time and a location at which its effects are felt. This is in contrast to congestion from demand peaks, which can generally cause longer running times and dwell times than the schedule sets out. Uniman (2009) refers to the result of congestion (and small regularly occurring anomalies) as *recurrent unreliability* in contrast to the aforementioned *disruption-related unreliability* and shows that the effects in terms of passenger travel times can be distinctly different.

The consequence of disruptions and (foreseeable) delays from congestion is a train service that deviates from the service plan, which will henceforth be referred to as a *disrupted service*. Although delayed trains are the most common form of disrupted service, there can also be early trains or trains which are completely missing from service (e.g., they become defective and are withdrawn). A good understanding of the different types of disruptions and service deviations is essential for studying service control since the nature of the deviation determines the boundary conditions for service control interventions.

Aside from the management of congestion and disruptions, *routine service control* is the task of monitoring and, as needed, intervening in the service to correct for the inherent variability in running times and dwell times throughout the day. For the remainder of this chapter, the focus will be on disruptions, as they can occur in a variety of ways, unlike the effects of congestion. The following section categorizes them by their cause, effect on the service and duration.

2.2.2 Causes

The cause of a service disruption can be endogenous or exogenous to the transit system. Exogenous causes can be, for example, a passenger operating the emergency alarm, an object on the tracks or weather (on open track sections.) These causes are basically beyond the direct control of the transit agency. The occurrence of disruptions with endogenous causes, on the other hand, can be influenced by the transit agency's maintenance procedures, accountability structures, and employee discipline policies. Examples of endogenous causes are defective trains, infrastructure problems (such as signal failures), staff communication errors or the unavailability of a driver or train unit.

Despite their manifold causes, disruptions manifest themselves on the train service in a finite number of ways:

Non-moving line blockage:

A blocked train is not able to leave the station it is berthed at or proceed beyond a certain point on the line. Since it is generally not possible for trains in rapid transit systems to pass each other, a train which is blocked on the line will not only be delayed, but will also cause following trains to queue behind it and thus be delayed.

Slow-moving line blockage:

This might be caused by a defective train which is only able to move at a reduced speed. Unlike the non-moving blockage, service from affected stations does not come to a standstill, but travel times are increased, not only for the defective train but also (as a function of the service frequency) for those following it.

Single train delay:

This type of delay is experienced by one single train without directly affecting other trains in the depot or on the line. It could, for example, be a train which pulls out of its depot late because no train operator was available at its scheduled pull-out (this is sometimes known as a "hold-in" (Moore, 2002)).

Train blocked in terminal:

A train might not be able to depart a terminal for various reasons. The effect on other trains depends very much on the specific terminal configuration, but in a typical stub-end terminal with more than one reversing track (see Figure 2-1), this will not result in a complete blockage. Instead, the capacity of the terminal is reduced and following trains are only able to reverse at a reduced frequency.

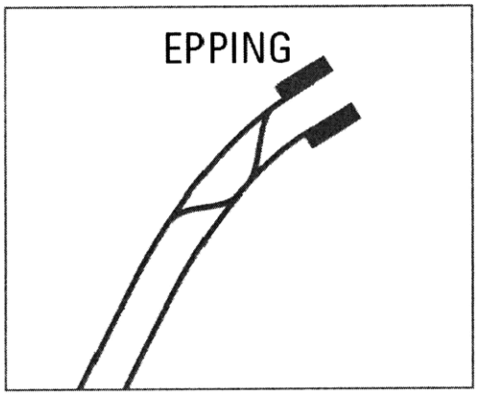


Figure 2-1: Example of a stub-end terminal (Source: London Underground Operations Track Overview)

Reduced infrastructure capacity:

Unlike the aforementioned disruptions, this phenomenon does not necessarily involve a train. Reduced infrastructure capacity can be caused by track, switch or trackside equipment failures, resulting either in temporary slow zones or sections of the line becoming unusable.

2.2.3 Effect on the service

The effect on the service can be defined from either the agency's or the passenger's perspective. This difference is especially important when considering high-frequency metro lines where passengers can be expected to arrive randomly, without referring to a published timetable (Jolliffe & Hutchinson, 1975). In this case, to the passenger, important variables of service quality are platform waiting time (which is a function of expected headways and headway variance) and on-train travel time, which is directly related to service speed, but not necessarily schedule adherence. However, from the agency's point of view, the degree of adherence to the operations plan (and thus to the crew and vehicle schedules) is of critical importance. Disruptions can cause both a deviation from the schedule (i.e., train or driver lateness) and a deviation from service quality standards. The two effects, although correlated, are not necessarily in a direct cause-effect relationship. For instance, it may be possible to maintain regular headways as scheduled on a line section despite trains running late due to an earlier disruption. That would be a deviation from the schedule, but a passenger waiting for a train would experience service at the expected headways, and there would be no deviation in service quality with regard to regular headways.

Fundamentally, a disruption can cause a gap in the service (possibly followed by bunched trains), an incorrect sequence of trains, general lateness or a combination of these effects. The possible impacts on passengers vary depending on the situation:

- A gap caused by a line blockage (whether non-moving or slow-moving) will often be followed by a group of delayed trains with short headways, depending on the duration of the blockage and the scheduled service frequency on the line. The primary impact on passengers is through long waiting times during the gap and crowding in the first train or series of trains after the gap. The first train or few trains after the gap may also experience longer than scheduled running times due to passenger congestion. Once these effects have "passed", passengers essentially experience re-regularized headways and might classify the service as "good". However, from the agency's point of view, there will be a series of trains on the line which are delayed with respect to the schedule, and this may be undesirable to the service controller for various reasons (cf. chapter 4).
- A single train delay, on the other hand, can cause a trip not to be covered, resulting in a gap of two headways between two other trains which are running on time. The delayed train might enter service later, out of sequence. To passengers, the gap has the same effect as

described above (increased waiting times, potential crowding on the first train after the gap), and the first train after the gap may experience increased running times. However, from the point of view of service control, the number of trains which are off schedule is smaller than in the above case (generally only one or two), and the spacing between trains after the gap is generally more consistent with the timetable than after a blockage.

- As mentioned above, a single train delay can cause a train to enter service out of sequence. On lines with branches, delays on one branch do not immediately affect the other branches, but, a different train sequence from that in the timetable can result when trains then merge onto the common trunk section. While this deviation from the schedule may not matter to passengers as long as regular headways are maintained, it may be problematic for the agency. The importance of train sequences and lateness with respect to the timetable will be explored in more depth in section 4.4. General lateness with respect to the timetable can occur if the scheduled running times are inadequate, for example during peak hours with high demand and passenger congestion. It can also result from disruptions where the throughput capacity of the line or a terminal is reduced. Therefore, passengers may experience longer overall travel times as a result, but general lateness need not be accompanied by any gaps in the service or changes in train sequence.

2.2.4 Duration of a disruption

The duration of a disruption is one of its key characteristics, and disruptions are often categorized as "small" or "large" depending on their duration (cf. for example London Underground Ldt. (2008) or Song (1998)). However, there appears to be no consensus as to how these "time categories" should be defined. During personal conversations with the author, service controllers at the London Underground defined disruptions of 5 minutes or less on the high-frequency Central line as "minor" since the resulting delays can usually be compensated for during the layover time at the terminal. Another categorization is offered by Song in his study of a high-frequency heavy rail line, where he classifies a disruption longer than 20 to 30 minutes as "large" based on the fact that this usually requires a partial line suspension and the implementation of an alternative operations plan on the remaining sections of the line (Song, 1998). The author believes that, while the duration of the disruption is certainly an important variable, it only makes sense in conjunction with a characterization of the effect on the service, as discussed in section 2.2.3. Given the multiple degrees of freedom, characterization of incidents as "small" or "large" is certainly not sufficient to judge their impact, and hence they will not be addressed in such terms in this thesis.

2.3 Service control interventions

During a disruption, the main service control tasks are to coordinate the responses to the disruption, maintain an adequate level of service on line sections not directly affected by it and avoid conflicts between trains. This task is known as *disruption management*. Along the same lines, service controllers work to manage the delays incurred through congestion and to ensure that train capacity is allocated effectively in order to meet spikes in demand, which is the task of *congestion management*. Once a disruption has cleared or peak demand has abated, service controllers move into the phase of *service recovery*, also known as *service restoration*, where they work with real-time information and a "toolbox" of changes they can make to the system to achieve an ultimate target state (typically the service plan).

Collectively these changes will be referred to as *service control interventions*. The controller's choice of intervention is driven by a set of objectives and priorities, which are either defined by agency policy or informally by the controllers. An intervention usually applies to an individual train. Together, the set the interventions which are performed in order to restore service to its target state constitute the *service recovery strategy*. It is important to note that when analyzing service recovery, the strategy must be considered as a whole – the purpose of individual interventions may not be clear unless one understands how they fit in to the overall strategy.

Another point to bear in mind is the strong linkage between the train schedule and the crew schedule. Although in the planning process the crew schedule is developed given the train schedule, in daily operations the two constantly need to be coordinated to deal with variations in service patterns and resource availability, in order to ensure that each train trip has a crew. A change to a train's trajectory (e.g., a diversion to a different destination) will always change a driver's trajectory and his or her original schedule.

The following sections present a complete list of train-related and crew-related interventions which service controllers can perform. Crew-related interventions affect crew usage but not train routings. Train-related interventions, on the other hand, affect trains and crews alike. In looking at the list, the reader will realize that, in order to achieve a certain outcome (e.g., put a late train back on schedule), there are many different possible interventions from which a controller can choose.

2.3.1 Train-related interventions

Change of schedule:

This intervention is the most elementary service control intervention and consists of modifying the scheduled running times of trains. In many modern signaling systems, the interstation travel speed can be modified. Otherwise, a controller would need to resort to holding trains at stations.

Holding at stations:

This intervention consists of delaying the departure of a train from a station beyond the end of its normal dwell time. Trains may be held to even out headways, to ensure on-time departures if they are running early or for connections with other trains. During blockages, trains can be held downstream of the blocked train in order to counteract the formation of a gap and they can be held upstream to avoid the formation of a queue.

Early dispatching:

At terminals and other points along the line with scheduled layovers, trains can be dispatched early, usually for evening out headways or to avoid delaying a late train behind the train which is dispatched early. Trains may also be dispatched early to ease congestion at terminals.

Short-turn:

A train which is short-turned is reversed before it reaches its scheduled destination. This can be done either to fill a gap in the opposite direction or to reduce its lateness by shortening its cycle time. Trains can also be short-turned upstream of a blockage in order to prevent a queue from forming. On-board passengers with destinations beyond the short-turn point need to alight and wait for the next train to their destination.

Withdrawal:

A train which is prematurely withdrawn from passenger service into a siding or depot is another form of intervention. This can happen at the terminal before the end of the block or at a siding along the line. The effect on passengers is the same as in a short-turn. This may be done because the train becomes defective, or because the train driver needs to be relieved but no relief driver is available. In addition, withdrawing trains can be used as a disruption management technique, as will be explained later in this section under "Trip Cancellation".

Diversion:

On lines with branches, trains with a destination on one of the branches can be diverted to serve another branch instead. The objective might be to fill a gap on the branch to which the train is being diverted or to withdraw it to a depot located at the end of that branch. However, on branches with different cycle times a diversion can also be performed in order to shorten the cycle time of a late train or to lengthen the cycle time of an early train. The impact on passengers to the original branch is the same as if the train were short-turned. If a diversion takes effect after the train departs a terminal, passengers traveling to stations on the original branch will need to alight and wait for the next train serving their branch.

Extension:

An extended train travels past its scheduled destination to a station located beyond it. There is no negative impact on passengers on board the train as all scheduled stops are served. The degree to which passengers to the "new" destinations benefit depends on when the extension is announced. An extension can be performed to fill a gap, withdraw a train to a depot or to lengthen its cycle time.

Track and platform assignment:

This type of intervention has minimal impacts on passengers and on the train service, but must be mentioned for completeness. In stations or on sections of the line where two or more tracks in the same direction are available, controllers can change the assignment of trains to platforms or tracks stipulated in the operations plan in order to avoid conflicts or bypass disruptions.

Expressing:

Expressing is a technique in which a train in passenger service skips one or more stops which it was scheduled to serve. Passengers on board that train are asked to alight if their destination is one of the skipped stops, and passengers waiting at those stops experience an increase in waiting time as they wait for the next non-express train.

Adding service:

Passenger service can be added in the form of unscheduled train trips which were originally not featured in the service plan. This may be achieved by using a spare train or a train which was withdrawn from another part of the line. Train service can be added either to supplement scheduled services if they prove to provide insufficient capacity to serve passenger demand or to fill gaps caused by disruptions. Trains which need to be moved from one depot to another can be run in passenger service if they do not have any major defects, thus providing a benefit to passengers at little extra cost.

Adding an out-of-service trip:

This intervention is effectively an unplanned deadheading, which refers to any train movement on the line which is not in passenger service. There is a large variety of reasons for such trips, but they are generally related to moving trains or drivers from a point in the line where they are no longer needed to their new site of operation or to the depot.

Trip Cancellation:

Unlike a short-turn, withdrawal or diversion, where a train does not operate part of the trip it was scheduled for, a canceled trip is not operated in its entirety. Service controllers can cancel a trip if there are not sufficient resources (trains and drivers) available to operate it. However, cancellations may also serve as a disruption management or service restoration technique. Controllers may temporarily remove a train from service and cancel one or more round trips, then insert it back into service at a later point. This can serve as a technique for putting late trains back on schedule, or it can be done to reduce congestion and facilitate real-time train management during disruptions. On high-frequency line sections where passengers arrive randomly, they will most likely not be aware of the missing train, but they will experience longer waiting times.

Renumbering:

When a service controller renumbers a train, the train unit's set number / trip number combination changes. This means that the train is now associated with a different vehicle block, and thus a different set of scheduled trips to be operated. Renumbering trains is typically associated with one of the aforementioned changes in routing (a diversion, an extension or a short-turn). In the event that a train is renumbered to a train/trip number with the same original destination, the controller's intent most likely is to assign it to a scheduled trip for which it is on time. It is important to note that renumbering trains in such a case without any other intervention only affects the lateness of a train unit with respect to the timetable; the driver's lateness does not change.

Train priority at junctions:

At any junction where branches merge, service control must establish a train priority scheme. The planned form of this scheme may be embedded in the automatic signaling system and can be as simple as "first come, first served". Trains can also be held at junctions, either to their scheduled sequence or as an intervention to establish a certain service pattern downstream. A more in-depth analysis of junction capacity can be found in Vescovacci (2003).

Train priority at terminals:

Train priority decisions at terminals are very similar to those at junctions. The standard principle of terminal operation is "first in, first out", but it can be modified to "first in, second out" (or n^{th} out, depending on the terminal configuration), thus modifying the sequence of trains and their layover times.

2.3.2 Crew-related interventions

Substituting a spare driver:

Replacing a rostered driver with a spare driver is one form of crew-related intervention. For example, such an intervention can be used to relieve a late-running driver who needs to step off but has not yet completed his/her driving assignment.

Dropping back:

Dropping back in a planned form is a crew scheduling technique which allows the train layover time at a terminal to be shortened without compromising the driver layover time. Specifically, every driver steps off his/her train at the terminal, takes a break and then departs on the following train. Dropping back can also be imposed by service controllers as an ad-hoc intervention measure to speed up the reversal process with the help of a spare driver.

Jumping up:

On a line where drivers are planned to drop back, the intervention called "jumping up" refers to a change back to every driver departing with the same train that he or she arrived on, i.e., a discontinuation of the policy of dropping back.

Switching drivers:

If trains need to be resequenced, this can easily be done by renumbering them, but more importantly the drivers may need to be resequenced – one driver needs to move to a train ahead of his/her train and the other needs to move to a train behind. If the two drivers are at a terminal at the same time or meet each other on the way, a cross-platform driver change is possible. This might be problematic since it compromises the layover time of one of the drivers. To avoid this, controllers can work with a spare driver, for example as follows. The drivers of trains 1 and 2, both for the same destination and where train 2 is following train 1, need to be switched. The two drivers shall be called driver 10 (on train 1) and driver 20 (on train 2). A spare driver steps onto train 2 a few stations short of the terminal. Driver 20 waits at that station and steps onto train 1 after train 1 has reversed. Driver 10 then steps off and waits for train 2, where he or she replaces the spare driver. Using this technique, driver 10 had a layover at the terminal, whereas driver 20 had a layover at the station where the spare driver stepped on.`

Stock and crew:

Stock and crew functions similarly to switching drivers; it is a technique for moving late trains (and late drivers) forward in the timetable without short-turning the train. However, it only works if there is a spare train and driver available at the terminal. While the late train is traveling towards the terminal in passenger service, the spare driver with the spare train departs the terminal at the time the late train was scheduled. The late train and the spare train meet at a station and the two drivers switch trains across the platform. The spare train then receives the number of the scheduled train and continues on its path whereas the train heading towards the terminal becomes the new spare train and is stabled by the spare driver. As with the crew changeover described above, a problem might arise because the original driver does not get a layover at the terminal.

2.4 The relation between service control and planning

In Figure 1-1, a flow of decisions from the management level to the planning and on to the operations level is shown, but at the same time, a flow of information is needed from the operational level back to planning and management. What these connections circumscribe is in fact a very strong bi-directional link between service control and operations planning. This interaction can best be understood by revisiting the operations planning process. The inputs used for scheduling are usually data on running times and dwell times, and assumptions or models of scheduling variables for which insufficient or no data are available, such as layover times, throughput capacity and junction and terminal capacity (Allen, 1981).

The validity of the assumptions or models can only be assessed with the help of data gathered during operations. Planners can also use operational data to identify unknown bottlenecks, for example by analyzing the origin of delays, the variance of certain variables in daily operations or levels of train impedance. However, it is very important to recognize that the system being analyzed is not a system of autonomous actors (i.e., trains, drivers and passengers), but rather a system which is controlled by a central, intelligent entity (service control) which has an appreciation of the state of the entire network and can influence individual actors within its control (trains and drivers) to change that state. That means that the data being observed (e.g., running times, dwell times) may tell an incomplete story without being linked to service control interventions, and that this, in turn, can distort the models used for determining scheduling variables. A hypothetical, illustrative example is a bottleneck analysis for investment planning. Suppose there is a junction where two branches of a rail line merge into a trunk section. During peak hours, the throughput capacity of the junction is insufficient for the number of trains scheduled to pass through that junction. Service controllers have developed a strategy to deal with this constraint by slowing down trains from one of the branches (perhaps the branch with lower demand). An analyst looking at the running times and dwell times on the branches without consulting service control may incorrectly conclude that the bottleneck is located on the branch instead of recognizing that it is a junction capacity issue.

The manner in which service control manages disruptions also provides valuable input to service planning. As already discussed, an operations plan represents the ideal operations on a line, but some disruptions on high-frequency metro lines are inevitable. By understanding how disruption management strategies are implemented as a function of its parameters (e.g., train frequency, scheduled recovery time), by investigating where spare network capacity such as reversing tracks or crossovers is used and what resources (e.g., spare trains and drivers) are required for disruption management, planners can try to accommodate service control better in the planning process. For instance, they can restrict the usage of these resources in the operations plan or make them easy to reallocate in the event of a disruption.

Furthermore, understanding service control can also help an analyst understand which parts of a operations plan are most vulnerable to disruptions. For instance, one can analyze which scheduled trips are most frequently changed and which stations are most frequently affected. This may lead to some surprises, as the effects of disruptions may be felt much less in the line sections where the disruptions occur than in the line sections where trains and personnel are removed by service controllers in order to deal with the disruption.

In summary, one can say that although the operations plan may represent the optimal situation, planners need to understand how that plan is affected by disruption management and how well assets and resources are utilized, as this directly affects future investments in capacity improvements.

2.5 Literature review

This section explores previous research in service control. As the reconstruction of operations on a rapid transit line is one focus of this thesis, section 2.5.1 reviews prior research on this subject. Afterwards, literature on service control strategies will be reviewed in section 2.5.2, and section 2.5.3 shows research on other aspects of service control not directly related to intervention strategies.

2.5.1 Reconstruction and analysis of operations

Wile (2003) explores how transit agencies can use automatically collected operational data for a variety of applications and establishes a range of agency functional needs which can benefit from such data. He makes the important point that the data needs of an agency are typically defined as a function of the specific application for which the automatic data collection system is being purchased, often with little or no consideration of possible benefits outside that field, and calls for agencies to take a holistic approach to the procurement of systems which generate data.

Dixon (2006) presents a tool to automatically reconstruct daily operations on heavy rail lines of the Massachusetts Bay Transportation Authority (MBTA) from data produced by the Operations Control System (OCS). The primary focus is on the extraction of scheduling variables, such as dwell time and running time distributions and schedule adherence values. Based on his experiences with the reconstruction of operations, he points out the value which can be gained from data available to transit agencies at little or no additional cost.

2.5.2 Research on service control strategies

Although research on service control and dispatching dates back to 1972, much of the early work assumed that the dispatcher had only limited or no real-time information on the position of vehicles along the line. Control strategies developed by those researchers generally examine dispatching or holding strategies at predetermined control points of the line (e.g., terminals or timepoints), taking as input the distance between a vehicle and its immediate neighbors. The most notable early research into this topic was by Osuna and Newell (1972) and Koffman (1978). Osuna and Newell described the variability in running times of vehicles on a transit route as independent, identically distributed random variables with a known distribution. Based on that, they presented the conflict between a transit agency's goal of minimum route time and maximum service reliability (which generally improves with increased running times) and developed a dynamic programming approach to determine optimal holding and dispatching strategies at a terminal. Early work in real-time dispatching by Koffman built upon a simulation model of a bus route; it was employed to compare real-time holding and skip-stop strategies with the benefits of implementing bus signal priority and reducing dispatch uncertainty.

Since these two initial research efforts, much work has been done on the holding problem without real-time line information, minimizing unreliability (and thus passenger delays) by optimizing the location of control points and slack times allocated in the timetable with the help of simulation, analytical and dynamic programming models. Practically all these studies, which include Abkowitz (1986) and Liu & Wirasinghe (2001), were performed on bus routes. Furthermore, an important characteristic of these studies was that they dealt with variability in running times under "normal" operating conditions, without considering service recovery from long delays (blockages) which cannot rely solely on holding or early dispatching. This task was referred to as routine service control earlier in this text.

The emergence of real-time information systems providing data from all vehicles serving a transit line has enabled recent research to take a broader approach to the holding and dispatching problem. For instance, Adamski & Turnau (1998) developed a bus line simulation model which they employed to test a set of dispatching strategies in which all stations were controlled. They alternatively monitored headway adherence and schedule adherence and used deviations from these variables as feedback in the control loop. The objective was to minimize those deviations.

Three of the first pieces of research on expressing and short-turning trains were done on the MBTA Green line by Macchi (1989), Deckoff (1990) and Soeldner (1993). They did not have real-time information on vehicle locations, so the strategies were focused on pre-determined, short sections of the line. Macchi's research on expressing established which passenger groups were affected by that strategy and then determined how to quantify the impacts on passengers (which, in this case, occur through changes in platform wait times). He then developed guidelines for expressing to minimize overall passenger impact. As no real-time vehicle location information was available for that research, he based his decision criteria on the preceding and following headways, and found that expressing was justifiable only when the following headway was short. Deckoff used the same subway line to study short-turning. Again, the available information was limited to the preceding and following headways and the short-turning decision for any given train was made at a control point near the line's end. Deckoff established the different groups of passengers affected by the short-turn, which also included "dumped" passengers, and developed a model to quantify the impacts. Based on results obtained with input data generated by a Monte Carlo simulation, Deckoff found two cases where short-turning one train provided benefits: either if two consecutive trains had very short headways or to prevent a large headway in the other direction. Soeldner connected the models of Macchi and Deckoff and compared the benefits of short-turning with the benefits of expressing trains on the MBTA's Green line to determine which of the strategies was favorable under which circumstances. Furthermore, Soeldner collected data on actual control decisions taken by local dispatchers and evaluated them in terms of the passenger impacts. The data was taken from manual logs from two days of Green line operations. The dispatcher's decisions were classified as “good” or “bad” decisions by comparison with model predictions. He found 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. While the works of Macchi, Deckoff and Soeldner are important first steps, their results are specific to the MBTA Green line and cannot be directly applied to other metro lines.

Eberlein (1995) studied the use of deadheading, expressing and holding strategies on a non-branching rail line. She formulated two nonlinear integer programming models, a simplified and a generalized one, which used the minimization of passenger wait time as the objective function. She concluded that the effectiveness of the control strategies was largely a function of the demand pattern, which is not surprising given the objective function, and that deadheading and expressing provided similar benefits. Furthermore, her modeling effort also found that the maximum benefit could be achieved if the strategies of holding and expressing were combined. Eberlein focused on routine control and did not include disruption management.

O’Dell (1997) developed a real-time decision support tool for service recovery on a rail line after a disruption. She approached the problem in a fashion similar to that of Eberlein, but derived a linear objective function in order to obtain closed-form analytical results. Furthermore, she included vehicle capacity, which had not been incorporated by Eberlein. The linear programming model included holding and short-turning as feasible strategies for responding to a disruption, and the overall objective function was to minimize passenger waiting time. Three holding strategies, a short-turning strategy and a do-nothing strategy were compared in a case study of the MBTA's Red line, and she found that holding a limited number of trains to even out headways behind and in front of a blockage had clear benefits over a do-nothing strategy or a hold-all strategy. Furthermore, she concluded that the effectiveness of short-turning was greatest when only a small number of stations were skipped.

Shen (2000) followed a similar path in model development as O'Dell, but generalized it to allow for any train to be held anywhere and combined this with expressing and short-turning strategies. The model was a deterministic, mixed-integer program that captured dwell-times. Again, the objective function was to minimize passenger delay, but passenger delays both at stations and on trains were considered and the trade-off was examined. One of the most important simplifications of the model was that the duration of the incident was known ahead of time. This simplification is critical given that the sensitivity analysis conducted by Shen revealed that the benefits of certain control strategies were sensitive to the duration of the disruption. Similarly to O'Dell, Shen found that the greatest benefits were achieved by actively holding a finite number of trains in front of and behind the blockage, combined with short-turning. The model was used and validated on another heavy rail metro system, Tren Urbano in San Juan, Puerto Rico, by Ortiz (2000), who used it to derive optimal control strategies and compare them to expected control strategies under different incentive/penalty clauses imposed on the private operating company through the private-public partnership contract.

Song (1998) analyzed the response to major disruptions, defined as blockages of 20 minutes or more, on the MBTA Red line. Song studied possible strategies such as one-track operations and shuttle services which could form the basis of a major disruption response system. Song also studied the terminal dispatching problem with real-time information about the location of all other vehicles and developed a heuristic model for holding or short-turning trains if a train would arrive at the terminal too late to reverse and depart on time. The objective of the model was to minimize passenger wait time and overcrowding in vehicles. The finding was that the effectiveness of a formalized strategy increased as the lateness of the trains became more severe.

Puong (2001) built on the models of O'Dell and Shen, developing a deterministic holding model with real-time information. However, Puong's model included terminal capacity and the delays incurred by passengers left behind by fully loaded trains. Again, the objective was to minimize in-train and on-platform delays to passengers, and Puong assumed the duration of the disruption to be known. A particular focus of his research was the effect of finite train capacity when that constraint became binding, in which case the model showed more benefits to a holding pattern with uneven rather than even headway sequences as suggested by Eberlein. However, the holding patterns proposed by Puong's model might be too complex for implementation in real systems, as they sometimes involve trains being held at multiple stations.

2.5.3 Other research on service control

Rahbee (2001) examined the analysis of operational problems on rail transit lines. His research was not strictly limited to service control, but rather took a comprehensive view towards improving service quality on rail rapid transit lines. He argued that a common research approach was to make assumptions on how the transit line operates without making sure that those assumptions were in fact correct. To break with that practice, he proposed a framework for how to identify strategies for service quality improvement; the study of service control is an integral part of that framework. In a first step, Rahbee demonstrated the usefulness of carefully analyzing real-time observational and system data with the help of space-time plots and other representations of system behavior. After a careful identification of patterns, Rahbee explained how analysts need to discuss their findings with the personnel making the decisions, including service controllers and operations planners alike. This process eventually yields a detailed understanding of the operations on a line along with information on individual operational problems which need to be addressed.

With respect to service management, Rahbee stated three main investigation goals which an analyst can pursue:

- To document objectives and constraints as they exist in the management of a particular line.
- To investigate to what degree service control decisions are being made according to the agency's objectives and guidelines.
- To investigate whether the agency's objectives and guidelines regarding service control are properly thought out.

An extensive piece of work building strongly on the practitioner's point of view is a manual written for the RATP (Froloff, Rizzi, & Saporito, 1989), the transit agency of the Paris metropolitan area. It focuses on bus service and was originally written in an effort to identify and systematically categorize objectives, constraints and techniques for bus service control in preparation for the design of a simulation system for controller training. The authors then present a study of three classical problems – gaps, bunching and general lateness. The focus on bus service makes their work not directly applicable to rail, but many of the topics addressed (such as the controller's decision environment) have never been published elsewhere and they form a solid foundation for studying service control on rail systems.

Barker (2002) investigated the allocation of responsibilities for service management on bus services of the CTA (Chicago Transit Authority). Furthermore, he considered the relationship between service management and information and communication practices. He found that there were benefits to a distribution of responsibilities between service controllers and dispatchers in the field and he looked at potential benefits of providing field dispatchers with handheld computers providing real-time information. Barker segments service control tasks into "predictable" (routine) and "unpredictable" ones depending on the nature of the disruption or delay, and argued that "predictable" routine dispatching is better performed in the field, while there are clear advantages to the centralized management of "unpredictable" control tasks in a control center.

Finally, Moore (2002) did research on the CTA's policies for the management of disruptions and for service restoration, mainly on their bus routes. She established a framework describing the variables which need to be known to service controllers when managing disruptions and developed a set of elementary measures for the impact of disruptions. She found that pre-planning and formalizing responses to certain types of disruptions could greatly improve an agency's ability to deal with problems as they occur.

Both Barker and Moore dealt with a case in which controllers worked with a set of different lines and might not immediately be familiar with the specifics of the line on which the disruption had occurred.

2.5.4 Appraisal of previous research

Overall, it must be said that little research has been done on service control considering all the "tools" a service controller has and taking into account real-time data. This may not be surprising given the complexity of the field. The research presented in the previous section shows that the study of service control originated in the study of local dispatching methods without real-time information and has moved towards considering the state of the entire line as real-time information systems have become available. However, all research which has worked with mathematical models of a line to find optimal control strategies has made strong assumptions about operations on a rapid transit line, which may limit its applicability in real-world situations. Research into other aspects of service control, such as by Barker and Moore, has focused on bus operations and has also tended to be agency-specific. The major exception is Rahbee.

The most important issues with previous research are identified below. Work on these aspects of service control could help move the entire field towards more realistic and implementable models.

1. So far, only Froloff, Rizzi and Saporito have attempted to describe the decision environment of service controllers, i.e., the complete description of variables, objectives and constraints which affect service controller decisions. The work of Froloff, Rizzi and Saporito is focused on bus service and lacks several aspects which are important in high-frequency rail operations. All other research so far has assumed that the only objectives in service control are related to passenger travel time or passenger wait time, and has largely focused on headway regularity, ignoring other important drivers of service control including schedule adherence issues. This is problematic, as it does not consider crew management, which is an integral part of service control. Depending on the situation and the line under consideration, it can enter the controller's decision as a constraint or as an objective. As will be described in more detail in chapter 4, there are several important aspects of crew management. Drivers have a maximum driving time which they cannot exceed, and any driver stepping off a train at a relief point needs to be met by a relief driver. Furthermore, some drivers have a hard time constraint on when they must step off in order to meet other obligations. The result is that in a system with drivers (i.e., in any system which is not fully automated), driver lateness is a large concern, and since drivers are tied to vehicles and both are assigned to schedules, this is the point where lateness enters the picture. Depending on the operational characteristics of the line, crew lateness may even be more important than train lateness since trains are interchangeable and the requirements of rolling stock management are generally not as binding as those of crew management.

   To illustrate the importance of crew management, one can think about how this affects the service control strategies investigated by several of the researchers cited in this chapter. Holding trains often means transitioning them into delayed state. In the case of trains behind a blockage, this exacerbates the delay they are already suffering through the blockage. Trains in front of a blockage, supposing that they would have been on schedule before the disruption, are added to the pool of late trains by being held. Thus, after the blockage has cleared, controllers need to deal with a set of *late crews*. On the other hand, short-turning trains causes them to be out of sequence, and thus the drivers are out of sequence too. A driver on a short-turned train may either pass the crew relief point too late or too early, in which case no relief driver may be available. As previously mentioned, crew management is certainly not the only important additional factor to consider in service management, and the decision environment of a service controller covers a variety of other areas.

2. Many of the research efforts so far dealt with simple line layouts and a restricted set of control interventions. Yet, the more complex a line and its service patterns, the more complex the controller's "tool box" becomes, and more types of interventions are available to controllers – for instance, extending trains beyond their destination, multiple points where trains can be short-turned or the possibility of swapping trains between branches with different cycle times to alter running times.
3. O'Dell, Shen and Puong compared the performance of their models with a do-nothing scenario to demonstrate the effectiveness of their model formulations. However, these models are only applicable on lines which have centralized service control, and it is unrealistic to assume that service controllers would "do nothing" (or if that were the case, they would probably have good reasons to do so), since that defeats the purpose of their job. While the time savings with regard to a do-nothing scenario may be useful for comparison across models, the *real* question is whether any of the models are able to provide an improvement over service control on the line as it exists in practice. Doing so obviously requires an analysis of operational data and a reconstruction of the service on a daily basis. Coupled with the same demand data used by the service control models, a comparison could then be drawn between them. However, it must be said that such a comparison would make little sense unless the various other decision factors which drive service control (as described above) are also included in the models.
4. Last but not least, summarizing the entire passenger experience in terms of travel (or waiting) time might be too crude. More research is needed into how passengers experience and value things like additional on-platform waiting time, waiting time on a held train, being forced to alight early due to a short-turn etc. An overview of past research efforts in the UK is provided by Wardman (2001). If total travel time is being measured, attention should also be paid to the effect of passenger volume – for instance, 100 lost passenger minutes can refer to 100 passengers being delayed by 1 minute each or 10 passengers being delayed by 10 minutes each. Presumably the irritation caused to the individual passenger differs between these two cases, but the tradeoff is not yet well understood. A possible research direction would also be to investigate what level of irritation different types of service control interventions cause to current and potential customers. As Deckoff (1990) puts it:

   "People, in general, remember service problems on a transit system in terms of the worst experience they have had on it. For many riders on the [MBTA] Green line, being on a short-turned train [...] is the most inconvenient experience of the system they have undergone.

Rarely does the relatively unobtrusive experience of a long wait engender such frustration." This is, of course, the view of an individual researcher and can certainly be debated, but in general terms, he points towards a very important aspect of measuring passenger impacts – the context of how a passenger encounters a delay appears to be an important determinant of customer satisfaction. In the author's experience, controllers tend to have a "gut feeling" about the level of irritation caused by different types of interventions, as will be discussed later in this thesis.

2.6 Research framework

The shortfalls of previous research in service control show that there is a need for an approach which integrates multiple perspectives and data sources in order to gain a better understanding of service control. This section presents the research framework which is proposed in this thesis and which provides a basis for remedying some of the problems highlighted in section 2.5.4.

In order to have a thorough understanding of service control (and therefore of operations) on a line, an analyst needs to consider the following elements: service control decisions (i.e., interventions), the reasons for which the decisions were taken, the level of service on the affected line sections and passenger travel times. These components and their relationships are shown in Figure 2-2. Furthermore, the two factors which lead to a service control intervention are included in Figure 2-2, disruptions and general operating and demand conditions. If data on those factors are available, they can be included in the analysis.

The interventions essentially represent deviations from the operations plan imposed by controllers, and measuring the level of service allows the analyst to quantify the impact of those service control strategies on the service. However, this is only an intermediate step since ultimately the concern should be with the impact on passengers. Once the level of service has been quantified, the link can be made to passenger travel times, either through calculation or measurement. This "cause chain" from interventions to the level of service to passengers, is marked with a solid arrow in Figure 2-2. Once it has been established, the analyst can directly infer the impact of service control strategies on passengers, as shown in Figure 2-2 with a dashed arrow.

Nonetheless, the picture is not complete without an understanding of the reasons for the decisions, which requires a more qualitative approach. As previously mentioned, the decision to perform a service control interventions is informed by the controller's knowledge of operating and demand conditions as well as disruptions to the service. It is important to note that a disruption is not a direct cause for an intervention. Instead, it is the root cause of a deviation of the service from internal service quality metrics or it potentially causes a conflict between trains, which then in turn leads a controller to decide whether and how to perform a service control intervention. In that sense, since a controller may feel that a small disruption does not warrant an intervention, the decision not to make any changes to the service in the face of a disruption is also a possible output of service control. This research focuses mostly on the reasons for intervention decisions, the identification of service control interventions and the assessment of their impact. Disruptions and the characterization of operating and demand conditions are only considered as necessary to understand the root causes for interventions.

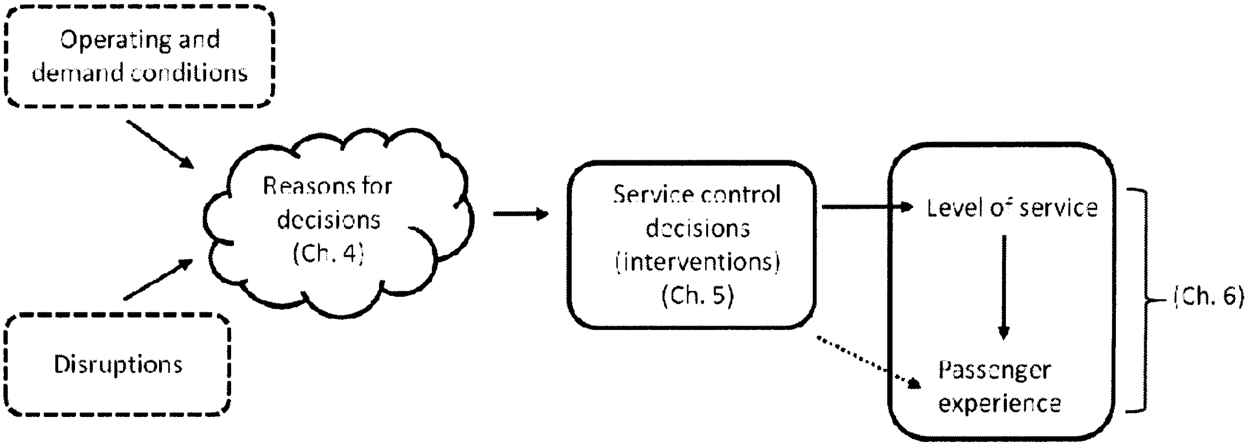


Figure 2-2: The components of service characterization on a rapid transit line

As mentioned in points 1 and 2 of section 2.5.4, neither the complete set of tools available to a service controller nor the full decision environment have been considered in previous research. While the list of interventions was presented in section 2.3, chapter 4 aims to present a complete picture of the factors which affect a service controller's decision. Point 3 of section 2.5.4 explained that any model aimed at improving service control must be compared to the current practice of service control, which in turn requires a reconstruction of operations on a line. This is the focus of chapter 5, which presents an algorithm for reconstructing operations on a line and for identifying service control interventions. Chapter 6 presents measures for assessing the impact of service control interventions. Although the issue raised in point 4 of section 2.5.4 is very broad and would require significant additional research efforts, chapter 6 shows how the use of data instead of models for quantifying the impacts of service control interventions can provide meaningful and transparent results. Finally, chapter 7 integrates the components of this framework in four cases to illustrate how that can be done and to draw some case-specific and some general conclusions.

As explained in this chapter, the research for this thesis was based on the London Underground Central line, which will be introduced in chapter 3. The algorithm and measures presented in chapters 5 and 6 are primarily relevant to the Central line and to the data which were available for it, but the process which is described is applicable to the study of any metro line, although the details of the implementation would vary. The reader who is not directly affiliated with the London Underground should therefore see the application to the Central line as an illustration of how the general idea behind the thesis can be put into action in a practical setting. The author is aware that the data available for the Central line are, compared to other metro systems, of high quality, and that several of the methods and measures which will be introduced depend on the availability of such data. Thus, an analyst wishing to apply this methodology to another metro line would need to adapt it in order to account for the characteristics of the data available in that particular case. Nevertheless, in the long run, the availability of automatically generated operations and passenger data will doubtless increase in virtually all metro systems, making it easier to conduct a study like the one presented here.

3 The London Underground Central line

This chapter gives a brief introduction to the London Underground Central line, on which this research is based.

3.1 TfL and the London Underground

The public transportation network of greater London is under the responsibility of Transport for London (TfL), a local government body created in 2000 as part of the Greater London Authority (GLA). TfL has the oversight and planning authority over all modes of public transport and the urban arterial street network within London, with the exception of most National Rail suburban services. TfL manages approximately 700 bus routes, the Docklands Light Rail (DLR), a small tram network (Tramlink), 11 heavy rail lines which constitute the London Underground network and a small portion of the National Rail network which is currently being integrated into the TfL family as the London Overground (TfL, 2009). Operations are generally contracted out to private companies, with the exception of the London Underground, which is operated by London Underground Ltd. (LUL), a subsidiary of TfL. The London Underground network is one of the largest metro systems in the world with a total length of 402 line-km, approximately half of which is below ground. It is comprised of 207 stations on 11 lines, with many of the lines having multiple branches. The official network map with all Underground, Overground and DLR lines is shown in Figure 3-1.

One of the most important aspects of London's public transportation network is its integrated fare system, which offers passengers fare products valid across all modes of public transportation managed by TfL. One of the largest successes of TfL in recent years has been the introduction of a smart card payment system branded "Oyster card". Introduced in 2003, it is currently used as a fare medium for over 70% of all trips made with TfL.

In 2007, 27.8 million "journey stages" were made on an average weekday in London. A journey stage is defined by TfL as a component of a trip between interchanges, i.e., it is made by a single mode of transportation[2]. Of the aforementioned 27.8 million journey stages, 10.8 million (38.8%) were made by public transportation, which further breaks down into 2.9 million trips by underground, 5.4 million trips by bus and tram, 2.3 million trips by rail and 0.2 million trips by DLR. Of the remaining 17 million journey stages, 10.5 million (37.8%) were made by private motorized

---

[2] There are, however, differences in the way journey stages are counted. On the bus network, every time a passenger boards a different bus (i.e. transfers), it is counted as a separate journey stage. This is not the case on the London Underground, where interchanges within the system are ignored. Thus, on the Underground, a journey stage is the same as one passenger trip from the moment that person enters the Underground system until he or she leaves it.

modes (car, motorcycle), 6.2 million (22.3%) on foot and by bicycle and 0.3 million (1.1%) by taxi. Public transportation ridership has seen a steady increase in the past 10 years. In 1997, only 7.4 million journey stages were made by public transportation on an average weekday, which accounted for approximately 30.7% of total of journey stages[3]. The bulk of the growth has been shouldered by the bus network, but since 1997 the Underground has also seen an increase in ridership of approximately 30%, leading to increased levels of congestion, especially during the peak hours.

