Characterizing Transit System Performance Using Smart Card Data

by

Lauren Tarte

Submitted to the Department of Civil and Environmental Engineering on May 21, 2015 in partial fulfillment of the requirements for the degree of Master of Science in Transportation

Abstract

Although automated data collection systems have been in existence long enough to be the subject of extensive research, they continue to transform both the agency and customer side of public transportation. Transit system performance is defined along three dimensions: the supply of trains or buses; the demand from passengers; and the product of the two, service performance. This research takes a broad view of these dimensions, and explores several means of measuring each one, ways in which they interact, and how this information can be valuable for a transit agency. This thesis focuses specifically on the London Underground, but the intent is that the types of analysis described herein can be applied to any rail transit system with similar data resources.

The research has three parts. First, a methodology is developed to estimate passenger volumes on the portion of a line between two adjacent stations within a defined time interval. This approach can be implemented using data from either smart card or entry and exit gates. It relies upon the outputs of London’s Rolling Origin Destination Survey to infer passenger route choice. The second component identifies the possible causes of poor service performance in terms of both supply and demand, and defines a framework for examining supply and demand data to identify which causes are contributors. This research suggests that a better understanding of capacity constraints can be gained by jointly analyzing AVL-based capacity measures, AFC-based demand measures, and AFC journey times as a measure of service performance.

Finally, this thesis explores the possibility of using smart card data in real time to estimate system state. This metric is defined as passenger accumulation, a measure of the number of passengers on a given portion of a line in real time. Building from the approach developed in the first section, this work designs a method to determine accumulation in real time at detailed levels of spatial granularity. The accumulation metric is then compared to Transport for London’s existing tools to assess whether accumulation can provide value as a real-time indicator of system state. Based on this analysis and current data availability constraints it cannot be concluded that passenger accumulation provides a reliable indicator of real-time system state.

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

Thesis Supervisor: John P. Attanucci
Title: Research Associate of Civil and Environmental Engineering
ACKNOWLEDGEMENTS

I am indebted to Nigel Wilson and John Attanucci for their mentorship through this process. Your guidance and insight has made this thesis infinitely better, and I have grown immensely for it.

I am also grateful to the Analytics, Insight, and Transport Planning groups at Transport for London for all their advice on the direction of this research and their answers to my many, many questions.

Thank you to the Transit Lab graduates of 2014, who set an example in perseverance and always assured me that my doubts, fears, and struggles were a part of the process.

To Transit Lab and Friends – your friendship, support, and nerdy discussions about transportation in our free time have been an invaluable part of my MIT experience. I cannot express my gratitude for all the times you’ve been there for me, especially in these last few weeks. Thank you for the hugs, hard drives, and cookies, and for letting me draft you as unpaid consultants on this thesis.

Thank you to Emily and Becca for all the things.

I cannot thank my parents enough for supporting me through the first round of my education and again through this second one. Thank you for listening to me wax didactic to an extent that only parents can, and for always telling me that I could do it and never telling me how to do it. Thank you to Mollie for your encouragement and for the endless cat pictures. And thank you to the rest of my indescribably amazing family.

To everyone at home in the Emerald City, I missed you every day.
# TABLE OF CONTENTS

1 INTRODUCTION .......................................................................................................................... 13
  1.1 Motivation ................................................................................................................................. 13
  1.2 Objectives ................................................................................................................................. 13
  1.3 Research Approach ................................................................................................................ 15
  1.4 Thesis Organization ............................................................................................................... 15

2 LITERATURE REVIEW ............................................................................................................... 17
  2.1 Measuring supply and demand variability ............................................................................. 17
  2.2 Understanding service performance ..................................................................................... 17
  2.3 Applications of smart card data ............................................................................................ 20

3 DIMENSIONS OF TRANSIT SYSTEM PERFORMANCE ............................................................ 21
  3.1 Measures of transit system performance .............................................................................. 21
  3.2 Current TFL use of baselines ................................................................................................. 23
  3.3 Available data sources .......................................................................................................... 26
    3.3.1 Oyster .............................................................................................................................. 26
    3.3.2 NetMiS ............................................................................................................................ 28
    3.3.3 Gate counts .................................................................................................................... 28
    3.3.4 RODS ............................................................................................................................. 28
    3.3.5 CuPID .............................................................................................................................. 29
  3.4 Data sets used in this thesis .................................................................................................... 30
  3.5 Determining passenger flow .................................................................................................. 31
    3.5.1 Flow as a measure of demand ...................................................................................... 31
    3.5.2 Detailed flow methodology ........................................................................................... 32
    3.5.3 Example flows on the Victoria Line ............................................................................. 36

4 ANALYZING THE RELATIONSHIPS BETWEEN PERFORMANCE DIMENSIONS ................. 41
  4.1 Drivers of service performance ............................................................................................. 41
  4.2 Identifying contributing causes of delay ............................................................................. 42
    4.2.1 Indicators of high demand ......................................................................................... 43
    4.2.2 Indicators of low supply ............................................................................................. 43
  4.3 Framework for analyzing system dimensions ....................................................................... 44
  4.4 Example Analyses ............................................................................................................... 48
  4.5 Summary ............................................................................................................................... 58

5 A REAL-TIME MEASURE OF SYSTEM STATE ...................................................................... 61
  5.1 Previous research on real-time applications of AFC data ..................................................... 61
  5.2 Expanding the passenger accumulation metric ................................................................... 64
    5.2.1 Developing an improved methodology ....................................................................... 64
    5.2.2 Applications of the proposed methodology ................................................................. 67
  5.3 Current TFL tools for indicating system state ....................................................................... 73
  5.4 Evaluation of passenger accumulation as a real-time indicator ......................................... 76

6 CONCLUSION .......................................................................................................................... 83
  6.1 Research summary ............................................................................................................... 83
6.2 Recommendations ...........................................................................................................................................85
6.3 Future research ................................................................................................................................................86

APPENDIX A: UNDERGROUND STATION CODES ..................................................................................................89
APPENDIX B: TERMS USED IN THIS THESIS .........................................................................................................92
APPENDIX C: TFL PERFORMANCE METRICS ....................................................................................................93
APPENDIX D: PICCADILLY SEGMENTS (FREEMARK) ............................................................................................94
BIBLIOGRAPHY ......................................................................................................................................................95
LIST OF FIGURES