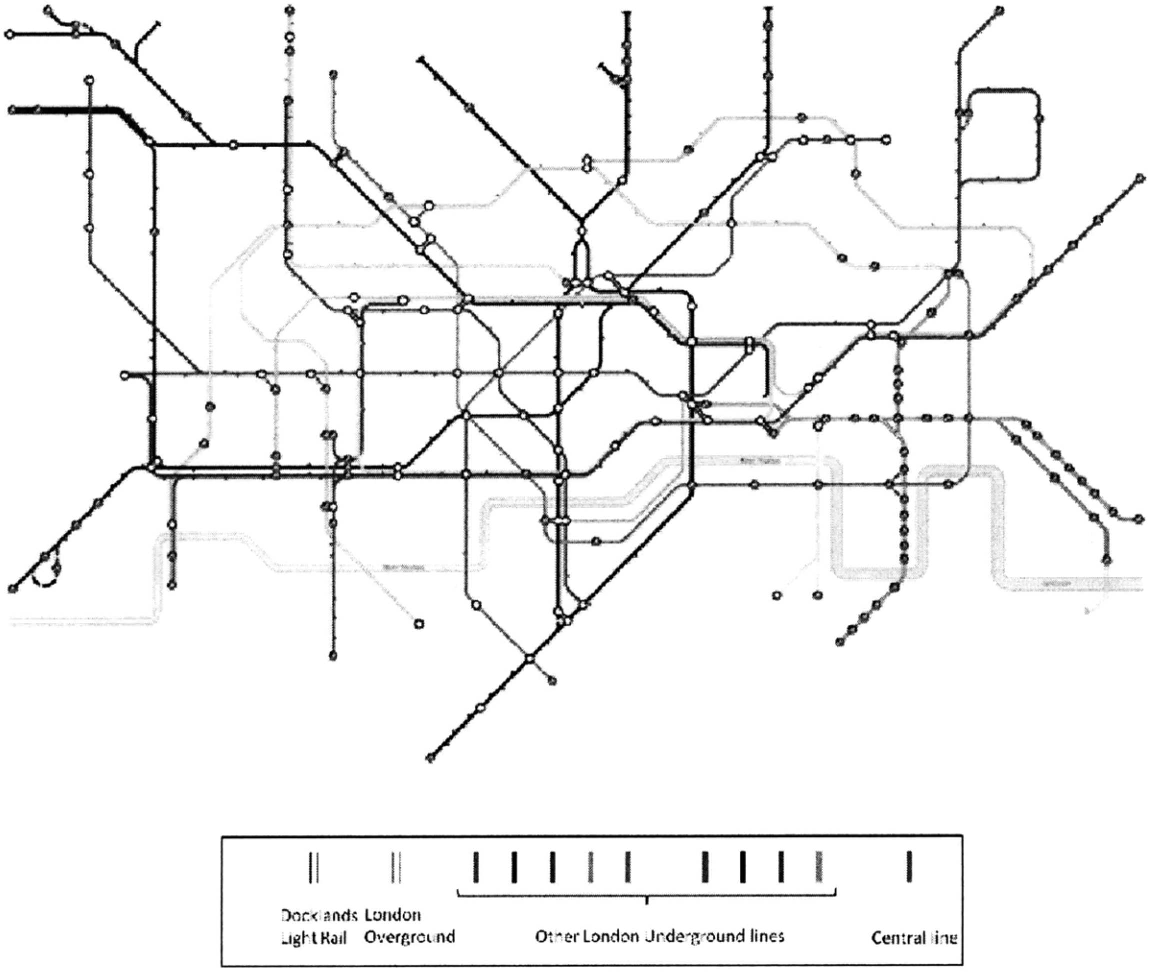


Figure 3-1: The TfL rail network (Source: TfL, 2009)

---

[3] In 1997, 24.1 million journey stages were made in London. Aside from public transportation, they were divided up as follows: Car and motorcycle – 11 million (45.6%), walking and bicycle – 5.6 million (23.2%), taxi – 0.2 million (0.8%). Due to rounding errors, these individual modes add up to slightly more than 24.1 million journey stages.

3.1.1 The Central line

At 74 line-km, the Central line is the longest line in the London Underground system. Initially opened in 1900, it runs east-west underneath London's central business district and branches out to serve suburbs in east and west London. In the network map shown in Figure 3-1, the Central line is indicated in red. The Central line serves a total of 49 stations, of which 22 are located in the central London trunk portion between Leytonstone and North Acton. On the western end, the line splits into two branches, terminating at Ealing Broadway and West Ruislip. On the eastern end, one of the branches terminates in Epping, while the other constitutes the "Hainault Loop". While the trunk portion is predominantly a deep tube line, many of the branch stations are located above ground.

Figure 3-2 shows a line map. The line's train depots with maintenance facilities are located at West Ruislip and Hainault (indicated on the map with "M"), with additional sidings at White City, Woodford and Loughton (indicated on the map with "S"). Aside from the depots, sidings and terminals there are reversing tracks at Northolt, North Acton, Marble Arch, Liverpool Street, Leytonstone, Newbury Park and Debden, which are marked on the map with purple dots. The Central line has five crew depots where crew reliefs take place, which are marked on the map with "C". Four of the crew depots (West Ruislip, Hainault, Loughton, White City) are located at train depots or sidings with train storage and reversing facilities, while one (Leytonstone) is located along the line.

In the late 1990's, the Central Line was the first line on the London Underground to be upgraded to an electronic fixed-block signaling system with automatic train operation and train protection (ATO/ATP). As part of the upgrade, service control was centralized at the Wood Lane control center, which was outfitted with a new operations control system (OCS). The most important components of the OCS are the central signaling servers and a cluster of six service controller workstations located in the control room, along with two large overview screens which show the train service on the Central line in real-time. From the OCS workstations, service controllers can monitor the ATO system and, if there is a need to make a change manually, access most of the signals (some are set up for automatic operation only) and all of the switches on the line and operate them by remote control. The communications equipment in the control system is comprised of the train radio, which allows one-to-one communication with train drivers, and the line telephone for communicating with crew and depot managers. The process of service control is described in more detail in section 4.2, and section 5.1 discusses the data which is available from the OCS.

New ATO-compatible rolling stock was introduced between 1993 and 1995 and the signaling system was commissioned between 1999 and 2001. The Central line has a total of 85 8-car trains with a maximum capacity of 1047 passengers per train (calculated assuming 5 standees per m^2) (London Underground Ltd., 2002). In normal operations, trains run in ATO mode, where the driver only operates the doors and issues a departure command to the train, which drives automatically between stations. In case of a failure of the ATO system and for shunting moves (short moves of empty trains within stations or between stations and depots), the trains can be operated manually, either with ATP ("coded manual operation") or without ATP and speed limited to 16 km/h ("restricted manual operation"). The line is centrally controlled from the Wood Lane control center, where the control room and the central signaling servers are located. Controllers do not have the ability to control all signals on the line manually; this is only possible in the so-called controlled areas, which are around depots, sidings, junctions, reversing tracks and crossovers. Everywhere else along the line, signals are controlled automatically. An interesting feature is that every train has a countdown clock in its cab. As it berths in a station, the clock counts down the seconds until the train's scheduled departure, but it serves only an advisory role and can be disregarded by the train operator. If the train is behind schedule, the countdown clock will simply count down 20 seconds from the arrival of the train without indicating that the train is late, as this is considered the minimum dwell time.

The central London trunk section of the line has connecting stations with all other London Underground lines, and one of the stations, Mile End, shares its platforms with the District line.

3.1.2 Passenger flows

In terms of passenger flows, the Central line has two distinctive markets. As it extends into the suburbs in east and west London, it provides service for suburban commuters to and from the central business district, but due to its alignment within central London, there is also high demand for short-distance travel within the trunk portion of the line, both from passengers making short trips in central London and from passengers transferring to/from other lines. The London Underground has pre-set time periods for operations on its lines, including the Central line, as shown in Table 3-1. These periods and terminology are adopted for the remainder of this thesis, unless otherwise noted.

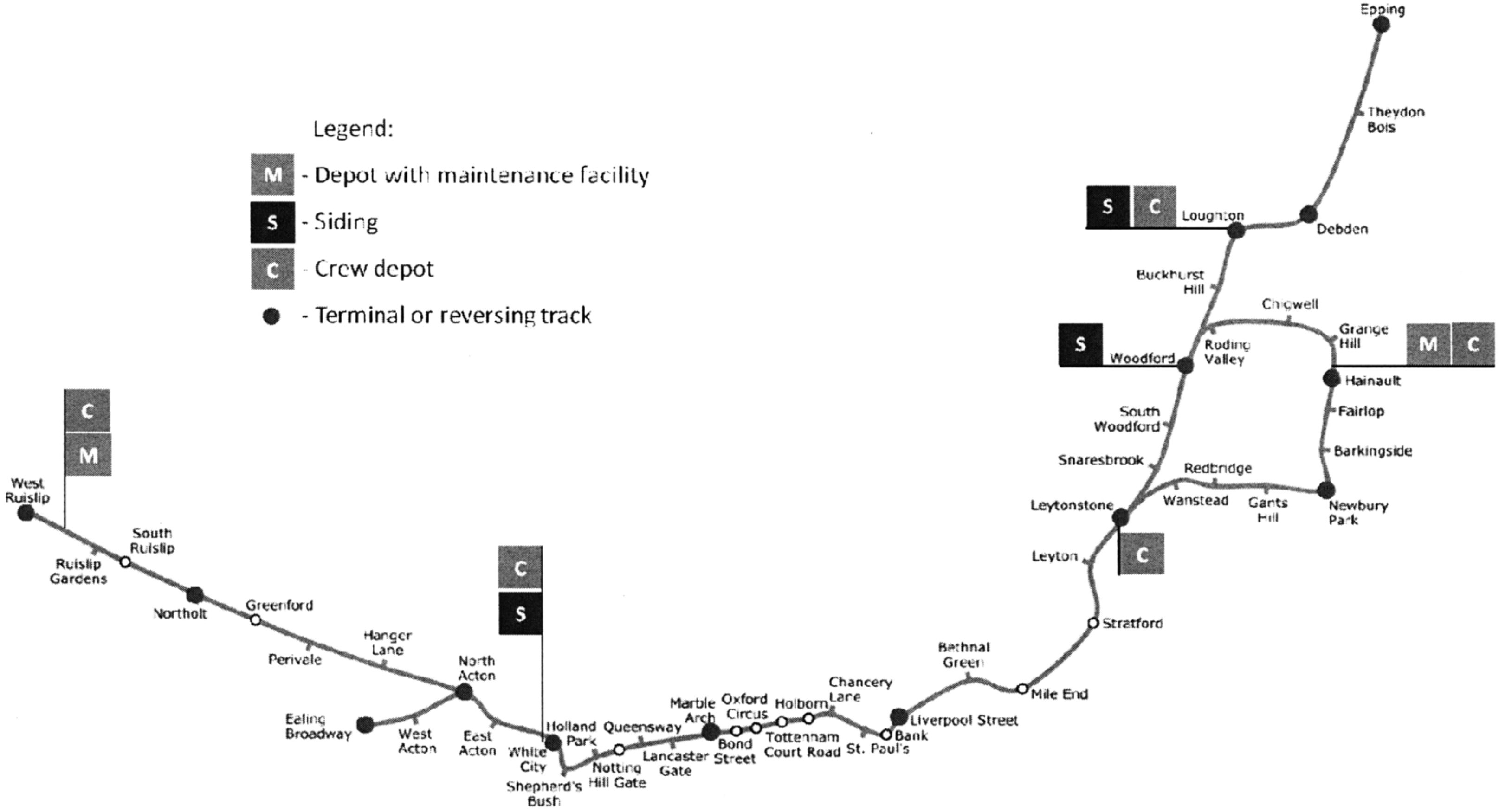


Figure 3-2: The Central line (adapted from Forrester (2006))

Period name	Times	Remarks
Early morning	Before 07:00	First train pulls out 04:53
AM peak	07:00 – 10:00	The peak of the AM peak is between 08:00-09:00
Interpeak	10:00 – 16:00	
PM peak	16:00 – 19:00	The peak of the PM peak is between 17:00-18:00.
Early evening	19:00 – 22:00	
Night	After 22:00	Last train pulls in 01:36

Table 3-1: Time periods on the Central line

Table 3-2 shows the average numbers of passenger trips per hour, allocated to the respective period by trip departure and by line section on a weekday. The trunk is defined as the part of the line between North Acton and Leytonstone; the western branches are to the west of North Acton whereas the eastern branches begin east of Leytonstone. The passenger counts presented in this thesis are based on TfL's 2008 rolling origin-destination survey (RODS) and include passengers who exclusively used the Central line as well as passengers who transferred between the Central line and another Underground line (the latter are determined using a path choice assignment model). In total, the Central line carries approximately 800,000 passengers per day, making it one of the busiest lines of the London Underground. One can see that throughout the day, the heaviest demand is incurred on the trunk section, with large passenger flows between the branches and the trunk in the AM and PM peak. During the interpeak, passenger demand is mostly concentrated on the trunk portion.

A sample graphical load profile for the entire AM peak is shown in Figure 3-3; for better orientation, this figure includes the rest of the Underground network, though without load profiles. Again, one can see that the heaviest link loads occur in the central business district. The busiest link, Holborn to Chancery Lane, exhibits a flow of approximately 320,000 passengers per day. Of those, 80,000 are during the AM peak and 90,000 during the PM peak.

Eastbound

	Western branches to trunk	Within trunk	Trunk to eastern branches	Western to eastern branches
Early Morning	1160	1996	135	4
AM Peak	6843	22654	1055	10
Interpeak	1647	14129	1101	24
PM Peak	2181	31985	5504	192
Early Evening	792	15345	2677	76
Night	206	7907	1016	19

Westbound

	Eastern branches to trunk	Within trunk	Trunk to western branches	Eastern to western branches
Early Morning	1266	3490	251	40
AM Peak	6736	28279	1774	210
Interpeak	1368	14724	1271	22
PM Peak	1361	26186	5203	48
Early Evening	462	12287	2842	4
Night	138	5082	1006	2

Table 3-2: Average passenger flows per hour on the Central line

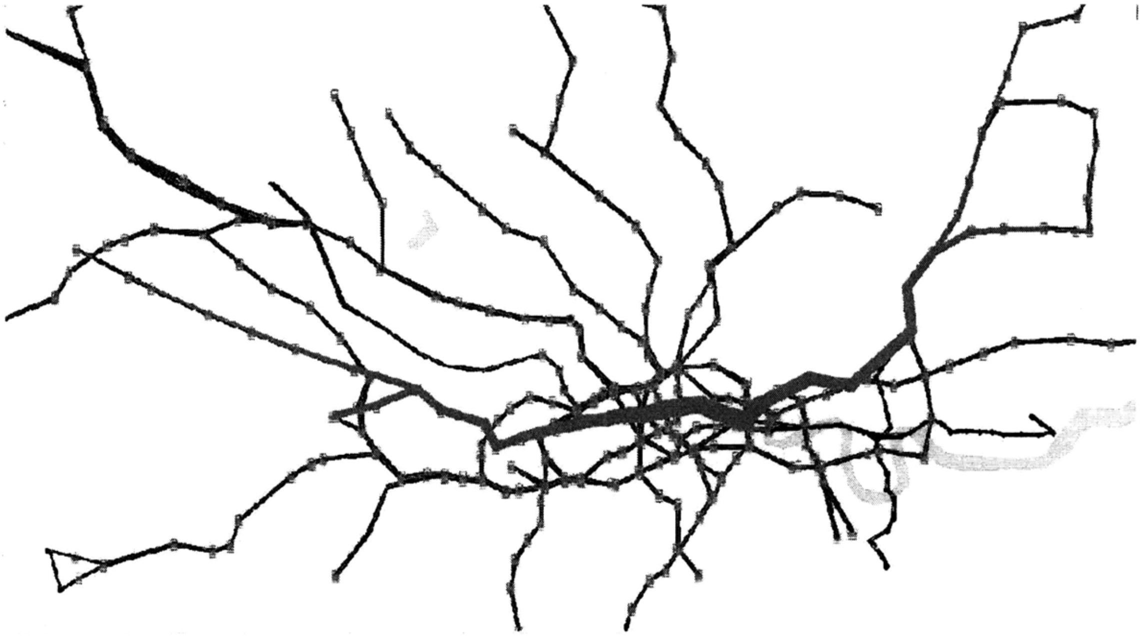


Figure 3-3: Central line load profile (total flows) in the AM peak

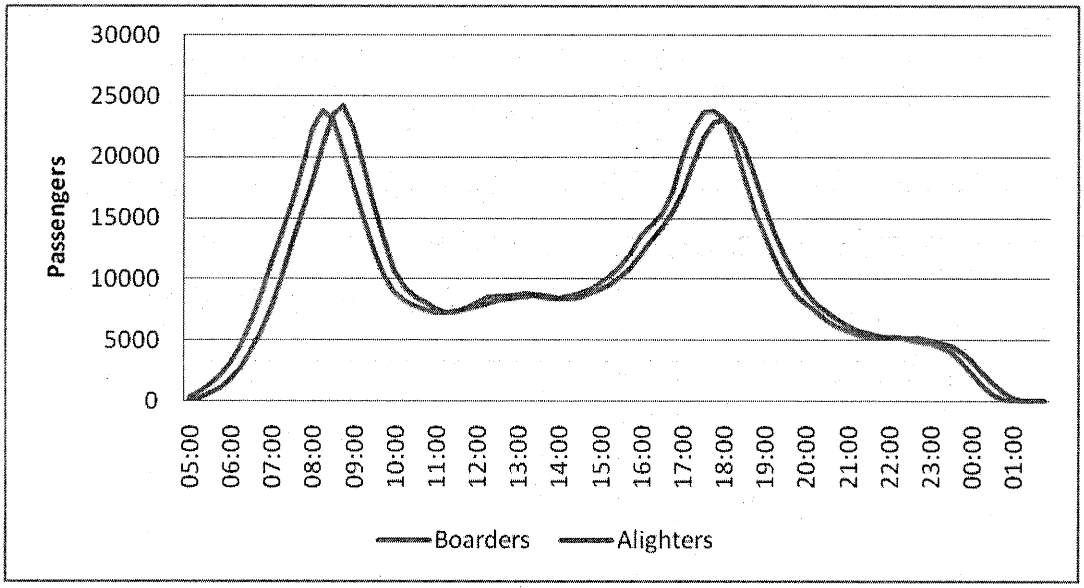


Figure 3-4: Average numbers of boarders and alighters on the Central line

Strong peaking is characteristic of weekday demand patterns throughout the London public transportation network, including the Central line. This is evident in Figure 3-4, which shows the average numbers of boarders and alighters on the Central line according to RODS by time of day.

Finally, Table 3-3 shows the 10 busiest stations on the Central line; they are all located within the central London section of the line. Not surprisingly, the 10 largest OD (Origin-Destination) pair flows on the line are virtually all between these stations.

Station name	Boarders	Alighters	Total movements
Oxford Circus	72156	78017	150173
Liverpool Street	71482	67546	139028
Bank / Monument	58202	63906	122108
Holborn	58085	63331	121416
Tottenham Court Road	53930	52184	106114
Bond Street	48689	43509	92198
Mile End	47480	49261	96741
Stratford	44989	46506	91495
Notting Hill Gate	29694	28134	57828
Chancery Lane	26173	28687	54860

Table 3-3: The 10 busiest stations on the Central line

3.1.3 The Central line timetable

The current timetable is Working Timetable 64, which came into effect on November 19th, 2006. It accounts for the passenger flow patterns described above by scheduling the maximum service of 30 trains per hour westbound through the central London trunk portion during the peak of the AM peak (serving passenger trips originating from the eastern end) and eastbound during the peak of the PM peak (serving passengers destined for the eastern end). The opposite directions are served by 27 trains per hour during the two peaks. In the interpeak, service through the trunk is 24 trains per hour. The tables in appendix B show train departures per hour as a function of time for all stations on the Central line. The scheduled reversal time at terminals is at least 10 minutes. However, according to estimates (London Underground Ltd., 2002), the time required for a driver to switch ends of a train is 4 minutes, which leaves 6 minutes recovery time to compensate for train lateness.

An interesting point about the timetable is that it is actually composed of two largely independent "core services":

- West Ruislip to Epping.
- Ealing Broadway to Hainault, with low frequency service to Woodford via Hainault.

Allen (1981) describes this concept as "self contained services" and explains that it is:

" [...] much favored by the [service] controllers since it ensures that the effect of delays to the service is kept to a minimum. For instance, a delay occurring [on one of the branches] will affect trains operating to and from that station but trains working between [the other branches and the trunk] will continue to run, largely unaffected; thus, the controller will have to attend to only half the total service."

This statement applies only to cases where a disruption on a branch does not cause a spill-over into the trunk, for example in the form of a queue of held trains upstream or conflicts between trains at junctions downstream of the delay. This concept will be revisited in section 7.2. The scheduled train set numbers on the Central line (as described in section 2.1) generally reflect this partitioning of the line, although the assignment is not strict – occasionally trains change their assignment from one core service to the other during the day. This is mainly due to the allocation of storage capacity along the line, which is not directly proportional to the numbers of trains required on each core service.

In total, the timetable requires 79 train units, of which 67 are scheduled to run throughout the interpeak. Twelve trains provide peak service only and are scheduled to stable during the interpeak. As the fleet is comprised of 85 trains, there are six spare trains to replace scheduled trains which become defective or need to undergo planned maintenance.

An important distinction is that not all trains operate from end to end (i.e., from West Ruislip to Epping or from Ealing Broadway to Hainault/Woodford). A subset of trains are scheduled to reverse at White City, North Acton or Northolt in the west and Newbury Park or Loughton in the east. Central line controllers refer to these trains as "local services" since they are designed primarily to serve the trunk section. Section 7.2 shows how this concept has been embedded in the way service controller deal with train shortages.

For all stations with the exception of those on the Hainault to Woodford section, schedules published for the general public contain the first and last trains and a range of train headways throughout the day. For example, for Notting Hill Gate station eastbound, the timetable states that from 06:00 to 24:00 trains depart every 2-6 minutes, without any differentiation by train destination. The train service between Hainault and Woodford runs at 10 to 20 minute headways throughout the day, and all departure times are published.

4 The service controller's decision environment

Chapter 2 outlined the task and responsibilities of service controllers and showed that so far, academic literature has not fully appreciated the decision environment in which rail service controllers operate. This chapter is intended to fill that gap. Following a short introduction to the service control process on the Central line in section, 4.2, section 4.3 will present the decision factors which were found to affect service control in the specific case of the Central line. It should be seen as a description of the current service control process on the Central line, derived from an informal observation process. However, all of these decision factors also apply to other metro lines and systems, albeit with differing degrees of importance. This generalization will be formally introduced in section 4.4.

4.1 Introduction

The results from this chapter are based on the notes from time which the author spent over the course of two weeks in January 2009 at the Central line control center. The goal was to build as comprehensive an understanding as possible of the decision factors and parameters in service control through observing the controllers manage the line in real-time, discussing their actions with them and understanding why a certain strategy was chosen over alternatives. This was an informal process. Beyond observing daily service management on the Central line, the author also reconstructed the response to past disruptions with the service controllers and discussed hypothetical situations. The results were then distilled into a list of decision factors which the author presented to service managers[4], asking them for their opinion and possible changes they would make to it. The author also had the opportunity to discuss the training process for new service controllers, although this confirmed the informal nature of service control, where most of the skills are learned on the job.

As discussed in chapter 2, the two main components of the operations plan are the working timetable and the crew schedule. A deviation from those plans can trigger a service control intervention, either to correct that deviation or to avoid a conflict. However, aside from the elements of train and crew management which are formalized in the operations plan, a range of other factors, which will be introduced hereafter, have an influence on service control decisions. Generally speaking, the constraints imposed on a service control decision can be related to these factors or to provisions in the operations plan, and there are complex interactions between all of them. In order to understand the rationale behind service control interventions, an understanding of these factors is necessary.

---

[4] The service manager is in charge of a shift of service controllers (see description in section 4.2).

In general terms, service control needs a set of simple, real-time performance measures through which the operations on a line are evaluated and compared to their target state defined by the service or operations plan. These performance measures largely define the decision rules and priorities for service regulation. Common supply-centric measures can be headway regularity, lateness of service or total missed trips. The last two measures define schedule adherence, and headway regularity is featured separately because it can be independent from schedule adherence, as explained in chapter 2. Passenger-centric measures might include total passenger delays or travel time reliability metrics, but they are generally much more difficult to calculate in real-time and to relate to the service variables which controllers can influence. To the author's knowledge, adherence to crew schedules is currently not tracked by the London Underground or any other major transit agency.

In many transit systems, service controllers are responsible for the implementation of the entire operations plan into daily operations (see, for example, Froloff, Rizzi, & Saporito, 1989). In the London Underground, the responsibilities for the train and the crew schedules are separated, with service controllers primarily being responsible for managing the train service (e.g., monitoring train schedule adherence), whereas the crew depot DMTs (Duty Manager - Trains) are responsible for crew management and making crew schedule operational. The Central line has five crew depots and hence five DMTs dealing directly with crew management. In addition, the management of rolling stock is only handled by service controllers for rolling stock outside depots. Within depots, it becomes the responsibility of the DDM (Duty Depot Manager). During the author's conversations with London Underground operations personnel and staff from planning and management levels, it became apparent that the decision environment of controllers was often described in a simplistic way. While service controllers would often emphasize that service control is driven by timetable adherence, some managers suspected that the primary driver was crew management, to the detriment of passengers. However, the issue is more complicated, and neither of the two statements can be accepted as such. It was interesting to observe that even the Central line's service controllers did not acknowledge the full complexity of their decision environment immediately, but would only mention the relevant factors one by one in the course of in-depth discussions of their decisions.

4.2 The process of service control on the Central line

Service control on the Central line is carried out by a group of controllers who work in shifts of four to six people, depending on the day of week and the time of day (less controllers work on weekends and in the evenings). There are three functions within the control center: the line controllers, the signalers and the information manager. The line controllers, of which there are usually two present in the control center, monitor the service in the OCS and on the overview screen and make decisions on when and where interventions in the service are needed. They are the primary point of contact for a driver or for a station supervisor when a disruption occurs. Furthermore, since they make the decisions on interventions, they need to be in contact with the DMTs and DDMs as crew or rolling stock management issues arise. The intervention decisions are then passed on to the signalers, who carry out the interventions by coding them in the OCS, setting signals and communicating with the drivers. When the control center is fully staffed, there are three signalers, each of whom is responsible for a section of approximately one third of the line. In addition, there is always an information manager present in the control center, who is responsible for communications with station staff and ultimately with the general public. If there are delays on the line, he or she informs station managers and arranges for announcements to be made to passengers. On the Central line, service controllers are trained for all three of these functions, and they usually decide at the beginning of a shift who takes which position, i.e., there is constant rotation.

In addition to the controllers, there is usually a service manager and a duty engineer present in the room. The service manager is in charge of the controllers. He or she is usually a senior service controller, has mostly administrative functions and does not actively manage the system. However, in case of large disruptions which require difficult service control choices, the service manager often participates in the decision making process. On the other hand, the duty engineer is an expert on the technical aspects of the rolling stock and infrastructure systems and assists the controllers with decisions in case of equipment failures.

4.3 Important drivers of service control

4.3.1 Manageability and uncertainty

First and foremost, recognition needs to be given to the importance of system manageability and uncertainty, as they permeate many if not all of the situations which service controllers have to manage and they were cited very frequently by Central line service controllers during discussions with the author. While it is very difficult to identify the *direct* influence of these factors on how train service is restored after disruptions, one can say that they define the overall approach controllers take to managing a problem. When a disruption occurs, controllers must usually react under severe time pressure and with uncertainty about the duration of the incident and what other problems might occur in addition to those immediately apparent. This leads them to embed a set of principles into their work described in what follows.

Robustness of intervention:

With every intervention, there is a possibility that it is not implemented as planned due to misunderstandings or unforeseen events. In deciding between different recovery strategies, several controllers stated that they would try to avoid timetable changes which they knew from experience might be misunderstood or disregarded by drivers. As an example, one controller explained that, in order to put a late train (and a late driver) back on schedule, he preferred short-turning the late train rather than performing a stock and crew operation using a spare driver. The reason given was that every crew changeover bears a risk of delay and a risk of misunderstanding, since the instructions for these interventions have to be communicated to drivers via train radio.

Management of workload:

The decision taken by the controller in the above example can also be understood in terms of workload management. It is interesting to observe how the Central line controllers are self-regulating in terms of workload. During "quiet" times they may have more time to dedicate to isolated interventions, for example by communicating with the drivers to discuss crew relief issues. However, when larger disruptions occur, particularly during peak hours, controllers need to manage their workload efficiently in order to tend to all needs while preserving the capacity to respond to unforeseen events or to new information regarding the disruption duration. The result is that controllers will tend to choose "simpler" intervention strategies, as described above, and narrow the scope of line management *during* a disruption to "keeping the service running", as many controllers formulated it. That can be understood as saying that the objective is to meet the level of service requirements (as outlined in the following section) with a relatively short time horizon while a disruption is affecting the line, and that other considerations or even a longer-range view may not enter the picture until the recovery phase. As it is often necessary to act fast when an incident occurs, the controllers do not have the time to think through multiple permutations of possible solutions. Their choices will therefore not necessarily reflect the optimal solution as could be determined by post-analysis, but they will generally choose a solution which meets all current constraints, is feasible under time pressure, manageable and flexible to deal with uncertainties.

Maintaining slack capacity in controller workload is a strategy to deal with the uncertainty in how a disruption will develop or whether other incidents will occur later. Another strategy, which will be discussed in section 4.3.5, is the maintenance of slack capacity on reversing tracks and at depots. Interestingly, this same principle did not appear to apply to spare crews and rolling stock, presumably because these aspects of the service are not within the full responsibility of service controllers, as explained earlier in this chapter. During the interviews, Central line controllers frequently cited uncertainty as the reason for limiting the time horizon of their decisions to approximately two trips of any train under consideration. However, it is plausible that this is also a matter of the natural limits of human processing capabilities.

4.3.2 Level of service

As stated above, the primary responsibility of Central line controllers is for the timetable and the train operations on the line, with the DMTs responsible for crew management. However, as these two areas are closely connected, their interdependencies still have a large influence on controller's decisions, as will be discussed subsequently.

Furthermore, section 4.1 introduced the notion that service controllers need to work with a set of real-time performance measures. In the case of the Central line, the overarching objective of the service controllers is schedule adherence, which can be broken down into the objectives of:

- Operating as many of the scheduled trips as possible.
- Operating those trips with as little lateness as possible compared to the timetable.

The objectives are based on the premise that the timetable is the optimal output for the system, even in the case of disruptions; the reader may note that controllers assume that “optimal service to passengers” is implied in these objectives and does not enter as a separate evaluation criterion. Furthermore, the author observed that the concept of headway regularity was not embedded in the service controller's decision process in any strict sense, but was rather handled as a “fuzzy” criterion which is not defined and handled in a unified way by controllers. The author suspects there are two reasons for this. Firstly, the complexity of the Central line's layout and of its traffic patterns limit the human capability to manage a service which deviates greatly from the plan, and secondly, the information conveyed in real-time to controllers via the operations control computer system contains the schedule adherence for trains, but no information on headways.

In daily operations on the Central line, infeasibilities arise very frequently due to disruptions, congestion, the unavailability of assets or occasional crew errors. In that case the controllers face the following decisions:

1. Should the timetable be retained as the target state of the system? If not, controllers move to the implementation of an ad-hoc alternative service which is intended to fulfill the performance targets as closely as possible. This will only occur in the case of major disruptions which require partial line suspensions.
2. If the timetable is retained as the objective but cannot be completely fulfilled, which trips will be covered and which not? This may involve a reassignment of vehicles and drivers and the insertion of unscheduled trips in order to re-link scheduled ones which are to be covered.

Hence, although timetable adherence is regarded as the primary objective, the result of the daily infeasibilities is that the ideal situation in which all service is run as scheduled is rarely achievable, and the Central line controllers frequently need to select which trips not to cover, where to insert additional trips and where and how to modify the timetable. In doing that, they were observed to follow a set of unofficial decision rules and priorities, which may differ among individual controllers and among controller shifts. Those related to the level of service can be summarized as follows:

- Lateness with respect to the timetable is generally tolerated as long as it can be recovered during reversal and does not cause any other conflicts. This means that interventions to correct minor lateness (thus also evening out headways) during otherwise "good" service are rarely performed.
- The primary way of evaluating the level of service (LOS) is by visual inspection on the large overview screens. Controllers are generally more concerned with *service patterns* rather than quantitative measures for LOS or strict headway regularity. A cause for this might be that headways are not calculated by the computer system and the overview screens are not to scale. Thus, in the absence of a disruption, gaps or train bunches are not readily detected until they either become obvious on the overview screen (which happens faster for higher frequency service, i.e., on the trunk portion of the line) or they are reflected in the lateness of individual trains. The main service patterns sought by controllers are:
    - Regular train destination sequence through the trunk portion – ideally, the branches should be served by alternating trains.

  - Regularity of service on the branches. Since the branches have lower service frequency than the trunk portion, controllers aim to cancel or delay as few trips as possible on the branches, giving them priority over trunk service. However, what is deemed to be a "regular service" in terms of headways depends on the individual controller's definition. In one of the disruptions observed by the author, the service controllers in charge did not define service regularity as even headways, but rather in terms of a maximum headway between trains, corresponding to the published maximum headway.
- Low-frequency sections of the route where the timetable is published, such as the northern part of the Central line loop between Hainault and Woodford, are special cases. The objective of covering all trips through that section as closely as possible to the scheduled time is seen as very important. Trains to Woodford via Hainault will thus be treated with priority above other services to the eastern section of the line.
- The sequence of branch destinations of trains serving central London is both a trigger (in case of imbalances) and a constraint for interventions. It influences the choice of trains to redirect. In practice, changes are often made to the routing of peak-hour only trains or trunk service trains (between White City and Newbury Park or sometimes Loughton; referred to by controllers as "local trains") in order to cover missing runs to branches or to fill gaps in the service. These trains are removed from the trunk portion, thereby reducing the service provided there.
- The time of day has an influence on how aggressively controllers will work to restore the service to timetable. After the PM peak, it is often preferred to stable trains early (or late) rather than to change their routing with short-trips, diversions etc.

The last scheduled trains on every section of the line have a special function. On one hand, they are required to run under their original number for the safety of infrastructure maintenance crews (cf. section 4.3.7), and on the other hand, they need to connect with the last trains on intersecting lines, as stipulated by the timetable. Therefore, the last trains must be run as scheduled.

Although passenger impact will be treated separately in section 4.3.6, it is worth reflecting on it at this point. Under totally random passenger arrivals, the expected passenger waiting time is a function of the average headway and the headway variance, as given by Larson & Odoni (1981):

$$E(w) = \frac{E(H)}{2}[1 + (Cov(H))^2] \tag{1}$$

Where:

E(w) = expected value of passenger waiting time,
E(H) = expected value of headways between vehicles, and
Cov(H) = coefficient of variation of headways between vehicles.

The points above show that the average headway and headway variance are treated with different priority on different sections of the line, with more importance attributed to them on the low-frequency (and low-demand) sections. Moreover, the variance of passenger waiting time is an important component of travel time unreliability, and some of the policies shown in the list above are bound to introduce irregularities into the service, such as the policy of using trunk service trains to cover services to the branch terminals, combined with little monitoring of headways on the trunk. This will be addressed in a case study in section 7.2.

4.3.3 Crew management

Without a doubt, crew management is one of the most complex aspects of service control, not only at the London Underground but in many metro systems worldwide. It has a direct impact on how the train service can be operated and is governed by numerous rules and regulations. This section is divided into three parts. In part a), the driving time regulations for drivers at the London Underground are explained. Part b) then shows how the Central line controllers were observed to include crew management issues into their decisions, and part c) explains how the availability of spare drivers influences the recovery from delays and disruptions.

a) Driving time regulations

London Underground (LU) drivers are subject to work rules which pose hard constraints on how they can be scheduled. Those rules include:

- Maximum driving time per day (8 h 30 min)
- Maximum driving time without a break (4 h 15 min)
- Maximum driving time within a four-week period (140 h)

- Minimum meal break (35 min)

In addition, drivers should begin and end their daily shift at the same crew depot. LU has no official rules about where the meal break can take place, but there is a strong preference among DMTs and drivers to have the meal break take place at the driver's "own" crew depot, presumably for reasons of management and convenience. A further constraint is that drivers who have two pieces of work within one day with a meal break in between must step off the first shift at the latest at the starting time of the second shift minus 45 minutes (meal break plus 10 minutes walking time), because otherwise they will be late for their second piece of work.

These constraints can all be accommodated in the original crew scheduling process but they can be challenging when the service is disrupted and trains need to be rescheduled. A driver who has exceeded the maximum permitted driving time is theoretically entitled to park the train and quit work. However, several controllers indicated that in the case of bad disruptions, drivers will generally be understanding if the driving time regulations are breached, as long as they can assume that controllers are working to minimize the excess driving time.

b) Inclusion of crew management considerations

During the conversations with Central line controllers, it was found to be difficult to assemble the individual statements of controllers into a coherent picture of how crew constraints influence their control decisions. It is rare for controllers to acknowledge directly the importance of crew management in service control; if asked directly, almost all controllers will state that their job is "to run the service" and that "staffing the trains is entirely the responsibility of the DMT". The fact that crew management is a major issue in service control only becomes clear when specific examples of interventions are discussed with controllers. The author strongly believes that this is not an intentional concealment of information. Instead, it is an indication that the structure of responsibilities within the London Underground may not adequately reflect the fact that crew and train management are closely intertwined.

Section 8.2.2 will address this question in detail and describe how the London Underground may want to review the responsibilities of DMTs and controllers in terms of crew management in order to better account for this close connection. Furthermore, it is possible that the way crew information is currently provided to controllers, i.e., in a static, printed form which is not updated as changes are made, is not optimal to support service control decisions. Instead, an electronic overview of crew locations with the most relevant information, which could be integrated into the OCS, may be a worthwhile investment to improve the real-time information provided to controllers.

The author's observations and conversations with controllers yielded the following summary of how crew management influences service control on the Central line.

Driver lateness:

One must avoid breaching the aforementioned rules on working time and drivers scheduled to step on and off a train at a crew relief point need to be matched up. If a driver is late and one of these objectives is at risk of not being met because a train is delayed, controllers can either maintain the original train routing and use a spare driver to relieve the late driver or they can perform a rerouting, e.g., short-turn the train, cancel its last round trip or divert it to a branch with less cycle time than its original destination. It is important to note the link between train lateness and driver lateness. The Central line's operations control system only reports train lateness, that is, the delay of a train with respect to its scheduled trip number. As long as a train is still running under its original number (i.e., has not been renumbered, as described in section 2.3.1), the reported train lateness will be the same as the lateness of the driver on board. In this case, a rerouting (such as a short-turn) which puts the train back on its original schedule will also guarantee that the driver's working time and relief constraints are met. However, if a train is renumbered, the train lateness and driver lateness are decoupled. In this case, it is possible to have a train which the OCS reports to be on time yet has a late driver on board which is at risk of breaching the work rules, and it may eventually be necessary to short-turn or otherwise reroute that train in order to bring the driver back to the crew relief point on time.

Central line crew schedules generally have some built-in slack at the end of a piece of work, so most of the Central line controllers stated that the lateness of a driver stepping off does not become a concern until the driver is about 20 minutes late, as long as he or she can be matched up with a relief driver. On the other hand, although none of the controllers spoken with acknowledged it directly, the author observed that controllers were very reluctant to have drivers step onto a train late when beginning work. The author suspects that this may have three reasons, two of which are related to the risk of delays during crew reliefs:

- Controllers cannot communicate directly with drivers at crew depots, and all changes would have to be communicated via telephone to the local DMT. This process is potentially unreliable (i.e., misunderstandings can arise) and lengthy.

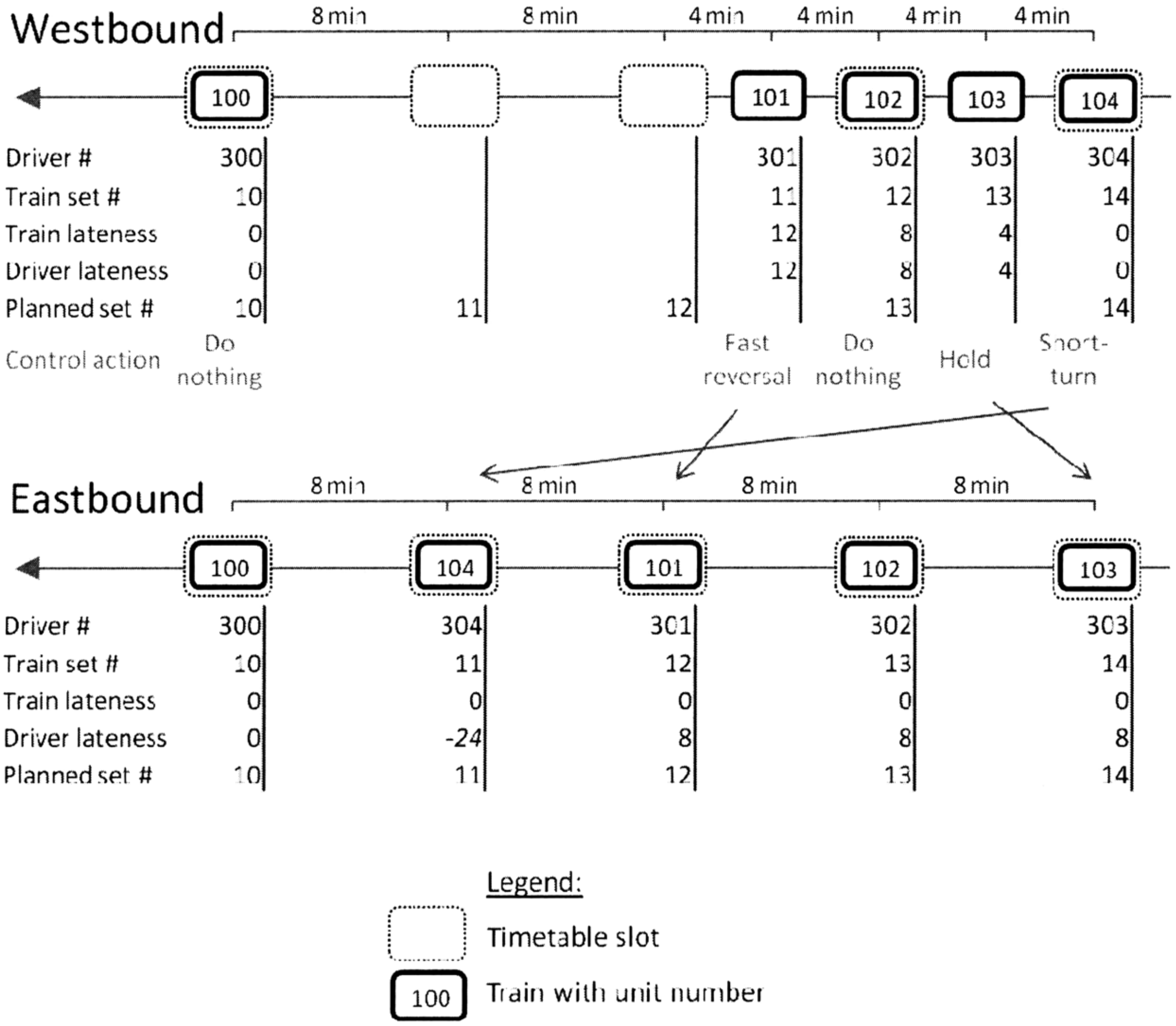
- The same holds for drivers: if their train is late, they will need to be informed by the DMT of their new pickup time as the drivers do not have an information system which informs them when "their" train is approaching. If there is a misunderstanding or if the DMT forgets to advise the driver of the change, there will be no relief driver waiting on the platform when the train pulls in – this is potentially very harmful to the service as it can create a blockage or even force the controller to withdraw the train from service if the relief driver cannot be located.
- If the relief driver is stepping on for his/her first piece of work and must meet a hard time constraint later during the day (for example the beginning of a second piece of work), it may be unwise to use all the aforementioned slack time up front, since this will immediately create a problem if there is a disruption later in the day which delays this driver.

Since in the case of minor delays, it is acceptable for drivers to step off slightly late but it is undesirable for drivers to step on late, Central line controllers use a combination of renumbering and rerouting delayed trains in order to put drivers back on time for a crew relief. Such a strategy is described in the following, idealized example, which is shown in Figure 4-1:

Consider a sequence of five trains, unit #100 through #104, traveling westbound to the same destination (e.g., West Ruislip). The set numbers are #10 to #14. The scheduled headway is 8 minutes and all five trains are on their penultimate trip before a crew relief. In between the trains, there may be trains scheduled to other destinations (e.g., Ealing Broadway) which will not be considered here. The "planned set number" in Figure 4-1 denotes the scheduled slot which the trains with the respective set numbers are supposed to be in.

Unit #101 suffered a blockage earlier and is 12 minutes late with respect to its scheduled slot, i.e., there is a 20 minute gap behind unit #100 and the three trains following #101 are bunched. The scheduled reversing time is 10 min and the minimum reversing time is 6 min, i.e., a train can compensate for 4 min of lateness at reversal. A controller needs to decide how to restore service while ensuring that the crew reliefs, which are scheduled on the next trip eastbound, can take place at the right time without passing on lateness to the relief drivers. Unit #100 is on time, so no action is taken with respect to that train. The controller renumbers the trains to assign them to a new timetable slot: unit #101 becomes set #12, unit #102 becomes set #13, unit #103 becomes set #14 and unit #104 is bumped up to set #11, making it 24 min late. With a fast reversal, unit

#101 can make its new scheduled slot on time. Unit #102 is now on time for its new slot and #103 is now 4 min early, so it needs to be held. Meanwhile, unit #104 is short-turned and inserted into the gap on time to make the eastbound set #11 slot. Now, traveling eastbound, all trains are on time again, but the drivers are not: driver #304 on unit #104, which was short-turned, is now 24 min early, while all drivers behind unit #104 are 8 min late. However, thanks to the renumbering of the trains this problem is solved when the crew reliefs take place, as shown in Table 4-1. The relief drivers step onto the train with the set number they were planned to drive, but the driver stepping off is not the one which they would have relieved according to the operations plan, which does not matter. Drivers #301 through #303 step off slightly late and driver #304 steps off early, but the relief drivers step on on time, onto a train which is running on schedule.



Westbound	100			101	102	103	104
Driver #	300			301	302	303	304
Train set #	10			11	12	13	14
Train lateness	0			12	8	4	0
Driver lateness	0			12	8	4	0
Planned set #	10	11	12		13		14
Control action	Do nothing			Fast reversal	Do nothing	Hold	Short-turn

Eastbound	100	104	101	102	103
Driver #	300	304	301	302	303
Train set #	10	11	12	13	14
Train lateness	0	0	0	0	0
Driver lateness	0	-24	8	8	8
Planned set #	10	11	12	13	14

Figure 4-1: Recovery from a blockage

Planned sequence		Changed sequence	
Driver stepping off:	Planned relief driver:	Driver stepping off:	New relief driver:
300	310	300	310
301	311	304	311
302	312	301	312
303	313	302	313
304	314	303	314

Table 4-1: Crew reliefs for recovery example

Crew management in the context of service recovery:

Among the most important influences of crew management on service control is its effect on *how* controllers restore the service after disruptions. The essence of many conversations with Central line controllers which touched on crew management was that when a controller performs an intervention to improve the level of service, as outlined in section 4.3.2, the choice of which train to perform the intervention on and when to do it is strongly driven by when and where crew reliefs are scheduled to take place. In choosing their recovery strategy for restoring the service to the timetable (the primary objective), one of the secondary objectives of controllers will be to meet crew management requirements. For example, if a train trip needs to be extended in order to close a gap on a branch, trains without an impending crew relief will be the most attractive candidates. As a matter of fact, several of the controllers spoken to described how they frequently employed recovery strategies which hinged on crew reliefs. crew reliefs would act as "fixed points" in the timetable, and controllers would reroute and renumber late trains to meet these departures as scheduled, because they know that in doing so, they are automatically restoring service to the timetable *and* meeting crew management constraints. That is, relief drivers step on on time, onto a train which is on schedule. An example of such a strategy was given above. These strategies don't explicitly focus on passengers who are on board the trains after a disruption, but they represent a "proven" way to restore the service to timetable and to crew schedule and they are intuitively understandable for controllers.

Crew problems:

As already indicated above, it must be mentioned that disruptions which are at the origin of a delay can be crew related, such as a driver not being in place for a relief. On the Central line, three of the crew depots are located along the line (White City, Leytonstone and Loughton), but the timetable does not schedule additional dwell time for crew reliefs. One of those crew depots

(Leytonstone) does not have the capacity to berth trains in case no driver is available. Both these factors can easily cause delays to the service if there are problems with a crew relief, which explains why Central line controllers are very attentive to the locations and times of crew pickups.

c) Utilization of spare drivers

Drivers need to step off at their crew depot if they are at risk of reaching the maximum driving time during their next round trip. In that case or when unscheduled trips are inserted, controllers have to rely on spare drivers. On the London Underground, these spares are assigned to individual depots and are dispatched by the DMT. The long round trip times on the Central line mean that a spare driver's remaining driving time often becomes a constraint on spare utilization. As an example, consider the Central line's crew depot at White City. The scheduled cycle time from White City to Epping and back is 2:12h, whereas to West Ruislip and back it is 57 min. Hence, a spare driver with less than about 2:20h of remaining driving time cannot cover a duty to Epping, but would still be available to work a train to West Ruislip if he or she has at least 1 h of remaining driving time. That is, the availability of spare drivers is a function of which non-spare duty needs to be covered and is dependent on how much driving a spare driver has already had to carry out during his or her shift. Moreover, the conversations with Central line controllers revealed that the utilization of spares seems to be a function of their availability. The larger the pool of spare drivers at a particular crew depot, the more controllers will be inclined to utilize them for interventions or for reliefs. Sufficient numbers of available spares give controllers additional flexibility in how they recover service after disruptions since they can circumvent the maximum driving time constraints by relying on spare drivers, thus reducing the need for service curtailments (such as short-turns or cancellations) because of late drivers. Due to the constraint of maximum driving time per month, the availability of spare drivers tends to decrease towards the end of a calendar month.

4.3.4 Rolling stock management and maintenance

If a disruption occurs along the line as opposed to at a terminal, it is often the case that trains are only blocked in one direction. However, it was observed in the Central line control room that in the case of a prolonged blockage in one direction, controllers would also stop the service in the opposite direction. The reason for this is what London Underground controllers refer to as the *stock balance*, which is important because an imbalance (for example if, after an eastbound blockage in central London, there is an oversupply of train units on the western part of the line but a shortage of train units on the eastern part of the line) in the distribution of train units on the line presents large problems for service recovery. This is especially important on the Central line with its long running times and high scheduled frequencies on the trunk portion, making it a lengthy and complicated process to redistribute the trains such that the timetable can be run again.

Stock balance is also of concern when trains stable in depots, due to limited depot capacity and due to the fact that the number of trains in a depot over night should equal the number of scheduled pullouts from that location at the beginning of service. In case defective trains have to be withdrawn from service to depots other than their "home depot", this is likely to cause unscheduled trips or cancellations later during the day in order to reestablish the stock balance.

A further trigger for such interventions can be the rolling stock maintenance schedule. On the Central Line, the maintenance facilities at train depots specialize in different types of maintenance work, and every morning the depot managers (DDMs) issue a list to the service controllers indicating which train units need to ***stable as booked*** (i.e., return to a particular maintenance facility) in the evening. They are then assigned to timetable runs such that they would end at the desired depot. However, when trains are rerouted or renumbered by controllers, they are removed from their planned trajectory, and there is no feature in the service control computers which retains the information about where a train needs to stable in the evening – it would need to be manually reconstructed. Most often, depot managers assess the position of the trains they need in the early evening and then communicate with service controllers, who will divert the trains as necessary.

In summary however, the requirements of rolling stock management are relatively straightforward and do not impose serious constraints on disruption management and recovery. All Central line controllers spoken to stated these constraints in a similar fashion.

4.3.5 Infrastructure capacity and track layout

Before considering the specific situation of the Central line, a few general remarks on the function of terminals should be made. They are based on Allen (1981) and Lee (2002). The capacity of terminals and reversing points has a direct influence on the propagation of delays throughout the line:

- The layover time scheduled at terminals helps modulate headways and put trains back into sequence or on schedule. Maximum layover time is a function of terminal capacity (number of reversing tracks and platform clearance time) and train arrival rate. Hence, if the scheduled arrival frequency at a terminal is increased or if trains become bunched due to delays, the maximum possible layover time is reduced and controllers potentially need to perform interventions in order to avoid congestion at terminals. For example, a controller might short-turn trains before they reach the terminal, divert them to another terminal or release berthed trains from the terminal early.

- A special case is reversing points (sidings) along the line which trains move into from a regular two-track station (Figure 4-2 shows an example). If the reversing train becomes delayed and the next train is following with a smaller than scheduled headway, this station can turn into a choke point, since agency policies usually stipulate that the train which is reversing needs to be inspected to ensure nobody is on board before it moves into the siding. This process, which is called *detraining* at the London Underground, can take several minutes. Since the following train does not have the ability to pass the reversing train, it risks being delayed. To avoid that, a controller might decide to extend the trip of the reversing train to a station with three or more tracks.

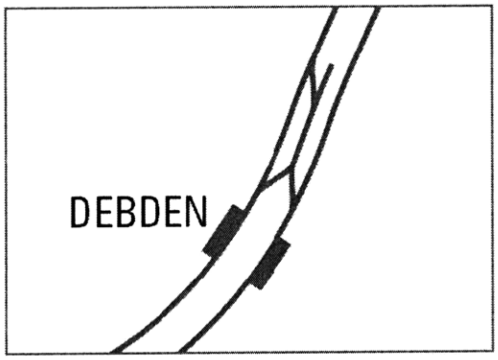


Figure 4-2: A reversing track at a station on the Central line (Source: London Underground Operations Track Overview)

On the Central line, terminal congestion happens most frequently during peak hours, and many crew reliefs are scheduled after the end of the peak. The tighter layover times at terminals become as frequencies are increased to serve peak demand, the more likely it is for delays from the peak hour to carry over into the crew relief period. These delays may then trigger control interventions at a later point, thereby negatively affecting the level of service in the off-peak.

During the conversations with Central Line controllers, several expressed their concern about the difficulties of recovering service under a very high frequency timetable, stating that the irregularities in headways and train sequences caused by disruptions would, among other effects, very quickly lead to train congestion at terminals. This lengthens the recovery process and, as stated above, causes controllers to rely more on interventions such as short-turns or trip cancellations for normalizing service. A further point made by the controllers is that additional reversing capacity along the line helps maintain a higher level of service on non-disrupted portions during long incidents, for example during line blockages, and that it is problematic to use this capacity for scheduled reversals. On the Central Line, it can be observed that during line blockages, the reversals upstream of the blocked section become choke points of the service. Therefore it has become standard practice among Central Line controllers to "thin out" the service on the rest of the line by removing a certain number of trains if a long disruption is anticipated.

The author's observations showed that on the Central Line, the limited capacity of existing terminals (especially on the Western part of the line) affects service on a near-to-daily basis. Particularly during peak hours, terminal congestion is common at Ealing Broadway and around White City.

One final observation is related to the use of reversing tracks and spare capacity at terminals for removing defective trains from service. The Central line service controllers appeared to be reluctant to use the entire spare capacity of a terminal or siding for a prolonged period of time to store a defective train, since that would reduce spare capacity to deal with unforeseen events requiring short-turns. A specific case which was observed was the following. In the late afternoon, a westbound train was withdrawn to the reversing track of North Acton station as it suffered a burst air line and its speed was restricted to 15 km/h. This occurred while a special timetable was in effect due to construction work, and North Acton station was not being used for reversals, so theoretically it would have been possible to leave the train there until the service thinned out sufficiently during the evening to withdraw the train to White City depot without causing delays. However, the service controllers appeared to become nervous about the train blocking a reversing track that might be needed to respond to a disruption or delay, so they preferred to withdraw the disabled train to White City immediately. The limited speed caused a slow-moving blockage, thus delaying several eastbound trains which followed it to White City. However, that delay was *controllable and of a known extent*, whereas the consequences of leaving the defective train in the siding were unknown and potentially more severe.

4.3.6 Passenger impact

As seen in section 2.5.4, academic literature often formulates minimal passenger travel time or waiting time as the main objective of service control. However, the conversations with Central line controllers revealed that passenger impact is a relatively fuzzy concept in daily operations management. Although the London Underground calculates performance metrics such as excess travel time and average wait time after the fact, these measures are not available to controllers in real-time, let alone with high enough precision to evaluate alternative service control interventions.

Therefore, controllers must work with assumptions and past experience. The observations in the Central Line control room showed that after disruptions, the main objective is to restore service as quickly as possible to run as scheduled in the operations plan; the implicit assumption is that, over the course of the entire day, passenger benefit is maximized by a service running at schedule for the maximum possible time. The impacts of disruption management on passengers *during* the recovery phase generally enter into the controller's decision process as a set of constraints, as outlined below. This list was compiled by the author and discussed with service managers on multiple occasions.

Crowding and congestion:

Overcrowding of platforms is a significant concern as the risk increases that passengers or their belongings may fall on the tracks, and train crowding is important because it leads to left-behinds, platform congestion and passenger illness on the trains. However, it was found that service controllers did not attempt to actively manage daily passenger congestion on the Central Line, as the line is already running very close to its nominal capacity – the options available to alleviate crowding are therefore severely limited. According to information provided by controllers, crowding is actively considered during service disruptions as it poses a constraint on what control interventions can be performed:

- Holding of trains in front of a disruption is rarely performed during peak hours because trains held in stations become so crowded that they do not provide any more capacity for downstream passengers until they reach their first major alighting point, and the high levels of crowding are likely to cause further delays. Furthermore, controllers explained that the holding of trains in front of incidents was a frequent reason for passenger complaints (see below).
- If a train needs to be short-turned or detrained, this can only be done at an uncrowded station where there is sufficient platform capacity to absorb the additional number of passengers. Similar considerations apply for diversions, if one must expect a large number of passengers to alight when the change in destination is announced.

- Along the same lines, it was observed that when controllers needed to divert or short-turn a train (for example to cover the trip of a defective train), they considered the load level of the next train to the original destination in order to ensure that there was sufficient capacity to absorb the additional passengers from the rerouted train.

Passenger complaints:

One of the shortfalls of the aggregate measures of passenger impact is that they generally don't account for the fact that additional waiting time and inconveniences such as crowding can be valued very differently by passengers depending on the individual situation, as mentioned in section 2.5.4. Although there is no easily deployable measure of "passenger (un)happiness", one can use the frequency and reasons for passenger complaints as a proxy. London Underground controllers are not informed about formal complaints received by other divisions of the Underground, but they are well aware of the (much more frequent) verbal complaints to station staff or – in extreme cases – assaults on staff by customers. The author's observations in the control room showed that many service controllers had a strong "gut feeling" about how different types of service control interventions or the lack thereof would cause customer discontent and complaints, and when making decisions they would weigh the alternatives accordingly. This process depends very much on the individual controller and it tends to be quite subjective.

Availability of alternatives:

As the Central line has multiple branches and successive trains through the trunk portion are usually bound for different destinations, the availability of alternatives to passengers affected by a change in service forms part of the considerations when controllers need to divert, short-turn or withdraw trains. As was explained by the Central line controllers, this is most important in the case of incidents or train defects which require a train to be rerouted on short notice, for example for withdrawal to a depot. Following trains will then need to be diverted or extended in order to re-balance service to the different destinations, thereby forcing some passengers to alight and change trains. Several controllers stated that the goal in such cases is to make passengers wait at most one or two headways until they can board a train to their original destination. In addition, the re-sequencing of trains needs to be done such that all passengers can take a train to their original destination which is following the one they are forced to alight from or which they can change to across a platform. For low-frequency branch services, this means that one train may need to be held at a station with three platforms until the following train pulls in.

Total excess journey time:

As noted above, there is no measurement of total excess journey time available to controllers, and it would not be easily implementable since it would involve a predictive component. However, observations in the Central line control room showed that in general controllers try to minimize the impact on passengers by performing interventions which are not time-critical (i.e., not in direct response to an incident) outside peak hours and on low-demand sections of the route, and by selecting the least-loaded trains for time-critical interventions, if there is a choice. Nonetheless, the details of how passenger impact is considered depend strongly on the individual controller.

On the other hand, a policy which was stated by all controllers is that the "incident train", i.e., the train which was originally blocked by a disruption, will be allowed to run through to its scheduled destination. Despite crew management or schedule problems that might arise, this was perceived as necessary since the passengers on board that train experienced the maximum on-train delay of all passengers affected by the disruption. In addition, this practice ensures that the on-platform waiting time of downstream passengers is not extended beyond the disruption-related gap by recovery measures.

4.3.7 Safety

Two safety-related decision factors in service control were observed during the author's visit to the Central line control room: Passenger safety and maintenance crew safety, as explained below.

a) Passenger safety

The two areas of concern which were touched on by Central line controllers when speaking about passenger safety were platform crowding, as already explained in section 4.3.6, and in-tunnel stops. If a train is blocked in a station, controllers will try to keep the following trains from being stopped inside a tunnel. This is due to the initial uncertainty about the duration of the blockage and because evacuating passengers from a train is significantly more problematic when the train is stopped in a tunnel rather than in a station. In addition, the temperature inside a crowded train with closed doors can quickly reach critical levels. If a train becomes blocked in a tunnel, controllers need to make a decision about whether or not to evacuate it after approximately 10 minutes, given the nature and projected duration of the disruption. This would involve either reversing it to the previous station or, if that station is already occupied by the following train, having passengers walk out through the tunnel. The strategy of avoiding in-tunnel stops means that upstream trains are likely to be held in stations in the case of a disruption instead of allowing them to close in behind the incident train. This has an influence on the nature of the timetable deviation observed after the incident and on the interventions necessary to "repair" the service. Instead of a train bunch with minimum headways behind the incident train, one would still observe train spacing greater than the minimum headway, but with lateness extending further back than if the trains had not been held.

b) Maintenance crew safety

As previously mentioned, the last trains serving any London Underground line must run as scheduled and with their correct train set number. These trips are published for track maintenance crews, who are allowed to request the discharge of track current and descend onto the tracks only when that train has passed. It was observed in the control room that having a train in place to cover the last run was a primary concern of controllers when dealing with disruptions in the late evening, especially because the set of available trains and drivers is fairly small towards the end of service, so the loss of one train is more serious than during the day. It may be noted that late-evening disruptions due to passenger action are most frequent on Fridays and Saturdays (when the presence of drunk passengers is high).

4.3.8 Infrastructure maintenance

The effect of infrastructure maintenance requirements on daily train service on the Central line is generally foreseeable, as maintenance work is usually planned in advance, allowing time to provide customer information and to develop an alternative operations plan. However, controllers mentioned two maintenance procedures which may have an unplanned impact on service and need to be coordinated by controllers in real-time: sandite application (see below) for combating leaf buildup and track de-icing. Both are only important on the open track sections during the autumn and winter months; the interference with service is limited, but can cause some cancellations or diversions:

- A subset of the Central Line trains is fitted with dispensers for antifreeze liquids. If temperatures near the ground approach the freezing point, these trains need to run over the open track sections regularly, including throughout the night. Scheduled trips requiring a de-icing train are noted in the timetable, and in the morning depot managers assign de-icing trains to these runs. However, controllers noted that disruptions during the day can put these trains on a different trajectory, in which case it would be necessary to divert them to their starting point once de-icing is required. The assignment of de-icing trains to specific trips is strongly driven by rolling stock balance considerations, since they are often in continuous operation during the night.

- Unlike de-icing trains, the Central Line's sandite train is a special maintenance train which does not operate in revenue service. Its only noticeable effect on operations is when it obstructs trains in revenue service in the vicinity of terminals, but controllers agreed that the delays to train service and to passengers are minimal.

4.4 The link between service control, the timetable and reliability

One of the most important points of the above description of the service controller's decision environment is that there is a very strong connection between crew management and service control. However, given this strong connection, it is often impossible to attribute an intervention solely to service management or solely to crew management. These two factors are often in effect simultaneously, as a change to a train's trajectory is also always a change to the trajectory of its driver.

It was also stated that gaps in the service and the evenness of headways are certainly not the only timetable deviations which controllers are concerned about; lateness of trains and wrong train sequences are just as important. Although level of service considerations, rolling stock management and safety issues may generally help explain why controllers intervene to correct for these deviations, the issue of crew management is virtually omnipresent in their decisions, as not only the times at which drivers pass their relief locations are important, but also their sequence (because driver reliefs are based on train numbers, not on the drivers stepping off, and relief drivers orient themselves by the sequence of trains passing the relief location). Furthermore, it could be seen from the list presented in section 4.4 that many of the factors which can trigger service control interventions (such as rolling stock management issues) are not necessarily related to disruptions, i.e., there is evidence that the reliability of the system depends on factors that have hardly been recognized so far, let alone been monitored or modeled. In performing interventions to account for these issues, controllers can actually *cause* gaps in the service and delays which might likely require remedies at a later stage. In other words, service control not only "repairs" unreliable service, it can also cause unreliable service. The multifaceted decision environment also may help explain why the Central line controllers have come to disregard the specification of the cause code in the OCS – the list provided by the drop-down menu only contains causes of disruptions which might then in turn cause delays and timetable deviations. This list only contains the primary causes of a limited set of interventions which immediately follow a disruption.

4.5 Relevance to other metro lines and systems

As the entire description of the service controller's decision environment so far has been focused on the Central line, this section explores how transferable the findings are to other metro lines.

4.5.1 Special characteristics of the Central line

The author believes that several characteristics of service control on the Central line are caused by the complexity of the line and especially the complexity of scheduled traffic patterns on it:

- A complex service is naturally more difficult to manage than a simple line, and the author believes that the limitations of human processing capability are a strong driver behind the Central line controller's emphasis on schedule adherence. Managing the line when, for example, every train is running either off schedule or on an unscheduled trajectory would hardly be feasible for controllers, even if the service to passengers (in terms of headways and destinations served) were very good. Therefore, the working timetable must serve as a reference point, even under circumstances in which its optimality is debatable.
- The line's layout on the east, with long branches, a loop structure and a low-demand segment, makes service control on that side of the line complicated. Therefore controllers employ different service control and recovery strategies on the east and on the west ends of the line.
- The variety in control interventions and control strategies available to deal with a specific problem on the line is in part a result of the layout of the line and of the OCS. In combination with the fact that there is little formal controller training and few official guidelines, this causes *controller variability,* i.e., variability in the way different controllers deal with the same disruption or timetable deviation.
- In section 3.1.3, an explanation was given as to why there is a preference at the London Underground to create timetables on branching lines as a combination of "self-contained services". However, this section has shown that this assumption (Allen, 1981) does not always hold. On a line where the self-contained services share trains and drivers, they will hardly be independent from the point of view of service control. Furthermore, on a high-frequency line, it is not true that delays on one of the self-contained services will not affect the others, as train impedance on the trunk portion, congestion at shared terminals and trains or drivers which are scheduled to switch between these services will necessarily transfer the

delays onto the other self-contained services. Pending further research, the author would like to put a question mark behind the principle stated by Allen.

4.5.2 General applicability

First and foremost, the modern and largely automated signaling system and OCS is a characteristic of the Central line which cannot be found in many other lines of the London Underground. This is true to a lesser extent for other metro systems across the world, but it can be expected that in the future, more and more metro systems will be moving to similar or more advanced signaling systems and OCS as the Central line. Generally speaking, the amount of resources (in terms of time and control personnel) consumed by service control tasks depends on the degree of automation of the system. The specific decision factors presented above are relevant not only for the Central line, but, albeit to varying degrees of importance, also for other metro lines.

As stated in section 4.3.1, uncertainty permeates all aspects of service control, and it was seen that there are two main issues. On the one hand, in the case of disruptions, controllers need to make decisions very quickly with the (uncertain) information available at that moment, and on the other hand, there is always uncertainty about whether an intervention will have the desired effect. The latter is, to a large degree, a function of the reliability of the rail line's equipment (e.g., train radios), the design of the OCS (e.g., how easy it is for controllers to hold a train) and the discipline of frontline staff. These are also the points at which a transit agency can have the most influence, although many of the possible improvements are capital-intensive.

Crew management, as highlighted in section 4.3.3, is a major concern in any metro system with drivers; leaving it out of the equation simply yields an incomplete picture of service control. Whether or not the workforce is unionized, drivers are always subject to driving time regulations, and furthermore, driver overtime due to lateness and the provision of a pool of spare drivers is a large cost factor for agencies, which they naturally attempt to minimize. Rolling stock management, as shown in section 4.3.4, is not nearly as complex as crew management, but is based on a set of considerations, such as stock balance, which apply more or less universally beyond the Central line. The example of the Central line has also shown that the complexity of the crew and rolling stock management tasks increases strongly with line length, running times and the number of crew and train depots along the line.

Two further issues are strongly influenced by the transit agency and its policies: level of service and safety. With respect to the level of service, the study of the Central line has shown that in the absence of an official agency policy, controllers will define their own understanding of what constitutes "good" service – in terms of schedule adherence, headway regularity, traffic patterns etc. However, these unofficial policies are not driven by binding, higher-level constraints; hence, an agency's management level has considerable leverage in this respect. The same holds for maintenance crew safety policies. The objectives with regard to passenger safety stated by the Central line controllers appear to be more or less universally applicable.

The infrastructure maintenance requirements described in section 4.3.8 are highly line-specific and only of importance because a large section of the Central line is above ground. However, the point which this illustrates is that by incorporating these requirements into the operations plan, it may be possible to limit the incurred delays and service control interventions. Similarly, infrastructure capacity constraints depend strongly on the line layout. While junctions and terminals are known potential bottlenecks, the example of the Central line has shown that reversing facilities along the line can become tight constraints on the ability of controllers to manage a line during disruptions.

As a last observation, an issue which was not mentioned by Central line controllers but is known to form part of the service control decision environment in other metro systems is energy management. Energy costs generally form a large enough part of a line's operating cost to be of concern to the transit agency. The most relevant decision a service controller would need to make would be about an early or late shut-off of traction current on an individual section of the line.

5 Reconstruction of operations and identification of interventions

There is clear value in improving the flow of information from a transit system's operational level to the planning and management levels, thus closing the feedback loop shown in Figure 1-1. Furthermore, the description of the service controller's decision environment using the example of the Central line has shown that this environment is significantly broader than has been acknowledged in academic literature to date, namely that service controllers deal with many issues which are of direct interest to operations planners. While it is clear that great value can be gained from speaking directly to service controllers about their decisions, an analyst first needs to work on thoroughly understanding the operations on a line, identifying service control interventions and getting a preliminary idea of how the two interact. The right pieces of information (e.g., logs of control interventions for service recovery), placed in the hands of service controllers or their managers, can then help the analyst complete the picture and understand the reasons behind specific interventions through direct discussions with controllers.

Manual logs of service control are likely to omit some of the most interesting cases or may be biased, as stated in section 1.3, so there is a need for automated reconstruction of operations and service control. This chapter describes the data sources available for the Central line and an algorithm to produce the data needed for analysis. This serves two purposes. First, to introduce the dataset with which the analyses of chapter 7 were conducted, and second, to demonstrate what the data needs are and how a relatively simple algorithm can be used to produce the necessary information.

This chapter first introduces the data sources which were available for reconstructing operations on the Central line, and it will show which elements were found to be the most important for that task. It then describes the methodology which was used to assemble the various data sources into one coherent dataset and to infer service control interventions from that dataset. Since the structure of the algorithm which was developed for that purpose directly follows from the methodology, this chapter describes the data processing in terms of the three most important steps of the algorithm and explains how to infer service control interventions with the help of the pseudocode for the algorithm. The reader will find that this analysis is strongly focused on service control interventions affecting the routing of trains, and that crew-related interventions are not examined in depth; the reason for this were issues with the data quality, which will be further explained in section 5.3.

5.1 Data sources

The Central line's automatic signaling system was provided by Westinghouse signals. At its heart is a set of central signaling servers which comprise the operations control system (OCS) and are located in the service control center at Wood Lane. The servers receive a continuous data stream from wayside signaling equipment and onboard computers, including information on track occupancies, signal and switch status and train information (identification numbers, measured loading factors[5]). It maintains an internal schedule which is based on the working timetable, as well as lookup tables of dwell times and running times for all Central line stations and track sections. Once the working timetable is read into the OCS, it is recalculated with the signaling system's own set of running times. In the course of the data assembly which will be described in section 5.2.1, it was found that the internal timetable differed from the working timetable by at most a few seconds, if it had not been changed manually by service controllers. The OCS matches up trains on the line with its internal timetable by set number and trip number, and uses the scheduled trajectory of a train to automatically set its road[6] by setting switches as needed and clearing signals for its passage. If a controller edits the departure time at one station for a scheduled trip or inserts an unscheduled trip, the train trajectory is automatically recalculated based on the programmed running times and dwell times; the internal timetable is dynamically updated to reflect the most recent changes coded by controllers.

At the end of a service day, the OCS server saves 19 data files (the so-called DMA[7] data) onto a hard drive. Appendix C provides a list of all available data files along with a brief explanation of their content where known. However, as the responsibility for the signaling infrastructure changed hands twice between 2000 and 2009 – from London Underground to a public-private partnership (PPP) contractor and then de facto back to London Underground as the private contractor went bankrupt – much of the knowledge about the data has been lost. As a result and because there is very limited available documentation, the meaning of much of the data remains unclear. For the analyses presented in this thesis, the DMA data files which were used were the so-called TDA ("Train Data") and TDL ("Traffic day") files, which are by far the most important files produced by the OCS. These are the only two data files which are regularly downloaded from the signaling server and permanently stored by a London Underground technical specialist. The other files are eventually

---

[5] The Central line trains (1992 tube stock) report their weight as a percentage of their crush load in real-time to the OCS (explained in detail later in this section).

[6] In London Underground terminology, a "road" is a sequence of track blocks which a train traverses between an origin and a destination.

[7] Originally: Data Management Application.

deleted from the server. While the meaning of the majority of data fields in the TDA and TDL files could be reconstructed by using the documentation and by inference from their contents, there remain several fields of data which could not be identified. The following sections describe these two files in more detail. All information contained in the TDA and TDL files is available to the Central line's service controllers in real-time.

Since the TDA and TDL data are used to reconstruct service control interventions, as described later in this chapter, it is necessary to describe how Central line controllers can make changes to the service, since these are the changes which then appear in these two files.

Central line controllers have several options available to make any change. For any short-turn, withdrawal, diversion or extension of a train which is running on the Central line with a particular set/trip number, a controller can:

- Edit the internal timetable of that set number/trip number and insert a new origin or destination.
- Renumber the train, in which case the scheduled origin and destination of the "new" set/trip number will be applied to that train.
- Insert a new set/trip number which is not in the timetable and specify its routing.

5.1.1 The TDA file

The TDA file is an event-driven log containing all train arrivals and departures at stations recorded throughout the day. Table 5-1 shows the data columns. The TDA file records train movements on the line and the shunt tracks adjacent to depots (i.e., tracks used for train movements between the line and the depots), but no movements are recorded within depots. However, movements on the shunt tracks are recorded as movements in the adjacent station, thus often creating multiple entries at such stations which only differ by timestamp.

The first data item, the expected departure time (EXPECTED_TIME), corresponds to the internal timetable referred to earlier in this section. The OCS calculates the observed train arrival and departure times by adding or subtracting 8 seconds to track occupancy/unoccupancy records (when the system was designed, 8 seconds were determined to be the average time needed by a train to pull in to or pull out of a station when under automatic operation). If a train entered and departed a station moving in the same direction (i.e., did not reverse), the OCS later assembles the arrival and departure logs with the two respective timestamps into one record. If a train reversed at a station, the arrival and departure are recorded separately, each with a single timestamp.

The routing information shows the scheduled origin and destination of the train (TRAIN_FROM, TRAIN_DEST) as found in the timetable for its set number/trip combination. If a controller edits the train's routing, then this will appear as "diverted from" (TRAIN_DIV_FROM) and "diverted to" (TRAIN_DIV_TO). However, if a train's destination changes by virtue of it being renumbered or receiving an unscheduled number, this generally will not appear in the "diverted from/to" fields of the TDA data, hence the remark in the table. While the train origin/destination is an internal variable, the "Passenger destination" (PAX_DEST) or, if non-null, "Passengers diverted to" (PAX_DIV_TO) is the destination indicated on the destination displays in stations. The train lateness is calculated as the difference between the expected departure and the recorded departure and displayed in seconds.

Field name	Content
EXPECTED_TIME	Planned departure time according to OCS
ARRIVAL_TIME	Observed train arrival time
DEPARTURE_TIME	Observed train departure time
STATION	Station
TRAIN_FROM	Train origin (correct if not renumbered)
TRAIN_DIV_FROM	Train diverted from (correct if not renumbered)
TRAIN_DEST	Train destination
PAX_DEST	Destination indicated to passengers
TRAIN_DIV_TO	Train diverted to (correct if not renumbered)
PAX_DIV_TO	Passenger destination if train diverted
TRAIN_SET_NO	Train set number
TRAIN_TRIP_NO	Train trip number
TRAIN_TRIP_INSTANCE	(Unknown)
SYSTEM_TRAIN_REF	(Unknown)
LATENESS	Lateness with respect to expected departure time
LOADING	Train loading
DIRECTION	Direction of travel
CAUSE_CODE	Cause code for manual changes to train schedule
REGULATION_DELAY	(Unknown)
TIMETABLE	Timetable number
ADDITIONAL_TRIP_IND	(Unknown)
PLATFORM	Platform number

Table 5-1: Data columns of the TDA file

Train loading (LOADING) reflects the train weight as measured at the train's axles before departure. It is displayed as a percentage of the crush load weight; it increases in steps of 3% and is cut off at 100%. It is not clear how the weight measurement was calibrated (i.e., what passenger numbers the weights translate to) and how reliable they are. The last data field of interest is the cause code (CAUSE_CODE). Whenever a controller makes a change to the timetable, the reason for that change (and therefore for the change in train routing) must be selected from a drop-down list. However, as the author's observations showed, these often appear to be selected at random and must be regarded as meaningless. The table also shows four fields (TRAIN_TRIP_INSTANCE, SYSTEM_TRAIN_REF, REGULATION_DELAY and ADDITIONAL_TRIP_IND) whose meaning could not be identified, as discussed earlier in this chapter.

5.1.2 The TDL file

Unlike the TDA file, the records of the TDL file are time-driven. Every thirty seconds, a "snapshot" of the line is created which records the position of every train along with the supplementary information shown in Table 5-2. The train position is stored as a track circuit[8] number along with the time at which the train crossed the track circuit boundary. The track circuit is presumably the one occupied by the first car. As with the TDA file, movements in depots and sidings are not recorded.

Several of the data items recorded in the TDA file are also found in the TDL file. However, the TDL file contains two additional, important pieces of information: the train unit, which is identified in the TDL file by the numbers of the four permanently coupled two-car sets it is composed of (TRAIN_UNIT_NO1 through TRAIN_UNIT_NO4), and the number under which the driver is logged in (TRAIN_CREW_NO). While writing the data assembly algorithm described in section 5.2.1, it was found that the data in the TDL files was more error-prone than the data in the TDA files.

5.1.3 The timetable and crew schedule

Further data sources used were the electronic versions of the working timetable and the crew schedule. The timetable comes in a format very similar to the TDA file. It specifies the departure time from every station for every train set/trip number, as well as the train's scheduled origin and destination. Furthermore, the timetable shows any restrictions that apply to a train (such as "run only when required") and whether the scheduled movement is in or out of passenger service.

---

[8] A track circuit is a wired circuit spanning the length of a block in the fixed-block signaling system. It is used to detect the presence of a train and to transmit information identifying it.

The crew schedule could only be obtained for weekdays. For every piece of work on board a train, the crew schedule indicates which crew number fulfills it as well as the boarding and alighting locations and times of that driver. It was originally intended to compare the crew schedule with the recorded crew IDs in order to identify crew management decisions. However, as discussed in section 5.3, doing so proved difficult.

Field name	Content
TIMESTAMP	Timestamp of line snapshot
TIME_TC_ENTERED	Time when train entered track circuit
TRAIN_SET_NO	Train set number
TRAIN_TRIP_NO	Train trip number
TRAIN_UNIT_NO1	Carriage number 1
TRAIN_UNIT_NO2	Carriage number 2
TRAIN_UNIT_NO3	Carriage number 3
TRAIN_UNIT_NO4	Carriage number 4
TRACK_CIRCUIT_NAME	Track circuit number
HIGH_LEVEL_TRAIN_POS	Nearest station
DESTINATION_ABBR	Current train destination
TRAIN_ROUTE_QUALIFIER	(Unknown)
IN_OUT_SERVICE_CODE	Train in / out of service
LATENESS	Lateness at last station departure
REGULATION_DELAY	(Unknown)
TRAIN_CREW_NO	Train crew currently logged in
TRAIN_MODE_MEA	Train driving mode
DEICING_STATE_MEA	Status of train's deicing equipment
TRAIN_LOADING	Train loading
RADIO_CHANNEL	Train radio channel
CREW_ID_PLANNED	(Unused)

Table 5-2: Data columns of the TDL file

5.2 Methodology

In the work presented here, the objective in reconstructing the operations on the Central Line was the creation of a dataset which would allow the most accurate identification of service control interventions. Furthermore, it should allow for forensic analyses of the operations on any particular day of interest. It was determined that the best procedure was to create a new dataset which matched up the timetable and crew schedule with the DMA data, thus allowing the user to quickly identify discrepancies between the timetable, crew schedule and the DMA data. The VBA program which was written for this purpose incorporates an algorithm, developed over the course of several months, which allows the creation of this combined dataset for any given day for which at least the DMA and timetable data are available (the use of the crew schedule is not imperative for the reconstruction of interventions to the train service). The resulting dataset is structured similarly to the timetable since that allows the most direct comparison, i.e., it features all arrivals and departures, both scheduled and observed, at the station level.

Table 5-3 shows an excerpt from the produced dataset, in which all timetable and DMA records which belong together are matched up, but timetable and DMA records without a counterpart are also inserted into the joined dataset. The column numbers in the first row are provided for reference. In this example, one can see the stations at which train set #41, trip #5 (as shown in columns 1 and 2) to Northolt (columns 15 and 16) stopped on its way through central London. Only the section between Liverpool Street and Holland Park is shown in this excerpt (the sequence of stations is in column 3). The train is running approximately 23 minutes late (column 19, indicated in seconds) and is being short-turned at North Acton (columns 17 and 18). The scheduled departure (column 8) is taken directly from the timetable, whereas the expected departure (column 12) has been calculated by the OCS based on its lookup table for running times.

Working timetable								Crew schedule	TDA file													TDL file	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
SET NUMBER	TRIP NUMBER	STATION	PLATFORM	ARRIVAL MODE	DEPARTURE MODE	SCHEDULED ARRIVAL	SCHEDULED DEPARTURE	SCHED CREW	EXPECTED DEPARTURE	ARRIVAL TIME	DEPARTURE TIME	ORIGIN	DIVERTED FROM	DESTINATION	PASSENGER DEST	TRAIN DIVERTED TO	PAX DIVERTED TO	LATENESS	LOADING	DIRECTION	CAUSE CODE	CREW ID	UNIT ID
41	5	LIS	5	P	P	08:48:25	08:48:45	409	08:48:44	09:09:38	09:10:48	HAI		NOR	NOR	NOA	NOA	1324	100	W	SIG	401	106
41	5	BAN	5	P	P	08:50:40	08:51:00	409	08:51:00	09:12:16	09:13:10	HAI		NOR	NOR	NOA	NOA	1330	97	W	SIG	401	106
41	5	STP	1	P	P	08:52:25	08:52:45	409	08:52:44	09:14:38	09:15:22	HAI		NOR	NOR	NOA	NOA	1358	94	W	SIG	401	106
41	5	CHL	1	P	P	08:54:25	08:54:45	409	08:54:44	09:16:44	09:17:28	HAI		NOR	NOR	NOA	NOA	1364	74	W	SIG	401	106
41	5	HOL	1	P	P	08:55:40	08:56:00	409	08:55:56	09:18:12	09:18:52	HAI		NOR	NOR	NOA	NOA	1376	65	W	SIG	401	106
41	5	TCR	1	P	P	08:57:25	08:57:45	409	08:57:44	09:20:06	09:20:54	HAI		NOR	NOR	NOA	NOA	1390	48	W	SIG	401	106
41	5	OXC	1	P	P	08:58:55	08:59:15	409	08:59:08	09:21:46	09:22:22	HAI		NOR	NOR	NOA	NOA	1394	42	W	SIG	401	106
41	5	BOS	1	P	P	09:00:10	09:00:30	409	09:00:30	09:23:16	09:23:52	HAI		NOR	NOR	NOA	NOA	1402	35	W	SIG	401	106
41	5	MAA	1	P	P	09:01:40	09:02:00	409	09:01:56	09:24:44	09:25:10	HAI		NOR	NOR	NOA	NOA	1394	26	W	SIG	401	106
41	5	LAG	1	P	P	09:03:25	09:03:45	409	09:03:44	09:26:44	09:27:06	HAI		NOR	NOR	NOA	NOA	1402	19	W	SIG	401	106
41	5	QUE	1	P	P	09:05:25	09:05:45	409	09:05:42	09:28:36	09:29:00	HAI		NOR	NOR	NOA	NOA	1398	16	W	SIG	401	106
41	5	NHG	4	P	P	09:06:55	09:07:15	409	09:07:10	09:30:04	09:30:34	HAI		NOR	NOR	NOA	NOA	1404	13	W	SIG	401	106
41	5	HOP	1	P	P	09:08:25	09:08:45	409	09:08:44	09:31:42	09:32:06	HAI		NOR	NOR	NOA	NOA	1402	10	W	SIG	401	106

Table 5-3: Excerpt from the aggregated dataset with the respective sources

It is important to note that such a dataset can serve many purposes beyond the ones for which it will be used here, as has been demonstrated by previous research, for instance by Rahbee (2001) and Dixon (2006). As an example, DMA data can be used on its own to determine running time and dwell time statistics as a function of time of day (and indirectly or directly as a function of demand, depending on whether demand data are available). However, there is much added benefit in comparing the DMA data with the schedule and thus comparing the observed state of the system with its target state at any point in time. This will be further discussed in section 5.2.2.

5.2.1 Assembly of the dataset

The following sections describe the algorithm as a set of steps which are taken to convert the raw data into a coherent, joined dataset and then infer service control interventions from it.

Step 1: Import and extract DMA data

In the first step, the raw TDA and TDL files are imported, provided with headers and the timestamps are transformed into a standard format. At this point, the data in the two datasets are not usable as they contain a large number of corrupted, incomplete and duplicate records. After removing corrupted records which cannot be used for the analyses, the algorithm focuses on two main tasks:

- As mentioned above, movements between stations and depots generate multiple records in the TDA file – although movements inside the depots are not recorded, movements on the so-called "depot roads" are. For instance, a train which arrives at Hainault station, then moves onto a depot road, reverses and pulls into the depot typically generates three records at the location "Hainault" which differ only by timestamp. By inference from the surrounding records, the algorithm attempts to identify such movements and deduce which of the records were created in the station and which were created in the depot roads. It then corrects the location labels and removes intermediate records created by shunting moves.
- The train unit identification numbers in the TDL file, which are pivotal for the identification of service control interventions, were found to be particularly error-prone. Every Central line train unit is a combination of four two-car sets which are permanently coupled. Each of these two-car sets has a unique identification number which is shown in the TDL file, and one train unit is a combination of four such identification numbers. The algorithm scans the TDL file and creates a list of known combinations, each of which is assigned a new, unique number identifying that particular train unit (henceforth referred to as the "unit ID"). The list is

dynamically updated to correct for errors and allow the identification of train units in spite of individual car ID numbers being missing or incorrect, and the original TDL file is augmented by one column which contains the new unit ID.

The aforementioned missing records and the exclusion of corrupted records (as mentioned at the beginning of step 1) may lead to a slight, though not serious, overcounting of total interventions. In the course of the analysis, it was observed that corrupted records were usually dispersed throughout the dataset, without any noticeable accumulation at any point. The algorithm for identifying service control interventions, which will be described in detail in section 5.2.2, is not sensitive to missing and corrupted records as long as they do not occur at the end of a train trip, i.e., at reversal. If that is the case, the algorithm may wrongly identify the disappearance of a train from the data as a service control intervention (e.g., a withdrawal) whereas in reality, the records of further movements on the line were corrupted. From experience in working with the data and analyzing program output, the author believes that these occasional errors are too small to compromise the quality of the output. However, an analyst using the program for a microscopic analysis should always check the recorded interventions for plausibility, as such misidentifications often involve train movements which are not possible given the track layout and are therefore easy to identify.

Step 2: Import the timetable and duty schedule

The import of the timetable is a straightforward task as the only changes which need to be made are the adjustment of time formats and station abbreviations to match those used in the DMA data. The timetable is augmented by one column which features the scheduled crew ID, and the respective crew number is added in from the duty schedule.

Step 3: Match up the TDA, TDL and timetable data

The third step consists of matching up the three data sources. For this purpose, an empty data table (referred to as the joined table) is created which contains columns for timetable data, TDA data, observed crew and unit IDs (from the TDL data) and the scheduled crew (within the timetable data). It is first filled with the complete TDA data, and then scanned through. For every entry, the algorithm searches for the correct timetable and TDL records. These are matched using the set and train number, as these are the primary identifiers which assign a train to a certain trip in the timetable. In the timetable, the algorithm searches for the unique combination of set/train number and station to match that in the TDA record. If a record is found, its information is copied into the joined table, otherwise those fields remain blank.

A similar procedure is followed for finding the crew and unit ID from the TDL data, except that the matching cannot be done through the station code since the TDL data is time-driven and does not necessarily contain an entry while a train was berthed at a station. Therefore the algorithm searches for the last TDL record within 120 seconds before the latest timestamp (either arrival or departure) of the TDA record. Finally, all timetable records which could not be matched with TDA records are copied into the joined table – these are potentially parts of the service which were not run due to service control interventions.

5.2.2 Identification of service control interventions

Before moving to steps 3 and 4 of the algorithm, this section provides explanations and definitions which are useful for understanding those last two steps. Steps 1 and 2 showed how the joined dataset is assembled, so it can be analyzed to identify service control interventions. The focus is on interventions which are clearly identifiable as deviations from the timetable; they are listed in Table 5-4 along with a brief explanation. The research objective for this thesis is to focus broadly on a set of interventions which are often used as part of disruption management and recovery. Of the service control interventions described in section 2.3, the two most prominent ones which are missing from Table 5-4 are expressing and holding trains. While expressing is not performed on the London Underground as a matter of principle, holding is a common strategy. However, more detailed information than is available would be required for discerning trains held by controllers from trains which experienced long dwell times for other reasons. The only exception is when trains are stopped at stations upstream of a blockage, in which case it can safely be assumed that the trains are being held due to the blockage and they can be identified manually.

Before proceeding to the identification of interventions, common definitions were needed since, to the author's knowledge, there is no universal definition of the interventions in question. The algorithm follows the definitions presented in section 2.3 and Table 5-4. In addition to the basic information on the routing of the train, two additional data items were extracted: where and when the first intervention was coded into the OCS by a controller and at what point in time the intervention first had an effect on the service. Table 5-5 shows how that time was determined as a function of the intervention type.