Figure 3-1. Weekly demand compared to the previous year ..................................................23
Figure 3-2. Journeys per week compared to prior years ......................................................23
Figure 3-3. Daily journeys and revenue compared to prior years .......................................24
Figure 3-4. Jubilee Line performance scorecard ...................................................................25
Figure 3-5. Components of journey time ............................................................................27
Figure 3-6. Differences between recorded and actual journey times ..................................30
Figure 3-7. Link flows on the Victoria Line (southbound 07:00 - 10:00) ...............................36
Figure 3-8. Link flows on the Victoria Line (southbound 08:30 - 08:45) ...............................36
Figure 3-9. Link flows on the Victoria Line (northbound 07:00 - 10:00) ..............................37
Figure 3-10. Link flows on the Victoria Line (northbound 08:30 - 08:45) ............................37
Figure 3-11. Link flows and capacity on the Victoria Line (southbound 08:30 - 08:45) ........38
Figure 3-12. Link flows and capacity on the Victoria Line (northbound 08:30 - 08:45) ........39
Figure 4-1. Piccadilly Line from east terminus to Central London ......................................45
Figure 4-2. Frequency at selected stations (Piccadilly Line, westbound, 9 Dec 2013) ............49
Figure 4-3. Standard deviation of headways (Piccadilly Line, westbound, 9 Dec 2013) ........50
Figure 4-4. Distribution of headways (Piccadilly Line, westbound, 9 Dec 2013) .................51
Figure 4-5. Distribution of journey times (Piccadilly Line, westbound) ...............................52
Figure 4-6. Typical and individual journey times (Piccadilly, BGR-HOL, 9 Dec 2013) ..........52
Figure 4-7. Typical and individual journey times (Piccadilly, CRD-HOL, 9 Dec 2013) ........53
Figure 4-8. Median link flows (Piccadilly Line, westbound) .............................................54
Figure 4-9. Flow by links ....................................................................................................55
Figure 4-10. Flow by time of day .......................................................................................55
Figure 4-11. Frequency of trains (Piccadilly Line, westbound, 27 Nov 2013) .......................55
Figure 4-12. Standard deviation of headways (Piccadilly, westbound, 27 Nov 2013) ..........56
Figure 4-13. Flow by links ................................................................................................56
Figure 4-14. Flow by time of day .......................................................................................56
Figure 4-15. Typical and individual journey times (Piccadilly Line, BGR-HOL, 27 Nov 2013) .57
Figure 4-16. Typical and individual journey times (Piccadilly Line, CAL-HOL, 27 Nov 2013) .57
Figure 4-17. Comparing flow to capacity (Piccadilly Line, westbound, KXX-RSQ, 27 Nov 2013) ....58
Figure 5-1. Absolute passenger accumulation on the Piccadilly Line between Acton Town and King's Cross .................................................................62
Figure 5-2. Excess passenger accumulation for the previous example ..................................63
Figure 5-3. Passenger accumulation by interval size (Victoria Line, southbound, 1 Nov 2013) .66
Figure 5-4. Passenger accumulation by segment (Victoria Line, southbound, 1 Nov 2013) .....67
Figure 5-5. Passenger accumulation by segment (Victoria Line, northbound, 1 Nov 2013) ....68
Figure 5-6. Passenger accumulation by link (Victoria Line, southbound, 1 Nov 2013) ..........69
Figure 5-7. Passenger accumulation across days (Victoria Line, southbound, Central segment) .70
Figure 5-8. Distribution of passenger accumulation (Victoria Line, northbound) .................71
Figure 5-9. Distribution of passenger accumulation (Victoria Line, northbound, Central segment) ....71
Figure 5-10. Real-time passenger accumulation (Victoria Line, northbound, 31 Oct 2013) ....72
Figure 5-11. Heartbeat frequency graph (Victoria Line, northbound at Oxford Circus, 31 Oct 2013) ....74
Figure 5-12. Heartbeat lateness graph (Victoria Line, northbound at Oxford Circus, 31 Oct 2013) ...74
Figure 5-13. Heartbeat waterfall diagram (Victoria Line, northbound, 31 Oct 2013) ......................75
Figure 5-14. TrackerNet (Piccadilly Line)...............................................................................................76
Figure 5-15. Real-time passenger accumulation (Victoria Line, southbound, 5 Nov 2013) ............77
Figure 5-16. Disruptions on the Victoria Line (5 Nov 2013).................................................................................................................................77
Figure 5-17. Real-time passenger accumulation (Victoria Line, southbound, Brixton segment, 5 Nov 2013) ........................................................................................................................................78
Figure 5-18. Headways at Oxford Circus (Victoria Line, southbound, 5 Nov 2013)..........................79
Figure 5-19. Frequency at Oxford Circus (Victoria Line, southbound, 5 Nov 2013) ....................79
Figure 5-20. Lateness at Oxford Circus (Victoria Line, southbound, 5 Nov 2013).........................80
LIST OF TABLES

Table 3-1. Sample station-to-station travel times (minutes).................................................................................33
Table 3-2. Sample station-specific time windows....................................................................................................33
Table 3-3. Sample RODS station movement records .................................................................................................34
Table 3-4. Comparison of calculated and RODS flow values (Victoria southbound 07:00 - 10:00)......................37
Table 3-5. Scheduled and actual service on the Victoria Line (14 Nov 2013).........................................................40
Table 4-1. Mapping indicators to causes of delay ......................................................................................................44
Table 4-2. Origin-Destination collections including peak load point .................................................................46
Table 4-3. Origin-Destination collections excluding peak load point.................................................................47
Table 4-4. Disruptions logged in CuPID (9 Dec 2013) .........................................................................................49
Table 5-1. Running time by segment .......................................................................................................................68
INTRODUCTION

Though no longer new, automated data collection systems continue to transform both the business and user side of public transportation. Previous research has described how this data may be leveraged to study the control environment (Ravichandran, 2012; Carrel, 2009) and to measure operational characteristics such as reliability (Frumin, 2010; Uniman, 2009) and disruption impacts (Freemark, 2013). This research takes a broader view, and rather than focusing on a particular function of a transit system, explores several key characteristics of a system, how they influence one another, and how this information can be valuable for a transit agency. This thesis focuses specifically on the London Underground, but the intent is that the types of analysis described here can be applied to any rail transit system with similar data resources.

1.1 Motivation

This research developed out of the abundance of data generated by automated data collection systems, and a desire to apply this data in a way that both improves the quality of passengers’ experiences and facilitates transit agencies’ planning and operational tasks. Taking a data-driven approach can identify opportunities where service can be improved without the cost of infrastructure upgrades. Transport for London already excels at incorporating analytics into their business processes, yet there remain areas that are unexplored. In particular, the enormous volumes of data produced from smart card systems such as the Oyster card provide unparalleled information about how passengers move through the system. Examining farecard data in conjunction with data about train movements from automatic vehicle location systems provides opportunities to draw inferences about the way these two aspects of the system—passengers and trains—act and interact.

This plethora of data provides a window into the dynamics of a transit system and its environment. A system’s performance is described along three dimensions: the supply of trains, demand from users, and finally the relationship between the two, service performance.\(^1\) This thesis aims to enhance the understanding of each of these dimensions and the relationships between them. Such information has the potential to inform service planning, performance management, operational strategies, and travel demand management tactics on scales ranging from long-term to real-time.

1.2 Objectives

The overarching goal of this thesis is to provide insight into the three dimensions of rapid transit system performance. By enhancing the understanding of how these dimensions vary and interact,

---

\(^1\) The distinction between the terms “system performance” and “service performance” is clarified in Appendix B.
this research aims to improve an agency’s ability to proactively and reactively respond to conditions. This thesis has three specific components:

1. **Develop a method for estimating typical demand at disaggregate spatial and temporal levels.**

This objective stems from the concept of quantifying transit system performance as experienced by passengers. A new method of estimating demand with great spatial and temporal precision can supplement an agency’s existing analyses. The results of this methodology should, in turn, allow a transit agency to answer questions such as the following:

   *How do conditions change across spatial and temporal parameters?*
   *How do conditions vary within the same spatial and temporal context?*

A methodology for defining demand can be used to estimate demand both on a particular day and across many days, providing a baseline for service across the network. Specific instances, such as a day or a peak period on a particular line, can then be compared to this baseline to determine how they vary from expectations.