Intervention		Description
Short-turned trip	-	*As described in section 2.3.1; train remained in service*
Diverted trip	-	*As described in section 2.3.1; train remained in service*
Extended trip	-	*As described in section 2.3.1; train remained in service*
Withdrawn en route	-	*Train ran only part of its schedule trip, was withdrawn into a siding or depot*
Diverted and withdrawn	-	*A diversion followed by a withdrawal into a siding or depot*
Extended and withdrawn	-	*An extension followed by a withdrawal into a siding or depot*
Reversed at trip origin	-	*The train departed in the opposite direction from where it was scheduled to go*
Trip cancellation	-	*As described in section 2.3.1*
Unscheduled trip	-	*As described in section 2.3.1*
Out of service trip	-	*As described in section 2.3.1*
Renumbering	-	*The train received a new set/trip number*
Reverse renumbering	-	*The train was renumbered again, to its original set/trip number*
Unidentifiable re-routing	-	*Error code*

Table 5-4: Service control interventions identified in the DMA data analysis

Intervention	First effect on service
Short-turns	When the train reversed
Withdrawals en route	When the train was withdrawn
Diversions (all)	When the train passed the junction
Extensions (all)	When the train passed the originally scheduled terminus
Trip cancellations	At the scheduled departure time from the origin
Unscheduled trips, OOS trips	When the train departed its origin
Renumberings, reversals at origin and unidentifiable re-routings	N/A

Table 5-5: Time of first effect on service of an intervention

Step 4: Find canceled trains

This step requires the data to be sorted by set/trip number, and within one set/trip number, chronologically by scheduled departure time. The algorithm scans the data and identifies set/trip numbers which were not operated in their entirety. Any trip which is operated partially will count as one of the interventions described in step 5.

Step 5: Find routing changes, unscheduled trips and renumberings

It should be noted that for the identification of routing changes, unscheduled trips and renumberings, the data in the joined dataset needs to be sorted by train unit ID, and, within the data for one train unit ID, chronologically by observed departure time (or arrival time, if a record has no departure time). This ensures that the algorithm does not lose track of trains when they are renumbered by controllers, and a train can be followed throughout the day regardless of the set/trip number assigned to it. This is in contrast to a first version of the algorithm, which was based on detecting interventions in data sorted by set and trip number – in other words, it relied on the idea that trains would generally retain their set number and that following a set number would reveal routing changes. However, it emerged that renumbering trains occurred much more frequently on the Central line than originally thought and this approach often lost track of trains. Therefore, the algorithm was rewritten to use data ordered by train unit ID.

At the beginning of each trip, i.e., immediately after its pull-out from the depot or upon reversal, the algorithm stores its set/trip number and reads its scheduled destination from the timetable. It then follows the train to its next reversal and compares the final destination with the scheduled destination. If they are identical, nothing is recorded. If they differ, it then checks whether the train reversed at that point and traveled in the opposite direction in passenger service within the next 60 minutes. If yes, then the intervention is classified as a diversion, extension or short-turn. If not, the assumption is that the train was withdrawn, so the intervention will be classified as a diversion and withdrawal, extension and withdrawal or a withdrawal en route. This difference was introduced in order to get a first-order approximation of which interventions were performed to withdraw defective trains from service and which had other causes. Wherever possible, double-counting of interventions is avoided. For instance, if a train is diverted from one branch to another, reversed and run back to the trunk portion of the line with an unscheduled trip number, this was only counted as a diversion, not a diversion and an unscheduled trip.

Section 2.1 explained that trains are intended to retain their set number from the moment they pull out in the morning until they stable at night. Hence, if the data shows that at any point during the day a train unit changes its set number, it is flagged as a renumbering ("reformation" in London Underground jargon). If the renumbering was subsequently reversed and the train received its old number back, it was classified as a "reverse renumbering" (or "reverse reformation").

Train trips along the line without any matching timetable data (excluding shunting moves around depots) were classified as unscheduled trips if they were in passenger service (i.e., if a destination was displayed to passengers) and as out of service trips if the destination shown was "out of service" or "special service".

Two more intervention types were added for cases in which it was not possible to identify automatically what controllers had done with a train:

- "Unidentifiable re-routing": recorded as an error code in the joined data set. This occurred most often around the Hainault loop with its many permutations of possible train trajectories.
- "Reversal at origin": recorded if the train departed its origin in the opposite direction from where it was scheduled to go. This is usually found if the train was assigned to the wrong timetable number and controllers instructed the driver over train radio.

The full list of interventions which was identified from the data can be seen in Table 5-6, which shows a sample intervention count for one day.

Total Interventions:	**111**
Short-turned trips:	16
Diverted trips:	6
Extended trips:	11
Withdrawn en route:	2
Diverted and withdrawn:	4
Extended and withdrawn:	1
Reversed at trip origin:	0
Trip cancellations:	18
Unscheduled trips:	5
Out of service trips:	3
Renumberings:	44
Reverse renumberings:	0
Unidentifiable re-routing:	1

Table 5-6: Sample output of the intervention identification for November 19, 2008

5.3 Identification of crew movements

The reader may have noticed that the assembled dataset, which is shown in Table 5-3, features both the planned crew ID and the recorded crew ID of a train. This would, in theory, provide an opportunity to study crew movements and work on establishing an empirical link between service control and crew management, which was one of the original goals of this thesis. However, it was found that the quality of the Central line crew data reflecting actual crew movements was insufficient for such an analysis, and the correctness of the recorded crew IDs was impossible to verify.

Occasionally, drivers would forget to log in upon reversal or when stepping onto a train, making it difficult to trace where the crew relief had taken place, or they would log in incorrectly. Moreover, conversations with a DMT in January 2009 revealed that there was no consistent policy on how spare crews should log in (which can be either with their spare number, with the number of the crew they are replacing or with a fantasy number) when they step onto a train. Although the number of wrong crew data may be relatively small, it becomes nearly impossible to determine whether a deviation between the planned and the observed crew ID is due to a false login or to an actual change in crew. Therefore, it was eventually decided not to pursue that research thread.

5.4 Conclusions

This chapter has described an algorithm which assembles data from various sources into one coherent dataset which juxtaposes the operations as planned (i.e., the timetable) with the operations that were observed on a particular day. While the described procedure is specific to the Central line and its signaling system, it shows that with a relatively simple algorithm, it is possible to create a powerful dataset. The majority of problems which were encountered during the development of this algorithm were related to inconsistencies in the original data and the fact that the data labels and structures had not been designed with the idea of being used for analysis. This brings up a very important point, which emphasizes what Wile (2003) concluded, namely that before procuring an automatic data collection system (or, for that matter, any system which produces operational data), the transit agency should clearly define its needs and possible uses of that data and ensure that the output data meets those requirements.

While the methodology presented here for the case of the Central line yields a highly detailed dataset, such a level of detail is not crucial to the study of service control. An analyst working with different types of data would of course need to make adjustments according to the characteristics of the data which are available, but he or she should at least aim to reconstruct the scheduled and the observed starting and ending point of every train trip, as this allows one to identify the most important service control interventions, which are trip diversions, extensions and short-turns. Furthermore, if it is possible to determine precisely which trip in the timetable a train unit is assigned to (in this case, it could be done with the set and trip number), it should be easy to identify unscheduled trips and trips which are completely missing from the operations data (i.e., cancellations). The most important element needed for identifying reroutings is a number which uniquely identifies each train on the line, so that it can be followed throughout the day and its trajectory can be compared to the timetable. If controllers do not have the possibility to renumber trains, then the aforementioned set and trip number should suffice, but if trains are occasionally reassigned between scheduled trips, then the analyst will need to develop a method for identifying individual train units, as was done here.

6 Measuring the impact of service control

Following the identification of service control interventions as described in chapter 5, a method needs to be developed to describe the impact of service control interventions since that allows the comparison of intervention strategies with the goal of establishing which are most effective in a specific situation. This chapter starts out with a brief review of the motivation and goals behind quantifying the impacts of service control interventions. Afterwards it describes the measures which were found to work best in the case of the Central line. An analyst wishing to extend such an analysis to other metro lines will be able to draw on these measures, although it will still be necessary to consider what measures work best in the context of that specific line. Finally, possible extensions to the measures are discussed.

6.1 Motivation and fundamental questions

Measuring the impact of service control interventions is not a trivial task and the work which is presented in this thesis is merely an initial effort to do so. There are two main reasons why a transit agency should be interested in measuring the impact of service control:

- Firstly, to quantify the impact of disruptions. The London Underground is a good example in this respect, since it has contracted out its infrastructure and vehicle maintenance to private entities under a PPP (Public-Private Partnership). These contracts include penalty schemes which prescribe financial penalties for the PPP contractor in the case of service disruptions which are caused by equipment failures. The amount of the penalty assessed is a function of the total passenger delays caused by the disruption, measured in *lost customer hours* (known internally as *NaCHs* – nominally accumulated customer hours (London Underground Ltd., 2006)), which are calculated with the help of the train service simulation model (TSM). However, the TSM model does not incorporate service control. Therefore, in terms of assessing the responsibility for delays more clearly, it would be desirable to know what the initial impact of a disruption was, what delays were caused by service control interventions to deal with the disruption and possibly what the delays would have been in the absence of service control interventions.
- Secondly, as Rahbee (2001) points out and as was discussed in section 1.3, service management techniques are often passed down by word-of-mouth to new controllers, with little formal training, and can vary considerably across lines and even across shifts on the same line. Any initiative to improve service controller training should focus on choosing

> among the different "tools" for restoring the service and would need to be based on an understanding of what interventions are most effective in any given situation. A first step towards this is the specification of impact measures.

Common performance metrics for transit lines often take an aggregate view of operations – generally the performance of the line as a whole is measured over time periods ranging from one day to several months (Wilson, 2008). However, the impact of service control interventions is typically limited to specific sections of the line and to small time windows, which means that a more disaggregate view is necessary. Given this, the analyst faces a dilemma – highly disaggregate data allows for a high-quality analysis of intervention strategies but makes it challenging to analyze service control over longer periods of time. Aggregate data, on the other hand, allows longer-term assessment but may hide the effects of service control interventions; finding the right balance is not easy.

6.2 Summarizing and visualizing results and impact measures

The following three sections will briefly present service control, level of service and passenger impact measures which were found to be useful in the study of the Central line, followed by a fourth section discussing how these measures are tied together. The measures are based on the data available for the Central line and were developed in the context of the applications which are presented in chapter 7.

6.2.1 Disruptions and the service control component

If a transit agency maintains a database on disruptions, it can be used to point the analyst towards situations which may have required service control interventions. The important characteristics of a disruption are its nature, location and duration (cf. section 2.2.1). Furthermore, an important question is whether the delay experienced by the blocked train had a direct effect on other trains or not. TfL maintains a disruption database which contains individual disruption reports. Those reports are filled in manually by service controllers and contain a detailed description of the disruption along with information on the affected train, the location and the initial delay, i.e., the delay suffered by the train which experienced the disruption. Any delays from disruptions and congestion incurred during operations and the service control interventions employed to deal with those delays can be reconstructed from the operations data (known in the Central line context as DMA data, as discussed in chapter 5). Section 5.2.2 explained what types of interventions were identified from the data, so at this point the focus will only be on the presentation of those interventions.

a) The starting point: Service control intervention counts

At an aggregate level, the number of interventions per day, as shown in Figure 6-1 for a sample week (Monday through Friday), was found not only to be a good proxy for the number and severity of schedule deviations for one day, but also to give a sense of the relative workload of controllers over a period of time. This helps the analyst select the days for which a more microscopic analysis will be conducted. Such a selection process is necessary due to the large time investment required by a microscopic analysis of operational data. However, at this stage, this selection can only be based on total intervention counts since the analyst does not yet know how those interventions were distributed throughout the day. Nonetheless, with a quick visual assessment the analyst can select between days with few interventions (presumably days with good operations), "normal" days, days which appear severely disrupted and days with unusually high levels of certain types of interventions. In Figure 6-1, November 12, 13 and 14 appear to be comparable in terms of level of interventions. In this context, they could be classified as days with average levels of service control interventions. November 10 and 11, on the other hand, have high levels of short-turns and cancellations. This is indicative of a larger delay or disruption during the day. None of the five days has exceptionally high levels of interventions. Depending on his or her interests (e.g., management of larger disruptions or routine service control on "normal" days), the analyst can select a number of days for further examination. Once those days have been selected, a more detailed list as shown in Table 6-1 and Table 6-2 helps the analyst move to a more disaggregate level of analysis and determine the time windows and specific interventions of interest.

By displaying the intervention counts as a function of time, as in Table 6-1, the intervention matrix provides a quick way of presenting the actions of service controllers throughout a day, and if the times of disruptions are known, it allows a quick visual assessment of the recovery strategy which was chosen. In this example, one can see what appear to be two distinct groups of interventions: in the morning from the AM peak until approximately 10:00 or 11:00 and in the afternoon, starting around 15:00. In the morning, there appears to have been a disruption or delay westbound with controllers restoring service by short-turning several trains. The canceled trips before that do not appear to be related to the disruption; it is possible that they are due to rolling stock shortages during the peak hours.

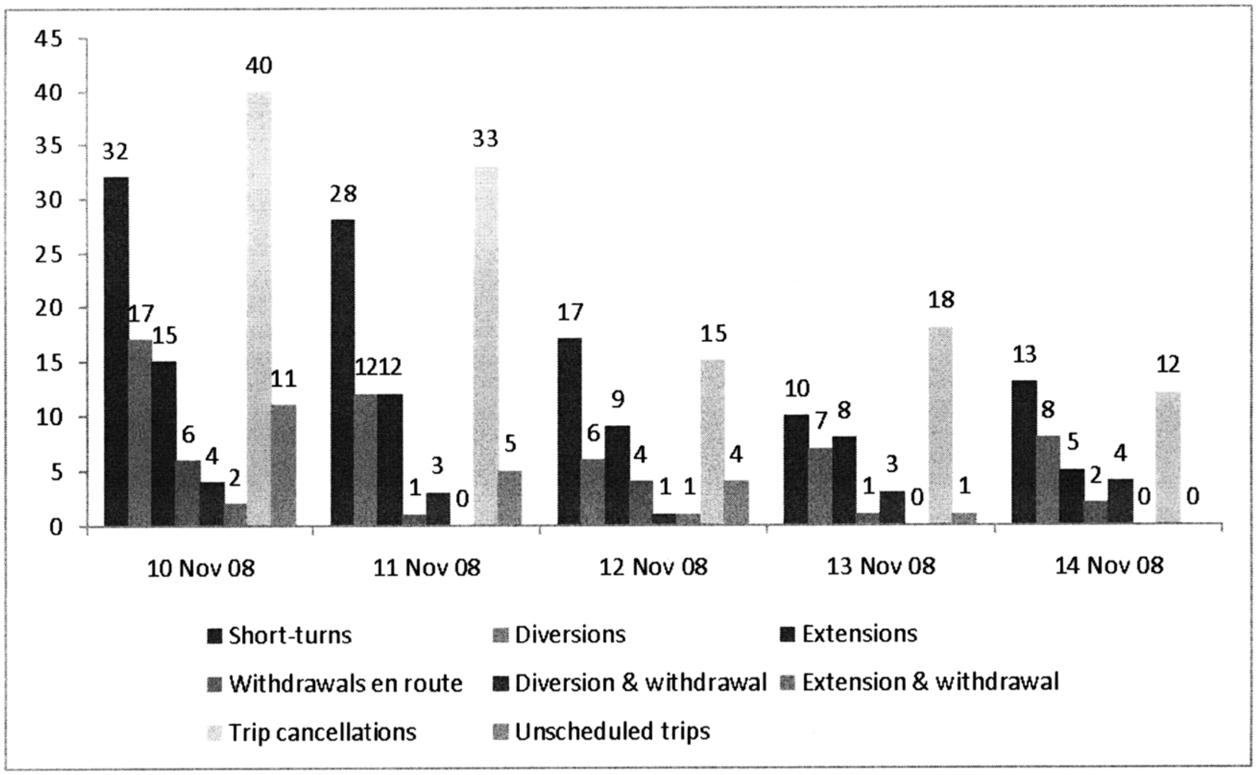
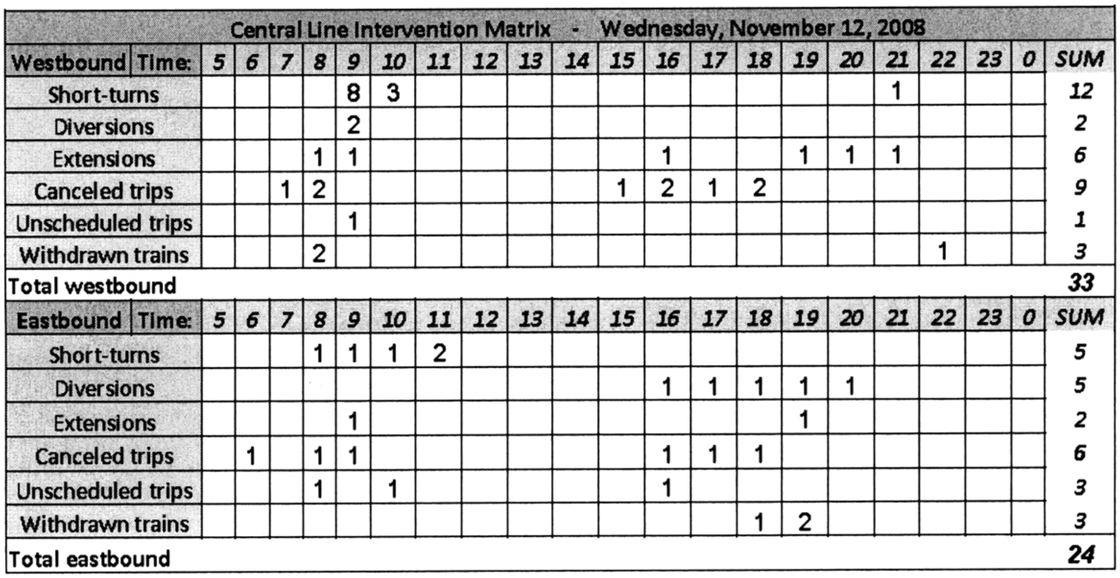


Figure 6-1: Total daily interventions by type for a week in November 2008

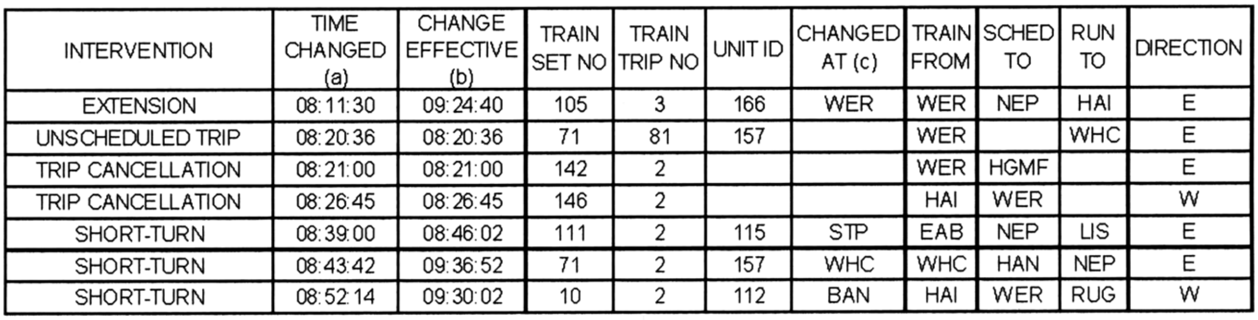
Table 6-2 shows the detail of some of the interventions between 08:00 – 10:00 : an extended trip which was originally scheduled to reverse at Newbury Park but was run to Hainault, two canceled trips and one unscheduled one between West Ruislip and Hainault. The last three interventions are short-turns which may have been performed in response to delays from a disruption. One can see that the decision to short-turn the three trains was taken within a short time, between 08:39 and 08:52. Two of the trains, units 115 and 112, had their destinations changed while they were in the middle of the line, at St. Paul's and Bank. The destination of the third one, unit 157, was already changed as it was at its origin, White City. One can also see how much time passed between the announcement of the change and the actual short-turn. Unit 115 reversed at Liverpool Street, barely 7 minutes after passengers had been informed of that change. The other two reversed close to their original destination, at Newbury Park and Ruislip Gardens.

In the afternoon, the intervention pattern shown in Table 6-1 is different: individual trains were canceled over several hours, and five eastbound trains were diverted. This suggests there were one or more trains which needed to be withdrawn since they became defective, causing controllers to cancel trips during the PM peak. A look at the list of interventions (not shown here) reveals that four of the five diverted trains were originally Hainault services, suggesting that there may have been problems in the area around Hainault, causing controllers to remove trains from that branch.



Central Line Intervention Matrix - Wednesday, November 12, 2008																						
Westbound	**Time:**	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*0*	*SUM*
Short-turns						8	3											1				*12*
Diversions						2																*2*
Extensions					1	1							1			1	1	1				*6*
Canceled trips				1	2							1	2	1	2							*9*
Unscheduled trips						1																*1*
Withdrawn trains					2														1			*3*
Total westbound																						***33***
Eastbound	**Time:**	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*0*	*SUM*
Short-turns					1	1	1	2														*5*
Diversions													1	1	1	1	1					*5*
Extensions						1										1						*2*
Canceled trips			1		1	1							1	1	1							*6*
Unscheduled trips					1		1						1									*3*
Withdrawn trains															1	2						*3*
Total eastbound																						***24***

Table 6-1: Intervention matrix for November 12, 2008



INTERVENTION	TIME CHANGED (a)	CHANGE EFFECTIVE (b)	TRAIN SET NO	TRAIN TRIP NO	UNIT ID	CHANGED AT (c)	TRAIN FROM	SCHED TO	RUN TO	DIRECTION
EXTENSION	08:11:30	09:24:40	105	3	166	WER	WER	NEP	HAI	E
UNSCHEDULED TRIP	08:20:36	08:20:36	71	81	157		WER		WHC	E
TRIP CANCELLATION	08:21:00	08:21:00	142	2			WER	HGMF		E
TRIP CANCELLATION	08:26:45	08:26:45	146	2			HAI	WER		W
SHORT-TURN	08:39:00	08:46:02	111	2	115	STP	EAB	NEP	LIS	E
SHORT-TURN	08:43:42	09:36:52	71	2	157	WHC	WHC	HAN	NEP	E
SHORT-TURN	08:52:14	09:30:02	10	2	112	BAN	HAI	WER	RUG	W

(a): Time when a controller entered the intervention into the operations control system
(b): Time when the intervention had its first effect on the service:
- In case of an extension, when the originally scheduled station was passed
- in case of a short-turn, when the train reversed
-in case of a cancellation or unscheduled trip, when the train trip started or was scheduled to start
(c): Location where the controller entered the intervention into the operations control system
Station abbreviations: BAN = Bank, EAB = Ealing Broadway, HAI = Hainault Station, HGMF = Hainault Depot, LIS = Liverpool Street, NEP = Newbury Park, RUG = Ruislip Gardens, STP = St. Paul, WER = West Ruislip, WHC = White City

Table 6-2: Excerpt of an automatically generated list of interventions for November 12, 2008

b) Impact by section

All interventions which cause a train to be rerouted or canceled can be characterized in terms of their impact on individual sections of the line. A train trip which was canceled or short-turned means that part or all of the line sections it was scheduled to serve were served by one train less. On the other hand, an extension or unscheduled trip means that a train was added to the service on some sections of the line. A diversion has both types of impacts on different line sections.

While the addition or removal of train service at individual stations would be the most precise way of looking at the impacts, the data becomes more manageable without much loss of information if it is aggregated at the section level. The sections should be chosen such that:

- The scheduled level of service is approximately constant throughout the section.
- There are few, if any, reversing tracks within the section which are known to be used by service controllers.

On complex lines, such as the Central line, these principles can still lead to a more disaggregate view than is desired by the analyst. Therefore, it was decided to partition the line into a total of eight sections, thus accepting that on some of the sections the scheduled level of service was not entirely constant throughout the section (i.e., there were scheduled reversals within some sections). Table 6-3 and Figure 6-2 show the sectioning of the Central line. A more detailed map of the sections with the station names, reversing tracks, sidings, train and crew depots is shown in Figure 7-1, page 125.

Section no.	From	To	Reversing tracks within section	Notes
1	Ealing Broadway	North Acton	--	--
2	West Ruislip	North Acton	1 (Northolt)	--
3	North Acton	Liverpool Street	2 (White City, Marble Arch)	Marble Arch reversing track only used during major disruptions.
4	Liverpool Street	Leytonstone	--	--
5	Leytonstone	Woodford	1 (Newbury Park)	--
6	Leytonstone	Hainault	--	--
7	Hainault	Woodford	--	--
8	Woodford	Epping	2 (Loughton, Debden)	--

Table 6-3: Central line sections for impact analysis

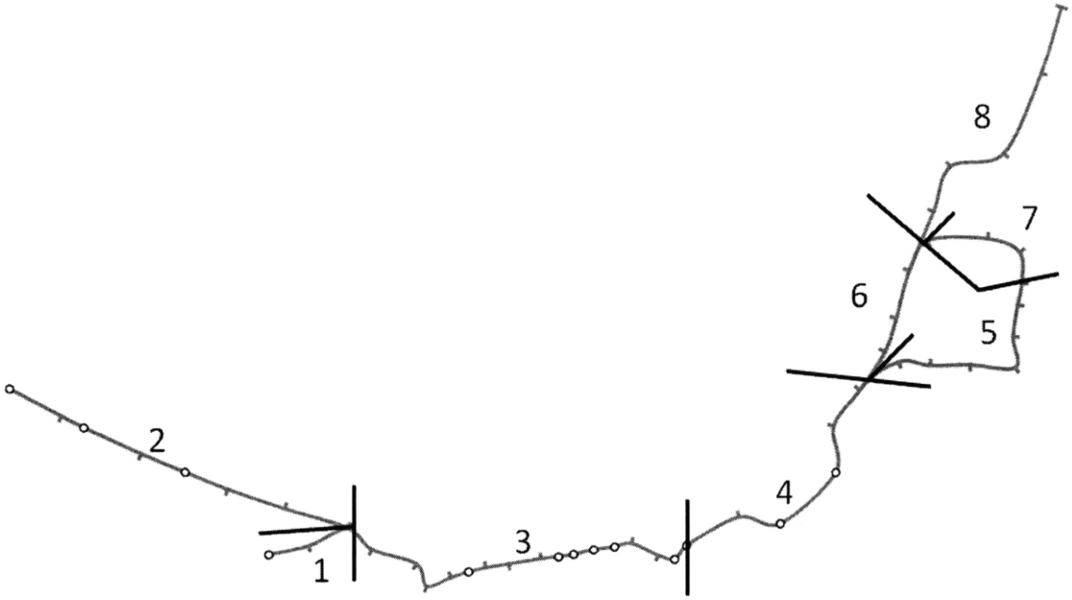


Figure 6-2: Central line sections

The intervention matrix shown in Table 6-1 can now easily be extended to show the impacts on individual sections of service removal or addition, as shown in Table 6-4 for westbound service on the sample day, November 12. As an example, one can see that between 15:00 – 16:00 two trains were canceled and one was extended. According to the matrix, the canceled trains lead to one missing trip (marked as -1) from section 8 and one missing trip from section 5, both to section 2. On the other hand, the extension occurred on section 2, where additional service was provided by that train.

As can be seen, there remain several sections with reversing tracks inside the section. The Central line presents a challenge in that virtually all reversing locations are only a few stations short of the terminal or junction; for example, White City is only two stations from North Acton junction, which is a "natural" section boundary. One had to strike a balance between creating small additional sections including only those few stations, thereby increasing the total number of sections and the complexity of the measures, or keeping some reversing tracks within a section and foregoing some detail in the analysis. Depending on the specific interests of the analyst, the partitions could be changed; for example, if the interest is to precisely attribute the impacts of disruptions to line sections, the number of sections should obviously be increased. For the purposes of this thesis, the sections in Table 6-3 were adequate, with the exception of section 6 and 8 which could have been merged since their scheduled service frequencies hardly differ[9].

Central Line Intervention Matrix - Wednesday, November 12, 2008

Westbound Time:	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	SUM
Short-turns					8	3											1				12
Diversions					2																2
Extensions				1	1							1			1	1	1				6
Canceled trips			1	2							1	2	1	2							9
Unscheduled trips					1																1
Withdrawn trains				2														1			3
Total westbound																					33
Section 1 (branch): Ealing Broadway to North Acton																					
Reduced service					-3	-2							-1				-1	-1			
Additional service					2												1				
Section 2 (branch): West Ruislip to North Acton																					
Reduced service			-1	-2	-5	-1						-2		-2							
Additional service				2	1							1			1	1					
Section 3 (trunk): North Acton to Liverpool Street																					
Reduced service			-1	-3	-6	-3					-1	-2	-1	-2			-1	-1			
Additional service				1	2												1				
Section 4 (trunk): Liverpool Street to Leytonstone																					
Reduced service			-1	-3							-1	-2	-1	-2							
Additional service																					
Section 5 (branch): Leytonstone to Woodford																					
Reduced service											-1	-1		-1							
Additional service																					
Section 6 (loop): Leytonstone to Hainault																					
Reduced service			-1	-2								-1	-1	-1							
Additional service																					
Section 8 (branch): Woodford to Epping																					
Reduced service												-1		-1							
Additional service																					

Table 6-4: Intervention matrix with effects by section for November 12, 2008

The most noticeable result of this choice of sections was that the attribution of impact, as shown in Table 6-4, became rather meaningless for section 3, since trains are often short-turned at White City, but that does not impact the trunk portion of the line with the exception of East Acton and North Acton stations. In all other line sections, given the location of reversing tracks relative to the positioning of the section, this problem had little influence on the significance of the attribution of impact.

---

[9] As the choice for these partitions was made at an early stage of the analyses and was deeply embedded in the program code, the choice was made to retain them.

6.2.2 Level of service

The level of service on any given section of the line is directly affected by service control interventions. As mentioned above, the net effect is either the removal of a train from service on that section or the addition of one. Therefore, the two variables of interest are the number of trains per hour (as this directly ties back to capacity on a high-frequency line like the Central line) and the distribution of headways. The latter can be analyzed in various ways, depending on what the analyst decides to focus on. In the case of the Central line it was initially decided to track the mean and variance of the headways.

However, in the course of the analyses, it was found that while trains per hour, mean headway and headway variance were good measures at a station level, the generalization to a line section was problematic since it involved averaging these measures over multiple stations where they were not necessarily equal. Instead, it was decided to introduce a measure which directly represented how much of the scheduled service was actually operated by refining a measure which is commonly used by the London Underground: Scheduled vs. operated service kilometers. To date, the London Underground uses this measure only on an aggregate daily level, and one of the criticisms of it is that it encourages controllers to operate services on the branches rather than on the trunk since that allows more distance to be covered in any period. It was decided to refine this measure and apply it at an hourly and section level. For every hour and every section of the line, scheduled and observed train-kilometers are calculated. The measures have the dimension of $\frac{km}{h}$. The disaggregation allows over- and underprovision of service to be pinpointed to individual parts of the line and time periods. (If such measures are adopted by the London Underground, they also have the advantage that they leave much less opportunity for controllers to "improve" the service quality measure by compensating for a lack of service on one section (e.g., the trunk) by running more service on another section, (e.g., the branches)). Furthermore, this disaggregation measures the net effect of an intervention strategy. For instance, if one train is short-turned but the one behind it is extended to cover its trip out to the branch terminal, this will show up as better service than if the first train had been short-turned without extending the following one. This is important since only looking at the intervention counts may convey the impression that larger numbers of interventions always result larger numbers of service curtailments. As the previous example shows, that is not necessarily the case.

Table 6-5 shows the scheduled weekday train service kilometers per section and per hour on the Central line under working timetable 64, which are used as the baseline for comparison with the observed service kilometers. Figure 6-3 and Figure 6-4 show the application of this measure on the line and section levels respectively for November 12, 2008, the same day for which Table 6-1 shows the service control interventions. The key to the colors is shown in Figure 6-5. Figure 6-3 reflects an aggregation over the entire line; service kilometers are only plotted as a function of time, not of line section. The number on the top of each bar indicates the percentage of scheduled service which was operated on the line during any given hour. While this makes the plots easier to read, they lose the information on the branches affected by the shortfall or overprovision. Nonetheless, they help pinpoint when problems occurred on the line. For instance, in Figure 6-3, one can see a clear drop in service between 08:00 and 09:00 westbound, which is consistent with the observations made in Table 6-4.

Figure 6-4 is the same plot as in Figure 6-3, but divided by line section. To preserve legibility, the percentages are not marked. One should also note that the scales of the y-axes vary by section. The two bar plots marked in Figure 6-4 with index (a) show the effect of the blockages which occurred westbound at 08:30 and 08:35 at Liverpool Street as well as the train bunch which passed afterwards (it appears as overprovision between 09:00 and 10:00). Furthermore, one can see the effect of the short-turns with which controllers responded to the delays, curtailing several services to West Ruislip and Ealing Broadway. This appears as a drop in the level of service on sections 1 and 2 west- and eastbound, which is indicated in the figure with indices (b) and (c). One can distinguish a gap caused by a short-turn from a gap caused by a blockage by the fact that the drop in level of service it is not followed by a train bunch, which would show up as overprovision in the following hour. Train service on the line normalized by approximately 10:00, as only slightly less train-kilometers than scheduled were operated on all sections. This is shown in the figure with indices (d) and (e).

There are two main disadvantages to these measures. Firstly, they do not capture how much service was provided for a given OD pair. Ideally, a level of service measure should be able to account for the fact that branch-destined passengers waiting for a train on the trunk section cannot take any train but must wait for one which is going to their destination. Thus, the level of service measured on the trunk section does not necessarily reflect the experience of all passengers. Secondly, it does not provide a measure for headway distribution. However, on a metro line with very high frequencies, such as the Central line's trunk section, this is not a serious problem: a delay will reduce the throughput and quickly translate into a service reduction since following trains will be delayed as well; some of them would then pass during the next time window following the delay, which will be reflected in this measure. Furthermore, if there were no train reversals within a section, the service kilometers operated are approximately proportional to the number of trains per hour on that section.

WESTBOUND	5:00	6:00	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	0:00	1:00
Section 1	13.2	26.4	33.0	33.0	33.0	29.7	29.7	29.7	29.7	29.7	29.7	29.7	39.6	34.8	31.2	24.9	19.8	21.3	19.8	16.5	0.0
Section 2	2.5	53.2	92.6	143.3	157.1	143.8	143.8	143.8	143.8	143.8	141.3	128.4	137.1	144.0	142.6	111.2	81.5	79.7	79.7	64.6	7.0
Section 3	27.3	171.6	332.0	412.3	384.3	331.3	331.3	331.4	331.2	331.3	355.7	370.7	376.3	372.4	295.5	230.6	203.5	174.6	170.4	109.7	0.0
Section 4	32.8	173.9	290.6	313.8	261.5	251.0	251.0	251.0	251.0	251.0	282.4	282.4	279.6	257.8	211.5	158.5	151.5	125.5	125.5	33.7	0.0
Section 5	17.1	45.4	66.7	70.1	54.8	56.0	56.0	56.0	56.0	56.0	58.9	62.5	65.4	60.7	56.0	37.4	28.0	28.0	26.2	14.0	0.0
Section 6	34.5	97.0	140.2	139.0	113.4	106.8	106.0	106.8	106.8	111.9	125.4	123.7	116.9	94.2	66.5	59.2	58.0	58.0	56.0	1.7	0.0
Section 7	0.0	20.9	28.1	33.0	18.2	18.2	18.2	18.2	18.2	18.2	26.8	19.6	23.2	18.2	18.2	18.2	18.2	18.2	18.2	6.1	0.0
Section 8	48.5	88.4	113.9	113.4	116.2	121.1	121.1	121.1	121.1	121.1	123.4	119.5	113.2	121.1	111.3	78.9	72.4	72.4	77.0	68.3	21.2

EASTBOUND	5:00	6:00	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	0:00	1:00
Section 1	9.9	23.1	33.0	31.2	34.8	27.9	29.7	29.7	29.7	29.7	29.7	29.7	38.1	36.3	33.0	24.6	21.6	19.8	19.8	19.8	3.3
Section 2	39.4	118.8	140.5	120.4	136.1	145.4	143.8	143.8	143.8	143.8	145.9	144.9	149.8	133.4	107.5	79.7	81.7	79.7	74.2	8.1	0.0
Section 3	46.4	199.3	305.0	371.8	376.0	351.2	331.3	331.3	331.3	331.3	336.0	373.7	412.3	373.2	306.1	236.0	203.5	174.0	170.4	76.9	6.3
Section 4	2.3	109.7	204.9	276.4	282.4	275.7	251.0	249.4	252.7	251.0	251.0	268.2	305.6	290.6	254.8	196.5	156.9	139.7	125.5	91.9	0.0
Section 5	7.5	22.1	46.7	60.7	56.0	54.5	55.8	56.0	56.0	56.0	54.5	56.0	68.8	62.0	51.4	34.5	28.0	28.0	28.0	28.0	0.0
Section 6	0.0	33.2	75.4	118.3	137.0	132.2	111.6	106.8	106.8	106.8	106.8	116.9	129.7	136.5	138.5	111.8	72.7	75.6	58.0	56.8	6.7
Section 7	0.0	12.1	18.2	18.2	18.2	17.1	18.2	18.2	18.2	18.2	16.9	20.7	18.2	18.2	17.1	18.2	18.2	18.2	13.3	6.1	0.0
Section 8	60.8	82.5	101.2	113.2	121.1	122.6	121.1	121.1	121.1	121.1	121.1	111.8	119.6	129.4	100.8	83.3	69.9	76.8	72.4	68.3	12.1

Table 6-5: Scheduled weekday service kilometers on the Central line

On low frequency sections, such as the Central line's Hainault to Woodford service which has scheduled 20 minute headways, the assumption that a gap or delay would immediately be visible in the operated train kilometers does not hold true. Even if a train experienced a large delay (e.g., 15 minutes), this would not necessarily affect other trains, and one would still observe the scheduled throughput (six complete trips per hour, three in each direction). The analyst would therefore be led to believe that perfect service was operated on that section, even though the three trains in each direction may have passed in a bunch with short headways.

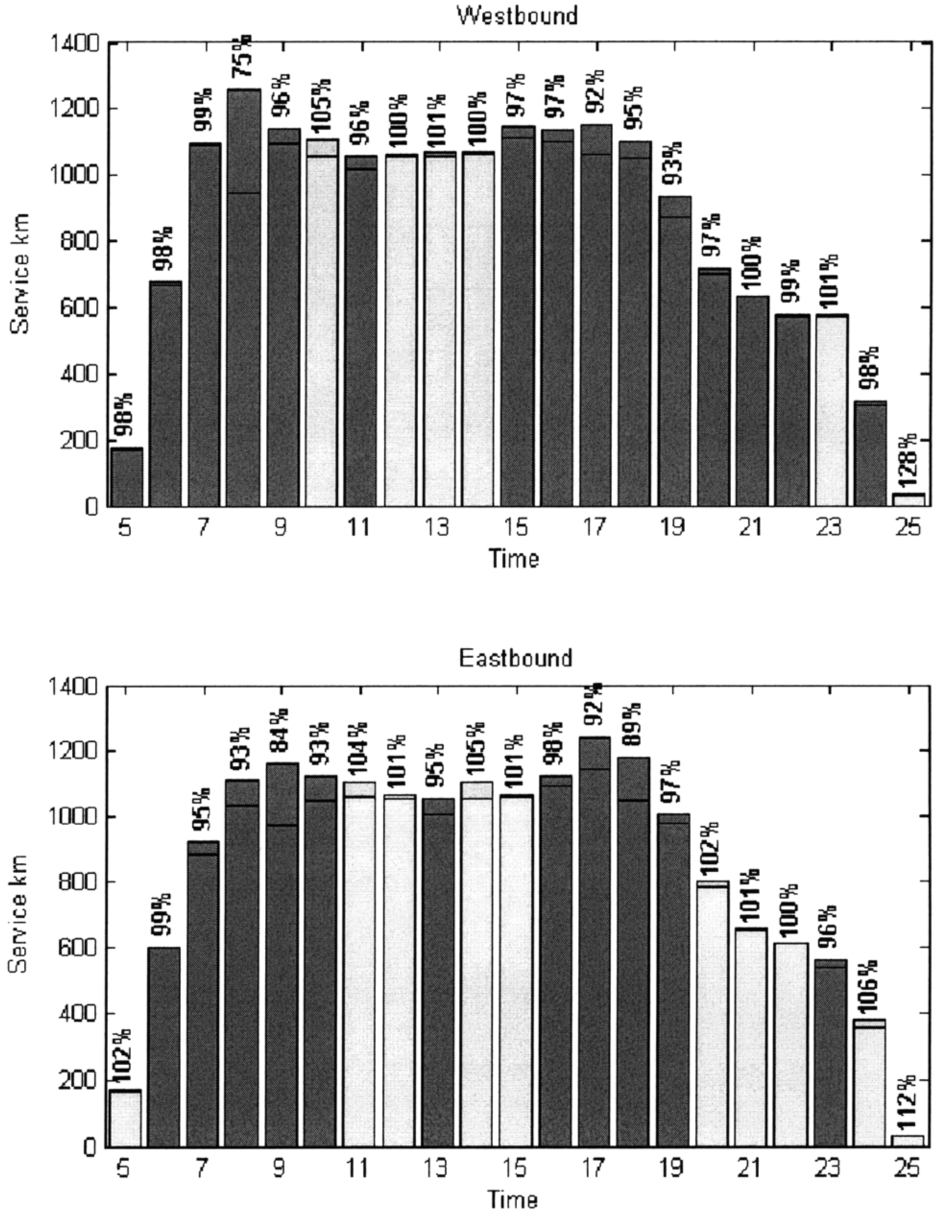


Figure 6-3: Service kilometers operated on the Central line on November 12, 2008

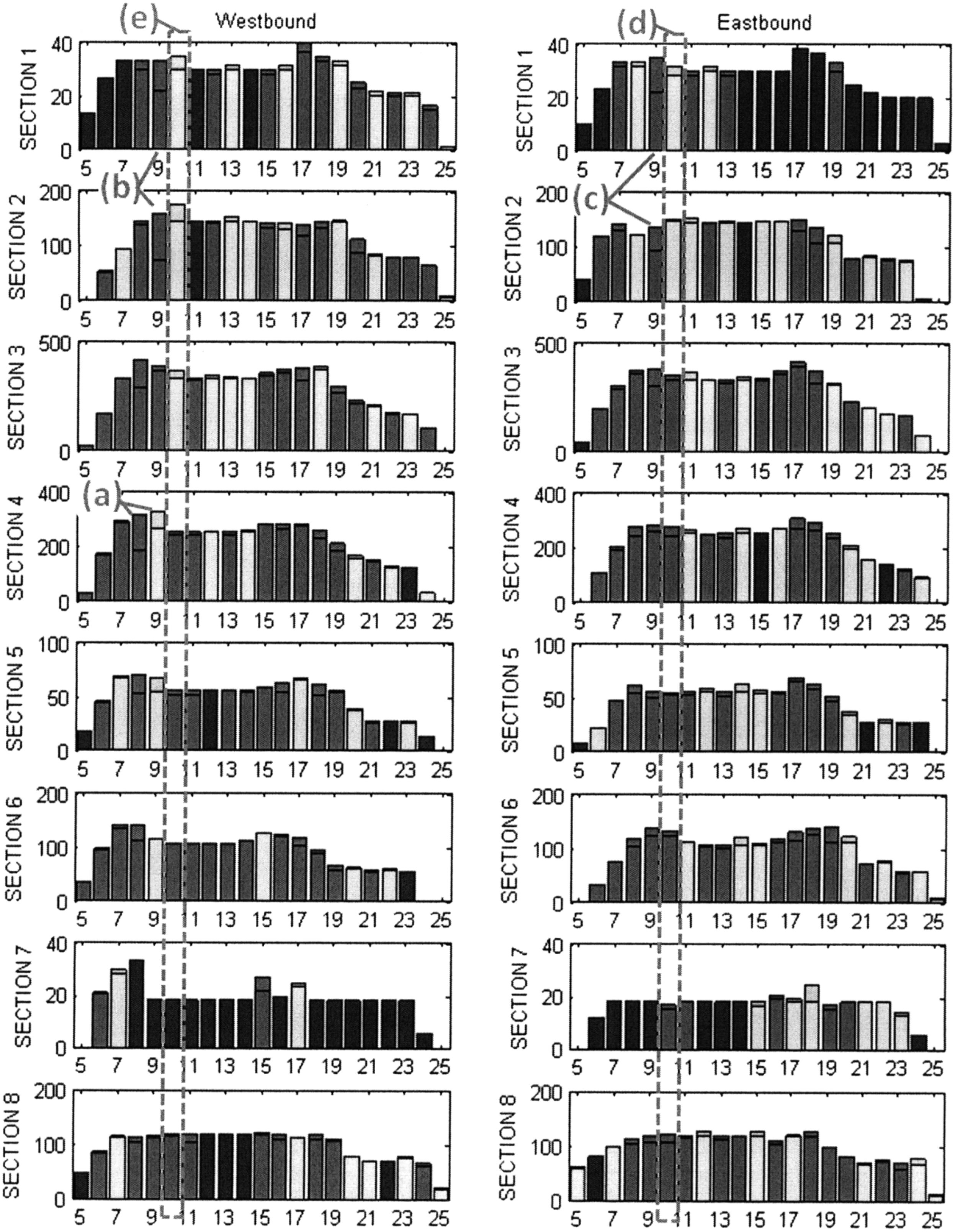


Figure 6-4: Service kilometers operated on November 12, 2008, by line section

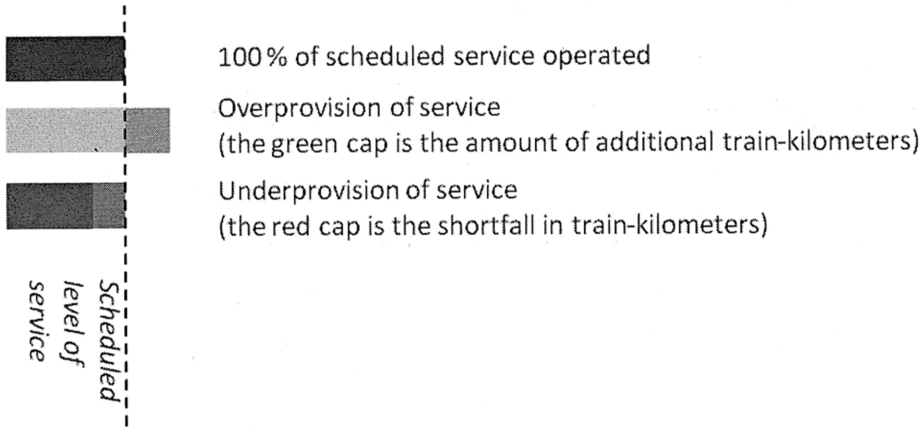


Figure 6-5: Key to service delivery plots

6.2.3 Passenger impact

The increasing use of smart card technology in the past decade is beginning to provide transit agencies with a wealth of information about passenger flows on their network. The availability of smart card data from Transport for London allows direct measurement of the impact of disruptions and service changes on passenger travel times. This section introduces the source and characteristics of the data and describes the measures for passenger impact which were chosen for the study of service control.

a) The Oyster Card data

For the analyses of passenger travel times, data from Transport for London's (TfL's) smart card system, the Oyster card, was used. The Oyster card was introduced in 2003 and has since obtained substantial market penetration, as many of TfL's fare products either are available only on Oyster or the use of Oyster provides a substantial discount over the cash fare (TfL, Your guide to fares and tickets, 2009). Currently, approximately 70% of all trips made on TfL services use Oyster. TfL has a zonal, time-differentiated fare structure: users are charged according to the number of zones traversed on their journey and there is a peak period surcharge. The fare calculation makes it necessary to collect both the origin and the destination of trips as well as the departure and arrival times. Therefore, Underground users are required to tap their smart cards at the entry and exit station of their journey. The result is an extensive dataset which is stored centrally and contains details for each Oyster trip made on the Underground.