A systematic, objective method of defining typical demand is useful both in ex-post and real-time contexts. Demand is already analyzed retrospectively on a regular basis, and would benefit from a precise baseline. Assessing how “normal” a particular day or time period was can help in evaluating the effects of a specific event or operational strategy. In real-time, conditions that are trending towards atypical may be a precursor to events that require mitigation.

2. **Develop a framework for analyzing the relationships between the three dimensions of transit system performance.**

While it is beneficial to analyze supply and demand independently, there is further value in examining how these two elements together contribute to service performance. This is a complex relationship, and difficult to capture, as supply and demand affect each other as well as producing service performance. This research proposes methods of exploring this relationship by linking explicit measures of supply and demand to measures of service performance. Variations in these dimensions (by endogenous or exogenous factors) are assessed for their impact on service performance. This research aims to demonstrate techniques that can be used to analyze the three dimensions in order to study how changes in supply and demand affect service performance. An improved understanding of this relationship has the potential to inform current travel demand management strategies, such as gate closures, passenger notifications, and service reductions or increases during planned events. Some of these strategies can also be applied in a short-term operational context. This research can suggest how to apply practices such as gate closures, passenger notifications, and service cuts during unplanned disruptions.

3. **Explore the application of automatic fare collection data as a real-time indicator of system state.**

The third part of this research centers on using Oyster data to describe service in real time. Passenger accumulation is a metric that uses Oyster data to determine the number of users in the system as a proxy for service performance (Freemark, 2013). Accumulation is compared against expected levels in order to evaluate system state.
Passenger accumulation as a real-time indicator is intended to help operational staff anticipate and respond to unusual conditions. For example, above average levels may prompt station staff to close gates in order to manage demand. The metric could also be used to inform passenger notifications.

1.3 Research Approach

The first component of this thesis concerns understanding existing practices to measure demand, supply, and service performance, and developing a new method for estimating demand. This approach takes its inspiration from past work and uses the same data sources as a previous effort, but approaches the solution in a unique way. The intent is to develop a method that can be applied across a range of spatial and temporal contexts, and that can describe demand at varying levels of granularity. The analyses presented in subsequent chapters build upon the methodology developed here.

The second part of this research establishes a framework for analyzing the relationships between dimensions. An ideal framework is first described, using normalized measures of each dimension. The relative values of these measures are then studied across shifts in temporal and spatial parameters. A limited version of this framework is then applied as an example.

The third part of this research explores the passenger accumulation metric. There are two components of this effort:

a. Expand the original methodology of the passenger accumulation metric to remove some of the previous limiting assumptions.

b. Evaluate whether this tool can provide better information – that is, quicker, more accurate, or more disaggregate – than what is currently available.

The first item involves developing a more complex, but more precise version of passenger accumulation than was originally designed. The second phase of this objective is addressed by analyzing, in retrospect, Oyster data as if it were in real time, and comparing it to the information provided by existing tools.

1.4 Thesis Organization

This thesis begins with a review of the existing literature in Chapter 2. Past research on measuring supply and demand is discussed, as is related work on the topics of service reliability and farecard-based performance measures. This section explores the advantages of prior approaches to these topics and identifies gaps in existing research.

Chapter 3 addresses the first objective of this research. It begins by explaining several prominent London Underground data sources, which are used throughout this thesis. The chapter then describes the development of the method for estimating demand and applies it to one London Underground line as an example.

While Chapter 3 observes each aspect of transit system performance individually, Chapter 4 focuses on the interactions between the dimensions. The analysis presented in this chapter explores the
interaction between supply and demand and their joint effect on service performance. Particular emphasis is placed on identifying the contributors to poor service performance, and whether these are primarily supply- or demand-based under various circumstances.

Chapter 5 explores an application of the automatic fare collection data used in the first two parts of this research. This chapter first extends the passenger accumulation metric beyond its original definition to more accurately represent system state. The second part of this chapter then evaluates passenger accumulation’s usefulness as a real-time indicator for the London Underground, by comparing it with tools that are already available.

Finally, Chapter 6 provides a summary of this study and its main conclusions. This chapter revisits this research’s potential applications at Transport for London, and notes the requirements necessary for such an implementation to be valuable to the agency. This thesis closes with an outline of areas that may benefit from further research.
This chapter reviews past work that is relevant to this thesis. It describes how the literature informs the direction of this thesis, and identifies gaps in the existing understanding of system performance that this research addresses. Section 2.1 describes work on measuring the variability in supply and demand. Section 2.2 describes the wealth of research into the factors that influence service performance and how this dimension can be measured. Section 2.3 describes several innovative methods of applying smart card data.

2.1 Measuring supply and demand variability

Van der Hurk et al. (2012) develop a method of predicting passenger volumes in the context of disruption management. The objective of this effort is to estimate passenger flows that will be affected by a disruption in order to inform response strategies. Their approach uses smart card data from the Netherlands Railways and applies an autoregressive integrated moving average model to predict demand in real time.

Morency et al. (2007) use smart card data to measure the variability of transit use in Gatineau, Québec. This approach examines the amount of spatial variation by measuring the number of unique bus stops each passenger used. A k-means clustering algorithm is applied to identify the times of day during which passengers take the most trips. Results are presented both for all users and segmented by card type.

2.2 Understanding service performance

Abkowitz et al. (1978) define reliability as “the invariability of service attributes which influence the decisions of travelers and transportation providers.” This definition discerns two types of reliability: the reliability of service provided by a transit agency and the reliability of service received from the passenger’s perspective. This thesis considers reliability of service provided as a supply-side measure, and the reliability of service as experienced by passengers as a service performance measure. Abkowitz et al. note that the former is most often defined as schedule adherence, and TfL follows this trend in its reliance upon train lateness (as will be discussed in Section 5.3). Abkowitz et al. recommend using measures of compactness to describe reliability rather than deviation from a scheduled value. This study emphasizes the need for reliability measures to reflect aspects of service that are important to passengers. The authors identify several factors that should be considered in the context of passenger-centric performance: passenger waiting time, in-vehicle time, the variability of total journey time, and seat availability.
One of the many further efforts on reliability is that of Uniman (2009). He characterizes performance by quantifying the extent to which individual travel times deviate from their typical values. This measure expands on the approach originally developed by Chan (2007). He defines this metric (shown in Equation 2-1) as the reliability buffer time (RBT), and describes it as “the amount of time passengers are required to allocate in order to complete a journey on-time with high probability.”

\[
RBT_{OD,T,P} = (t_{Nth} - t_{Mth})_{OD,T,P}
\]

Equation 2-1

Where \( RBT_{OD,T,P} \) = Reliability Buffer Time

- \( t_{Nth} \) = the \( N \)th percentile travel time, where \( N \) is an upper percentile
- \( t_{Mth} \) = the \( M \)th percentile travel time, where \( M \) the median

and each of these variables is specific to a given OD pair in time interval \( T \) over sample period \( P \)

Using travel times derived from AFC data, Uniman applies the metric to the London Underground with values of \( N = 95^{th} \) percentile, \( T = \) the three-hour AM peak, and \( P = 20 \) weekdays. He further explores a method to spatially aggregate the RBT measure to assess service performance for an entire line. He finds the RBT for every within-line OD pair (that is, both the origin and the destination are within the line of interest), and calculates the line-level RBT as an average of the OD-level RBTs weighted by the volume of passengers on each OD pair.

The RBT metric describes reliability from the passenger perspective by depicting a simple concept that influences passengers’ expectation of service. As a performance measure that is both passenger-centric and uses units of time, it can be easily used to quantify the benefits or costs to passengers of service changes. However, it has several drawbacks due to its nature as an aggregate measure. RBT evaluates service performance for the typical passenger, and not at the individual level. It is meant to be a long-term descriptor of service quality, and is not intended to evaluate reliability on a single day (or similarly limited sample period). In addition, RBT is relative to the scheduled headway, and so cannot be compared across lines (or across periods with different service levels on the same line).

Ehrlich (2010) defined a variant of RBT that uses AVL data instead of AFC, and Schil (2012) expanded this metric further. Both focused on an AVL-based version of RBT in order to apply the concept to buses, where exits do not require fare transactions. Wood (2015) applies this metric to the rail system in Hong Kong operated by MTR. He defines the AVL-based RBT as the Platform-to-Platform RBT (PPRBT), and notes that a chief advantage of the AVL approach is that it removes the effects of behavioral variation among passengers, such as walking speed, different access and egress distances, and in-station activities. Wood advances this metric by using a train loading model, based on an OD matrix and route choice data, to incorporate the impacts of denied boardings.

Another passenger-centric measure of service performance is Excess Journey Time (EJT), the difference between a passenger’s scheduled and actual journey time. The operating units of TfL use EJT as a performance metric (see Section 3.2), with actual journey times derived from AVL data, a series of models, surveys, and other sources, but does not use AFC data in their estimation. Frumin (2010) explores the concept of EJT in depth, and develops a method for calculating EJT at the individual passenger level using AFC data and applies it to the London Overground. Frumin explains EJT as a performance measure “that strikes useful balance between the passenger’s and operator’s..."
perspectives” in that it measures passengers’ experiences relative to an expectation, and the degree to which the operator’s meets service delivery standards (the schedule). He notes that, like any application that relies upon AFC results, Oyster-based EJT is limited in its application to those OD pairs where the path taken by passengers is either known or can be reasonably inferred by an assignment model. Furthermore, EJT values are limited in their temporal applicability, as they cannot be compared across different timetables (which are usually updated annually in London).

Freemark (2013) proposes using journey time variability to assess the impacts of disruptions. He measures the percent variation in Oyster travel times of trips that begin and end within a defined segment of a line as shown in Equation 2-2.

\[
\nu = \left\{ \sum_{i=0}^{m} \left[ \sum_{j=0}^{n} \frac{(t_j - T_j)}{T_j} \right] \right\} / m
\]

Equation 2-2

Where \( \nu \) = Relative variability for trips \( t \) in line section

\( m \) = number of minutes during the representative sample period

\( n \) = number of sample trips

\( t_j \) = actual journey time for trip \( j \)

\( T_j \) = normal journey time for trip \( j \) (characteristic of the OD pair of the trip)

Freemark’s analysis identifies times and areas (though not specific locations) where service performance was impacted. He also demonstrates how these effects change on different parts of the line over time, suggesting a means for tracking the propagation of delays through the system. By integrating the journey time variability \( \nu \) over the time interval during which it is above a predefined threshold, he quantifies distinct impacts on service performance.

While Freemark’s variability measure is intended to evaluate supply-side events, it also captures demand-driven effects due to its use of AFC data. The author compares journey time variability impacts with supply and service performance measures used by TfL, and notes that while there is some compatibility, his measure identifies additional instances of poor service performance. These impacts are likely due to demand-side influences and interactions between demand and supply.

A notable limitation of Freemark’s analysis is the baseline he uses to determine “normal” journey times. He selects five weekdays with low Excess Platform Wait Time\(^2\) scores (one of TfL’s many service quality measures). Not only is this a small sample size, but using only a single indicator is a simplistic means of characterizing a complex dimension of service. Freemark also developed the original work on the passenger accumulation metric, as will be discussed in Section 5.1.

Similar to the concept of RBT, which focuses on the additional time that passengers must allow for as an element of total travel time, Furth & Muller (2006) explore the concept of budgeted wait time. They argue that mean passenger wait time (explained in Section 3.1), frequently used by agencies to quantify service performance, underestimates the disutility to passengers of poor service reliability. They propose measuring the amount of time passengers plan to spend waiting on platform above

\(^2\) See Appendix C for the definition of Excess Platform Wait Time.
the mean wait time as a reliability (and therefore a service performance) metric. The budgeted wait
time is defined as the difference between the 95th percentile wait time and the mean wait time. The
authors go on to develop a cost function to estimate the impact of unreliability on passengers,
providing agencies with a means to directly quantify the benefits to passengers of investments that
improve reliability.

2.3 Applications of smart card data

Smart card data offers a particularly powerful tool for analysis, as the sheer volume of data collected
yields a greater number of observations across space and time than any other source (Agard,
Morency, & Trépanier, 2006; Bagchi & White, 2005). Pelletier et al. (2011) conduct a broad review
of the benefits of smart cards in public transportation. They divide the possible uses of smart card
data into three arenas: strategic, comprised of long-term planning and customer behavior
characterization; tactical, involving schedule adjustments and longitudinal ridership patterns; and
operational, which encompasses ridership statistics and performance indicators.

Smart card data can provide detailed ridership statistics across spatiotemporal parameters at a very
disaggregate level. Precise load profiles can be generated, allowing for heavy schedule
customization (Pelletier, Trépanier, & Morency, 2011). Passenger boardings and alightings also yield
detailed origin-destination matrices (Chan, 2007), and allow for a more detailed study of
interchanges. Understanding when and where passengers choose to transfer can help agencies
make long-term planning decisions regarding their network structure and schedule to best meet their
customers’ needs (Hoffman, Wilson, & White, 2009; Gordon, 2012).

Digges la Touche (2015) develops a method for determining delays in real time for the purposes of
improving passenger communications, and applies it to both the London Underground and to MTR’s
system in Hong Kong. Her method assigns individual journeys a delayed status if the journey time is
more than a given number of minutes (the delay threshold) above the average journey time for that
OD pair. Because delays are determined as a relative difference between the individual journey time
and the OD pair average, journeys can be aggregated spatially. Three types of delays are identified:
station entrance delays, based on all journeys entering at a station, station exit delays, and line
delays (though only same-line entry and exits are included for all three types). If the number of
delayed journeys at a given exit time is above a predefined minimum, the line or station is
categorized as delayed. This research presents a novel method of identifying delays. However, as the
approach is entirely reliant on journey time, the information it provides about system conditions has
a slight time delay. Journey times are not available until after a passenger has exited, yet they
describe conditions while the passenger was traveling. To be precise, “real-time” journey time data in
this study describes very recent conditions, rather than current ones.
This chapter addresses the first objective of this research: understanding the dimensions of transit system performance and developing a methodology for estimating demand. This methodology establishes a means of creating baselines for demand against which an individual case (a particular day, week, or time period) can be compared, providing a valuable first step for many types of analyses.

Section 3.1 describes the dimensions of transit system performance and the measures available to capture them. The current practices at TfL relating to baselines are covered in 3.2. Sources of automatically collected data and the nuances of TfL’s systems are discussed in 3.3, and Section 3.4 enumerates the specific data sets that are used throughout this thesis. Section 3.5 explains the method used to estimate passenger flow, one critical measure of demand.

### 3.1 Measures of transit system performance

This thesis explores transit system performance across three dimensions: the demand from passengers, the supply of trains, and service performance. Service performance is the result of the interaction between the first two dimensions; that is, it is a measure of how well the supply serves the demand.