The Oyster card is used as a fare medium on the London Underground, London Overground, Buses, Docklands Light Rail and parts of the National Rail network serving London. However, some fare products are available either optionally or exclusively on magnetic stripe paper tickets. Aside from one-day travelcards which are available on either a magnetic stripe ticket or on Oyster, the largest group of magnetic stripe ticket users on the Underground is passengers who interchange between National Rail and the Underground. They often have National Rail travelcards which are available on magnetic stripe tickets only and are valid both on National Rail and on the Underground. The only data available on magnetic stripe tickets are total entry/exit gate counts at a 15-minute level of detail. A further shortfall which needs to be accounted for is with respect to incomplete journeys where a passenger failed either to tap in or to tap out. However, they represent only approximately 2% of all trips.

For the purposes of reconstructing travel times, five data items are of interest (Uniman, 2009): the card ID which allows trips to be linked, the tap-in location and tap-in timestamp as well as the tap-out location and tap-out timestamp. Only complete Oyster journeys can be used for this analysis, and given absence of transfer times, the set of trips which can be used to calculate travel time between any given OD pair on the line is reduced to the trips which both originate and terminate at those particular origin and destination stations on the Central Line. However, this is not problematic as long as one is only interested in the travel time distribution, assuming that the complete Oyster trips for any OD pair represent an unbiased sample of the full OD passenger volume.

In order to quantify the total passenger travel time, an OD matrix is needed which allows comparisons across days. It would be problematic to use daily OD matrices for that purpose, since it would be unclear whether to attribute observed variations in total travel times to changes in passenger volumes or to changes in travel times. Therefore, it was decided to use a static OD matrix, which was constructed as follows:

For every hour throughout the service day, a seed OD matrix was constructed from the available Oyster dataset, using only trips which started and ended on the Central line. For every OD pair on the line, the passenger counts were averaged over the entire four-week analysis period in one-hour intervals. A trip was always entered into the time bin in which it began, i.e., a trip starting at 08:50 and lasting 20 minutes would fall into the 08:00 – 09:00 OD matrix. This represents the initial Central line-only, Oyster-only OD matrix. It still needed to be scaled up to account for magnetic stripe ticket users, incomplete records and transfer trips. This was done with the use of on/off counts from TfL's RODS (Rolling Origin-Destination Survey) and an IPF (iterative proportional fitting) process. RODS is an annual survey of a subset of the London Underground's stations to determine

OD flows, route choice and transfer volumes within the Underground network, which are then scaled to total gateline entry and exit counts for November of each year[10]. Therefore, RODS boarding/alighting figures include an approximation of the number of transfer trips and magnetic stripe ticket users which must be added to the counts of Central line-only Oyster users. For a more detailed description of RODS, see Chan (2007). For every station, the total RODS entries and exits were extracted, which were assumed to be the correct counts and served as control totals. With the seed matrix as described above and the control totals from RODS, IPF was performed in order to scale up the OD matrix such that the sums of entries and exits from stations in the OD matrix matched the control totals from RODS.

Furthermore, the closure of Shepherd's Bush station had to be taken into account. Obviously, there were no recorded entries and exits at that station during the analysis period due to this closure, and therefore it was decided to split the entries and exits from RODS evenly and attribute them to the two neighboring stations, White City and Holland Park.

b) Types of passenger impacts

Not all service control interventions have the same impact on passengers, and different passenger groups will experience the effects of an intervention differently. A difficult task in assessing passenger impact is the projection of the (dis-)benefits into the future, since:

- In daily operations, dispatching decisions are made continuously. Some interventions are performed in order to mitigate the negative impacts of earlier interventions, so the impact of a single intervention can often not be assessed without considering the overall strategy.
- Due to the stochastic nature of the environment in which the service operates, unforeseen events (disruptions) and the inherent variability of dwell times and running times quickly obfuscate the effects of an individual intervention. Conceptually speaking, performing an intervention (e.g., adding a train to the service) momentarily moves the system from one discrete state to another. However, from that point, it can develop in multiple different ways, depending on the stochastic outcomes that arise.

---

[10] November is the month with the highest average travel demand in London. Every year, a RODS matrix is estimated based on November data, and those passenger flows are assumed to be constant throughout the year for planning purposes. Therefore, RODS slightly overestimates ridership for the other months of the year, especially for the summer vacation months. The applications in chapter 7 are based on DMA and Oyster card data from April and November 2008. Since no major holidays are included in the April data, it was assumed that the differences in passenger volumes due to seasonality effects could be neglected.

In brief, when quantifying passenger impacts, it is advisable to limit the analysis to passenger delays (or travel time improvements) which can be clearly attributed to an intervention. The impacts of service control interventions on passengers include:

1. Increased waiting time at the trip origin: A reduction in the number of trains serving a specific OD pair means that on average, the waiting time of passengers at their origin increases. However, they may not necessarily know that this is due to a service control intervention.
2. Decreased waiting time at the trip origin: The opposite of the above, when a train is added to service on an OD pair.
3. Train transfer and additional waiting time during the journey: This is experienced by passengers who are on board a train which is short-turned or diverted. Theoretically, their additional waiting time is equal to that of passengers under point 1 above (cf. section 2.5.4).
4. Increased in-vehicle travel time: This is incurred by passengers on board trains which are held at stations, need to travel at reduced speed between stations or make in-tunnel stops (for example to allow the preceding train to be short-turned).

While it is important to bear in mind these different types of passenger impacts, identifying them individually with Oyster data is difficult if not impossible. This might be a direction for future research, as noted in section 8.3.

Therefore, the analysis is restricted to the study of a sample of travel times aggregated over a certain time period and pertinent OD pairs. The resulting travel time distribution captures all trips that started within that time period, distributed over all trains which served that OD pair. A travel time distribution is usually skewed to the left since there is a technical limit on the minimum possible travel time. The distribution's right "tail" includes riders who either experienced the worst service during that time period or whose travel times include activities other than traveling, such as waiting for somebody after tapping in or taking the wrong train. For a more in-depth discussion of travel time distribution, the reader is referred to Uniman (2009).

As an illustration, Figure 6-6 shows two travel time distributions for a high-volume OD pair on the Central line, Nottinghill Gate to Liverpool Street. On April 8 (blue plot) service was good, which causes a tight travel time distribution. On April 1 (red plot), service was disrupted and there were interventions, causing variability in travel times and higher maximum travel times, as can be seen from the much wider range of the travel time distribution. Both distributions show an irreducible minimum travel time of 19 min. However, it is not possible to distinguish whether the delays which are measured in the distribution for April 1 are caused by disruptions or by service control interventions; gaps in the service can be the consequence of either of those two events. As a matter of fact, one could argue that the passenger may not even care about the origin of the gap, but to the transit agency, it is of great interest.

Section 7.1 will show how the problem of distinguishing the origins of a delay can be overcome in a practical application. With the detailed knowledge about the operations which the analyst can gain from studying the signaling data, it is often possible to establish which passenger groups were affected by the disruption and which were affected by the service recovery process. As an example, if a train is blocked in a tunnel and a gap develops in front of that train, all passenger delays due to the gap can be defined as disruption-related delays. On the other hand, delays caused by short-turns of trains which were stuck behind the blocked train would be service control related delays. In order to separate the two passenger groups who experienced these delays, an analyst needs to define them in terms of location and entry time. However, separating the impacts with these two variables still requires certain assumptions and simplifications; for instance, in the example above, controllers may have held trains in front of the blocked train to prevent a large gap from forming. In that case, the size of the gap was both a consequence of the disruption and of service control. The analyst would need to define up front which delays are of interest, what is captured in the travel time distribution and what is not.

In this thesis, the focus shall be on disruptions and service control interventions with a negative impact (delays to passengers). In the specific case of a high-frequency metro line like the Central line, the temporal aggregation will capture passengers of multiple trains, but oftentimes, only a few trains are affected by a delay. For instance, of 12 scheduled trains per hour one or two may be short-turned. Assuming even demand, only $\frac{1}{12}$ to $\frac{1}{6}$ of the passengers to destinations beyond the short-turning point would be affected, either because they had to alight and wait for a later train or because they experienced a gap of one or two scheduled headways. Although the average travel time during this time period is an important measure, this example makes it clear that it is insufficient for capturing the effects of small disruptions and service control interventions, as it has a limited sensitivity towards the (often small) subset of passengers who experienced worse service. However, one should expect to see a change in the shape of the travel time distribution, as said subset of passengers experiencing worse service is pushed to the right of the distribution and the distribution thus becomes wider. This ties directly to the issue of *service reliability*, which is of great concern to transit agencies. Before making the trip, a single passenger will not know whether his/her train is the one which is going to be short-turned or otherwise delayed – his/her travel time might fall anywhere within the travel time distribution.

This calls for an additional measure (or measures) capturing the breadth of the distribution and thus the unreliability in travel time. The *Reliability Buffer Time (RBT)* is such a measure which captures the effects of service control well during the Central line analysis. Various authors have discussed this type of measure, among them Furth (2006), Chan (2007) and Uniman (2009), who introduced the name Reliability Buffer Time. The following section will review the RBT.

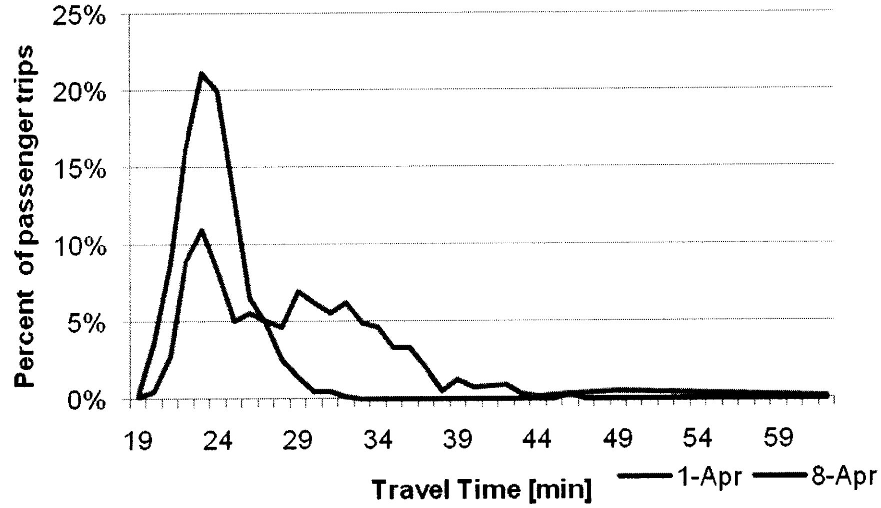


Figure 6-6: Two travel time distributions for Nottinghill Gate to Liverpool Street

c) The reliability buffer time (RBT)

The RBT is defined as (Uniman, 2009):

$$\text{Reliability Buffer Time} = (N^{th}\text{ percentile travel time} - M^{th}\text{ percentile travel time}) \qquad (2)$$

Uniman proposes M = 50th percentile as this is the median which is an indicator of the typical travel time. This represents the travel time that a frequent traveler on the system would base his/her expectations on. The 50th percentile is less sensitive to outliers than the mean. Furthermore, Uniman states: "*The variable N is an indicator for the threshold of certainty for reliable service*". In other words, if a passenger wishes to reach his or her destination on-time with 90% probability, he or she must budget the travel time which is indicated by the 90th percentile of the cumulative travel time distribution. Thus, the RBT is not an additional travel time experienced by all passengers, but for passengers who need to arrive at their destination on time with some degree of confidence, it is the amount of time they would need to budget in addition to the median travel time in order to reach their destination on time with the specified probability. A passenger who budgets the 90th percentile travel time to make an appointment but arrives in the median time might spend the additionally budgeted time unproductively before the appointment begins, and thus in this case one could argue that the RBT represents "lost passenger time".

d) Further travel time measures

Aside from the RBT, two further measures are required in order to put the reliability measure into context: The total travel time weighted by passenger demand and the total RBT weighted by passenger demand. The total travel time (TT_{total}) is calculated as the sum over all OD pairs of the product of the average travel time on an OD pair and that OD pair's passenger volume, that is:

$$TT_{total} = \sum^{OD} V_{OD} \cdot TT_{OD} \tag{3}$$

Where:

V_{OD} = passenger Volume by OD pair,
TT_{OD} = average travel time by OD pair, and
TT_{total} = total travel time

The calculation of total travel time can be performed at a line section (by origin and destination station) and time period level (by departure time) and represents the total time which all passengers traveling on any given OD are expected to spend traveling, without accounting for the unreliability of the system (because it is based on the average OD pair travel times).

The total RBT (RBT_{total}) is calculated in the same way as the total travel time, except that the average travel time by OD pair is replaced with the RBT for that OD pair:

$$RBT_{total} = \sum^{OD} V_{OD} \cdot RBT_{OD} = \sum^{OD} V_{OD} \cdot (90^{th}\ pct. - 50^{th}\ pct.)_{OD} \tag{4}$$

Where:

V_{OD} = passenger Volume by OD pair

RBT_{OD} = RBT by OD pair

RBT_{total} = total RBT

pct. = percentile

The total RBT is most meaningful when compared to the total travel time, as it represents the total time which all passengers would need to budget in addition to their expected travel time in order to arrive at their destination by their desired arrival time with 90% certainty. To avoid double counting, passengers must be categorized as a function of their origin and their destination. For instance, one would calculate the total RBT for passenger traveling within section 2, between sections 2 and 3, within section 3, and so forth. The unreliability enters through the RBT which is added to the median travel time.

A measure which was initially considered but eventually rejected for the study of service control is the average waiting time at stations, which could be calculated with train headway data under the assumption of random passenger arrivals. However, this measure may be misleading as it does not account for any additional waiting time incurred by passengers in locations other than their origin as a consequence of service control interventions like short-turns or diversions.

e) Shortfalls of measures based on travel time distributions

While travel time distributions are a powerful way of representing how passengers experience service on a line, it has some limitations. First, the travel time distribution does not explain the individual components of a trip, including access and egress time, waiting time and in-vehicle travel time. Second, any action of a passenger which is not directly related to the originally intended trip, for example waiting for somebody behind a gateline, taking the wrong train or changing travel plans while en route such that the traveled path is not the best for the new OD pair, enters into the travel time distribution despite its irrelevance to the quality of service. As a matter of fact, those passengers are likely to show up in the higher percentiles of the distribution, thus giving a false impression of service unreliability. Unfortunately, there is currently no way of filtering out such passengers, so the analyst depends on the assumption that either any inexplicable passenger behavior results in travel times beyond the 90th percentile or that this subset of passengers is sufficiently small and does not severely affect the summary statistics.

6.2.4 Tying the three measures together

On a conceptual level, it is straightforward to understand the connection between delays (disruptions), service control interventions, the level of service and passenger travel times. However, in a practical application, it may become very difficult to attribute changes in the level of service or in passenger travel times to individual train delays and interventions or even to link passenger travel times to the level of service. For studying service control with the goal of building a better understanding of it, there is no alternative to microscopic analyses of the DMA data for individual days. On the other hand, the aggregate measures can be used to:

- Find patterns in service control strategies and assess their frequency.
- Identify time windows and line sections of interest for detailed analysis.

How the four components of service monitoring are related depends very much on the interest of the analyst, who can either pursue a "top-down" or a "bottom-up" approach. In the latter case, the analyst starts with an interest in a specific disruption or known control strategy and begins with a microscopic analysis of the associated operational data. Afterwards, it is possible to move to a more aggregate level and try to answer the following questions:

- How often is this strategy employed?
- How does it affect the service and passengers, and is there a large variation in LOS and passenger impacts across different days on which this strategy was employed? What drives these variations?

In the top-down approach, the analyst begins with the intervention matrices introduced in section 6.2.1, with the goal of identifying patterns in service control strategies, or examines the LOS and passenger travel time measures in an effort to find unexplained passenger delays and drops in level of service. Having established the time periods and line sections of interest, the analyst can then move towards a more forensic analysis with the help of operational data. With either of these approaches, the author believes that the maximum benefit can be obtained from the analysis only by speaking to the controllers who made the decisions, in order to understand their objectives and constraints.

In the Central line analysis presented in this thesis, the author was most interested in specific responses to delays and therefore proceeded top-down by manually inspecting the intervention matrices and train-kilometer measures (as introduced in sections 6.2.1 and 6.2.2) to identify patterns over time windows and line sections for microscopic analysis of the DMA data. The disruption log provided background information after the identification of interesting patterns on the nature, time and exact location of the disruptions that led to the observed delays and service control interventions. Once the microscopic analysis of service control had been conducted, the appropriate OD pairs and time windows were defined for the passenger impact analysis.

6.3 Application to other metro lines and systems

This chapter showed the set of measures which were developed for the study of the Central line. Depending on available data, they may not be directly transferable to other metro lines. Nevertheless, an analyst can build on these measures and adapt them in light of the available data. The following list discusses how that could be done for the individual elements presented in this chapter:

- As discussed in section 5.4, the most important service control interventions can be identified by comparing the scheduled and the observed starting and ending point of a train trip. This information alone allows the analyst to generate a list or matrix as shown in section 6.2.1; although certain data items, such as the time at which an intervention took effect, may be missing, they are not crucial to understanding service control as long as they do not refer to the routing of the train. However, as will be seen in chapter 7, the microscopic analysis of disruptions and service recovery strategies can be greatly enhanced by information on the routing of crews, the dwell times of individual trains and the point along the line where controllers changed its destination.
- The availability of a disruption log is definitely a large benefit, as it helps explain the initial cause for a delay or intervention. However, since the disruption log does not connect to the rest of the data on a technical level, there are no hard requirements as to its form or content. Although a vital part of understanding the origin of problems on the line would be missing, service control can even be studied to some extent in the absence of a disruption log.
- The measures for level of service which were presented here are particularly useful for high-frequency metro operations, where a gap in the service will immediately show up as a drop in

level of service. In lower frequency services, the analyst may want to include information on headways. The author believes that, given the nature of the effects of service control interventions, average headway or the headway variance are not good measures as they may not be sufficiently sensitive to such types of problems. Instead, one may want to analyze the maximum headway, the number of headways beyond a certain threshold or headway distributions at control points along the line.

- Any type of automatic fare collection systems where entry and exit times are recorded, such as the Oyster system, would allow the application of the passenger impact measures which were presented here since they only track the time every user spends in the system. However, the use of these measures depends on such information being available. In systems where only entry times are recorded, an analyst may need to resort to models which help calculate passenger travel times as a function of the operations.

7 Applications

This chapter presents four cases where the framework presented in section 2.6 and all its elements are applied to analyze operations and service control decisions on the line. It is intended to serve two purposes. First, it provides concrete results to the London Underground which indicate where the agency may want to consider changes to current practices or policies. Second and on a more general level, it shows what results can be obtained in practice by looking at an operational question from multiple perspectives and integrating various types of data. It should be noted that the potential uses of the dataset assembled in chapter 5 are certainly not limited to the applications shown here, and by further developing and integrating the impact measures discussed in chapter 6, it may be possible to achieve even more powerful insights into daily operations on a line. In any event, if the analysis reveals operational or service control related problems, the next step should be to discuss the results with controllers directly and from a neutral point of view as elaborated on by Rahbee (2001) in order to better understand the observed behavior. Since references to stations and sections of the Central line are made throughout this chapter, the line map from chapter 3 is reproduced with line section numbers in Figure 7-1.

The starting point for the analyses presented hereafter was an exploratory throughput analysis which was performed on the Central line trunk portion using operations data for the AM and PM peaks of three weeks in March 2008. For all stations on the trunk portion of the line between Leytonstone and White City, the number of trains which served those stations westbound from 08:00-09:00 and eastbound from 17:00-18:00 (i.e., during the peaks of the peaks) was calculated. Since this was originally done using the trains per hour as a simple measure for level of service, the analyzed part of the line did not include East Acton and North Acton as the scheduled level of service from those stations is lower than on the rest of the trunk.

Those were then averaged over the 20 stations in order to obtain daily figures for the throughput during the two peak hours with the results shown in Table 7-1. There appears to be a consistent underdelivery of service – on average, 18% less service was operated in the AM peak and 16% less service in the PM peak compared to the scheduled throughput, which is 29.9 trains per hour westbound and 29.6 trains per hour eastbound. It was initially hypothesized that this was the effect of congestion. However, assuming recurring daily levels of congestion, the daily variability of the throughput and the fact that on two occasions a throughput of over 29 trains per hour was observed (eastbound on March 6 and westbound on March 14) suggested that congestion alone could not be responsible. Instead, discussions with London Underground staff revealed that it was presumably the combined effect of delays from disruptions and of service control interventions dealing with delays and train shortages; the author was told that March 2008 had been a particularly "bad month" in terms of rolling stock defects and the need for unscheduled train maintenance, which implied that it may not be representative of normal operations on the line. Therefore, and since there was no Oyster data available for this period, it was not further analyzed. However, it raised several interesting questions, namely how service controllers manage and restore service if delays occur during the peak hours, how they deal with train shortages and whether high scheduled frequencies have an impact on the recoverability of a timetable. All of the cases presented in this chapter are related to these initial questions and show how an analyst can proceed to answer them.

	3-Mar	4-Mar	5-Mar	6-Mar	7-Mar	10-Mar	11-Mar	12-Mar	13-Mar	14-Mar	17-Mar	18-Mar	19-Mar	20-Mar	Average	Difference
Westbound 08:00 – 09:00																
Trains per hr scheduled	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9		
Trains per hr observed	21.9	24.3	25.1	26.3	24.6	23.8	25.9	25.2	26.7	29.8	22.9	26.7	27.4	25.5	25.4	-18%
Missing trips	4.6	3.3	2.7	2.7	3.1	6.3	2.0	2.3	2.1	1.0	4.9	2.1	2.0	2.7		
Unscheduled trips	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		
Eastbound 17:00 – 18:00																
Trains per hr scheduled	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6	29.6		
Trains per hr observed	24.0	26.4	26.3	29.8	25.4	26.6	26.0	26.7	27.2	25.7	23.0	20.9	25.3	25.6	25.6	-16%
Missing trips	5.4	3.7	4.4	2.4	4.5	3.4	2.9	3.9	2.6	2.9	6.1	7.1	2.9	4.5		
Unscheduled trips	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0		

Table 7-1: Throughput through Central line trunk portion

The Oyster data which was available for the analysis covered four weeks in 2008, specifically March 30 - April 12 (April 10 and 11 were missing from the operational data for unknown reasons) and November 8 - November 23, 2008. To ensure comparability across days, weekends were excluded from the sample, since on weekends the line sometimes operates on a special timetable, due to engineering work and because passenger volumes are much lower, posing problems with sample sizes.

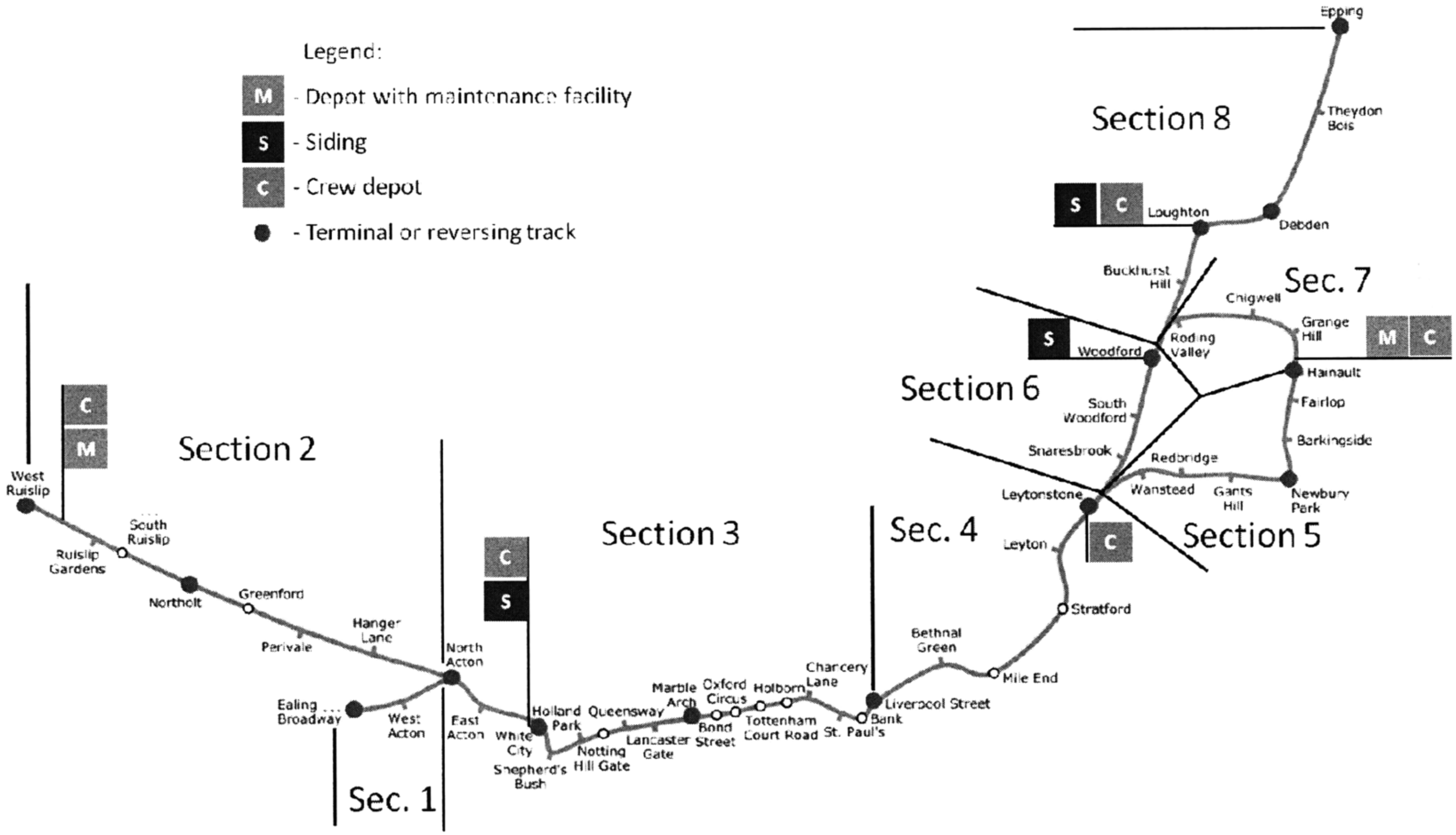


Figure 7-1: Central line map with section numbers (adapted from Forrester (2006))

The following four example cases are presented in this chapter:

- The policy of primarily canceling trunk service trains in the case of rolling stock shortages is examined in section 7.2, with the help of data from one day where that strategy could be observed over the course of three hours without any disturbances.
- Following up on the discussion of the influence of crew management on service control in section 4.3.3, section 7.3 looks at a specific example of two days on which short-turning was used as a service restoration technique and examines how crew management may be one reason for the observed differences.
- Section 7.4 examines how a specific disruption during the AM peak period was managed and quantifies the delays encountered because of service control interventions. Furthermore, it suggests ways to address the problems which are identified.
- The final example, in section 7.5, looks at the workload of controllers directly before and one year after a timetable change and considers how scheduling variables might have contributed to change in the observed number of service control interventions.

7.1 Specifics of methodology

Following the introduction of measures to quantify the impact of service control in chapter 6, several issues related to sample size are discussed here to operationalize the measures. If an analyst wants to quantify the total passenger travel times and RBTs on a line section, the first step is to query the Oyster dataset for all trips which originated and ended on that line section. The passenger trip data returned by this query will represent a set of OD pairs which does not necessarily cover all possible OD pairs within that section, but only those on which at least one passenger was observed. The RBT is a measure at the OD pair level which is calculated from percentiles of a travel time distribution, and in order to calculate the 50th and 90th percentile of a distribution accurately, a certain minimal sample size is required. In this analysis, it was decided to use only OD pairs with a minimum of 10 Oyster trips within the period of interest, since this allows the 90th percentile to represent a travel time which was actually measured – if the travel times were ordered by their magnitude, it would be that of the 9th passenger. The minimum of 10 trips per OD pair had to be relaxed in only one case, noted in section 7.3. Since three of the four case studies presented hereafter use a comparison between days to quantify delays, it needed to be ensured that every OD pair which went into the analysis of travel times was represented on both days, with sufficient sample size in both cases. Fortunately, as travel patterns did not vary greatly across days, most OD pairs had a sufficient sample size on both days. Furthermore, as will be seen in more detail later during this chapter, it was decided to define the 50th percentile travel time of a day with good, i.e., undisturbed operations, as the baseline for the travel time expected by passengers; the RBT is therefore calculated as the 90th percentile on the disrupted day minus the 50th percentile on the "baseline day", and it captures the spread of the travel time distribution with respect to an undisrupted day.

One very interesting measure would have been the number of passengers on the disrupted day which exceeded the 90th percentile of each OD pair travel time distribution on a "good day", which would be based on the assumption that on a "good day", passengers budget their travel time according to that 90th percentile, and therefore, if the travel time on the disrupted day remained below that boundary, they would not suffer any delay beyond their usual tolerance. This measure relies on large sample sizes, which proved to be problematic on the Central line, especially in the study of off-peak service. Therefore this measure was dropped from the analysis, but in any future work, it should be considered whether it would be appropriate since it would provide a very meaningful measure for the number of passengers who were severely impacted by an event on the line.

It was noted during the analysis that the data contained outliers with very high travel times which were impossible to explain simply through operations (it was not uncommon for some travel times to be around 80 or 90 minutes on an OD pair only a few stations apart, even under good operating conditions). It had to be assumed that these travel times were due to individual passenger behavior and were not linked to the performance of the Central line, making it necessary to exclude them. In all cases presented hereafter, a simple exclusion rule was chosen, based on the assumption that the maximum delay experienced by passengers on a specific OD pair were caused by the maximum gap between trains serving that OD pair. This implies that there was no large variability in train running times. The validity of that assumption was ensured in every case by sampling the running times of trains on the line sections of interest and checking that they were within a narrow range without any outliers.

Specifically, a time allowance was calculated and added to the median travel time within a certain timeband. Travel times beyond that time allowance were excluded from the dataset. For every timeband which was analyzed, the maximum headway was determined and 10 minutes were added for access/egress time. This was then rounded up to the nearest 5 minute-step. In one of the cases (section 7.3) this resulted in a time allowance of 25 minutes beyond the median, whereas in the two other cases the time allowance was 20 minutes. Generally, this resulted in only very few trips (less than 3% of the entire dataset) being excluded.

Section 6.2.4 explained how the OD matrix which was used for these analyses was derived. It was available at an hourly level of detail. In order to ensure comparability between the days which were studied and exclude seasonality effects, the same set of hourly OD matrices was applied to all days, independently of whether they were in April or November.

The plethora of variables which influence daily operations on the Central line (or any other metro line, for that matter) makes comparisons across days difficult, since at a certain level of detail, every situation is unique. The analyses presented here had a fairly small sample size (18 days) to draw on. Every effort was made to find the most representative examples with respect to the author's experience in the Central line control room, but given the small pool of data, this only amounted to one or two examples for each of the intervention strategies examined.

7.2 Case 1: Cancellation policy

7.2.1 Motivation and hypothesis

Section 4.3.2 mentioned that in case of rolling stock shortages, for example when an insufficient number of trains is available due to maintenance requirements or when trains have to be withdrawn during the day for some reason, Central line controllers must choose which train trips to cancel. During the author's visits to the control room, several controllers stated independently that their policy in that case was to cancel trunk service trains ("local trains"). For example, in case a train unit becomes defective in service, and it was covering a branch service, this leads to service controllers assigning a different train unit originally serving the trunk to replace it. In view of the explanations given by service controllers and service managers on the Central line, one can safely assume that this is a common strategy.

Given the main demand patterns on the Central line, this strategy seems counterintuitive, especially during the interpeak when the demand for trips within the trunk is significantly higher than for trips between the branches and the trunk (the passenger counts in Table 3-2 give a sense of the actual volumes). Therefore, the question was raised by the author whether this policy causes noticeable total delays to trunk passengers due to increased average waiting times whereas passengers from the outer branches experience less significant or no delays since more, or all, of the scheduled departures are served. What speaks for this hypothesis is that Central line service controllers were observed not to even out headways on the trunk when no other delays were reported, which in this case would result in occasional gaps of approximately two scheduled headways on the trunk. As an example, a branch train serving West Ruislip needed to be withdrawn due to a defect. A trunk service train was reassigned to the West Ruislip service and was therefore missing a train from its originally scheduled trajectory. That trunk service train would not have served the West Ruislip branch anyway, so a passenger arriving at the beginning of the gap left by the trunk service train and waiting for a train to West Ruislip would have waited the length of that gap under "normal" circumstances, until a West Ruislip train arrived. On the other hand, a passenger traveling to a destination on the trunk could have boarded any train regardless of the destination (i.e., including the trunk service train), so he or she ends up waiting longer than anticipated.

What speaks against this hypothesis is that even when a few trains are missing, the interpeak trunk frequency is still very high, which may make the delays negligible. The baseline frequency for the interpeak is 24 trains per hour, which corresponds to an average 2.5 minute headway. Even with 4 trains missing, the *average* headway would be only 30 seconds larger.

7.2.2 Analysis procedure

Knowing about this strategy from the conversations with controllers, the author used the intervention matrices to identify days on which a steady number of cancellations had been performed repeatedly over a period of several hours. It was found that often these cancellations were performed simultaneously with other service control interventions, but on one of the days (April 1, 2008), there was a three-hour period in the afternoon from 13:00 - 16:00 with several cancellations, yet no other event occurred to disturb service on the trunk in any noticeable way. After that, the disruption log and service kilometer plots were used to ensure that there were in fact no other unobserved major delays.

Table 7-2 shows the service kilometers operated on a section level for this period on April 1 as well as the interventions which occurred during that time. The color coding for this table and all following ones of the same type in this chapter is shown in Figure 7-2, and the service delivery plots for April 1 can be found in appendix E. The impact of the missing trains and the resulting uneven small gaps in the service can be seen in Table 7-2. From 13:00 – 16:00 service on the trunk sections (sections 3 and 4) remains consistently below schedule, with the only exception being section 3 (North Acton to Liverpool Street) westbound during the 13:00 – 14:00 period. On the western branches (sections 1 and 2) service delivery generally hovers around 100%, although there are some drops in service on the West Ruislip branch (section 2) eastbound, which can be explained by the fact that two of the defective trains were withdrawn to West Ruislip depot located at the end of section 2 and therefore only served the westbound trip. The westbound diversion and extension noted in the table in the 13:00 – 14:00 timeband are such a case. On the eastern branches (sections 5 through 8), the level of service delivery is more variable. On several occasions full service is provided, but especially on section 5 (Leytonstone to Hainault) there are periods of underprovision of service since many of the canceled train trips were scheduled to reverse at Newbury Park, located within section 5. This is, however, not critical since the primary interest is the effect on the trunk portion.

		Train kilometers operated (% of scheduled train kilometers)								Interventions					
		West branches		Trunk		East branches									
	Section no.	1	2	3	4	5	6	7	8	Short-turns	Diversions	Extensions	Cancellations	Unscheduled	Withdrawn
Westbound	13:00-14:00	33 (111%)	141.2 (98%)	330 (100%)	235.7 (94%)	53.2 (95%)	103.1 (97%)	18.2 (100%)	116.9 (97%)		1	1	2		
	14:00-15:00	29.7 (100%)	137.5 (96%)	300.8 (91%)	227.9 (91%)	56 (100%)	99.2 (89%)	18.2 (100%)	121.1 (100%)				2		
	15:00-16:00	29.7 (100%)	141.3 (100%)	341.6 (96%)	262.6 (93%)	58.9 (100%)	118.3 (94%)	23.2 (87%)	115.2 (93%)				2		
Eastbound	13:00-14:00	33 (111%)	127.6 (89%)	321.1 (97%)	227.8 (90%)	56 (100%)	110.4 (103%)	18.2 (100%)	116.7 (96%)				2		
	14:00-15:00	29.7 (100%)	155.6 (108%)	320.6 (97%)	244.3 (97%)	53 (94%)	97.7 (92%)	18.2 (100%)	115.1 (95%)				1		
	15:00-16:00	29.7 (100%)	135.7 (93%)	299.8 (89%)	213.1 (85%)	57.6 (106%)	96.1 (90%)	18.2 (108%)	127 (105%)	1			2		

Table 7-2: Service delivery and interventions on April 1, 2008

% of scheduled kilometers operated
< 95%
≤ 95% and < 100%
≥ 100%

Figure 7-2: Color code for service delivery matrices

Table 7-3 shows the 11 canceled set/trip numbers during that time period, none of which served the outer branches. Furthermore, one can see that for several set numbers, two or three sequential trips were canceled. Sets 101, 102, 105 and 145 are trunk services between White City and Newbury Park (Hainault for set 145). There was originally a train assigned to these sets, but when a different train on a branch service (three Ealing Broadway – Hainault trains and one Northolt – Loughton train, as can be seen in Table 7-3) became defective and there were no spare trains to replace it, the controllers needed to decide between canceling the remaining trips of the defective train's set number or assigning a train from a less important (in the eyes of the controller) set number to it and canceling the remaining trips of that less important set. Controllers chose the second option and assigned the trains from sets 101, 102, 105 and 145 to those branch services, as is shown in Table 7-3. Conversely, the train which was originally assigned to set 110, running between Loughton and Northolt, became defective in service and was withdrawn. Since this was already one of the less important set numbers (again, in the eyes of the controller), no train was reassigned to it.

Set	Trip	Departure	Origin	Destination	Remarks
101	8	13:11:00	White City	Newbury Park	The train unit originally assigned to set 101 replaced a defective train on an Ealing Broadway - Hainault service.
101	9	14:13:30	Newbury Park	White City	
101	10	15:11:00	White City	Newbury Park	
102	6	13:31:00	White City	Newbury Park	The train unit originally assigned to set 102 replaced a defective train on an Ealing Broadway - Hainault service.
102	7	14:33:30	Newbury Park	White City	
105	8	13:33:30	Newbury Park	White City	The train unit originally assigned to set 105 replaced a defective train on a Northolt - Loughton service.
105	9	14:31:00	White City	Newbury Park	
105	10	15:29:30	Newbury Park	White City	
110	8	13:41:45	Loughton	Northolt	The train unit originally assigned to set 110 became defective, was withdrawn and not replaced.
110	9	15:01:15	Northolt	Loughton	
145	6	15:52:00	Hainault	White City	The train unit originally assigned to set 145 replaced a defective train on an Ealing Broadway - Hainault service.

Table 7-3: Canceled train trips between 13:00 and 16:00 on April 1

Finally, a baseline day was needed for comparison, since the passenger travel times observed on April 1 can only be interpreted by comparison to travel times on a "very good" day. As a reference day, November 21 was chosen; in the entire dataset, it is the day with the lowest overall number of service control interventions and the best service delivery measured in terms of service kilometers. Table 7-4 shows the service kilometers operated and the interventions on November 21. The corresponding service delivery plots can be found in appendix E.

		Service operated (% of total scheduled service)								Interventions					
		West branches		Trunk		East branches									
	Section no.	1	2	3	4	5	6	7	8	Short-turns	Diversions	Extensions	Cancellations	Unscheduled	Withdrawn
Westbound	13:00-14:00	29.7 (100%)	143.8 (100%)	326.1 (98%)	252.7 (101%)	56 (100%)	105.6 (99%)	18.2 (100%)	119.3 (98%)						
	14:00-15:00	29.7 (100%)	151 (105%)	338.6 (102%)	249.4 (99%)	56 (100%)	106.5 (95%)	17.1 (94%)	119.6 (99%)				1		
	15:00-16:00	29.7 (100%)	135.2 (96%)	341.3 (96%)	268.1 (95%)	58.9 (100%)	124.1 (99%)	24.5 (91%)	124.4 (101%)		1				
Eastbound	13:00-14:00	29.7 (100%)	138.2 (96%)	319.9 (97%)	251 (99%)	53.2 (95%)	105.5 (99%)	18.2 (100%)	122.1 (101%)						
	14:00-15:00	29.7 (100%)	149.3 (104%)	343.7 (104%)	240.3 (96%)	52.7 (94%)	111.6 (105%)	18.2 (100%)	123.6 (102%)			1			1
	15:00-16:00	29.7 (100%)	138.8 (95%)	327.6 (98%)	263.4 (105%)	60.7 (111%)	105.6 (99%)	14.6 (86%)	117.7 (97%)						

Table 7-4: Service kilometers and interventions on November 21, 2008

Subset of passengers	**Definition**
Trips within trunk	Passengers with origins and destinations on sections 3 and 4 of the line.
Trips between trunk and western branches	• Passengers with origins on sections 1 or 2 and destinations on sections 3 or 4. • Passengers with origins on sections 3 or 4 and destinations on sections 1 or 2.
Trips between trunk and eastern branches	• Passengers with origins on sections 5, 6, 7 or 8 and destinations on sections 3 or 4. • Passengers with origins on sections 3 or 4 and destinations on sections 5, 6, 7 or 8.

Table 7-5: Definition of passenger subsets for the analysis of the cancellation strategy

The data of interest were the passenger travel times within the trunk and to/from the branches, with the precise definitions shown in Table 7-5. There is some fuzziness in the structure of the query, which is related to the problems in partitioning the line as explained in section 6.2.1. Specifically, passengers to and from East Acton and North Acton are counted as trunk passengers although strictly speaking, they cannot board a train which is scheduled to reverse at White City. However, as only three trains per hour are scheduled to reverse at White City, it was decided that the effect of this problem is limited and thus could be ignored.

7.2.3 Results

Table 7-6 shows the results of the analysis, calculated using the same OD matrix for both days as discussed in section 6.2.3. One can see in Table 7-6 that the interpeak passenger volumes on the trunk are markedly higher than the branch volumes, which is also reflected in the number of OD pairs excluded because their sample size was too small to calculate the RBT (cf. section 7.1); approximately 40% of all trips to and from the branches are excluded due to the small sample sizes. However, this is not problematic for the question at hand, since the same OD pairs were sampled on April 1 and November 21 (ensuring that both had a sufficient sample size) and the analysis is primarily focused on the difference between the two days. In terms of total travel time calculated using equation (3) from section 6.2.3, one can see that the difference between the two days is minimal, which is not surprising. The scheduled frequency on the trunk during the interpeak is high (24 tph) and on average, there were two trains per hour missing (cf. Table 7-2). On a side note, the total travel time hardly changed between April 1 and November 21 despite the service running at 24 tph on the latter day and approximately 22 tph on the former. This suggests that the marginal benefits of such a high off-peak frequency may be small, and that the issue lies much more with the variance of the headways (see below).