Each of these dimensions can be gauged using a variety of metrics. Demand can be measured by the number of entries and exits at a station, the number of journeys made on an origin-destination (OD) pair, or the flow (number of passengers in a given time interval) on a particular section of a line.

Measures of supply describe train movements. Service can be broken down into two characteristics: speed and capacity. Running times express how fast (or slow) service was between two stations. Capacity can be expressed as frequency (trains per hour) or its inverse, the headways between consecutive trains at a station.

Service performance measures are those attributes that are a function of both demand and capacity. There is some overlap between measures of supply and measures of service performance. While speed is primarily a supply measure, it can also be considered a service performance metric as high levels of demand can affect running times.

An important aspect of service performance is regularity. Regularity can be expressed as wait time. Wait time is related to supply measures, as it is derived from headway data, but, like speed, can be influenced by high demand. Another measure of performance is journey time, which includes both speed (as in-vehicle time) and wait time at the stop or platform. Consequently, both the supply and demand characteristics affect journey time, through the speed and wait time components.
TfL calculates wait time (referred to as Platform Wait Time, or PWT)\(^3\) using the standard formula for expected wait time on high-frequency transit service, shown in Equation 3-1 (Transport for London, 2010; Osuna & Newell, 1972).

\[
E[w] = \frac{\sum_i H_i^2}{2 \cdot \sum_i H_i} = \frac{\mu_H^2 + \sigma_H^2}{2 \cdot \mu_H} \tag{Equation 3-1}
\]

Where \( E[w] \) = Average wait time  
\( H_i \) = Observed train headway \( i \)  
\( \mu_H \) = average headway  
\( \sigma_H \) = standard deviation of headways

This definition of wait time assumes that passengers board the first train that arrives, and that passengers do not time their arrival to catch a specific train; that is, arrivals are random and independent of the train schedule (Osuna & Newell, 1972). The first assumption is true when capacity is not limiting, but is violated when denied boardings occur. In practice, the second assumption is considered valid when headways are 10 minutes or less (Welding, 1957; Frumin, 2010; Transport for London, 2010).

The London Underground relies on a suite of performance metrics to track service quality. Many of these, such as Excess Journey Time\(^4\), the percentage of scheduled kilometers run, and the Reliability Index\(^5\), are based on train movement data. Thus, these metrics can be categorized as measures of both supply (in that they are derived solely from supply-side data) and service performance (in that they describe the resulting quality of service).

In recent years, there has been a trend towards using more passenger-centric measures of performance, as discussed in Section 2.2. One such measure is passenger journey time, as calculated from AFC (Oyster) data. AFC-based journey time is one simple measure of a trip's quality from the passenger perspective, and it provides detailed spatial and temporal data about trips at the individual passenger level.

The level of crowding is another measure that represents the passenger's experience; however crowding measures are difficult to estimate accurately. LU estimates crowding penalties from general demand levels and supply-side data as part of the Journey Time Metric\(^6\) (Transport for London, 2012b). When crowding is extreme enough that it results in denied boardings, the effects manifest in longer journey times, but AFC data does not capture situations in which crowding levels are high but not severe enough to cause denied boardings. Better measures of crowding would be based on comparing passenger loads on individual trains to capacity. Current research is focused on estimating passenger loads on trains through passenger-to-train assignment models (Zhu, 2014; Paul, 2010), and this approach may provide richer measures of crowding in the future.

---

\(^3\) See Appendix C for more information on the terms “wait time” and “Platform Wait Time.”  
\(^4\) See Appendix C for the definition of Excess Journey Time (EJT).  
\(^5\) See Appendix C for the definition of the Reliability Index.  
\(^6\) See Appendix C for the definition of the Journey Time Metric.
3.2 Current TfL use of baselines

TfL often compares demand figures from Oyster data against past data for comparable periods. For example, weekly demand may be compared against the previous week or the same week the previous year (see Figure 3-1) or years (see Figure 3-2).

**Headlines**

LU journeys up 3.8% on the equivalent week in 2013; bus journeys up 2.4%; rail journeys up 16.8%.

Last week, was probably the busiest week on London Underground for 2014 so far, and the busiest week since the middle of last December, with over 26.9 million passenger journeys made – 'probably' as official figures will be confirmed later. This increase in demand doesn't appear to be driven by any particular event or station; just continued growth.

The number of journeys on London Overground are much higher when compared last year this due to the severe disruption and line suspensions caused by a derailed freight train at Camden Road.

**Weekly Flash Stats**

- Est. Total Underground passenger journeys: 26.9m (+0.3m on last week)
- Est. Total Bus passenger journeys: 48.9m (-0.2m on last week)
- Est. Total Rail passenger journeys: 5.6m (+0.03m on last week)

![Underground journeys by week](image)

**Figure 3-1. Weekly demand compared to the previous year (Transport for London, 2014b).**

**Figure 3-2. Journeys per week compared to prior years (Transport for London, 2014b)**

TfL also produces reports that tally daily journeys and fare revenue (see Figure 3-3). Demand may be totaled at the daily, weekly, or period\(^7\) levels and compared against the same point in time the previous week, period, or year. These reports are used to measure overall trends, reevaluate forecasts, and track progress against budget. Reports also note both planned and unplanned events (such as sporting events or major disruptions) that may impact demand. For example, Figure 3-1 mentions a freight derailment that occurred the preceding year, making the current year’s growth figures look artificially high.

---

\(^7\) See Appendix B for the definition of an operational period.
Daily Underground journeys were 4.0m in Period 9, slightly up from 3.9m in Period 8.

Underground journeys per day were up 4.4% year-on-year in Period 9.

Daily Underground revenue stood at £7.2m in Period 9, up from £7.1m in Period 8.

Underground revenue per day was up 6.6% year-on-year in Period 9, up from growth of 6.0% in Period 8.

Revenue per journey was up 2.1% year-on-year in Period 9, up from 1.2% growth in Period 8.

Figure 3-3. Daily journeys and revenue compared to prior year (Transport for London, 2014e)
Many of these comparisons are limited in terms of insight provided by the use of small samples as the baseline. Comparing a single Tuesday to the Tuesday of the same week in the previous year may not be helpful if either day is not representative. This thesis aims to fill this gap by exploring methods to determine baselines from a wider sample of data.

On the supply side, TfL’s ex-post analyses of operations focus primarily on the effects of disruptions. Performance metrics, as explained in Section 3.1, often describe both the supply and service performance dimensions. On the operations side, performance is often communicated using scorecards all the way from the system level down to the station level. Figure 3-4 is an example of a line-level scorecard. These performance metrics are mainly based on AVL data. Scorecards, and many other supply-side analyses, compare service to the average over the previous operational period. Scorecards also look at overall trends and the average of all periods in the current financial year.

![Figure 3-4. Jubilee Line performance scorecard (Transport for London, 2014d)](image)

Measures are also often compared against delivery targets. These goals may be set by a variety of means. Targets that are set at the network level, such as LCH and EJT (shown in Figure 3-4), are based on performance forecasts that are developed as part of long-range planning. The exact value may be adjusted within the projected range by LU management, who balance the drive for continuous improvement against the need for realistic goals. Executive commitments, such as the mayor’s directive to cut Lost Customer Hours by 30%, may also steer target-setting. High-level objectives trickle down to individual lines, and further down to their operational subcomponents. Simpler targets are often based on the previous year’s performance, and may take into account

8 See Appendix C for definitions of all the measures in Figure 3-4.
9 As part of his reelection campaign in 2012, Mayor Boris Johnson promised a 30% reduction in tube delays by the end of 2015. TfL measures this goal through Lost Customer Hours accumulated by the London Underground (Transport for London, 2012d).
subjective input on what is a realistic goal for a particular part of the system (for example, a line known to have poor reliability will have lower reliability targets).