The analysis of the travel time reliability shows only a slightly different picture. The RBT for each OD pair for November 21 was calculated as shown in equation (2) in chapter 6, using the 90th percentile and the median of travel times on November 21, whereas for April 1, the 90th percentile was calculated from the travel times of April 1 and the median was based on the travel times on November 21. The total RBT was calculated according to equation (3) in chapter 6. For all three passenger groups, the total RBT is between 10 – 20% of total travel time, and the differences between November 21 and April 1 do not appear to be very large. Nevertheless, it is worth noting that the total RBT increases for the trips within the trunk and those between the trunk and the western branches, whereas it decreases for the trips between the trunk and the eastern branches.

	Within trunk		Between trunk and western branches		Between trunk and eastern branches	
	1-Apr	21-Nov	1-Apr	21-Nov	1-Apr	21-Nov
OD pairs sampled (total OD pairs)	310 (411)		96 (307)		129 (351)	
Passenger trips sampled (total passenger trips)	85269 87877		5990 8052		5350 10461	
Total travel time [min]	1170430	1169327	159715	159254	116359	116907
Total travel time difference [min]	1102		461		-548	
Average difference per passenger [min]	0.01		0.08		-0.10	
Total RBT [min]	211921	192703	24363	22394	20332	21992
Average RBT over OD pairs [min]	2.542	2.268	3.917	4.135	3.860	4.132
Std. dev RBT over OD pairs [min]	1.654	0.922	1.961	2.009	2.120	1.598
t-statistic	2.5498		-0.7635		-1.1607	

Table 7-6: Cancellation policy analysis results

The time period which is being analyzed, 13:00 to 16:00, was free of any significant delays and did not experience any demand peaks, so the assumption that all additional travel time in the form of RBT was experienced as on-platform waiting time and not as in-vehicle travel time is realistic. This, along with the assumption that the level of service in terms of the distribution of headways is constant within any line section (cf. section 6.2.2), allows the RBT values from the two days to be treated as independent random variables because they are not a function of any service characteristic other than the headways (and thus the waiting time). The last line of Table 7-6 shows the t-statistic which was calculated as the result of a two-sided, two-sample t-tests not assuming equal variances with the following null hypothesis:

H_0: The means of the two RBT distributions (April 1 and November 21) are equal.

As one can see, the t-statistics suggest a significant positive difference between the trunk travel times on April 1 and November 21 and a (less significant) negative difference regarding the branch travel times. This appears to confirm the hypothesis that passengers on the trunk experienced less reliable service, while showing that passengers on the branches actually received better service than on the baseline day, the failure to reject the hypothesis notwithstanding. However, a look at the numbers themselves shows that these differences are in the order of a few seconds in all cases; on April 1, the average RBT on the trunk was 17 seconds higher than on November 21, putting it at 2:33 min. This is the average additional time that a passenger must add to his/her typical journey time in order to be 90% sure of arriving at the destination on time. On the branches on both ends of the line, the RBT on April 1 was approximately 19 seconds lower than on the baseline day. Presumably these differences are hardly enough for a passenger to take notice, at least on average. Also, one should note that for trips between the western branches and the trunk, the average RBT decreases, but the total RBT increases, which suggests that the OD pairs with less reliable service on April 1 are those with higher passenger volumes.

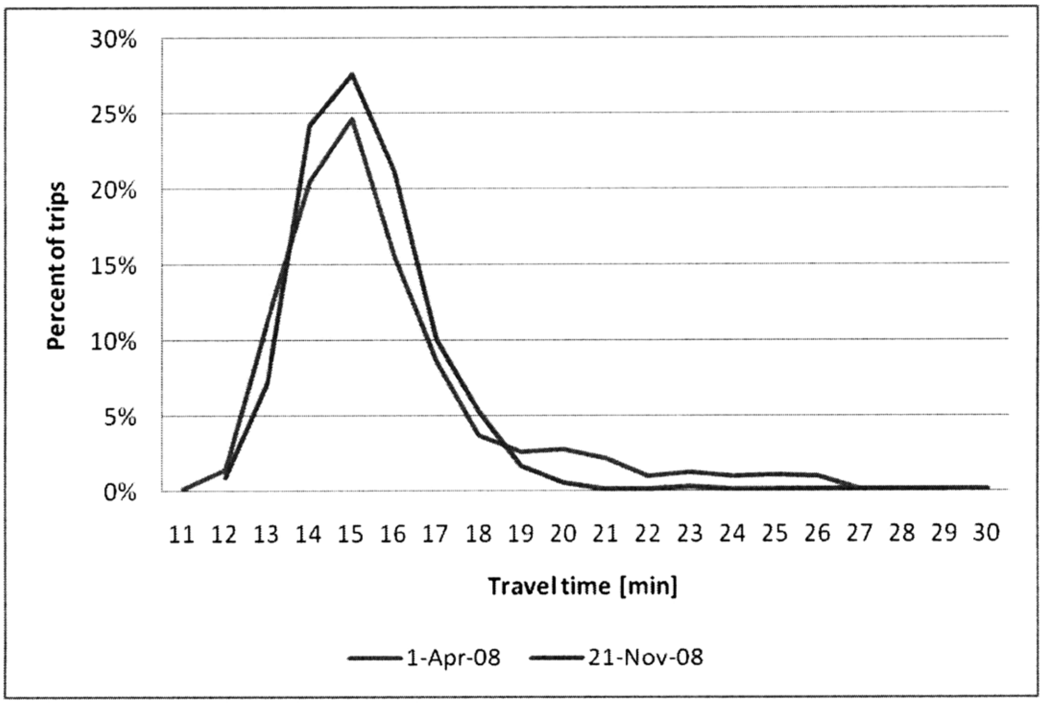


Figure 7-3: Travel time distributions between Liverpool Street and Oxford Circus

A different perspective is offered by Figure 7-3, which shows the travel time distributions for a sample trunk OD pair, Liverpool Street to Oxford Circus. This OD pair was selected because it has high passenger demand, thus ensuring that the distribution would include passengers who experienced the gaps, and it is representative of other trunk OD pairs. The scheduled trip time, excluding the dwell times at the Liverpool Street and Oxford Circus stations, is 9:55 min. The difference between the two distributions shows how the unreliability manifests itself. While the bulk of the trips fall within the 12 to 19 minute range on both days, the distribution for April 1 shows a clearly distinguishable tail beyond a travel time of about 19 minutes. In line with the hypothesis stated at the beginning of this section, these would be passengers who arrived on the platform at Liverpool Street station during one of the gaps left by the canceled trains.

For the Central line's service controllers, the strategy in question is for dealing with a shortage of rolling stock. These results show that, although their strategy does distribute the impacts of rolling stock shortages unevenly and perhaps mostly penalizing the trunk section with its heavy demand, its overall impact is small enough that one can argue that this does not call for immediate attention. However, in the long run, one may think of ways to better align service control decisions with passenger demand patterns in the event of rolling stock shortages. It is understandable that controllers are reluctant to remove trains from branch services with low headways, since that would cause unreasonably large gaps in the service, so a first set of options for remedying this dilemma (i.e., aside from increasing the rolling stock size) would lie in timetable design, where there are three possibilities:

1. To design alternative timetables for reduced numbers of available rolling stock which can be put into effect if trains need to be withdrawn from service. This is only realistic if the timetable design process is automated and can be repeated under different boundary conditions with a small marginal investment of resources, which is currently not the case for the London Underground.
2. To design certain train trips in the regular working timetable as "discardable" trips such that the impact of canceling them is minimal. This is more realistic than the above option, and to the author's knowledge has been discussed at various points in the design process for working timetable 65, which will take effect in January 2010.
3. To generally reduce the off-peak frequency in the working timetable, such that the constraint on rolling stock availability is eased and trains need to be canceled less often due to stock unavailability, as was suggested above. The minimal difference in total travel times between April 1 and November 21 suggests that the impacts of a frequency reduction might be very small.

The fourth option, of course, is to purchase additional rolling stock, and the fifth and last option is a do-nothing strategy in view of the fact that, despite its statistical significance, the small amount of added unreliability may not warrant large changes to the service. Which of these options is chosen is ultimately a strategic decision.

Finally, the question of how to better coordinate service control with the timetable (or the intentions behind it) can also be addressed from the perspective of service control, where the Underground could implement policies aimed at better monitoring and regulating headways even in the absence of disruptions. Specifically, the author would especially like to encourage considering the first two options for timetable design since they would not involve any capital investments and would not reduce service in any way but would help redistribute it efficiently when rolling stock shortages occur – which happens frequently on the Central line.

7.3 Case 2: Short-turning

7.3.1 Motivation and hypothesis

Short-turning is a very common control strategy on the Central line. (Figure 6-1 gave a sense of how often short-turns are performed in daily operations.) Within the analysis period, on average 21.5 trains were short-turned daily, although the day-to-day variance of this number is considerable. Furthermore, as discussed in chapter 4, a train can be short-turned for various reasons. It may be a strategy for dealing with train or crew lateness, but it may also be used to meet level-of-service targets or to balance the resources by reallocating a train (and its driver) to a different part of the line. This section will highlight the use of short-turning as a service recovery strategy after two very similar disruptions on the line. The motivation for analyzing this case is to investigate whether variability in response to similar disruptions could be explained with any of the decision factors outlined in chapter 4 or whether it needed to be attributed, at least in part, to "controller variability". Short-turning was selected as a focus for two reasons. Firstly, because it is a very common disruption recovery strategy and secondly because of all service control interventions, it is probably the one which is the most frustrating for passengers and leads to the largest number of passenger complaints.

7.3.2 Analysis procedure

Given the question outlined in section 7.3.1, the approach was to start by analyzing the intervention matrices to detect clusters of short-turns which would indicate the management of a group of late trains, presumably from an earlier disruption. After potential cases were identified, the incident log was consulted to select those which were similar in delay, and the level of service measures were used to compare the effects of those short-turns on the line. This allowed the author to narrow down the focus to two days of interest, April 3 and November 12, 2008. Both days had very similar delays, and the service was restored with similar numbers of short-turns, but the level of service measures showed that there were considerable differences in the effects on the line. The events on those two days are summarized below:

- April 3, 2008: A westbound train (set #114, trip #4) sat at Queensway station, located in section 3 of the line, for 14 minutes, from 10:27:34 until 10:41:44, owing to a passenger emergency alarm. When the disruption cleared, there was a cluster of delayed trains behind it. Table 7-7 shows the sequence of trains which was following set #114, along with information on what service control interventions were performed. The key to colors and numbered remarks as shown in Table 7-8; the key to the station codes is provided in appendix A. In total, 9 westbound trains were short-turned or diverted from the West Ruislip to the Ealing Broadway branch. In order to compensate for the resulting large gap, one train (set #106) out of the group of delayed trains, which was scheduled to reverse at White City, was extended to West Ruislip. Table 7-9 shows the service delivery and intervention matrix for that day; the corresponding plots can be found in appendix E.
- November 12, 2008: A westbound train (set #10, trip #2) suffered a delay on the approach to Liverpool Street station in the morning rush hour because of a loss of pilot light[11] "due to severe overcrowding on the train", as the incident log states. It appears that the train operator was not able to obtain a pilot light at Liverpool Street either, since the controllers decided to detrain the unit and run it empty to West Ruislip Depot for inspection. The detraining process, presumably onto a heavily crowded platform, caused further delays. A closer analysis of the operational data revealed that the information on the disruption stated in the disruption log was partially incorrect, which confirms the statement made in section 1.3 that during stressful times in the control center, the information recorded in manual logs is not very reliable. Train set #10 departed Bethnal Green station at 08:28:32. According to the running time estimate of the signaling system (3 min 14 sec from Bethnal Green to Liverpool Street), the train should have reached Liverpool Street at approximately 08:32, but it did not pull in until 08:45:46. The detraining process then delayed the departure until 08:50:30, at which time a cluster of delayed trains had built up behind train set #10, the first of which was train set #6. Table 7-11 shows the following sequence of trains and shows that in total, 10 trains were short-turned in order to reduce cycle time and compensate for the delays. A

---

[11] The pilot light indicates to the train operator that all doors are safely closed.

further two trains, sets #16 and #145, had their destinations swapped. Again, the key to colors and numbered remarks is in Table 7-8. Table 7-10 shows the service delivery and intervention matrix for November 12; the corresponding plots can be found in appendix E.

Set number	Unit number	Headway at Queensway East [min]	Origin	Planned Destination	Changed Destination	Change coded at	Impending crew relief (planned)	Crew relief observed	Remarks
114	165	00:00	LES	NOR			LES EAST	As planned	(1)
24	125	01:26	EPP	WER	EAB	HOP	LES EAST	As planned	
26	128	02:12	EPP	WER			NONE	--	
56	150	03:32	HAI	EAB	WHC	LAG	WHC EAST	See remark	(4)
27	104	01:20	EPP	WER	NOR	LAG	NONE	See remark	(2)
45	141	01:26	HAI	EAB	WHC	NHG	LES EAST	As planned	
106	162	01:24	NEP	WHC	WER		NONE	--	
30	129	01:28	EPP	WER	EAB	QUE	LES EAST	As planned	
72	140	01:24	HAI	EAB	NOA	BOS	WHC EAST	As planned	
115	166	01:16	LOU	NOR	EAB	OXC	WHC WEST	As planned	
31	130	01:58	EPP	WER	NOR	NHG	NONE	--	
57	114	--:--	HAI	EAB	MAA	BOS	NONE	--	(3)
32	131	07:32	EPP	WER			NONE	--	
End of recovery phase – westbound service returned to normal after these interventions									

Table 7-7: Service control interventions following the disruption on April 3, 2008

(1)	Train which suffered initial delay
(2)	Reversed and relieved by a spare driver at Debden
(3)	Served Woodford via Hainault on next trip
(4)	Crew relief at White City upon reversal
(5)	Crew put back on time with spare driver #315
(6)	Crew put back on time with spare driver #319
(7)	Crew was relieved by a spare driver at West Ruislip
(8)	Crew relief and short-turn at Leytonstone eastbound
RED	Short-turned trip
ORANGE	Diverted trip
GREEN	Extended trip

Table 7-8: Legend for Table 7-7 and Table 7-11

		Service operated (% of total scheduled service)								Interventions					
		West branches		Trunk		East branches									
	Section no.	1	2	3	4	5	6	7	8	Short-turns	Diversions	Extensions	Cancellations	Unscheduled	Withdrawn
Westbound	10:00 - 11:00	23.1 (78%)	87.4 (61%)	262.7 (79%)	261.3 (104%)	56 (100%)	111.6 (105%)	19.6 (107%)	122.9 (102%)	3	2	3			
	11:00 - 12:00	29.7 (100%)	131.2 (91%)	390.1 (118%)	264 (105%)	56 (100%)	102.4 (97%)	18.2 (100%)	111.8 (92%)	5	2	1			
	12:00 - 13:00	29.7 (100%)	157.4 (109%)	325.1 (98%)	219.1 (87%)	51.4 (92%)	96.7 (91%)	14.6 (80%)	108.3 (89%)						
	13:00 - 14:00	33 (111%)	134.9 (94%)	273.7 (83%)	169.3 (67%)	32.7 (58%)	76.6 (72%)	9.7 (53%)	85 (70%)	2			4		
Eastbound	10:00 - 11:00	21.6 (77%)	132.2 (91%)	319.2 (91%)	253.7 (92%)	49.8 (91%)	130 (98%)	15.8 (92%)	130 (106%)				2	1	
	11:00 - 12:00	27.9 (94%)	90 (63%)	278 (84%)	194.9 (78%)	60.7 (109%)	84.3 (76%)	19.5 (107%)	111.8 (92%)	1		1			
	12:00 - 13:00	29.7 (100%)	152.8 (106%)	301.4 (91%)	230.4 (92%)	42 (75%)	85.9 (80%)	16.9 (93%)	96.9 (80%)	2	2				
	13:00 - 14:00	31.5 (106%)	151.7 (106%)	352.9 (107%)	221.9 (88%)	41.8 (75%)	77.7 (73%)	12.1 (67%)	80.2 (66%)	4		3	1	1	2

Table 7-9: Service delivery and interventions on April 3, 2008

		Service operated (% of total scheduled service)								Interventions					
		West branches		Trunk		East branches									
	Section no.	1	2	3	4	5	6	7	8	Short-turns	Diversions	Extensions	Cancellations	Unscheduled	Withdrawn
Westbound	08:00 - 09:00	29.7 (90%)	137.1 (96%)	285.5 (69%)	187.9 (60%)	53.2 (76%)	110.1 (79%)	33 (100%)	106.9 (94%)			1	2		2
	09:00 - 10:00	21.6 (65%)	74.8 (48%)	365 (95%)	322.6 (123%)	67 (122%)	114.3 (101%)	18.2 (100%)	113.4 (98%)	8	2	1		1	
	10:00 - 11:00	34.5 (116%)	174.1 (121%)	366.6 (111%)	243.9 (97%)	51.4 (92%)	104.5 (98%)	18.2 (100%)	116.2 (96%)	3					
	11:00 - 12:00	29.7 (100%)	143.8 (100%)	320.7 (97%)	243.3 (97%)	51.4 (92%)	105.4 (99%)	18.2 (100%)	104.9 (87%)						
Eastbound	08:00 - 09:00	33 (106%)	121.2 (101%)	356.4 (96%)	243.4 (88%)	54.5 (90%)	101.5 (86%)	18.2 (100%)	107 (95%)	1			1	1	
	09:00 - 10:00	21.6 (62%)	92.4 (68%)	301.6 (80%)	256.2 (91%)	49.6 (88%)	123.4 (90%)	18.2 (100%)	109.2 (90%)	1		1	1		
	10:00 - 11:00	31.2 (112%)	149.1 (103%)	329.1 (94%)	241.8 (88%)	53.2 (98%)	121.5 (92%)	15.8 (92%)	109 (89%)	1				1	
	11:00 - 12:00	27.9 (94%)	152.6 (106%)	362.3 (109%)	264.3 (105%)	53 (95%)	111.7 (100%)	18.2 (100%)	116.2 (96%)	2					

Table 7-10: Service delivery and interventions on November 12, 2008

Set number	Unit number	Headway at Queensway East [min]	Origin	Planned Destination	Changed Destination	Change coded at	Impending crew relief (planned)	Crew relief observed	Remarks
6	104	00:00	EPP	EAB			WHC WEST	As planned	(1)
1	101	01:40	EPP	NOR	WHC	TCR	WHC EAST	See remark	(4)
111	115	02:06	NEP	WHC			WER	As planned	
52	100	01:32	HAI	NOR			WHC WEST	See remark	(6)
5	107	01:42	HAI	EAB			WHC WEST	See remark	(5)
11	114	01:26	DEB	WER			WER	As planned	
4	105	05:54	GRH	EAB	WHC	LIS	LES EAST	As planned	
12	116	01:50	HAI	WER	NOR	BAN	NONE	--	
41	106	01:38	HAI	NOR	NOA	LIS	HAI	As planned	
13	117	01:54	HAI	WER			WHC EAST	See remark	(6)
110	110	03:34	EPP	EAB			WHC EAST	As planned	
77	109	02:02	EPP	NOR			NONE	--	
75	141	01:24	DEB	EAB			NONE	--	
16	118	01:16	HAI	WER	EAB	BAN	WHC WEST	As planned	
7	111	04:46	EPP	EAB	WHC	STR	WHC EAST	See remark	(4)
117	171	01:18	GRH	WHC	MAA	LES	NONE	--	
145	121	01:20	HAI	EAB	WER	BEG	WHC WEST	As planned	
15	120	01:18	EPP	WER			WER	As planned	
76	161	01:20	HAI	NOA	WHC	MIE	NONE	--	
53	149	01:24	DEB	EAB	NOA	LIS	NONE	--	(3)
42	140	01:22	HAI	EAB			WHC EAST	See remark	(5)
104	162	01:28	HAI	NOR	WHC	STR	WHC WEST	See remark	(4)
17	124	01:30	HAI	WER			WHC WEST	See remark	(7)
14	119	01:38	EPP	EAB			NONE	--	
20	125	01:58	NEP	WER			LES EAST	See remark	(8)
112	167	01:46	DEB	NOR			NONE	--	
54	150	01:36	HAI	EAB			WHC EAST	See remark	(5)
43	142	01:28	HAI	EAB	NOA	WHC	NONE	--	
22	128	02:42	DEB	WER			NONE	--	
End of recovery phase – westbound service returned to normal after these interventions									

Table 7-11: Service control interventions following the disruption on November 12, 2008

The reader may have noted that, despite a similar number of short-turns on the two days, the patterns were different. On April 3, practically all trains after set #114 were short-turned or diverted, but all those trains were in a compact group. An observer standing at Queensway station (i.e., downstream from the disruption) would first have seen the gap caused by the disruption, followed by the train which was originally blocked and then the group of trains which was later short-turned and diverted. Between the first and the last train of that group approximately half an hour passed, after which time the service returned to normal. One may also note that most of the interventions were coded very late, as the trains were already on the western part of the trunk section when the changes were announced, and even the extension of train set # 106 from White City to West Ruislip was not announced until Lancaster Gate. On the other hand, on November 12, the short-turned and diverted trains are much more dispersed, with a maximum of three trains in sequence being short-turned. To the observer at Queensway, the short-turned and diverted trains passed over a period of approximately one hour in small groups in between trains which were run to their original destinations. Furthermore, many of the changes were coded while on the eastern portion of the trunk section. That is, most of the trains passing through central London were showing their correct destination.

A close look at the two service delivery matrices in Table 7-9 and Table 7-10 shows these differences. On November 12, the gap is clearly noticeable since it occurs during the peak hours, thereby strongly reducing the throughput capacity. Between 08:00 – 09:00 on the trunk (sections 3 and 4) westbound, only 60 – 70% of scheduled train-km were operated. The short-turns then affect sections 1 and 2 between 09:00 – 10:00 west- and eastbound. On April 3, a reduction of service on section 3 westbound between 10:00 – 11:00 can be seen, as well as the effects of the short-turns on sections 1 and 2 west- and eastbound during that same time period as well as 11:00 – 12:00. This illustrates one of the shortfalls of the service delivery measure. On November 12, many of the trains were short-turned at White City or North Acton, thereby completely avoiding the branches. On April 3, there were more diversions to Ealing Broadway and more trains were short-turned on the branches, leading April 3 to be classified as better service than November 12. This is one of the reasons why it is essential to approach this type of analysis from different perspectives and at the section level of disaggregation, using a variety of measures and datasets.

Operational data from these two days were then analyzed in detail in order to establish possible factors behind these differences, and finally the passenger impact was evaluated. The results are presented subsequently in sections 7.3.3 and 7.3.4 respectively. As a baseline day, November 21 was used again, as described in section 7.2.2. The service delivery and intervention matrix for November 21 is shown in Table 7-12.

		Service operated (% of total scheduled service)								Interventions					
		West branches		Trunk		East branches									
	Section no.	1	2	3	4	5	6	7	8	Short-turns	Diversions	Extensions	Cancellations	Unscheduled	Withdrawn
Westbound	08:00 - 09:00	34.5 (105%)	150.1 (105%)	393.1 (95%)	270.3 (86%)	67.2 (96%)	119.7 (86%)	25.6 (78%)	108.8 (96%)			1	2		
	09:00 - 10:00	31.5 (95%)	142.4 (91%)	372 (97%)	262.9 (101%)	56 (102%)	112.6 (99%)	18.2 (100%)	119.1 (102%)						
	10:00 - 11:00	31.2 (105%)	136.8 (95%)	349.2 (105%)	276.9 (110%)	60.7 (108%)	106.8 (100%)	23.2 (127%)	125.1 (103%)	1					
	11:00 - 12:00	29.7 (100%)	152.4 (106%)	327.9 (99%)	249.4 (99%)	56 (100%)	105.1 (99%)	16.9 (93%)	122.9 (102%)						
	12:00 - 13:00	29.7 (100%)	136.6 (95%)	339.4 (102%)	244.5 (97%)	56 (100%)	109.7 (103%)	19.6 (107%)	121.1 (100%)						
	13:00 - 14:00	29.7 (100%)	143.8 (100%)	326.1 (98%)	252.7 (101%)	56 (100%)	105.6 (99%)	18.2 (100%)	119.3 (98%)						
Eastbound	08:00 - 09:00	34.5 (111%)	119.5 (99%)	360.7 (97%)	251.5 (91%)	54.5 (90%)	102.1 (86%)	19.3 (106%)	104.9 (93%)						
	09:00 - 10:00	33 (95%)	125 (92%)	370.4 (99%)	303.5 (107%)	57.6 (103%)	135 (99%)	17.1 (94%)	117.7 (97%)				1		
	10:00 - 11:00	29.7 (106%)	134 (92%)	339.5 (97%)	253 (92%)	49.8 (91%)	130.9 (99%)	16.9 (99%)	130.8 (107%)						
	11:00 - 12:00	29.7 (100%)	153.9 (107%)	334.6 (101%)	250.6 (100%)	58.9 (106%)	110.4 (99%)	18.2 (100%)	124.4 (103%)		1				1
	12:00 - 13:00	29.7 (100%)	143.8 (100%)	333.8 (101%)	259.9 (104%)	62.3 (111%)	109.3 (102%)	18.2 (100%)	117.5 (97%)						
	13:00 - 14:00	29.7 (100%)	138.2 (96%)	319.9 (97%)	251 (99%)	53.2 (95%)	105.5 (99%)	18.2 (100%)	122.1 (101%)	1					

Table 7-12: Service delivery and interventions on November 21, 2008

7.3.3 A possible decision factor: Crew management

In both cases, the disruption caused a gap in the service and a cluster of late trains. In line with what was established in section 4.3.2, this does warrant a set of interventions in order to improve the level of service on the line. However, controllers performed these short-turns much more aggressively on April 3 than on November 12, and it was eventually found that crew management provided a very likely explanation for these differences. The term “likely explanation” is used here because this could not be verified with the service controllers, although the data strongly suggests it. Table 7-7 and

Table 7-11 show when and where each train between the first and the last short-turn had its next crew relief (which included more trains on November 12 than on April 3). In both cases, there is one short-turned train which can clearly be attributed to a level-of-service constraint: set #53 on November 12 and set #57 on April 3 were short-turned because they needed to cover a Woodford via Hainault service, which operates at low frequencies and with published departures, as noted in section 4.3.2.

On April 3, of the nine trains which were diverted or short-turned, six were on their last or penultimate trip before a crew relief. Three of the crew reliefs were scheduled to take place at Leytonstone eastbound, two at White City eastbound and one at White City westbound. Furthermore, one driver was replaced by a spare driver at Debden, which is not a regular crew depot, thus suggesting he or she was at risk of breaching a driving time constraint. It may be understandable that service controllers become nervous at the idea of a train arriving late for a crew relief at Leytonstone, since a problem with a crew relief at Leytonstone, such as a relief driver not being in place, can create a significant blockage since there is no possibility of storing trains at that station. On November 12, on the other hand, only five of the ten short-turned and diverted trains had an impending crew relief, and only one was located at Leytonstone eastbound. A further three were White City reliefs (1 westbound, 2 eastbound) and one was scheduled at Hainault. Given the concentrated distribution in time of the short-turns/diversions, it appears that on April 3, the larger number of impending crew reliefs was the cause, whereas on November 12, the crew relief constraints were not as binding, resulting in better spacing between short-turns/diversions. The difference in the tightness of crew relief constraints might be because the disruption on November 12 occurred at approximately 08:30, while the disruption on April 3 occurred at approximately 10:30, when many of the drivers who had stepped on for the morning peak hour were being relieved. The data shows that trains were short-turned at White City even for White City reliefs, which supports the statement in section 4.3.3 that controllers are often reluctant to delay a driver stepping on. As these pieces of work were during the morning, it is possible that some of the drivers were also rostered to work later in the day and they therefore had to meet a hard constraint in terms of when they should step off.

Moreover, upon inspection of the trains which were not short-turned or diverted, one finds that controllers appear to have used spare operators on November 12 to avoid having to short-turn many of the trains following the gap despite some impending crew reliefs, but did not do the same on April 3. Two drivers, #315 and #319, appear to have been used as spare operators as they appear on five different train units only for short trips, for instance taking a train from White City to Ealing Broadway and back. Although the data do not allow a complete reconstruction of crew movements and do not offer insights into the availability of spare crews, this raises the question of why spare crews were not used on April 3 in a similar manner – it may either have been because no spare drivers were available or because the service controllers decided against it for some reason.

Unfortunately it is not possible to verify the exact causes for the choices which were made on these two days, as the data are limited in this respect. These findings would have been more powerful if the analyses had been performed shortly after the days with the short-turns, in which case the analyst could have shown the results to the service controllers and initiated a discussion to better understand why those decisions were made.

7.3.4 Results for passenger impact

Quantifying the impacts of short-turns is not a straightforward task, as is illustrated with the following example. Suppose a train is traveling on the trunk section of the line and is scheduled to run to the terminal of a branch but is reversed by controllers just before it enters the branch. The effects on passengers traveling to the branch are easily defined. Passengers already on board the train experience increased journey time since they have to alight before the train is reversed and wait for the next train to their destination. Passengers waiting at stations on the branch experience longer average waiting times if the short-turned train leaves a gap in the service. However, the effects on passengers in the opposite direction on the trunk section are more difficult to define. The short-turned train was presumably scheduled to serve the trunk section anyway, but at a later point in time – one would need to predict how the service would have operated under a do-nothing scenario and compare it with the observed service. The issue becomes even more complicated as the train might not be covering its originally scheduled trip, but may have been renumbered to fit into the timetable slot in which it ends up traveling in the opposite direction.

To avoid these complications of dealing with passengers traveling within the trunk portion of the line, the decision was made to analyze only the impact of the service control decisions on stations to the west of the first short-turning point, White City. Since on both days many of the trains were short-turned at White City and completely omitted the branches, the impact on passengers traveling between stations to the west of that point and the trunk (in both directions) was clear: they would experience gaps in the service and, in some cases, be forced to alight and wait for a different train to their destination. Furthermore, unlike the analysis in section 7.2, a rolling time window was defined, because the interest was in capturing the delays attributable to the service control interventions and not the delays encountered by passengers due to the disruption and the subsequent gap in the service. The procedure in defining the time window was as follows for both days.

Only trips between the trunk and stations west of White City were of interest, since those are the trips which are directly affected by short-turns in the form of missing train trips. A time window of 2.5 h was defined for each station as the sampling window. The choice of a 2.5 h time window was made to ensure a large enough sample size without including the impacts of other delays on the line which occurred later. The time window was defined differently for stations upstream and downstream of the station where the blockage occurred, as shown in Table 7-13. It was assumed that as soon as the blockage cleared, the queue behind the blocked train started moving again. Therefore, for all stations upstream of the blockage, the time window was defined as starting at the time the blockage cleared (i.e., all stations upstream of the blockage point had the same sampling time window). On April 3, this was the case for all stations between Queensway and Leytonstone, whereas on November 13, it was for all stations between Liverpool Street and Leytonstone. For stations downstream of the blockage, the interest was in capturing all passengers boarding immediately after the originally blocked train passed, in order to avoid measuring the impact of the gap in front of that train.

As is also shown in Table 7-13, the moving time window was defined in the westbound direction as starting at the moment the first westbound train in passenger service after the disruption-related gap passed a station and ending for that station 2.5 h later. Furthermore, it was decided to start the sampling time window for eastbound trips at the same time as for westbound trips. This was somewhat arbitrary, though the choice was informed by the consideration that, as soon as the first westbound train after the incident-related gap passes, controllers have the possibility to short-turn trains into a gap in the opposite direction, i.e., the number of trains traveling in the eastbound direction can be controlled. The observations from the operations data confirmed this choice, as in both cases, the gap is only clearly recognizable as it travels west to the two terminals, Ealing Broadway and West Ruislip. In the opposite direction, it quickly becomes hard to distinguish because of short-turns and trains which are held in order to even out headways.

The selection of specific OD pairs for sampling passenger travel times was based on two criteria:

- Sufficient sample size, as explained in section 7.1. The original requirement of a minimum of 10 trips per OD pair was retained for eastbound trips, but for westbound trips it had to be lowered to a minimum of 8 trips because otherwise, all OD pairs touching the West Ruislip branch would have been excluded due to small passenger volumes.

- The OD pairs had to be represented with sufficient sample size on all three days to allow the comparison of RBTs.

	April 3			November 12		
	Start time	End time	Station	Start time	End time	
Upstream of blockage on April 3	10:41	13:11	LES	08:51	11:21	Upstream of blockage on November 12
	10:41	13:11	LEY	08:51	11:21	
	10:41	13:11	STR	08:51	11:21	
	10:41	13:11	MIE	08:51	11:21	
	10:41	13:11	BEG	08:51	11:21	
	10:41	13:11	LIS	08:51	11:21	
	10:41	13:11	BAN	08:54	11:24	Downstream of blockage on November 12
	10:41	13:11	STP	08:56	11:26	
	10:41	13:11	CHL	08:58	11:28	
	10:41	13:11	HOL	09:00	11:30	
	10:41	13:11	TCR	09:02	11:32	
	10:41	13:11	OXC	09:04	11:34	
	10:41	13:11	BOS	09:06	11:36	
	10:41	13:11	MAA	09:07	11:37	
	10:41	13:11	LAG	09:09	11:39	
	10:41	13:11	QUE	09:11	11:41	
Downstream of blockage on April 3	10:43	13:13	NHG	09:13	11:43	
	10:45	13:15	HOP	09:14	11:44	
	10:46	13:16	SHB	09:16	11:46	
	10:50	13:20	WHC	09:19	11:49	
	10:53	13:23	EAA	09:22	11:52	
	10:54	13:24	NOA	09:24	11:54	
	11:00	13:30	HAL	09:28	11:58	
	11:02	13:32	PER	09:30	12:00	
	11:04	13:34	GRE	09:31	12:01	
	11:06	13:36	NOR	09:34	12:04	
	11:09	13:39	SOR	09:37	12:07	
	11:11	13:41	RUG	09:39	12:09	
	11:20	13:50	WER	09:41	12:11	
	10:57	13:27	WEA	09:27	11:57	
	11:04	13:34	EAB	09:30	12:00	

Table 7-13: Sampling time windows for April 3 and November 12, 2008

The results of this selection process are shown in Table 7-14. The eastbound passenger volumes are generally higher, and the OD pairs are well distributed among the different line sections. On the other hand, for westbound trips, out of a total of 22 OD pairs which satisfied the criteria, only four had destinations located on the West Ruislip branch due to low passenger demand. These pairs represented trips to Northolt (2 OD pairs), Greenford and Hanger Lane (1 OD pair each). Therefore, westbound results might be biased towards the Ealing Broadway branch and they do not include trips beyond Northolt.

Westbound: From trunk stations to ...	**# of OD pairs**
Ealing Broadway - West Acton	14
West Ruislip - Hanger Lane	4
North Acton - East Acton	4
Total	**22**

Eastbound: To trunk stations from ...	**# of OD pairs**
Ealing Broadway - West Acton	16
West Ruislip - Hanger Lane	22
North Acton - East Acton	13
Total	**51**

Table 7-14: OD pairs by line section sampled for the analysis of short-turning strategies

Row		**Nov 12 W**	**Nov 12 E**	**Apr 3 W**	**Apr 3 E**
1	OD pairs sampled	22	51	22	51
2	(total OD pairs)	(136)	(191)	(158)	(200)
3	Trips sampled	2206	7627	2324	3718
4	(total trips)	(4153)	(11367)	(5215)	(5166)
5	Total travel time on Nov 12 / Apr 3[h]	937	3218	1068	1818
6	Total travel time on Nov 21 [h]	926	3121	984	1526
7	Total travel time difference [h]	11	97	84	292
8	Travel time per passenger on Nov 12 / Apr 3 [min]	25.49	25.32	27.58	29.35
9	Travel time per passenger on Nov 21 [min]	25.19	24.55	25.41	24.63
10	Total RBT on Nov 12 / Apr 3 [h]	189	777	324	862
11	Total RBT on Nov 21 [h]	154	423	192	185
12	Average RBT on Nov 12 / Apr 3 [min]	5.16	6.01	8.21	14.15
13	Average RBT on Nov 21 [min]	4.33	3.25	4.25	3.03
14	t-statistic vs. Nov 21	0.86	7.35	5.47	17.34

Table 7-15: Comparison of short-turning strategies on April 3 and November 12 with November 21

For all three days, the total travel time and the total RBT were calculated (using equations (3) and (4), respectively) and compared to the baseline day, November 21, with the results shown in Table 7-15. This example also illustrates an issue which makes travel time comparisons between strategies difficult – it is very unlikely that two similar disruptions will happen at the same time, forcing the analyst to apply different time windows to the two travel time calculations. However, passenger volumes and travel patterns vary throughout the day, and the analyst faces the following choice. Either, one can use the same OD volumes for both samples, which ensures comparability between the days but distorts the total delay estimates, or one can use the actual OD volumes, in which case the total delay estimates reflect the situation as faced by controllers, but makes a comparison between days difficult as there are two independent variables involved in the calculation.

In this analysis, the same OD matrix was used for all days, so the OD volumes varied only as a function of time, but not between days. Hence, the choice was made to retain the time-dependent OD volumes but to use total travel time only for comparison between the two disrupted days and the baseline day independently. Given that the intent of this case study is to compare the two short-turning strategies, several other values were calculated, which will be presented below. First, however, sampling issues need to be discussed.

Rows 1 through 4 in Table 7-15 show the results of the application of the exclusion rules presented in section 7.1. Row 2 shows the total number of OD pairs represented in the samples. A large portion of all OD pairs had only 1 or 2 trips within the analysis period, making the calculation of the RBT or even the mean with a reasonable degree of accuracy impossible (in the latter case, if the median were calculated based on only 1 or 2 trips, it would not be possible to verify how representative of actual OD travel times the results is). The exclusion of OD pairs with less than 10 (or 8) trips and the limitation to OD pairs which are represented in all three samples reduces the sample size considerably, since passenger volumes to and from the branches are low during the off-peak. A comparison of the results in rows 3 and 4 gives a sense of how the travel patterns differ between the two samples: on November 12, where the time window begins directly after the AM peak at 08:51, there are appreciably more eastbound trips from the branches to the trunk than on April 3, where the time window does not begin until 10:41. On both days, the exclusion rules allow only about 45 to 50% of westbound trips to be sampled, in contrast to the approximately 70% of eastbound trips which are captured. Even though the calculations are based on a scaled-up population[12], the total travel times lose some of their representativeness through this. Rows 5 through 9 show the results. What is interesting is the comparison of the differences between November 12 and November 21 with the differences between April 3 and November 21. It should be noted that the time window for November 21 is not the same for the two examples; it is adjusted to the one for the day it is being compared to. The increase in total travel time between November 21 and November 12 is very small, between 0 and 3% with respect to November 21. Specifically, westbound passengers on November 12 felt little if any impact and eastbound passengers felt a minimal impact. On the other hand, on April 3, there is an obvious impact. Total westbound travel times are 8.5% higher and total eastbound travel times 19% higher with respect to November 21. The differences between the west- and eastbound directions might either be the result of longer trips or of passenger behavior (see below). However, this difference was not examined in detail as it was beyond the scope of this research.

Regarding the differences in total travel times (row 7), caution should be used in interpreting these values. The samples were taken over the span of 2.5 h to ensure sufficient sample size, whereas the short-turns occurred within a time window of approximately 30 min to 1 h. Therefore, only a subset of all passengers had to bear the bulk of the delays shown in row 7, whereas the others experienced travel times closer to the times they would usually expect.

Rows 8 and 9 show the average travel times per passenger. One should remember that the westbound data is biased towards trips to the Ealing Broadway branch, and in the case of April 3, several trains were diverted from the West Ruislip to the Ealing Broadway branch. Thus, some of the negative impacts of the diversions may not have been captured.

Rows 10 and 11 give a sense of the magnitude of change in total RBT; the total RBT on November 21 is the unreliability experienced by passengers on a day with good operations. This is now compared to the unreliability of service on the disruption day, as captured by the total RBT of April 3 and November 12. As explained in section 6.2.3 c) and shown in equations (2) and (4), the RBT is calculated as the 90th percentile minus the median for an OD pair. On November 21, both these measures are based on the data for November 21. On April 3 and November 12, the 90th percentile is calculated from the travel times of the disrupted day, while the median is based on the travel times in the same time window on November 21. The total RBT is the cumulative amount of time which

---

[12] On the OD pairs which were included, the passenger volumes are based on RODS and are therefore scaled up to include transfers and magnetic stripe ticket users, but the OD pairs with insufficient sample size are excluded from the analysis.

passengers on any given OD pair needed, in addition to their expectations, in order to arrive at their destination with 90% certainty. It is, in that sense, a measure of the lost customer time which frustrates passengers the most. Since passengers do not know a priori whether their train may be short-turned, to a first approximation it is fair to ascribe the additional RBT to all passengers on an OD pair under disrupted conditions[13].

That said, rows 10 and 11 again show that the majority of delays due to unreliable service are experienced eastbound, and that the unreliability is much worse on April 3. Despite the fact that approximately 40% less trips are sampled on April 3 than on November 12 due to the drop in demand after the AM peak, there is a 23% increase in total RBT in both directions, from 966 passenger hours to 1185 passenger hours.

Finally, the average RBT for the 22 westbound OD pairs and the 51 eastbound OD pairs is reported in rows 12 and 13. These numbers allow a direct comparison between the three days. First, a remark on the average RBT values for November 21. One can see that the average westbound RBTs are higher than the eastbound ones. This result is intuitive since westbound trains are reaching the end of their journey, and one can expect more variance in the headways than for eastbound trains which are only a few stations down the line from their origin. A comparison between rows 12 and 13 shows the differences in average RBT for November 12 and April 3. Although the average RBTs for westbound trips are higher on both disruption days compared to November 21, they are less than one minute higher on November 12 and less than 4 minutes higher on April 3. These values are surprisingly small in light of the fact that on both days, there were gaps in the service in the order of 12 minutes. There are two possible explanations:

- Due to the small passenger volumes, the delays caused by gaps are not detectable in the RBT. For instance, on any given OD pair with 10 passengers over the course of 2.5h, it might only have been one passenger (or none at all) who happened to experience the gap.
- This may be influenced by passenger behavior. Passengers facing a potential 12 minute wait might be unwilling to stay in the station and either search for alternative routes, alight early or switch modes, which would reduce the sample to passengers who were affected less severely.

---

[13] There is of course the issue of how affected passengers are distributed within the time window which is being analyzed. However, such a level of detail is beyond the scope of this research, and it is not problematic since the interest is mainly in the differences between days, and it is more important that the measure is being applied similarly to the three days.

For eastbound journeys, the picture is different. These numbers are more reliable due to the larger sample size, and they clearly show a difference between November 12 and April 3, which can be attributed to the different short-turning strategies. On April 3, the average eastbound RBT (14.15 min) is 57% of the travel time per passenger on the baseline day (24.63 min).

The difference in RBT distribution is shown graphically in the probability density functions in Figure 7-4 and the cumulative density functions in Figure 7-5. The reader should note that these distributions include both westbound and eastbound trips and that they show the RBT by OD pair *unweighted* by passenger demand. As an example, on November 21, the majority of OD pairs had an RBT between 1 and 6 min, with very few OD pairs having RBTs above 6 minutes. This is shown in the green probability density function and cumulative density function. The OD pairs are the same as those used for calculating the results in Table 7-15, i.e., 73 OD pairs in total for all three days (22 westbound and 51 eastbound). The two figures confirm what was found above. On November 12, many of the OD pairs had an RBT between 1 and 7 min, which is similar to November 21, with a noticeably "fatter tail" to the right of the distribution. The RBT distribution on April 3 is much more spread out, with passengers on some OD pairs experiencing RBTs of up to 18 min, or even, in one case, 22 min.

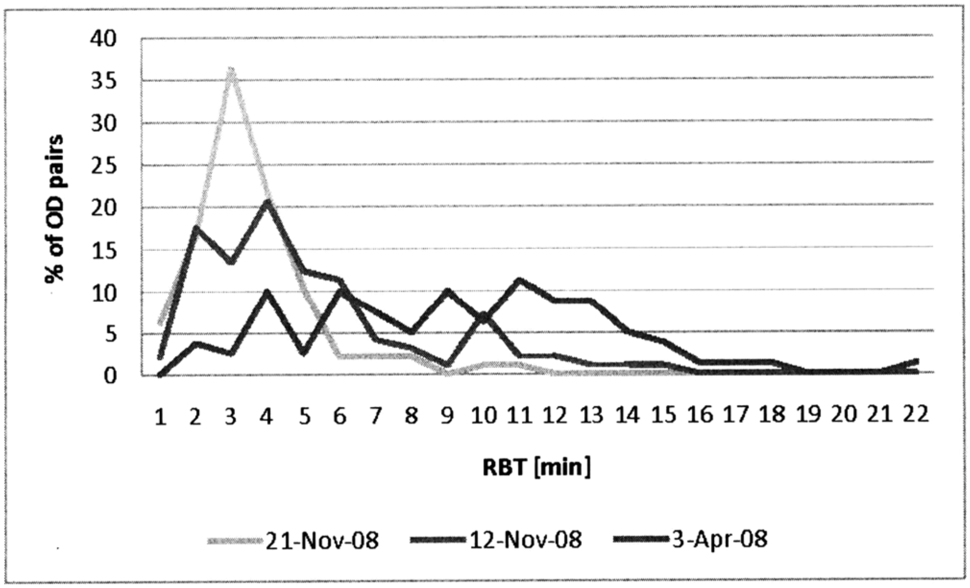


Figure 7-4: Probability density functions for the RBT on the three days of interest

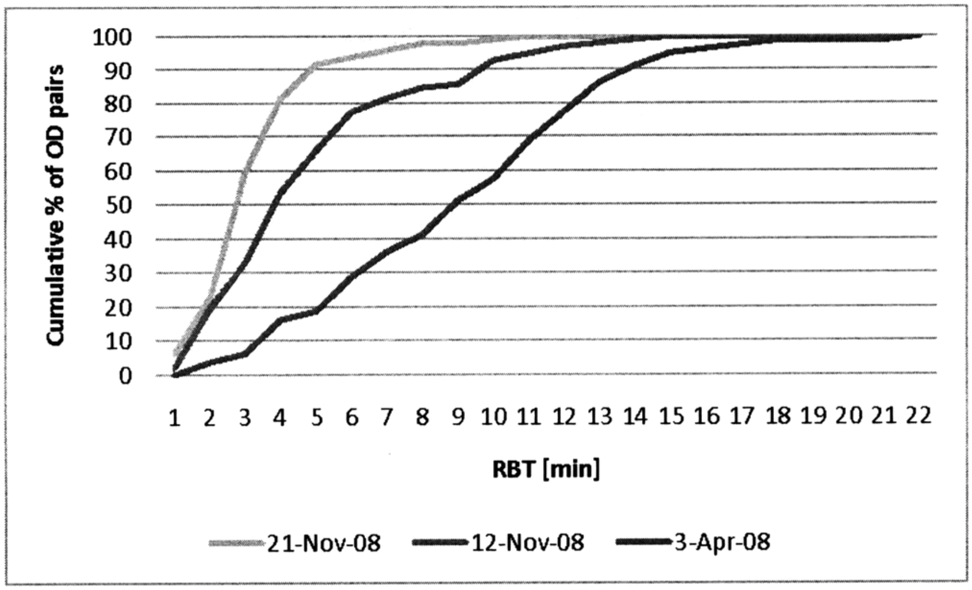


Figure 7-5: Cumulative density functions for the RBT on the three days of interest

The difference between westbound and eastbound RBTs along with the high RBT for April 3 eastbound suggests that further research is needed into the behavior of passengers under such circumstances. It is possible that eastbound passengers who are just beginning their journey are more willing to wait out the gap – thus contributing to the high RBT on April 3 eastbound – whereas westbound passengers who are "dumped" off their train because of a short-turn may be less willing to do so and instead have a higher tendency to leave the system. However, this is currently only a hypothesis without any support, and further research into it would be needed. While it may be possible to confirm or disprove this hypothesis through the analysis of Oyster data, this thread was not further pursued in this research.

Finally, a t-test was conducted to test the null hypothesis:

H_0: The means of the two RBT distributions (April 3 or November 12 and November 21 eastbound and westbound) are equal.

The assumptions of the test were exactly the same as in section 7.2.3, and the respective t-statistics are shown in Table 7-15. Aside from westbound trips on November 12, the difference in RBT between November 12 and April 3 is highly significant.

Despite the challenges in clearly quantifying differences between these two short-turn days, the bottom line of the analysis is that service to passengers to and from stations west of White City was significantly worse on April 3 than on November 12, both in terms of travel times and reliability of service, despite the fact that on both days, service controllers were dealing with similar disruptions. The analysis of operational data suggested that the difference in response to the delay may have been caused by crew management issues. Furthermore, based on the operations data, it can be stated that:

- All trains which were short-turned or diverted on April 3 were cumulatively 2h 37min late as they passed Shepherd's Bush station, immediately before reaching White City where the first ones were short-turned. Assuming that every train would have been able to compensate for 6 min of lateness at reversal, the cumulative lateness eastbound (when most of the crew reliefs would have taken place) would have been 1h 49 min.
- The lower bound for total "excess" passenger delays, i.e., delays suffered by passengers on April 3 but not on November 12, amounts to 268 passenger hours, not controlling for the differences in demand.

Assuming that train lateness reflects driver lateness, this would mean that 2h 37min (or a projected 1h 49min) of driver time were traded off against at least 268 hours of passenger delay.

However, this result is not meant to denigrate the service controllers who were on duty on April 3. Lacking complete information on the actual decisions they were facing, it must be assumed that they were acting to the best of their knowledge and in a way to best accommodate the *perceived* constraints and objectives. It is those constraints and objectives which may need to be reviewed in order to avoid such situations in the future, notably:

- This case illustrates that there would be clear advantages to better monitoring the utilization of spare crews and to adjust spare crew rosters to provide maximum availability during times in which service controllers are known to perform many interventions, for example after the AM peak or directly before crew pickups.
- Crew pickup locations might also warrant a review, especially the location of the crew depot at Leytonstone which appears to pose some operational challenges to service controllers (compare to section 4.3.3 b).
- Last but not least, further research would be needed into the design of crew schedules, knowing that train and driver delays occur frequently. For instance, one might want to review how tight maximum driving time constraints should be made in the crew scheduling process and especially what the Underground's policies are towards crews stepping off late. For

> instance, one might want to consider whether every driver's schedule should include slack time at the end of a piece of work for which the driver were to be paid as if he or she were driving, but where the driver would already have stepped off the train under ideal circumstances. In the case of driver delays, that time would be available as additional driving time at no marginal cost.

From the point of view of service control, the management of crews is also a matter of information management. In the case of disruptions, a controller needs to keep track of multiple late drivers, and there is often not enough time to look into each of the driver's specific situation regarding driving time constraints and lateness. Chapter 4 noted that gathering information on driver lateness and spare driver availability is a time-consuming process for a service controller since the information needs to be requested from the DMTs by telephone and in some cases from the drivers by train radio. However, the number of stations along the line where trains can be short-turned is limited and not short-turning a train on which the driver is at risk of hitting a maximum driving time constraint can cause a severe problem at a later point in time.

On April 3, the time which passed from the moment the disruption cleared until the time at which the first short-turning point (White City) was reached was not even 8 minutes, and the next short-turning point (North Acton) may not have had enough capacity given the number of delayed trains. Therefore, it must be assumed that service controllers did not have enough time to get a complete picture of crew management constraints, and their decision to short-turn virtually all trains after the disruption may have been driven by a lack of information on which ones really needed to be short-turned to meet these constraints. On the other hand, as the disruption on November 12 occurred at Stratford, controllers had much more time to find out about crew management issues and to organize spare drivers since the earliest possible short-turning point again was White City, which was not reached by the first of the delayed trains (set #6) until approximately 28 min after the blockage had cleared. This finding leads to the recommendation in section 8.2.2.

7.4 Case 3: Disruption management

7.4.1 Motivation and research question

This section, which will not go into as much detail as 7.2 and 7.3, aims to show a disruption management strategy and a potential chokepoint on the line possibly warranting further research. Section 4.3.5 stated that:

[due to infrastructure capacity constraints,] "it has become standard practice among Central Line controllers to "thin out" the service on [...] the line by removing a certain number of trains if a long disruption is anticipated".

However, when to begin those cancellations and how many trains to remove is largely up to the judgment of the individual controller.

In the course of the analyses of Central line operations, it was observed that in the case of disruptions, North Acton junction, White City and to a lesser degree Leytonstone junction can become congested and thus cause further delays to trains independent of the disruption. According to information from Central line controllers and depot staff at White City with whom the author was in contact, these problems were exacerbated by the fact that a crossover between the center track and the eastbound track in White City[14] was unusable for some considerable time, forcing reversing trains or trains entering or leaving the depot to perform additional movements which could interfere with scheduled through train service.

A standard practice for alleviating such congestion problems, as mentioned above, is the cancellation of trains. A particular example of this procedure, which was observed in the data for April 1, 2008 as part of the analysis of case 1 (section 7.2), will be presented here. Specifically, the question is whether it is possible to quantify the delays caused by this disruption management strategy, as it left gaps in the service, and how those delays compared to those encountered due to the disruption itself.

7.4.2 Analysis procedure

As mentioned, this disruption was discovered during the analysis of the train cancellation strategy which was employed during the afternoon of April 1 and was presented in section 7.2. Specifically, while looking at average travel times on one particular OD pair (Notting Hill Gate – Liverpool Street) throughout the day, the author noticed an interesting delay pattern during the morning of April 1, between 08:00 and 09:00. The Underground's incident log was consulted and operational data for the eastbound service during that time window was subjected to a more detailed analysis, which showed that the disruption seems to have caused the congestion problems at White City that controllers had referred to during the author's visit in January 2009. The author is aware of at least two other days during the analysis period, November 18th and November 20th, 2008, on which a disruption led to trains being severely delayed around North Acton junction, Leytonstone junction

---

[14] This refers to points 2715/2716. As of May 2009, they were usable again.

and White City station. Based on the explanations given by controllers, it can be assumed that the strategy of canceling trains to avoid congestion, which was observed in dealing with these problems on April 1, was in fact common.

7.4.3 Operational analysis

The service delivery and intervention matrix is shown in Table 7-16. The corresponding service delivery plot can be found in appendix E. The disruption occurred westbound between 08:00 – 09:00 on section 3. This caused two gaps in the service (as will be discussed below), leading to a drop in service provision on sections 3 and 4, and trains were held upstream of the disruption between Lancaster Gate and White City, both of which are in section 3. The reduced service on section 2 between 08:00 – 09:00 cannot be explained by the effects of the disruption, and a closer look at the operational data reveals that there appear to have been a few missing or late trains from that branch before the disruption. The analysis of the operational data showed the following sequence of events. Eastbound service on the Central line was operating without disruptions until approximately 08:30. Table 7-17 shows the eastbound headways of trains passing White City station. As one can see, the headways were somewhat irregular, falling between 1:06 min and 5:42 min. The latter was a gap between 08:09:44 and 08:15:26, which was caused by a train departing West Ruislip late, although nothing was registered in the incident database. At approximately 08:29, an eastbound train (set #141) suffered an in-tunnel delay between Lancaster Gate and Marble Arch stations due to a track circuit failure. The blockage was cleared at approximately 08:35, and train set #141 departed Marble Arch station at 08:36:44. Due to the blockage, trains were held upstream at all stations until White City. Presumably this was in part driven by passenger safety concerns as outlined in section 4.3.7. The train which was held in White City station was set #30, which had arrived at 08:31:36 from Ealing Broadway and was held by service control until 08:40:00. During that time, service controllers decided to cancel two trains and withdraw them to the White City sidings. Those were set #114, which was reversing at White City, and set #105, which was directly behind set #30 and had arrived from West Ruislip. Set #105 was approximately 70% loaded, according to the TDA data. Neither train appears to have been defective since they were kept in White City sidings and re-inserted into service for their scheduled departures westbound, i.e., one round trip from White City was canceled. Unfortunately it is not possible to reconstruct the exact reasons for this decision, as the data does not convey the full picture, but the two following reasons are possible:

- Congestion: Service controllers may have been concerned about train congestion building up between White City and North Acton which would have persisted even after the disruption had cleared due to dwell times at White City. This might have been exacerbated by the defective crossover.
- Lateness: Canceling two trains may also have been a way of removing them from the pool of late trains (and drivers), thus reducing the workload of controllers and positioning two "spare" trains at White City which could be used later for service restoration.

		Service operated (% of total scheduled service)								Interventions					
		West branches		Trunk		East branches									
	Section no.	1	2	3	4	5	6	7	8	Short-turns	Diversions	Extensions	Cancellations	Unscheduled	Withdrawn
Westbound	08:00 - 09:00	29.7 (90%)	108.5 (76%)	358.7 (87%)	279.6 (89%)	73.1 (104%)	111.7 (80%)	33 (100%)	123.6 (109%)		1	1	3		
	09:00 - 10:00	33 (100%)	156.3 (99%)	353.8 (92%)	213.2 (82%)	49.8 (91%)	73.5 (65%)	13.3 (73%)	96.7 (83%)	1		2	3		2
Eastbound	08:00 - 09:00	33 (106%)	109.9 (91%)	302.3 (81%)	219 (79%)	52.7 (87%)	95.5 (81%)	18.2 (100%)	113.1 (100%)	1		1			
	09:00 - 10:00	36.3 (104%)	129.5 (95%)	315.8 (84%)	217.7 (77%)	45.4 (81%)	97.7 (71%)	15.8 (87%)	80.7 (67%)	3		1	2	1	

Table 7-16: Service delivery and interventions on April 1, 2008

The result of these two cancellations was that there was a gap of 7:20 min in the service after the disruption had cleared. Table 7-17 shows the train headways at departure from White City station after 08:40:00. One can see that after train set #30, there was a 7:20 min gap caused by the two missing trains which were canceled, and after that, trains passed at short headways. What is problematic is that this was during the peak of the peak, such that train capacity became a problem: Train set #56, which was the first train after that gap, departed White City at crush load since it had absorbed all passengers from #105, and it remained at crush load until Holborn. The following train, set #31, had also reached crush load by Marble Arch and remained at that loading until Holborn. The removal of two trains from service during the peak hour, when demand and supply regularly tend to "bump up" against each other on the Central line, had reduced throughput capacity by approximately 2100 spaces in that short period of time (two trains with a capacity of 1047 passengers each). As a result, there was very little additional capacity for downstream passengers, and presumably many passengers were not able to board the crush loaded trains, thus causing them to be left behind and experience longer waiting time than could have been attributed to the gap alone.

Set no.	Trip no.	Scheduled Departure	Observed Departure	Headway [min]	Origin	Destination	Loading at departure [%]	Unit ID
20	4	08:00:40	08:02:24	–:--	WHC	NEP	16	121
22	5	08:02:56	08:03:40	01:16	EAB	DEB	48	124
45	2	08:05:10	08:05:22	01:42	WER	WOH	48	142
24	2	08:07:10	08:07:20	01:58	EAB	EPP	48	127
70	4	08:09:26	08:09:44	02:24	NOA	HAI	23	156
25	2	08:11:40	08:15:26	05:42	WER	EPP	74	128
111	2	08:13:56	08:16:56	01:30	EAB	NEP	55	166
113	2	08:16:10	08:18:16	01:20	WHC	LOU	19	168
141	4	08:18:26	08:19:22	01:06	WER	HAI	39	172
26	2	08:20:40	08:20:54	01:32	EAB	EPP	58	129
46	4	08:25:10	08:25:50	04:56	EAB	WOH	65	143
27	4	08:27:10	08:27:44	01:54	WER	EPP	65	130
71	2	08:29:26	08:30:14	02:30	NOA	HAI	32	157
30	2	08:31:40	08:40:00	09:46	EAB	EPP	68	132
105	3	08:33:56	Canceled	–:--	WER	NEP	70 (a)	135
114	3	08:36:10	Canceled	–:--	WHC	LOU	0 (b)	126
56	3	08:38:26	08:47:20	07:20	EAB	HAI	100	152
31	4	08:40:40	08:49:00	01:40	NOA	EPP	48	133
72	2	08:45:10	08:50:30	01:30	EAB	HAI	65	153
142	2	08:42:56	08:52:00	01:30	WER	HAI	81	173
32	3	08:47:10	08:53:34	01:34	NOR	EPP	45	131
47	4	08:49:26	08:55:16	01:42	EAB	WOH	45	146
33	2	08:53:56	08:57:12	01:56	WHC	EPP	6	107
106	4	08:51:40	08:58:22	01:10	WER	NEP	32	158
57	2	08:56:10	08:59:50	01:28	EAB	HAI	39	154
115	4	08:58:26	09:01:38	01:48	NOA	LOU	10	169

(a) - Loading at arrival

(b) - Scheduled to reverse at White City

Table 7-17: Departures from White City eastbound on April 1, 2008 between 08:30 and 09:00

7.4.4 Passenger impact

For the measurement of the passenger impact of the disruption, only eastbound passengers on the trunk section were of interest as they were the ones directly affected by the disruption and the following gap. As explained in section 7.3, a stationary time window was defined for stations upstream of the disruption (i.e., upstream of Marble Arch, since the disruption occurred between Lancaster Gate and Marble Arch) and a moving time window was defined for stations downstream of it, as shown in Table 7-18. Passenger travel times were compared with the same time periods on November 21. Timeband 1 included the 30 minutes before the end of the disruption, including the disruption itself while Timeband 2 captured the 30 minutes immediately after the disruption had cleared.

	Timeband 1		Timeband 2		
	From	**Until**	**From**	**Until**	
NOA	8:05	8:35	8:35	9:05	Upstream of blockage
EAA	8:05	8:35	8:35	9:05	
WHC	8:05	8:35	8:35	9:05	
SHB	8:05	8:35	8:35	9:05	
HOP	8:05	8:35	8:35	9:05	
NHG	8:05	8:35	8:35	9:05	
QUE	8:05	8:35	8:35	9:05	
LAG	8:05	8:35	8:35	9:05	
MAA	8:05	8:35	8:35	9:05	
BOS	8:07	8:37	8:37	9:07	Downstream of blockage
OXC	8:09	8:39	8:39	9:09	
TCR	8:10	8:40	8:40	9:10	
HOL	8:12	8:42	8:42	9:12	
CHL	8:14	8:44	8:44	9:14	
STP	8:16	8:46	8:46	9:16	
BAN	8:18	8:48	8:48	9:18	
LIS	8:20	8:50	8:50	9:20	
BEG	8:23	8:53	8:53	9:23	
MIE	8:25	8:55	8:55	9:25	
STR	8:29	8:59	8:59	9:29	
LEY	8:32	9:02	9:02	9:32	
LES	8:35	9:05	9:05	9:35	

Table 7-18: Time windows for analysis of passenger travel times on the morning of April 1

The results are shown in Table 7-19. The results are very interesting in that they show the impact of the disruption management strategy in comparison to the delays caused by the disruption itself. Firstly, the travel times per passenger hardly differed between the two timebands, and one can see that they are both higher than on the baseline day, November 21. As a matter of fact, passengers after the disruption effectively experienced the same service degradation with respect to November 21 as that during the disruption, as is evident from the travel time differences per passenger. Aside from the delays to eastbound passengers on the trunk portion (sections 3 and 4) which are quantified here, there were additional passenger delays not captured by this analysis which occurred later in the westbound direction since the two canceled trains were missing for an entire round-trip. Secondly, in terms of the unreliability of the service, the timeband including the disruption does not show a large increase compared to November 21, amounting to slightly more than 3 min. In other words, passengers during the disruption experienced longer overall travel times than under normal peak-hour conditions, and the service was somewhat less reliable than usual. However, the situation is more severe after the disruption, where the total additional RBT (1880h – 744h = 1136h) amounted to 32% of the total travel time under undisrupted conditions or slightly less than 6 minutes per passenger. Not only did the interventions cause additional waiting times of the same magnitude per passenger as the delays caused by the disruption itself, but the service became considerably less reliable. This is presumably due to the two gaps left by the removed trains as well as to passengers left behind by the two crush-loaded trains mentioned earlier.

	Timeband 1 (before and including disruption)		Timeband 2 (after disruption)	
	1-Apr	21-Nov	1-Apr	21-Nov
ODs sampled	81		81	
(total number of ODs)	182		181	
Trips sampled	8963		12231	
(total number of trips)	11860		16378	
Total travel time [h]	2844	2472	4014	3509
Total travel time difference [h]	372		505	
Travel time per passenger [min]	19.04	16.55	19.69	17.21
Travel time difference per passenger [min]	2.49		2.48	
Total RBT [h]	1098	591	1880	744
RBT per passenger [min]	7.35	3.96	9.22	3.65
RBT difference per passenger [min]	3.39		5.57	

Table 7-19: Results of passenger impact analysis for the morning of April 1

For more detail, Figure 7-6 shows the average passenger travel times by departure time for one particular OD pair, Lancaster Gate to Liverpool Street, along with the train departures from

Lancaster Gate on April 1 and November 21. This OD pair was chosen because it has high passenger volumes[15] and has its origin at the location where the disruption occurred. It is therefore very easy to know when the blockage affected trains from this station. One can associate the variability in average travel times on April 1 before 08:28 with the irregularity of service before the disruption (as reflected in the large gap in train departures between 08:19 and 08:25). Starting with the steep increase in average travel times at 08:28, when the disruption occurred, average travel times show markedly higher variability than before the disruption, but once it has cleared and the first train after the disruption passes at 08:38, travel times fail to normalize and remain highly variable due to the cancellations as discussed above. (In the April 1 dataset, two outlier travel times of more than twice the baseline average travel time (i.e., of more than 55 min) were excluded since it was felt that those represented passenger behavior not directly linked to operations on the line.)

Admittedly, this plot also raises several questions which cannot be answered here. For instance, one can see in the plot that there was a 7-minute gap starting at 08:52 on November 21 (marked with (a) in Figure 7-6). It causes a spike in travel times, which is marked in the graph with (b). On April 1, there was gap of similar duration at approximately the same time, but the magnitude of the corresponding increase in travel times is larger than on November 21 (marked with (c)). This cannot be explained by operations alone, and further research is needed in order to determine the actual causes. On a cautionary note, although the sample sizes of the two days are approximately the same, the distribution of the entries between 08:00 – 09:00 was not verified, i.e., some of the variability of the travel times on April 1 may be due to a small sample size at any given minute.

Notwithstanding this, a possible explanation for the high variability of travel times on April 1 may be that information about the delays communicated to passengers through public announcement systems prompted some passengers to change their behavior in at least one of two ways:

- Use an alternative route by taking the Central line in the opposite direction and transferring to the Circle line at Nottinghill Gate. TfL's official travel planner estimates a travel time of 34 min for this trip, so it is plausible that some of the passenger trips shown in Figure 7-3 actually reflect that route.
- Return to the gateline in order to obtain information from the station supervisor. This too can lead to a noticeable delay since access from the gateline to the platforms at Lancaster Gate is

---

[15] On November 21, the sample includes 194 trips (on average 3.2 entries per minute). On April 1, it includes 187 trips (on average 3.1 entries per minute)

via an elevator, and a passenger may easily lose 5 – 10 min by going from the platform to the ticket hall and back.

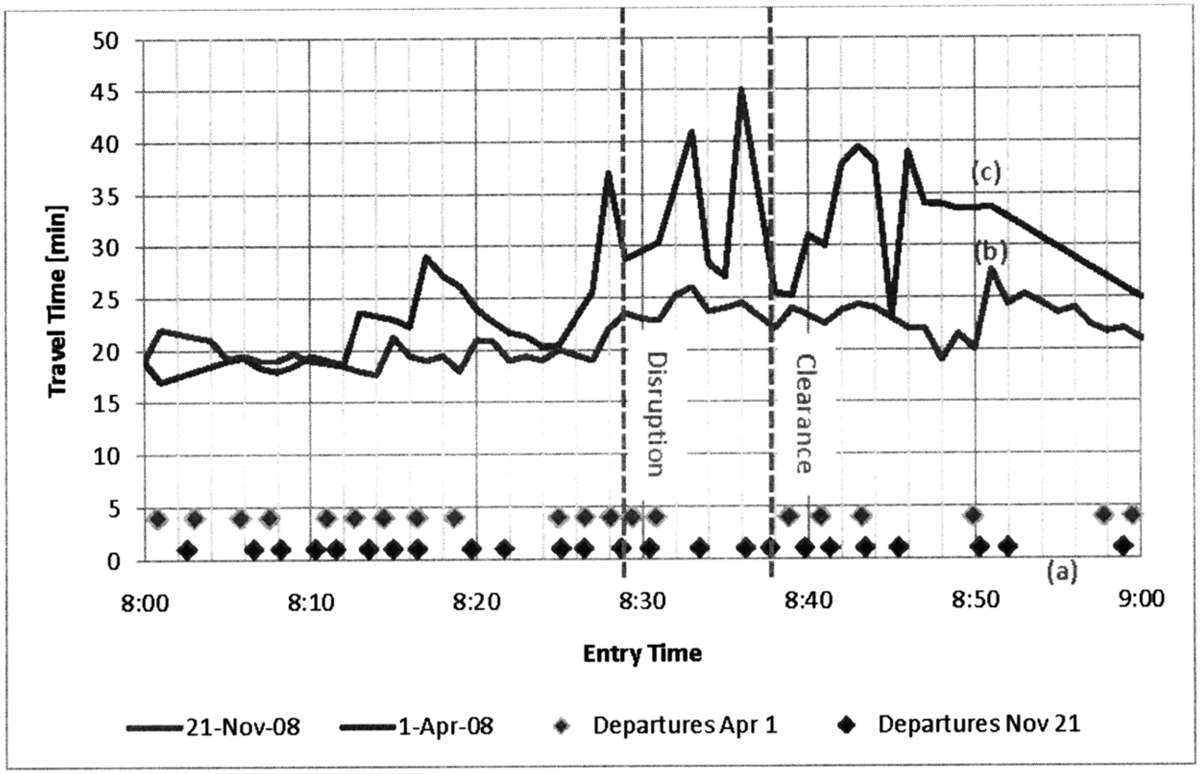


Figure 7-6: Average passenger travel times by entry time from Lancaster Gate to Liverpool Street

7.4.5 Discussion

Two possible reasons were suggested for the controllers cancelling the two trains. The first reason was congestion, and the question must be raised whether congestion at White City could have been avoided in the first place by taking appropriate measures. If westbound trains had been held not only between Lancaster Gate and White City but also between White City and the terminals at Ealing Broadway and West Ruislip, it is likely that passenger delays after the disruption would have been markedly lower. The second possible reason was crew lateness (i.e., controllers were worried that, in view of the delays, too many drivers were going to be late for their reliefs, so by canceling two trains they produced two spare trains plus two drivers which were not late), in which case it would be questionable whether controllers were following the optimal priorities, as the tradeoff of passenger delays and driver delays would presumably be even more out of balance than the one calculated in section 7.3. The fact that only trains inside the tunnel sections of the line's trunk section were held might be partially linked to the design of the OCS, as it does not allow the holding of a batch of trains on a branch but rather requires controllers to insert a hold on every train separately, and in addition only allows holding at certain designated stations ("controlled areas"). As a result, controllers will hardly dedicate more time than necessary to holding trains when their full attention is required by other events, such as an ongoing disruption.

As there might be further operational issues which have not been identified at this point, the author believes that this case merits further research based on the insights presented here. However, this case does illustrate that there is a need for more unified policies regarding not only service recovery but also the disruption management process. Furthermore, the author believes that it would be worth conducting targeted research into the causes for delays at the Central line junctions (North Acton and Leytonstone) and between North Acton and White City, both under disrupted and normal conditions.

7.5 Case 4: Timetable parameters

7.5.1 Motivation and analysis

This final section investigates the influence of timetable variables on the frequency of service control interventions. Specifically, it examines how the timetable changes made in response to rising passenger demand on the Central line have affected the workload of service controllers.

On November 19, 2006, the Central line switched from working timetable 63 to working timetable 64, which was still in effect as of May, 2009. There were only two notable changes introduced by working timetable 64. Most importantly, the off-peak baseline frequency through the trunk was increased from 21 to 24 trains per hour. Appendix C shows the differences in scheduled trains per hour for a few selected stations on the Central line. Aside from that, the running times were slightly increased in working timetable 64, by approximately one minute in each direction. However, according to conversations with London Underground staff, most of that additional time was added to the branches and not to the congested trunk portion of the line. In addition to the slight increase in running time, the increase in frequency essentially resulted in higher rolling stock requirements during the off-peak, i.e., reduced numbers of spare vehicles for recovering from lateness incurred during the peak hours.

The question was whether and how that was reflected in the number of service control interventions performed on the Central line. To answer these questions, operational data were analyzed from two months: October 2006 and October 2007. Unfortunately, no Oyster data was available for these two time periods in order to analyze changes in passenger travel times as was done in previous cases.

According to information obtained from London Underground staff, demand during that year increased by approximately 10% on the Central line, leading to increased passenger congestion during peak hours. October 2006 was the last full month before the timetable change, and as a comparison, October 2007 was chosen for two reasons, namely to control for seasonal effects and to avoid the adaptation period directly after a timetable change.

Operational data was obtained for weekdays of both months to avoid weekends on which engineering work took place. Unfortunately, a further two days from the October 2006 dataset had to be excluded due to data errors, leaving 20 days in October 2006 and 23 days in October 2007 for this comparison.

Appendix F shows a plot comparing the two periods, and Table 7-20 summarizes the average number of interventions during these two periods. Two days stand out in the plot: October 20, 2006 and October 16, 2007, both of which experienced severe disruptions with more than 100 canceled train trips, which is why the summary statistics in Table 7-20 excludes them. By visual inspection of the plot, it appears that not only the average number of daily service control interventions increased between 2006 and 2007, but also the frequency of days with high numbers of short-turns and cancellations (even excluding the two severely disrupted days). This is confirmed by the summary statistics in Table 7-20.

	Oct 2006 Avg	Oct 2006 StDev	Oct 2007 Avg	Oct 2007 StDev	t-statistic
Total interventions per day:	46.9	35.7	69.9	39.3	1.96
Short-turned trips:	18.5	18.0	25.4	19.7	1.16
Diverted trips:	5.6	3.9	8.2	6.3	1.59
Extended trips:	6.4	4.9	9.4	3.6	2.22
Trip cancellations:	7.4	7.6	15.5	8.6	3.21
Unscheduled trips:	4.8	5.5	8.5	7.8	1.78
Out of service trips:	4.2	2.1	2.9	2.2	-1.94

Table 7-20: Differences in service control interventions between Oct. 2006 and Oct. 2007

	Oct 2006	Oct 2007
Days with 20 or more short-turns:	5	11
Days with 20 or more cancellations:	1	6
(Total number of days analyzed:)	(19)	(22)

Table 7-21: Frequency of days with more than 20 short-turns or cancellations

The total number of interventions appears to have increased significantly from October 2006 to October 2007. The table shows that this difference is mostly due to increased numbers of short-turns and, on average, more than double the number of canceled trips per day in 2007 than in 2006. The t-statistic for the difference in short-turns may be misleading; it is relatively low due to the large standard deviations. For that reason, Table 7-21 shows how many days of the sample (19 days in 2006, 22 days in 2007 excluding the severely disrupted days) had more than 20 short-turns and more than 20 cancellations. Again, the results confirm the increase.

7.5.2 Conclusion

Despite the slight increase in running times, it appears that the increase in number of trains scheduled to run on the line led to less “slack” for recovery from disruptions without major changes to the service. Of course it may also be that the increased number of trains caused more train congestion and late-running in the first place since a disruption in the off-peak would perturb service more easily at higher frequencies.

The cause of the disruption on November 12 at Stratford (loss of pilot light due to severe overcrowding), which was presented in section 7.3, suggested that there is a certain link between congestion and the occurrence of disruptions on the Central line. This may be due to an increased probability of passenger problems as more people are on the trains, additional strain on rolling stock or the fact that peak service rolling stock requirements are very high, compromising train maintenance schedules. Whatever the cause, it is common that at the end of the morning peak, service controllers are often faced with defective trains which need to be withdrawn and late running trains which need to be put back onto schedule for various reasons. A lower level of service in the interpeak as in working timetable 63 makes it easier for controllers to manage trains such that late ones and ones which need to be withdrawn can be renumbered and thus “switched over” with trains scheduled to stable after the peak hour. In that sense, even if the same numbers of disruptions occurred before and after the timetable change, under working timetable 63 their effects could often be dealt with using the aforementioned method, with less need for interventions such as short-turns, diversions or cancellations.

Generally speaking, this analysis also shows that in view of the reliability of the rolling stock, the maintenance reserve for the Central line is probably too small for the requirements of working timetable 64. Whether there are other issues involved, and in what way passengers were impacted, would need to be the subject of further research, which should partly focus on how the distribution of service control interventions throughout the day changed between working timetables 63 and 64.

Nonetheless, these results correlate well with the experiences of Central line service controllers as described to the author as well as to other London Underground staff. That is, there is a tradeoff between the degree of utilization of line capacity, both in terms of infrastructure capacity and rolling stock/crew resources and the stability of the service including the number and nature of service control interventions needed for managing it.

7.6 Possible improvements to impact measures

This chapter applied the measures introduced in chapter 6 in a practical setting. Although they were found to provide useful results in the analysis of the Central line presented here, various possible improvements were noted which could enhance the effectiveness of these measures as discussed in what follows.

Generally speaking, one of the most difficult tasks when analyzing service control lies in dealing with large quantities of information from different sources, which ultimately need to be connected. Hence, the challenge lies not only in procuring the data, but (perhaps even more importantly) in aggregating it and consequently presenting it. Given the variety of information sources, it is desirable to present as much information as can be reasonably presented in one plot or table. Therefore, the author believes that future research on service control will still need to be based on visual inspection of data and measures. It would be beneficial to tie together all the information visually in one or two plots, as this might reveal patterns or connections which are difficult to see if the data derived from separate sources are inspected separately. One excellent way of doing this would be to integrate the disruption and intervention log into a space-time plot.

Furthermore, the passenger travel time metrics would need to be developed into a sufficiently disaggregate, continuously monitored line- or section-level variable, which could then be graphically integrated with the measures of level of service and ultimately into the space-time plot.

The author would also like to suggest the following possible improvements to space-time plots which could be used in future research:

1. The experience with Central line DMA data showed that it is very helpful for an analysis of service control to follow the movements of individual vehicles throughout the network.

Incorporating vehicle linkages at terminals in space-time plots, as originally proposed by Froloff, Rizzi and Saporito (1989), helps detect train priority decisions visually or explain the withdrawal of trains from service, if it is known what they were used for subsequently.

2. Along similar lines, to the author's knowledge no researcher has yet attempted to include crew data in space-time plots. Although this may be a challenging task, it can potentially help explain the linkages between crew management and service control. For example, it would be easy to detect short-turns which occurred on the penultimate trip of a driver before being relieved.
3. A further enhancement to space-time plots would be the incorporation of the planned trajectory of a train, which would be a graphical rendition of the principle which was used in the aggregation of the Central line dataset, i.e., the juxtaposition of planned with observed train movements.

The reader wishing to pursue work on the aforementioned graphical improvements is referred to Tufte (1983).

A final point that merits attention is the measurement of the level of service. So far, the train service kilometers have been calculated using fixed time windows. However, there are possible benefits to using this or an alternative measure with rolling time windows, as this might relate better to the momentary level of service provided at any given point in time. Before deploying such a measure, preliminary research would be needed into determining whether it provides an improvement over fixed time windows in quantifying the supply side of the train service.

7.7 Summary

This chapter has demonstrated how the integration of various data sources for the study of service control interventions can provide important results and point out areas where future research is needed in order to understand the operations on a line and the effect they have on passengers. Specifically in the case of the Central line, two areas have been identified in which a review of organizational structures and policies may provide benefits: crew management and train holding during disruptions. Furthermore, it has provided some insights into the implications of increasing the scheduled utilization of resources (in terms of trains, crews and infrastructure capacity) on an otherwise unchanged line which is operating near capacity. Not only does that lead to higher overall levels of service control interventions, but controllers also need to devise strategies to deal with train shortages since this decreases the number of spare trains available. In the absence of official policies, controllers are likely to address such problems by using rules of thumb which do not necessarily minimize passenger impact. Furthermore, several areas for possible future research have been identified, the most prominent of which is the behavior of passengers under disrupted conditions.

In addition, this chapter has also shown that conducting such research at an aggregate level remains a complicated task and that is it often difficult to isolate the effects of individual events or decisions from other noise in the data. For that reason, several improvements to the measures presented in chapter 6 are suggested.

Despite those problems, the type of results presented in this chapter has, to the author's knowledge, not been obtained before and therefore represents an important step towards closer integration of operational analysis with passenger data. Institutionalizing such an analysis process can be a powerful tool for evaluating observed controller decisions, for discussing their implications and designing improved decision rules.

8 Final remarks

The first section of this chapter, 8.1, summarizes the thesis and emphasizes its main findings and conclusions on a general level. The following section, 8.2, presents the main recommendations which are made to the London Underground as a result of the Central line service control analysis. It is divided into three areas: data collection, the structure of responsibilities within the London Underground and the design of the operations control system. Finally, section 8.3 describes what the author believes to be the most interesting threads for future research arising from this thesis.

8.1 Summary and conclusions

This thesis has taken a broad approach to the topic of service control on a high-frequency metro line. Since the specifics of the research and applications were focused on the London Underground Central line, several of its findings are specific to that line. However, the overall process, the description of the service controllers' decision environment, the measures for assessing the impact of interventions and the principal idea behind the intervention reconstruction algorithm are not specific to the Central line. Therefore, many of the conclusions are applicable to other metro lines and systems as well, as discussed in the respective sections of chapters 4 through 6. Of course, an analyst wishing to study a different line would need to adapt them to the characteristics of that line and to the available data, but the overall approach could largely follow the one described here. Moreover, as other transit agencies may have better crew data available, the limitations encountered in the analysis of crew management on the London Underground may not apply to the analysis of lines on other metro systems.

The first part of this thesis introduced the basics of service control and reviewed previous research and literature in this area. The review showed that to date, research into service control has virtually all been focused on individual components of service control and control strategies for special situations, and has been heavily based on modeling. Furthermore, previous research has always assumed that the primary objective of service control is the minimization of passenger travel time in the face of unreliability caused by external factors. While that may be an important aspect of service control, it by no means constitutes all the objectives and constraints that face a service controller. The author spent approximately two weeks in the control center of the Central line, and the result, chapter 4 of this thesis, provides a much more complete set of decision factors which were observed to cause service controllers to perform interventions or which influenced interventions performed for other reasons. The main drivers of service control which were observed were considerations about the level of service to passengers, crew management, rolling stock management, safety and infrastructure maintenance.

Aside from these considerations, virtually all decisions are influenced by uncertainties regarding the outcome of an intervention and concerns about the manageability of the service. It was seen that the reliability of the system depends on many factors which are endogenous to it and which may previously not have been recognized. In the absence of official policies or effective decision support, the management of these factors is often governed by rules of thumb. As a matter of fact, it can be stated that service control not only works to manage unreliability caused by exogenous events but can also be the cause of unreliability as controllers work to meet other objectives and constraints.

Therefore, any effort to improve service control (and thus, operations in general) on a specific metro line must build on a solid understanding of how that line operates. While this may previously have been a task which was achieved mainly with the help of models, the roles of modeling and data analysis in transit operations are shifting in light of the increased availability and accessibility of automatically collected operations and passenger travel data. This thesis aimed to point out the significant value such data can provide to a transit agency in helping to better monitor operations on a line with the goals of:

- Extracting and representing data from multiple sources, allowing the analysis of an operational question from various perspectives.
- Developing an integrated framework for studying service control interventions which takes into account what was learned from the control center visit.
- Building a better understanding of service control and its interaction with line characteristics and scheduling variables.

Two of the main elements of the developed framework are (i) a procedure which shows how a transit agency can make use of signaling system data to reconstruct daily operations on a line and infer service control interventions by using relatively simple algorithms, and (ii) a description of the service controller's decision environment with all objectives and constraints which were observed during the author's visit to the control center. In addition, a set of measures is proposed for capturing the main ingredients of monitoring operations on a line – service control, the level of service and the passenger experience. In a series of cases on the London Underground Central line, the framework was applied in order to demonstrate its implementation in a practical setting, gain insights into three common service control strategies and assess the impact of a timetable change on the way service control is performed on the line. The results could ultimately improve the scheduling process and the way line performance is monitored.

Overall, these applications showed that although the Central line service controllers have established strategies for dealing with certain frequent problems such as, for example, congestion around one of the depots, they are not always to the benefit of passengers. They also showed that simply counting interventions does not suffice for monitoring service control – the temporal and spatial distribution of interventions are equally important variables which define how a single intervention fits into a broader strategy. In part, the problem that service control strategies can sometimes have negative impacts on passengers could be remedied by implementing common policies and standards which frame the overall priorities of service control and avoid or diminish these negative impacts. However, such policies alone cannot provide the solution. One of the examples showed how service control decisions are influenced by crew management (and thus ultimately by the design of the crew schedules). Another example made a link to the design of the operations control system (OCS), showing that the difficulty of holding trains in the OCS may influence a controller's decision to cancel trains instead. Therefore, sustained improvements can be made only by accompanying said policies with changes in vehicle and crew scheduling and, in the long run, in the OCS design.

8.2 Recommendations to the London Underground

This section provides a set of recommendations to the London Underground. They are divided into three topics: Data collection, organizational structure and the design of the OCS.

8.2.1 Data collection

Based on an in-depth and comprehensive use of multiple sources of data in conjunction with the applications discussed above, several recommendations arise:

1. In line with the findings of Wile (2003), the design of databases used to store automatically collected operational data is important and requires thinking about possible uses before their deployment. Simple design attributes, such as one set of meta-data to ensure that codes, abbreviations and formats are consistent across different datasets can save many hours of effort at a later stage. In the course of this research, the author found that the availability of event-driven logs is more important than the availability of time-driven logs. And, a good level of detail in the log recordings is necessary around depots and in stations with multiple

platforms; ideally, the data should make it possible to discern train movements to or from siding tracks from movements at passenger platforms.

2. As preliminary work towards improving the modeling of line behavior after disruptions with TSM, which could ultimately be extended to include a rudimentary set of service control interventions, incident reporting forms should be improved. From an operational point of view, the categorizations which were laid out in sections 2.2.2 and 2.2.3 would be important, i.e., the type of disruption and the immediate effect on the line.
3. In the course of this analysis, it was found that two data items deserved particular attention in terms of potential improvement. First, the indication of whether a train is in or out of service, though already recorded in the Central line data, appears to be unreliable. More reliable information would help make a clearer distinction between trains which provided passenger service and those which didn't. Second, and more importantly, the Underground could gain significant value from more reliable crew data (cf. section 5.3). This could be achieved either by enforcing operator logins more strictly and establishing clear rules on how spare drivers log in or by the deployment of a personal chip card which drivers must insert into a reader in order to enable train controls. The author would like to point out the positive experiences of the Swiss Federal Rail company, SBB, which found that by storing the driver's seat adjustments and driving console settings on a chip card, the minimum turnaround time of mainline trains could be reduced from 4 minutes to 1.5 minutes if combined with a driver relief (Bosshard, 2008).

8.2.2 Structure of responsibilities

Section 4.3.3 showed how strongly linked train service control and crew management are and described strategies which have been developed by the controllers on the Central line to tie crew management constraints into their decision-making process. While it is clear that those strategies represent a viable way of combining the two, the question which needed to be asked was whether the resulting recovery patterns could be improved from the passengers' perspective. The case analyzed in section 7.3 suggested that there is indeed room for improvement.

The author believes that the problems surrounding crew management are deeply related to information management. In addition to the continuous stream of real-time information on the train service which controllers compare to (mostly static) timetable information to make service control decisions, the provision of information on crew movements and crew schedules adds another level of complexity which can be very difficult to handle simultaneously. At the same time, controllers do not have direct authority over drivers, which are primarily the responsibility of the DMT (Duty Manager - Trains). Yet, since every change to a train trajectory is also a change to a driver's schedule, the situation in which controllers have insufficient information about and a lack of authority over crews places an unnecessary constraint on their flexibility when managing the service.

This problem is exacerbated by the fact that management of train service is centralized whereas crew management is decentralized and there is no central tracking of drivers, driver lateness and spare driver availability. The information about these variables is spread out over the five crew depots on the Central line, without a continuous feed of information into the control center. It was observed that even at the level of the individual crew depots, DMTs often only track their drivers and driver lateness in the immediate vicinity of their crew depot. This is in part caused by the design of their information system which in standard view only displays trains within 20 minutes of the crew depot. Beyond those "boundaries", drivers are often the only ones who know whether they are late with respect to the crew schedule and whether they are at risk of violating driving time constraints. If controllers need to get a reliable picture of driver lateness and driver availability for unscheduled train movements, they generally have to request the information by telephone from the DMTs or by train radio from the drivers. These communications are time consuming. Since many interventions need to be made under significant time pressure, this is a limiting factor which may cause a controller to choose a solution which is robust in terms of crew management (i.e., there is a small probability that the driver or DMT will veto it, and more crew reliefs are met than might have been necessary) or which minimizes the need for communication altogether.

A possible improvement would be the introduction of a "Control Center DMT" who would be placed in the control center, have authority over the individual crew depot DMTs and be accountable to the service controllers. It would be his or her responsibility to coordinate among the DMTs, to keep an overview over the greater picture of crew utilization and to assist both controllers and DMTs in their decisions, thus creating a single point of contact between these two sides.