As on the demand side, these approaches offer an opportunity to examine supply measures using a larger sample size as the baseline. There is no assurance that the previous period is representative, and not all performance reports make mention of unique events (such as severe disruptions) that may affect the current (or previous) period’s metrics. The current suite of performance metrics is also limited in the level of detail provided, as most metrics cannot be disaggregated temporally or spatially. Some are available only for the four-week operational period, or across the whole day, and may not be calculated for an area more narrowly defined than an entire line. This research aims to supplement current retrospective supply analyses by examining service at a more disaggregate level, in order to provide greater flexibility in the spatial and temporal scope at which supply is studied.

Furthermore, evaluating service performance using AFC data provides a different view than the existing AVL-centric methods. Comparing demand and supply conditions to systematically derived baselines and evaluating their relationship to service performance can provide a different understanding of the system, as will be explored in Chapter 4.

3.3 Available data sources

This analysis focuses on Transport for London and the London Underground system, and is reliant upon TfL’s automated data collection systems. While AFC and AVL data is widely available at transit agencies throughout the world, each system brings its own nuances that influence analyses on which they are based. This section describes the specific TfL data sources that are used in this thesis.

3.3.1 Oyster

The Oyster card is Transport for London’s smart card and the source of the agency’s AFC data. Oyster cards were introduced in 2003, and are used on an array of modes across the London transportation network, including the London Underground, London Overground, DLR, buses, and on some National Rail services. Approximately 80% of all journeys made in the TfL network use an Oyster card (Transport for London, 2012c). Oyster penetration rates are lower at National Rail termini and stations frequented by tourists. Passengers without Oyster cards pay using magnetic tickets, which do not store transaction data and thus cannot be used to extract information about individual journeys.

Cards store transactions, and thus a journey on the (closed-system) rail services involves both an entry transaction, or tap-in, and an exit transaction, or tap out. A complete journey is formed by joining one card’s tap-in to a subsequent tap-out, thus providing information about the card user’s origin station and start time, destination station, and end time. (Certain other data, such as the card type and payment details, is also available, but is not used in this research.) Transactions are recorded to the second, allowing for precise calculation of journey times.

AFC journey times represent several components of a passenger’s journey, not merely the time spent on a train. Oyster journeys are comprised of access, egress, platform wait, and in-vehicle travel time (see Figure 3-5). Journeys that involve transfers also include interchange time and additional phases of platform wait and in-vehicle time (not shown below). Because Oyster measures the time from
faregate to faregate, it excludes the portions of access and egress times that fall outside the gates. At most stations, this lost time is relatively small, though some journeys may involve longer pre-gate line access times due to time spent purchasing tickets, or, in congested situations, queuing for gates.

While access, egress, and (in journeys that use multiple lines) interchange times vary across passengers, they are relatively constant for a large sample, and are largely unaffected by the level of supply. Platform wait time and on-train time, meanwhile, depend upon the speed, frequency, and regularity of trains, and are the primary source of variation in journey times.

![Diagram](image)

Figure 3-5. Components of journey time (Chan, 2007)

One important nuance of smart card data is that journey times vary not only with the characteristics of the trip, but also with characteristics of the user. A passenger who greatly dislikes crowded trains may take a different route, or choose not to board the first train that arrives, resulting in a longer travel time than a passenger who values speed above all else. Additionally, tourists, as unfamiliar users, may take an inefficient path through the system or become lost during their journeys. These all affect journey time, a key measure of service performance in this thesis (Wood (2015) provides a more in depth discussion of cross-passenger journey time variation and strategies to control for this variation). While no evaluation of performance based solely on journey time is perfect, use of a median for a sufficiently large sample ensures a reasonably accurate representation of performance for most passengers (Lam & Small, 2001). This is especially true if analyzing the AM peak period, when tourists are a small percentage of all users.

Further information on Oyster and TfL payment mechanisms is provided in Gordon (2012), Uniman (2009) and Chan (2007).

In December 2012, TfL began accepting contactless bank cards as an alternative means of fare payment on buses. TfL expanded the use of contactless cards to include the Underground, Overground, DLR, most National Rail services in London, and Croydon Tramlink in September 2014. Contactless payment cards (including debit, credit, and prepaid cards) function like an Oyster card using pay-as-you-go (Transport for London, 2014f). Like Oyster, the time, station, and a user ID are recorded for every tap in and tap out, allowing journey time to be determined for each trip. In March 2015, contactless payments accounted for 14% of all pay-as-you-go journeys, which equates to about one million taps per day (Transport for London, 2015).

While it is currently possible to query Oyster data is available in real time, the underlying system was not designed to report at this level of immediacy. A degree of latency exists between the capture of
transactions at station readers, and the availability of data in the central system. Thus, when conducting analysis in real-time or near real-time, most, but not all of the latest transactions will be visible.

### 3.3.2 NetMIS

Train movement data for the London Underground is available through NetMIS, the Network Management Information System and London Underground’s source for automatic vehicle location (AVL) data. Each NetMIS record represents a train’s appearance at a station, and includes the line, station, direction, arrival time, departure time, destination, and identifiers for the trip and lead car. From this data, it is possible to derive the headways and dwell times at a station, running time between stations, and the progress of a train along a line.

The quality of NetMIS data varies widely. Whereas newer lines, such as the Jubilee, have relatively good NetMIS data, lines with older signaling infrastructure, such as the Piccadilly, have frequent gaps. On lines with poor NetMIS quality, many records are missing values for one or more fields. Occasionally, there will be no record of a train at a particular station, though it appears at the preceding and following stations. The trip and/or lead car identifier fields are among the most common missing data, though they will often repopulate – with a different value – at a later station. Recent research has explored a new method to reduce holes in the data by identifying trips made by the same train and rejoining them, which may facilitate the use of NetMIS data in future analyses (Babany, 2014).

For a more detailed description of NetMIS and the train tracking system that generates the data, see Hickey (2011).

### 3.3.3 Gate counts

The turnstiles at LU stations provide counts of all passengers using the system. While gate counts, unlike Oyster, include passengers who pay with magnetic tickets, there is no way to link entry and exits for individual passengers. A handful of stations are either ungated, or allow for a behind-the-gate interchange between Underground and National Rail services. When gateline data either cannot be collected automatically or is unreliable, manual counts either supplement or replace automatic data at these locations (Transport for London, 2013). More information on gateline data, particularly regarding procedures at stations that are not fully gated, can be found in Chan (2007) and Gordillo (2006).