8.2.3 OCS design

As previously noted, the design of the OCS can have an influence on how controllers choose to manage the service since some interventions may be easier to code in the system than others. Moreover, it is evident that the type of information provided to controllers by the OCS is a determinant in how and when controllers decide to take action. The following points describe functionalities of an OCS system which are not currently implemented in the Central line OCS (or any other OCS of the London Underground, to the author's knowledge), but which the author feels would be very helpful to controllers and would provide benefits to the agency.

1. Currently, monitoring headways is not a straightforward task. Central line controllers need to calculate individual headways from the departure times indicated by the OCS and the visual representation of the line is not to scale. A simple addition to the OCS, which is already deployed in many other systems worldwide, would be software which continuously tracks headways between trains and automatically alerts controllers to potential problems. This could support policies shifting more attention to the passengers' experience and would be helpful to controllers when running the service with a reduced number of trains.
2. The Central line OCS is a memoryless system; all changes which are made are recorded by the controllers only with a wax pencil in the timetable. This has two major disadvantages. First, there is a constant need for controllers to communicate service changes to their colleagues; some information gets lost in the process, especially given the inevitably high levels of stress in the control center. Secondly, the OCS does not show controllers where trains are missing due to cancellations or other service control interventions, i.e., where there is potentially a gap in the service which is due to interventions and which needs attention. It also does not show which trains already had changes made to them and therefore may not reach a scheduled crew pickup point or stabling depot for the train on time. As an alternative, the author suggests the implementation of an OCS which tracks changes. Most importantly, canceled train trips should be shown in the OCS, and for every train unit it should indicate previous renumberings and special requirements for the train (e.g., stable as booked) or the crew. The latter could be coded into the system by the control center DMT, as suggested above.
3. The implementation of better predictive functions could help controllers better anticipate conflicts and expand their time horizon for decisions, thus reducing the uncertainty and the number of interventions which are made at short notice or which are performed in a way that may cause problems downstream. On one hand, the system should be able to project a train's trajectory when a controller is considering a change, so that the controller gets feedback on whether or not his/her intended outcome will be achieved. On the other hand, a system which anticipates conflicts and alerts controllers ahead of time can help them make decisions at an earlier stage, allowing more time for mitigating the impact of service control interventions

and for planning a coherent strategy. The author believes that this could best be achieved with a graphical representation of operations, such as a space-time plot which is updated in real-time and features a capability for projecting train trajectories and exhibiting them graphically.

4. Last but not least, asking controllers to provide information in the OCS on their reasons for intervention (as is already done in the Central line OCS) is a good idea, but it needs to feature more realistic reasons than are currently available. A simple yet very interesting use for such a list would be to quantify how many canceled trains are due to defects and how many cancellations are made for other reasons, such as crew or train lateness.

8.3 Future research

8.3.1 Immediate extensions to this thesis

Based on the findings of this research in terms of the decision factors of service controllers and the assessment of the impact of interventions, it would be possible to extend predictive operations models and assess their performance, which can help move towards better scheduling and disruption management support tools. Also, with some applied research into methods for institutionalizing and streamlining the analysis process using the framework presented in this thesis, one can ultimately work towards improving service controller training and providing controllers with better guidelines for daily operations management.

8.3.2 Larger areas of interest

Section 2.5.4 showed that there is still large potential for research in the field of service control; basically, all short-comings of previous research efforts which were noted in that section are possible fields for advancement. This section, however, aims to point out two subjects which the author believes would provide a particularly large benefit to the agency conducting the research: passenger behavior and crew utilization.

To date, it is practically unknown how passengers react to delays on the line or to service control interventions such as their train being short-turned. Smart card data provide many options to examine this problem more closely. For instance, it would be possible to track how many passengers leave the system or shift modes (e.g., tap in on a bus) as a function of the type and duration of a delay affecting their line. This could lay the foundations for quantifying the costs of disruptions associated with passenger inconvenience and lost revenue. Furthermore, although it is difficult to track the route choice of riders once they have entered the Underground system, it could be possible with a set of appropriate models and assumptions to infer from travel times or destination stations whether passengers chose to use different lines within the Underground network. In the longer run, one may also be interested in how passengers adjust to daily levels of service unreliability, e.g., how frequent problems on a line section can influence passengers' route or mode choice.

One must also bear in mind that the decisions of passengers are influenced not only by their experiences (e.g., a long gap), but also by service updates which are provided in real-time. For example, on the London Underground, information is distributed over various channels classifying the service on a line as "good service", having "minor delays", "major delays", and being partially or completely suspended. While the author believes that this information has a strong influence on passenger behavior, he also made the observation in the Central line control center that its accuracy and timeliness are sometimes questionable. Research into the relationship between passenger information, passenger behavior and line operations could prove very beneficial in improving such systems and understanding how passengers react to them.

The second area of research is focused on crew utilization. In this research, it was found that the quality of operational data did not allow a precise reconstruction of crew movements in the same manner that train movements were reconstructed. However, this is not to say that these problems cannot be overcome with some research into how to track crews and especially crew lateness – a set of assumptions and heuristic rules may go a long way towards this goal. By approximating the lateness of crews, one could undertake a larger study into the effects of driver lateness on service control interventions, which could aim at quantifying these impacts on an aggregate level rather than the microscopic level used in this research study.

Moreover, better data on the utilization of spare crews could provide insights into the fundamental tradeoff between the provision of spare drivers and spare trains on a line and the changes which service controllers need to make to the operating plan on a daily basis to meet operational targets. A first research effort in this area does not necessarily require a redesign of the OCS or the large-scale deployment of chip cards. It may be possible to conduct a limited-time experiment with drivers in which they are informed of the goals of the experiment, clear policies about logins are established, and drivers are asked to fill in additional information on questionnaires in order to validate observations from the data. Of course, such an experiment would not be possible without the full cooperation of drivers, hence the emphasis on the limited duration, e.g., two to four weeks.

Bibliography

Abkowitz, M. (1986). Optimal control of headway variation on transit routes. *Journal of Advanced Transportation, 20*, 73-88.

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

Allen, D. (1981). First Steps in Schedule Compiling. London: London Underground Ltd. (unpublished).

Barker, D. (2002). Communication, Information and Responsibility Distribution Strategies for Effective Real-Time Transit Service Management. *MSc Thesis*. Massachusetts Institute of Technology.

Bosshard, H. (2008, Aug 10th). 200 neue Doppelstockzüge statt neue Bahnstrecken. *Neue Zürcher Zeitung am Sonntag (newspaper article)*.

Ceder, A. (2007). Public Transit Planning and Operation. Oxford: Elsevier.

Chan, J. (2007). Rail OD Estimation and Journey Time Reliability Metrics Using Automated Fare Data. *MSc Thesis*. Massachusetts Institute of Technology.

Deckoff, A. (1990). The Short-Turn as a Real-Time Transit Operating Strategy. *MSc Thesis*. Massachusetts Institute of Technology.

Dixon, M. (2006). Analysis of a Subway Operations Control Database: The MBTA Operations Control System. *MSc Thesis*. Northeastern University.

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

Forrester, J. D. (2006). *Central Line*. Retrieved June 9, 2009, from Wikimedia Commons: http://commons.wikimedia.org/wiki/File:Central_Line.png

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

Furth, P. G. (2006). Service Reliability and Hidden Waiting Time: Insights from Automatic Vehicle Location Data. *Transportation Research Record, 1955*, 79-87.

Horsey, C. (2009, January). Personal contact at London Underground Ltd.

Jolliffe, J., & Hutchinson, T. (1975). A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop. *Transportation Science, 9*, 248-282.

Koffman, D. (1978). A simulation study of alternative real-time bus headway control strategies. *Transportation Research Record, 663*, 41-46.

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

Lee, D. (2002). Understanding Capacity and Performance of Urban Rail Transit Terminals. *MSc Thesis*. Massachusetts Institute of Technology.

Liu, G., & Wirasinghe, S. (2001). A simulation model of reliable schedule design for a fixed transit route. *Journal of Advanced Transportation, 35*, 145-174.

London Underground Ltd. (2008). Central line training manuals. *A collection of informal training material maintained by the Central line service managers*. Unpublished.

London Underground Ltd. (2006). NACHS 2006 - BCV Trains: NACHS Tables explained. Unpublished.

London Underground Ltd. (2002). Operations Plan: Central Line. Unpublished.

Macchi, R. (1989). Expressing Trains on the MBTA Green Line. *MSc Thesis*. Massachusetts Institute of Technology.

Moore, A. (2002). A Process for Improving Transit Service Management. *MSc Thesis*. Massachusetts Institute of Technology.

O'Dell, S. (1997). Optimal Control Strategies for a Rail Transit Line. *MSc Thesis*. Massachusetts Institute of Technology.

Ortiz, I. (2000). Analysis of Real Time Operations Control Strategies for Tren Urbano. *MSc Thesis*. Massachusetts Institute of Technology.

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

Puong, A. (2001). A Real-Time Train Holding Model for Rail Transit Systems. *MSc Thesis*. Massachusetts Institute of Technology.

Rahbee, A. (2001). Rail Transit Operations Analysis: Framework and Applications. *MSc Thesis*. Massachusetts Institute of Technology.

Shen, S. (2000). Integrated Real - Time Disruption Recovery Strategies: A Model for Rail Transit Systems. *MSc Thesis*. Massachusetts Institute of Technology.

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

Song, W. (1998). Real Time Dispatching Control in Transit Systems. *MSc Thesis*. Massachusetts Institute of Technology.

TfL. (2009). *Transport for London*. Retrieved June 10, 2009, from Modes of transport: http://www.tfl.gov.uk/corporate/modesoftransport/2794.aspx

TfL. (2009). Your guide to fares and tickets. London: Transport for London.

Tufte, E. (1983). The visual display of quantitative information. Cheshire: Graphics Press.

Uniman, D. (2009). Service Reliability Measurement using AFC Smart Card Data – A Framework for the London Underground. *MSc Thesis*. Massachusetts Institute of Technology.

Vescovacci, M. (2003). Junction Capacity and Performance in Rail Transit. *MSc Thesis*. Massachusetts Institute of Technology.

Wardman, M. (2001). A review of British evidence on time and service quality valuations. *Transportation Research E 37*, 107-128.

Wile, E. (2003). Use of Automatically Collected Data to Improve Transit Line Performance. *MSc Thesis*. Massachusetts Institute of Technology.

Wilson, N. (2008). Public Transportation Systems. *Course material.* Massachusetts Institute of Technology.

Appendix

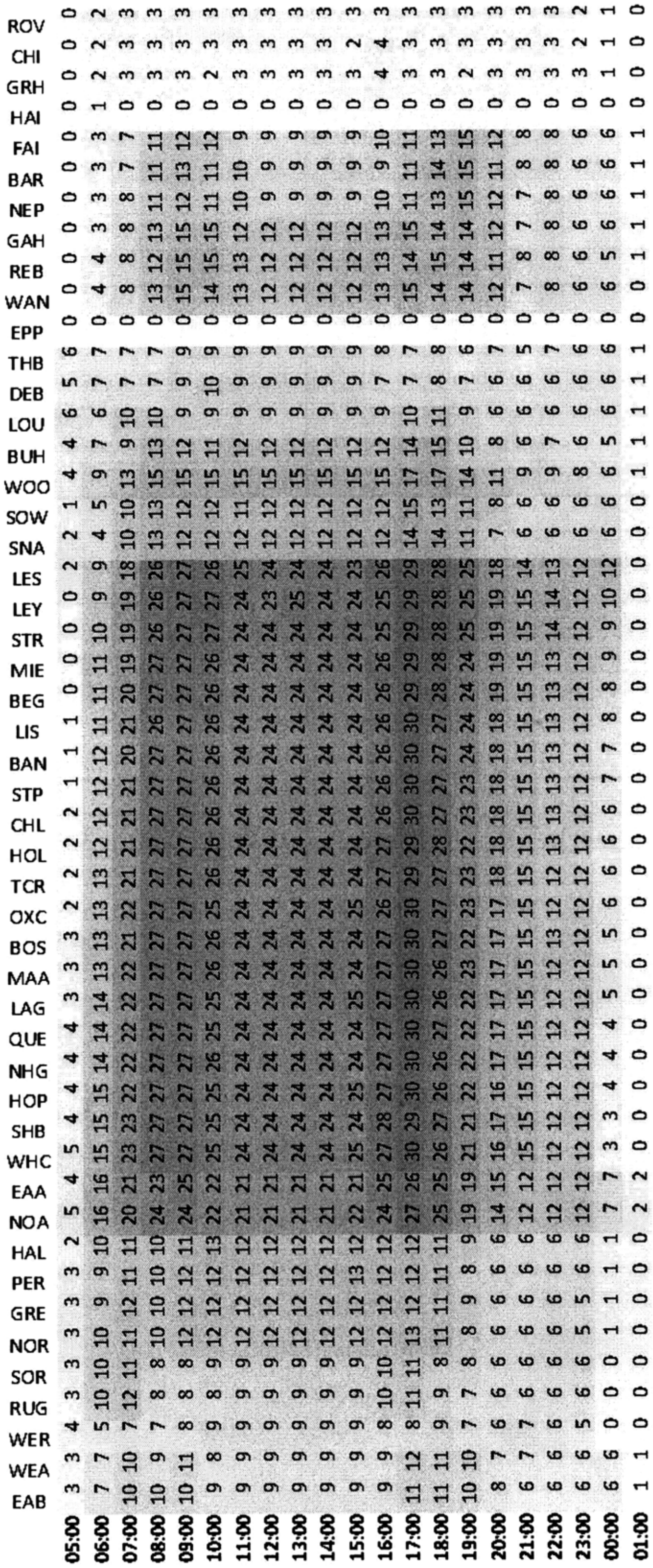
A. Three-letter Central line station codes

West Ruislip	WER
Ruislip Gardens	RUG
South Ruislip	SOR
Northolt	NOR
Greenford	GRE
Perivale	PER
Hanger Lane	HAL
Ealing Broadway	EAB
West Acton	WEA
North Acton	NOA
East Acton	EAA
White City	WHC
Shepherd's Bush	SHB
Holland Park	HOP
Notting Hill Gate	NHG
Queensway	QUE
Lancaster Gate	LAG
Marble Arch	MAA
Bond Street	BOS
Oxford Circus	OXC
Tottenham Court Road	TCR
Holborn	HOL
Chancery Lane	CHL
St. Paul's	STP
Bank	BAN

Liverpool Street	LIS
Bethnal Green	BEG
Mile End	MIE
Stratford	STR
Leyton	LEY
Leytonstone	LES
Snaresbrook	SNA
South Woodford	SOW
Woodford	WOO
Buckhurst Hill	BUH
Loughton	LOU
Debden	DEB
Theydon Bois	THB
Epping	EPP
Wanstead	WAN
Redbridge	REB
Gants Hill	GAH
Newbury Park	NEP
Barkingside	BAR
Fairlop	FAI
Hainault	HAI
Grange Hill	GRH
Chigwell	CHI
Roding Valley	ROV

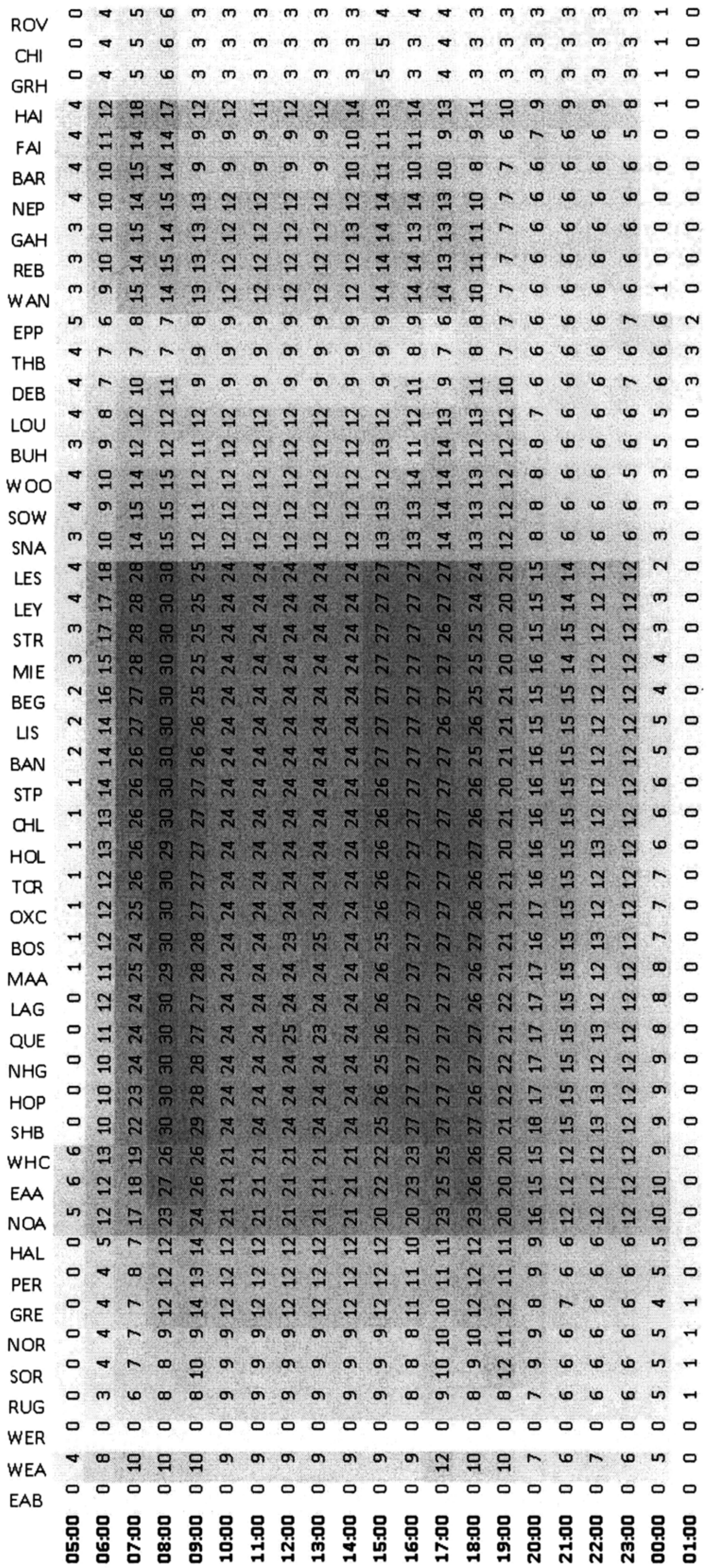
B. Scheduled trains per hour in working timetable 64

Scheduled trains per hour eastbound, shown as a heat diagram for better legibility.



	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00	01:00
ROV	0	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	1	0
CHI	0	2	3	3	3	3	3	3	3	3	2	4	3	3	3	3	3	3	2	1	0
GRH	0	2	3	3	3	2	3	3	3	3	3	4	3	3	2	3	3	3	3	1	0
HAI	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FAI	0	3	7	11	12	12	9	9	9	9	9	10	11	13	15	12	8	8	6	6	1
BAR	0	3	7	11	13	11	10	9	9	9	9	9	11	14	15	11	8	8	6	6	1
NEP	0	3	8	11	12	11	10	9	9	9	9	10	11	13	15	12	7	8	6	6	1
GAH	0	3	8	13	15	15	12	12	12	12	12	13	15	14	14	12	7	8	6	6	1
REB	0	4	8	12	15	15	13	12	12	12	12	13	14	15	14	11	8	8	6	5	1
WAN	0	4	8	13	15	14	13	12	12	12	12	13	15	14	14	12	7	8	6	6	0
EPP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
THB	6	7	7	7	9	9	9	9	9	9	9	8	7	8	6	7	5	7	6	6	1
DEB	5	7	7	7	9	10	9	9	9	9	9	7	7	8	7	6	6	6	6	6	1
LOU	6	6	10	10	9	9	9	9	9	9	9	9	10	11	9	6	6	6	6	6	1
BUH	4	7	9	13	12	11	12	12	12	12	12	12	14	15	10	8	6	7	6	5	1
WOO	4	9	13	15	15	15	15	15	15	15	15	15	17	17	14	11	9	9	8	6	1
SOW	1	5	10	13	12	12	11	12	12	12	12	12	15	13	11	8	6	6	6	6	0
SNA	2	4	10	13	12	12	12	12	12	12	12	12	14	14	11	7	6	6	6	6	0
LES	2	9	18	26	27	26	25	24	24	24	23	26	29	28	25	18	14	13	12	12	0
LEY	0	9	19	26	27	27	24	23	25	24	24	25	29	28	25	19	15	14	12	10	0
STR	0	10	19	26	27	27	24	24	24	24	24	25	29	28	25	19	15	14	12	9	0
MIE	0	11	19	27	27	26	24	24	24	24	24	26	29	28	24	19	15	13	12	9	0
BEG	0	11	20	27	27	26	24	24	24	24	24	26	29	28	24	19	15	13	12	8	0
LIS	1	11	21	26	27	26	24	24	24	24	24	26	30	27	24	18	15	13	12	8	0
BAN	1	12	20	27	27	26	24	24	24	24	24	26	30	27	24	18	15	13	12	7	0
STP	1	12	21	27	27	26	24	24	24	24	24	26	30	27	23	18	15	13	12	7	0
CHL	2	12	21	27	27	26	24	24	24	24	24	26	30	27	23	18	15	13	12	6	0
HOL	2	12	21	27	27	26	24	24	24	24	24	27	29	28	22	18	15	13	12	6	0
TCR	2	13	21	27	27	26	24	24	24	24	24	27	29	27	23	18	15	12	12	6	0
OXC	2	13	22	27	27	25	24	24	24	24	25	26	30	27	23	17	15	12	12	6	0
BOS	3	13	21	27	27	26	24	24	24	24	24	27	30	27	22	17	15	13	12	5	0
MAA	3	13	22	27	27	26	24	24	24	24	24	27	30	26	23	17	15	12	12	5	0
LAG	3	14	22	27	27	25	24	24	24	24	25	27	30	26	22	17	15	12	12	5	0
QUE	4	14	22	27	27	25	24	24	24	24	24	27	30	27	22	17	15	12	12	4	0
NHG	4	14	22	27	27	26	24	24	24	24	24	27	30	26	22	17	15	12	12	4	0
HOP	4	15	22	27	27	25	24	24	24	24	25	27	30	26	22	16	15	12	12	4	0
SHB	4	15	23	27	27	25	24	24	24	24	24	28	29	27	21	17	15	12	12	3	0
WHC	5	15	23	27	27	25	24	24	24	24	25	27	30	26	21	16	15	12	12	3	0
EAA	4	16	21	23	25	22	21	21	21	21	21	25	26	25	19	15	12	12	12	7	2
NOA	5	16	20	24	24	22	21	21	21	21	22	24	27	25	19	14	12	12	12	7	2
HAL	2	10	11	10	11	13	12	12	12	12	12	12	12	11	9	6	6	6	6	1	0
PER	3	9	11	10	12	12	12	12	12	12	13	12	12	11	8	6	6	6	6	1	0
GRE	3	9	12	10	12	12	12	12	12	12	12	12	12	11	9	6	6	6	5	1	0
NOR	3	10	11	10	12	12	12	12	12	12	12	12	13	11	8	6	6	6	5	1	0
SOR	3	10	11	8	8	9	9	9	9	9	9	10	11	8	8	6	6	6	6	0	0
RUG	3	10	12	8	8	8	9	9	9	9	9	10	11	9	7	6	6	6	6	0	0
WER	4	5	7	7	8	9	9	9	9	9	9	8	8	9	7	6	7	6	5	0	0
WEA	3	7	10	9	11	8	9	9	9	9	9	9	12	11	10	7	7	6	6	6	1
EAB	3	7	10	10	10	9	9	9	9	9	9	9	11	11	10	8	6	6	6	6	1

Scheduled trains per hour westbound, shown as a heat diagram for better legibility.



	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00	01:00
ROV	0	4	5	6	3	3	3	3	3	3	4	4	4	3	3	3	3	3	3	1	0
CHI	0	4	5	6	3	3	3	3	3	3	5	3	4	3	3	3	3	3	3	1	0
GRH	0	4	5	6	3	3	3	3	3	3	5	3	4	3	3	3	3	3	3	1	0
HAI	4	12	18	17	12	12	11	12	12	14	13	14	13	11	10	9	9	9	8	1	0
FAI	4	11	14	14	9	9	9	9	9	10	11	11	9	9	6	7	6	6	5	0	0
BAR	4	10	15	14	9	9	9	9	9	10	11	10	10	8	7	6	6	6	6	0	0
NEP	4	10	14	15	13	12	12	12	12	12	14	14	13	10	7	6	6	6	6	0	0
GAH	3	10	15	14	13	12	12	12	12	13	14	13	13	11	7	6	6	6	6	0	0
REB	3	10	14	15	13	12	12	12	12	12	14	14	13	11	7	6	6	6	6	0	0
WAN	3	9	15	14	13	12	12	12	12	12	14	14	14	10	7	6	6	6	6	1	0
EPP	5	6	8	7	8	9	9	9	9	9	9	9	6	8	7	6	6	6	7	6	2
THB	4	7	7	7	9	9	9	9	9	9	9	8	7	8	7	6	6	6	6	6	3
DEB	4	7	10	11	9	9	9	9	9	9	9	11	9	11	10	6	6	6	7	6	3
LOU	4	8	12	12	12	12	12	12	12	12	12	12	13	13	12	7	6	6	6	5	0
BUH	3	9	12	12	11	12	12	12	12	12	13	11	14	12	12	8	6	6	6	5	0
WOO	4	10	14	15	12	12	12	12	12	12	12	14	14	13	12	8	6	6	5	3	0
SOW	4	9	15	15	11	12	12	12	12	12	13	13	14	13	12	8	6	6	6	3	0
SNA	3	10	14	15	12	12	12	12	12	12	13	13	14	13	12	8	6	6	6	3	0
LES	4	18	28	30	25	24	24	24	24	24	27	27	27	24	20	15	14	12	12	2	0
LEY	4	17	28	30	25	24	24	24	24	24	27	27	27	24	20	15	14	12	12	3	0
STR	3	17	28	30	25	24	24	24	24	24	27	27	26	25	20	15	15	12	12	3	0
MIE	3	15	28	30	25	24	24	24	24	24	27	27	27	25	20	16	14	12	12	4	0
BEG	2	16	27	30	25	24	24	24	24	24	27	27	27	25	21	15	15	12	12	4	0
LIS	2	14	27	30	26	24	24	24	24	24	27	27	26	26	21	15	15	12	12	5	0
BAN	2	14	26	30	26	24	24	24	24	24	27	27	27	25	21	16	15	12	12	5	0
STP	1	14	26	30	27	24	24	24	24	24	26	27	27	26	20	16	15	12	12	6	0
CHL	1	13	26	30	27	24	24	24	24	24	26	27	27	26	21	16	15	12	12	6	0
HOL	1	13	26	29	27	24	24	24	24	24	26	27	27	27	20	16	15	13	12	6	0
TCR	1	12	26	30	27	24	24	24	24	24	26	27	27	26	21	16	15	12	12	7	0
OXC	1	12	25	30	27	24	24	24	24	24	26	27	27	26	21	17	15	12	12	7	0
BOS	1	12	24	30	28	24	24	23	25	24	25	27	27	27	21	16	15	13	12	7	0
MAA	1	11	25	29	28	24	24	24	24	24	26	27	27	26	21	17	15	12	12	8	0
LAG	0	12	24	30	27	24	24	24	24	24	26	27	27	26	22	17	15	12	12	8	0
QUE	0	11	24	30	27	24	24	25	23	24	26	27	27	27	21	17	15	13	12	8	0
NHG	0	10	24	30	28	24	24	24	24	24	25	27	27	27	22	17	15	12	12	9	0
HOP	0	10	23	30	28	24	24	24	24	24	26	27	27	26	22	17	15	13	12	9	0
SHB	0	10	22	30	29	24	24	24	24	24	25	27	27	27	21	18	15	13	12	9	0
WHC	6	13	19	26	26	21	21	21	21	21	22	23	25	26	20	15	12	12	12	9	0
EAA	6	12	18	27	26	21	21	21	21	21	22	23	25	26	20	15	12	12	12	10	0
NOA	5	12	17	23	24	21	21	21	21	21	20	20	23	23	20	16	12	12	12	10	0
HAL	0	5	7	12	14	12	12	12	12	12	12	10	11	12	11	9	6	6	6	5	0
PER	0	4	8	12	13	12	12	12	12	12	12	11	11	12	11	9	6	6	6	5	0
GRE	0	4	7	12	14	12	12	12	12	12	12	11	10	12	12	8	7	6	6	4	1
NOR	0	4	7	9	9	9	9	9	9	9	9	8	10	10	11	9	6	6	6	5	1
SOR	0	4	7	8	10	9	9	9	9	9	9	8	10	9	12	9	6	6	6	5	1
RUG	0	3	6	8	8	9	9	9	9	9	9	8	9	8	8	7	6	6	6	5	1
WER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WEA	4	8	10	10	10	9	9	9	9	9	9	9	12	10	10	7	6	7	6	5	0
EAB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

C. Comparison of scheduled trains per hour

Scheduled trains per hour in working timetables 63 and 64. Green cells denote an increase, orange cells a decrease in frequency.

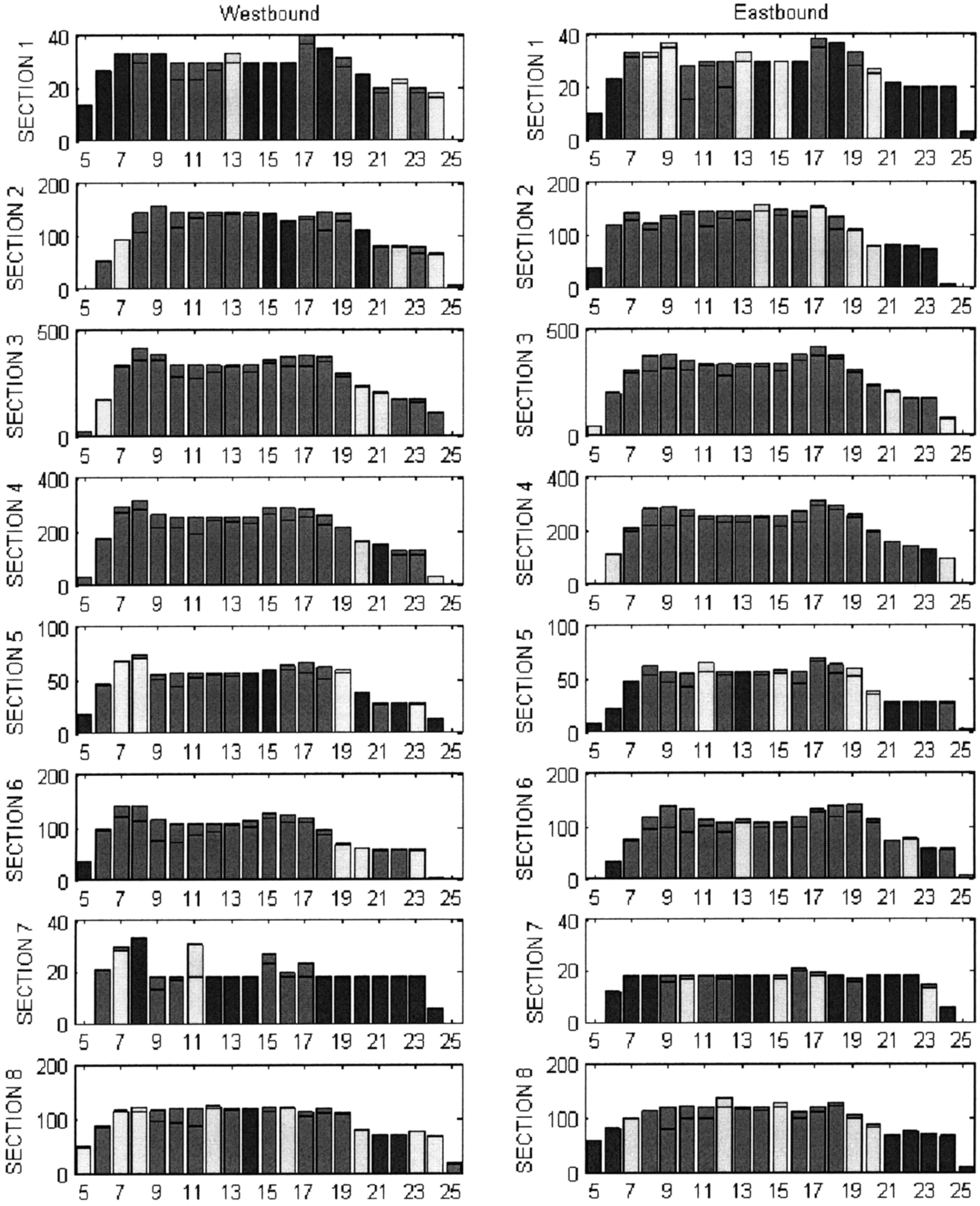
	Eastbound - Working timetable 63								Eastbound - Working timetable 64							
Time	EAB	WER	NOR	WHC	LIS	LES	LOU	NEP	EAB	WER	NOR	WHC	LIS	LES	LOU	NEP
5	3	4	3	5	1	2	6	0	3	4	3	5	1	2	6	0
6	7	5	10	15	11	8	6	3	7	5	10	15	11	9	6	3
7	10	7	11	23	20	19	10	8	10	7	11	23	21	18	10	8
8	10	7	10	27	27	25	10	11	10	7	10	27	26	26	10	11
9	10	8	10	25	26	27	6	14	10	8	12	27	27	27	9	12
10	9	9	9	23	24	25	6	14	9	9	12	25	26	26	9	11
11	9	9	9	21	21	21	6	9	9	9	12	24	24	25	9	10
12	9	9	9	21	21	21	6	9	9	9	12	24	24	24	9	9
13	9	9	9	21	21	21	6	9	9	9	12	24	24	24	9	9
14	9	9	9	21	21	21	6	9	9	9	12	24	24	24	9	9
15	9	9	12	22	21	21	6	9	9	9	12	25	24	23	9	9
16	9	8	12	27	25	24	8	10	9	8	12	27	26	26	9	10
17	11	8	13	30	30	29	11	11	11	8	13	30	30	29	10	11
18	11	9	10	26	27	28	10	13	11	9	11	26	27	28	11	13
19	10	7	8	21	23	24	9	13	10	7	8	21	24	25	9	15
20	8	6	6	15	18	19	6	12	8	6	6	16	18	18	6	12
21	5	6	6	15	15	14	6	7	6	7	6	15	15	14	6	7
22	6	6	6	12	13	14	6	8	6	6	6	12	13	13	6	8
23	7	5	5	12	12	12	6	6	6	5	5	12	12	12	6	6
0	6	0	1	2	7	11	5	5	6	0	1	3	8	12	6	6
	Westbound - Working timetable 63								Westbound - Working timetable 64							
Time	WEA	NOR	WHC	LIS	LES	LOU	EPP	NEP	WEA	NOR	WHC	LIS	LES	LOU	EPP	NEP
5	4	0	6	2	4	4	5	4	4	0	6	2	4	4	5	4
6	8	4	13	15	18	8	6	10	8	4	13	14	18	8	6	10
7	10	7	19	26	28	12	8	14	10	7	19	27	28	12	8	14
8	10	9	26	30	30	12	7	15	10	9	26	30	30	12	7	15
9	10	13	26	25	23	10	6	11	10	9	26	26	25	12	8	13
10	9	9	18	21	21	9	6	13	9	9	21	24	24	12	9	12
11	9	9	18	21	21	9	6	12	9	9	21	24	24	12	9	12
12	9	9	18	21	21	9	6	12	9	9	21	24	24	12	9	12
13	9	9	18	21	21	9	6	12	9	9	21	24	24	12	9	12
14	9	9	18	21	21	9	6	12	9	9	21	24	24	12	9	12
15	9	9	19	24	25	9	6	15	9	9	22	27	27	12	9	14
16	9	8	22	27	27	10	6	15	9	8	23	27	27	12	9	14
17	12	10	25	27	27	13	7	13	12	10	25	26	27	13	6	13
18	10	10	26	25	24	13	8	10	10	10	26	26	24	13	8	10
19	10	11	20	20	19	10	7	8	10	11	20	21	20	12	7	7
20	7	8	14	15	15	8	6	7	7	9	15	15	15	7	6	6
21	6	6	12	15	14	6	6	6	6	6	12	15	14	6	6	6
22	6	6	12	12	12	6	6	6	7	6	12	12	12	6	6	6
23	6	6	12	12	11	6	6	5	6	6	12	12	12	6	7	6
0	6	5	9	4	2	5	6	0	5	5	9	5	2	5	6	0

D. Files produced by the Central line OCS

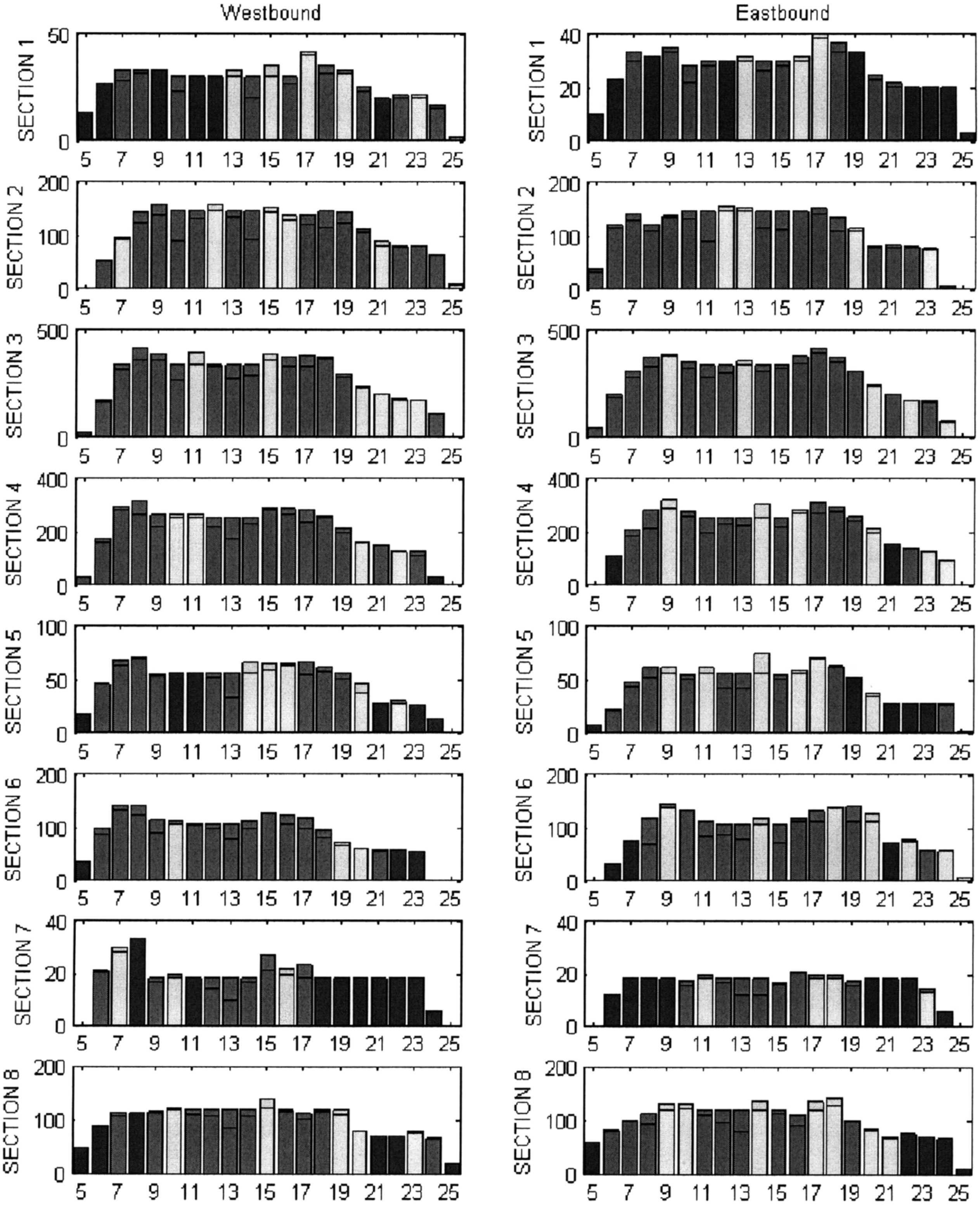
File abbreviation	Full name	Remarks
BTL	Block temporary speed restrictions	
CEL	Communications errors	
CTL		(Unknown)
EDL	Early departures	
MWL		(Unknown)
OCL	Operator commands	
PML	Automated passenger messages	
TAL	Timetable allocations	
TCO	Track circuit operation	Unused
TDA	Traffic day	See text
TDL	Train data	See text
TEH	Timetable edit headers	
TEL	Timetable edit elements	
TPL	Train passing	Unused
TTE		(Unknown)
TTH		(Unknown)
TTU		(Unknown)
UML	Unit mileage	
WL	Warnings	

E. Service delivery plots by section

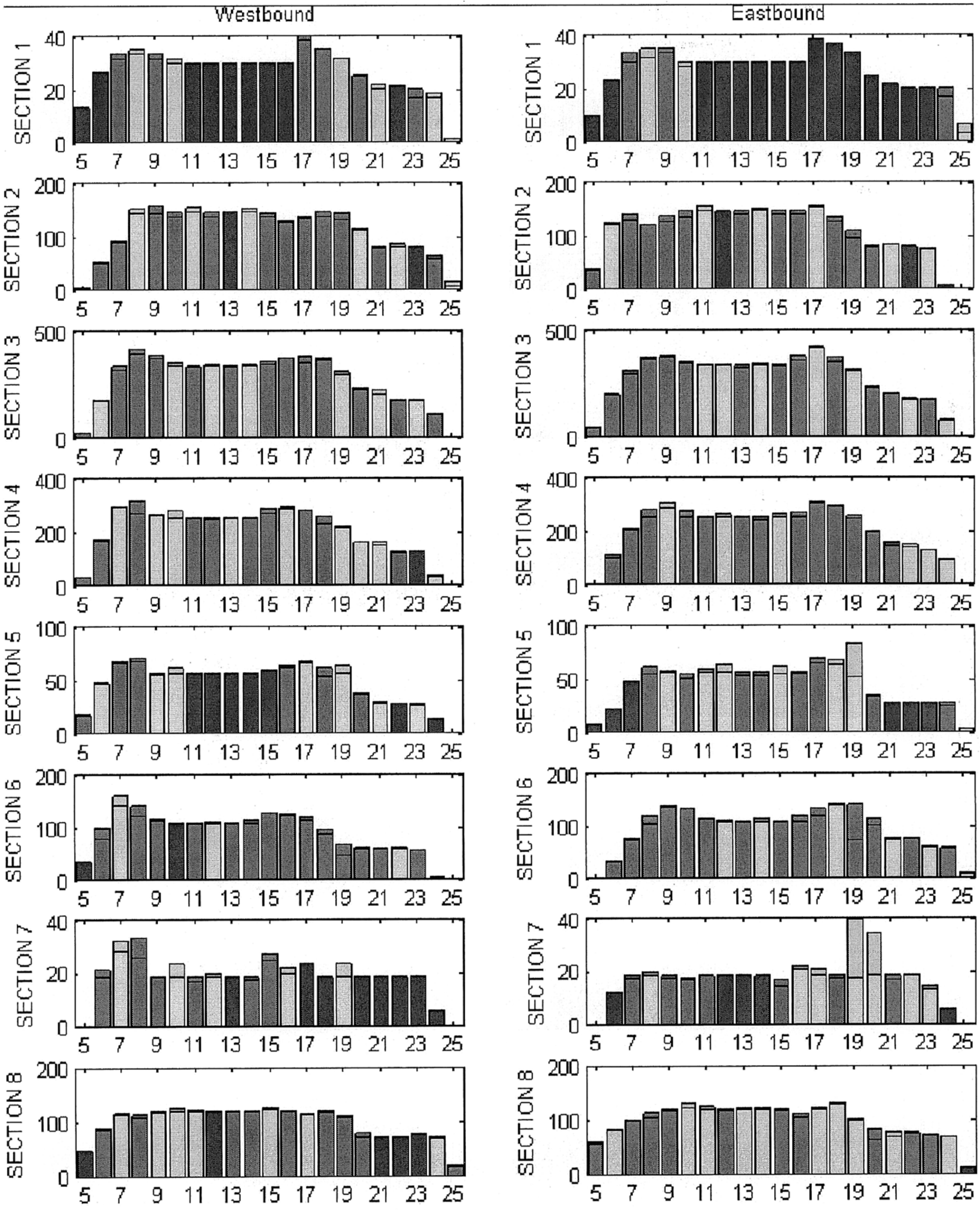
Service delivery plot for April 1, 2008



Service delivery plot for April 3, 2008



Service delivery plot for November 21, 2008



F. Interventions per day – October 2006 and October 2007