### 3.3.4 RODS

The Rolling Origin and Destination Survey (RODS) is an annual survey conducted on passenger travel patterns. Around 20 stations are surveyed each November, when demand is at its annual high point (Transport for London, 2012c; Transport for London, 2014c). Passengers are asked to complete and return a survey asking about the time and length of their journey, origin and destination, route, trip purpose, and frequency of travel, as well as some basic demographic characteristics. These responses are then scaled up to represent the whole network using gateline entry and exit totals, Oyster origin-destination figures, and the trip frequency data reported in the survey. Each year’s results are added to the cumulative database, dating back to 1998. For stations that have been surveyed multiple times, each year’s results are weighted according to their age.
Many different types of information are available as output of RODS. The outputs include details on journey times and distances, summaries of the number of interchanges passengers make, route choice information, and a “true” origin-destination matrix based on the postcodes of users’ ultimate starting and ending locations (rather than the locations at which they enter and leave the LU system). Most results are available at a variety of spatial and temporal granularities. One of the most useful outputs is the detailed interchange volumes at each station; a category of information that can’t be captured from Oyster. Chan (2007) discusses in detail how OD matrices and reliability measures based on RODS differ from those developed using Oyster data.

The final outputs of RODS are used in many applications. These include train service planning, aiding in decision-making on major infrastructure upgrades, identifying historical demand trends and forecasting, and as an input to network performance measures such as Lost Customer Hours (Transport for London, 2012c; Transport for London, 2012a).

RODS provides several kinds of information that are not available elsewhere, and this thesis is reliant upon its data on interchanges at stations in order to infer passenger flows. However, RODS does have several weaknesses that should be noted when applying it in any analysis. While RODS is largely accepted as accurately describing general trends, it may not be appropriate for describing passenger movements in a specific context (such as single day or time period).

Perhaps most importantly, the original surveys that make up the RODS database represent only a very small sample of all users of the LU network. RODS surveys are handed out to a sample of all riders at selected stations, and of the surveys distributed, response rates are typically in the range of 20-30% (Chan, 2007). Thus, the process of extrapolating RODS survey responses into a representation of all passengers in the network often involves large expansion factors. For outputs that are highly specific in geographic and temporal scope (for example, the number of passengers traveling on one OD pair in a 15-minute interval), the results may be based on a single response. Secondly, the rolling data collection process means that some stations may not have been surveyed for several years. Thus, using RODS to estimate passengers’ travel paths entails relying upon data that may be outdated. In addition, the exit totals that RODS provides at stations may not always be correct. The expansion process ensures that RODS station entries match the actuals from gateline data, but does not do the same for exits (Chan, 2007). RODS is also limited by the fact that it cannot account for seasonal variation in travel patterns.

Despite these limitations, RODS is still a key data source for this research and many other applications. While Oyster and gateline counts can give general information on how demand fluctuates over the year, no other source provides the same level of detail on route choice as RODS. Moreover, solving the interchange problem is not the primary goal of this effort. The intent is to develop an overall methodology for which route choice is an input, and for this purpose RODS is an acceptable data source. In the future, new technologies – such as data from mobile phones – may provide more extensive information on passenger movements. These data could then replace RODS as an input to the analyses conducted in this thesis.

3.3.5 CuPID

The Contract Performance Information Database (CuPID) is a tool used by TfL to record all identifiable disruptions to service. A myriad of details are recorded; including the time and location
that the disruption occurred, a narrative of the original event that caused the disruption and the resulting effects to service, the duration (measured as the start time until the time at which all trains are back on schedule). Every effort is made to identify the root cause if it is not immediately clear. TfL’s official procedure is log all disruptions that last two minutes or longer.

### 3.4 Data sets used in this thesis

The most heavily used data in this thesis are Oyster records from the autumn of 2013. Due to the immense size of Oyster data, TfL’s current policy is to store only eight weeks of data at a time. This research uses a 100% sample from October 20 to December 14, 2013, providing 40 weekdays for analysis. Bus journeys and non-travel transactions (such as topping up) are excluded, as are a variety of invalid transactions; including incomplete journeys, journeys that begin and end at the same station, journeys with a travel time of zero minutes, and certain staff passes. This time period conveniently encompasses the month of November, when RODS (the survey) is conducted.

An important limitation of the Oyster dataset is that at the time it was collected, recorded Oyster transaction times were truncated to the minute.\(^{10}\) Thus, the journey time calculated from an entry and exit pair has a margin of error of ±59 seconds (see Figure 3-6).

![Figure 3-6. Differences between recorded and actual journey times (Chan, 2007)](image)

AVL data is the most easily available of all the data used in this research. AVL data can be pulled for nearly any date range in recent years, for any station. AVL data is primarily used in conjunction with Oyster data for the corresponding dates and time periods.

Gate counts are used from the five weeks spanning November 4 – December 6, 2013. Gate counts are not available by individual day, but as weekday and weekend averages across this span of time. These values include National Rail passengers at most stations with shared gates. Nine stations use

---

\(^{10}\) As noted in Section 3.3.1, Oyster times are now recorded to the second. TfL began logging Oyster transactions to the second in early 2014, in preparation for the introduction of contactless payment cards on the Underground.
counts from a previous year, adjusted by the average change between years. Most of these stations are locations where automatic counts are not available, and use manual counts from past years.

The process to turn raw RODS survey results into complete, network-wide data takes nearly a year. This thesis uses RODS results from the most recent year available, 2013. (As described in Section 3.3.4, however, every year’s outputs include data from earlier years’ surveys.)

CuPID data is used for specific days which appear as examples in this thesis. This information is used in an explanatory function and not in quantitative analysis.

3.5 Determining passenger flow

Oyster data provides the number of entries and exits at stations, and the number of passengers on the London Underground’s some 72,000 unique origin-destination pairs. However, in order to link demand to supply, characteristics such as running time, headway, ridership are needed in a format that is spatially compatible. Train service on a line affects, and is affected by, an accumulation of OD pairs, and each OD pair affects many line segments. Therefore, demand is better represented as flow, or the number of passengers using a line segment in a specified period of time.

3.5.1 Flow as a measure of demand

This section describes a methodology for estimating passenger flow on each link of a line; that is, for the line segment between two adjacent stations. This thesis defines flow as the number of passengers that use a link during a specified time interval. For instance, the flow between Russell Square and Holborn on the Piccadilly Line between 8:15 and 8:30 is the number of passengers that enter this link during those 15 minutes. Each passenger is counted only once on each link; however, depending on the length of the time interval, a passenger may travel on several links, and thus can be counted multiple times at the line or segment level. An example is provided for the southbound Victoria Line during the AM peak using the data described in Section 3.4.

This approach can be used with either gate counts alone or with Oyster counts scaled to gate count totals to determine the number of passengers traveling during the specific date and period. The use of gate count totals ensures that passengers with magnetic tickets are included. Previous efforts have used Oyster data and expanded the number of passengers on each OD pair to the gateline totals (Ravichandran, 2012) to account for magnetic tickets while keeping the level of spatial detail afforded by Oyster. Because this method relies on entries and exits, rather than passengers on a specific OD pair, this approach scales entries and exits separately (and not by OD pair) when Oyster data is used (see Equation 3-2).

\[
e_{n,i}(H) = \frac{\sum G_{n,i}(H)/k}{\sum O_{n,i}(H)/k} \cdot O_{n,i}(H)
\]

Where \( e_{n,i}(H) \) = total entries (or exits) at station \( n \) on day \( i \) during hour \( H \)

\( G_{n,i}(H) \) = gate entries (or exits) at station \( n \) on day \( i \) during hour \( H \)

\( O_{n,i}(H) \) = Oyster entries (or exits) at station \( n \) on day \( i \) during hour \( H \)

\( k \) = number of links

Within this section and the next, “period” or “time period” refers to one of TFL’s standard daily time periods (such as the AM peak). See Appendix B for clarification.
\[ O_{n,i}(H) = \text{Oyster entries (or exits) at station } n \text{ on day } i \text{ during hour } H \]
for all weekdays in the sample \([0, k]\)

While it is simpler to use gateline totals directly, gate counts are not available to the same level of temporal detail as Oyster (see Section 3.3.3). There are two limitations of the gateline data used in this thesis: one, gateline count data is only available for 15-minute intervals, and so cannot account for minute-to-minute fluctuations in demand. Secondly, only averaged gateline data is available (as \[\frac{1}{k}\left(\sum_{0}^{k} G_{n,i}(H)\right)\]) and thus it cannot be used to determine single-day flows. The following section describes the estimation of flow using gate counts alone. Either approach can provide flow for any multiple of the 15-minute intervals defining the gateline counts. If flow is needed for an individual day, scaled Oyster entries and exits should be used. If finer temporal detail is desired than the 15-minute level, either scaled Oyster taps or the method developed by Ravichandran (2012) should be used.

The actual line segments used by a passenger are unknown from their Oyster journey and may not necessarily be easily inferred from their entry and exit taps, as many OD pairs may have several plausible paths. Some information about their route is therefore required from another source. Guo (2008) and Paul (2010) both explore the use of disutility models for path choice on the Underground, to which RODS is one of many inputs. This approach uses route choice information obtained directly from RODS data.

### 3.5.2 Detailed flow methodology

RODS provides passenger counts at various levels of spatial and temporal aggregation that are representative of a single day. This methodology converts these counts to proportions of entries or exits. The gate counts furnish the total number of passengers entering and exiting at each station. The percentages from RODS are then scaled by the gateline totals. This methodology calculates flows by 15-minute intervals on every link along the line in a single direction.

**Determine station-specific time windows**

**Step 1.** For a single time interval, this method determines the flow on every link during that interval. For a given station (the base station), the calculation process uses data about passenger movements at upstream stations before the time interval, and at downstream stations after the interval (these are the offset stations). This requires travel times between all possible permutations of base and offset stations (see Table 3-1). Because this method is direction-specific, travel times in the reverse direction are shown as negative.

Table 3-1 and Table 3-2 are abbreviated versions of the full tables for the southbound Victoria Line. Both tables show the base station in the left column, and the offset stations across the top row. The base station for each link is the link's origin station (e.g., Oxford Circus is the base station for the Oxford Circus–Green Park link), and thus, every station except the last (Brixton in this example) is a base station. For each base station, every station is an offset station, resulting in a \((n - 1) \times n\) matrix for \(n\) stations.
Table 3-1. Sample station-to-station travel times (minutes)

<table>
<thead>
<tr>
<th>Base Station</th>
<th>Preceding/following station (southbound)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Walthamstow Central</td>
</tr>
<tr>
<td>Walthamstow</td>
<td>0</td>
</tr>
<tr>
<td>Central</td>
<td></td>
</tr>
<tr>
<td>Oxford Circus</td>
<td>-23</td>
</tr>
<tr>
<td>Stockwell</td>
<td>-32</td>
</tr>
</tbody>
</table>

Table 3-2. Sample station-specific time windows

<table>
<thead>
<tr>
<th>Base Station</th>
<th>Time interval at base station</th>
<th>Preceding/following station (southbound)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Walthamstow Central</td>
</tr>
<tr>
<td>Walthamstow</td>
<td>8:45 - 9:00</td>
<td>N/A</td>
</tr>
<tr>
<td>Central</td>
<td></td>
<td>...</td>
</tr>
<tr>
<td>Oxford Circus</td>
<td>8:45 - 9:00</td>
<td>8:22 - 8:37</td>
</tr>
<tr>
<td>Stockwell</td>
<td>8:45 - 9:00</td>
<td>8:13 - 8:28</td>
</tr>
</tbody>
</table>

For each base station and 15-minute interval, every offset station is assigned a time window that is shifted by the travel time between it and the base station. This creates a matrix of time windows. A sample is shown in Table 3-2 for the time interval from 8:45 to 9:00. These time windows are used to select data in later steps; estimating the flow on the Oxford Circus–Green Park link (for example) from 8:45 to 9:00 requires data from Walthamstow Central in the time interval 8:22 – 8:37. If estimating flow for more than one 15-minute interval, such a matrix is constructed for every interval.

**Calculate RODS proportions**

The RODS results provide a file of station movements; that is, the number of passengers that make each possible access, egress, or interchange movement at every station. These numbers are smoothed when the raw RODS data is processed. A sample of this file is shown in Table 3-3.

The station, type of movement (access, egress, or interchange), and descriptions of the starting and ending locations are given. The number of passengers that make each movement is provided for each 15-minute interval and summed for each time period.
Table 3-3. Sample RODS station movement records

<table>
<thead>
<tr>
<th>Start platform</th>
<th>End platform</th>
<th>Station NLC</th>
<th>Start station</th>
<th>AEI</th>
<th>AM peak</th>
<th>0715-0730</th>
</tr>
</thead>
<tbody>
<tr>
<td>STOCKWELL T H</td>
<td>STOCKWELL NORTHERN NB</td>
<td>716</td>
<td>Stockwell</td>
<td>A</td>
<td>953</td>
<td>123</td>
</tr>
<tr>
<td>STOCKWELL T H</td>
<td>STOCKWELL VICTORIA NB</td>
<td>716</td>
<td>Stockwell</td>
<td>A</td>
<td>1959</td>
<td>254</td>
</tr>
<tr>
<td>STOCKWELL T H</td>
<td>STOCKWELL NORTHERN SB</td>
<td>716</td>
<td>Stockwell</td>
<td>A</td>
<td>364</td>
<td>23</td>
</tr>
<tr>
<td>STOCKWELL T H</td>
<td>STOCKWELL VICTORIA SB</td>
<td>716</td>
<td>Stockwell</td>
<td>A</td>
<td>197</td>
<td>2</td>
</tr>
<tr>
<td>STOCKWELL VICTORIA NB</td>
<td>STOCKWELL EXITS</td>
<td>716</td>
<td>Stockwell</td>
<td>E</td>
<td>92</td>
<td>3</td>
</tr>
<tr>
<td>STOCKWELL VICTORIA SB</td>
<td>STOCKWELL EXITS</td>
<td>716</td>
<td>Stockwell</td>
<td>E</td>
<td>1273</td>
<td>45</td>
</tr>
<tr>
<td>STOCKWELL NORTHERN NB</td>
<td>STOCKWELL NORTHERN NB</td>
<td>716</td>
<td>Stockwell</td>
<td>I</td>
<td>809</td>
<td>106</td>
</tr>
<tr>
<td>STOCKWELL VICTORIA SB</td>
<td>STOCKWELL NORTHERN SB</td>
<td>716</td>
<td>Stockwell</td>
<td>I</td>
<td>200</td>
<td>13</td>
</tr>
<tr>
<td>STOCKWELL VICTORIA SB</td>
<td>STOCKWELL NORTHERN SB</td>
<td>716</td>
<td>Stockwell</td>
<td>I</td>
<td>4470</td>
<td>194</td>
</tr>
<tr>
<td>STOCKWELL VICTORIA SB</td>
<td>STOCKWELL NORTHERN NB</td>
<td>716</td>
<td>Stockwell</td>
<td>I</td>
<td>631</td>
<td>23</td>
</tr>
<tr>
<td>STOCKWELL NORTHERN NB</td>
<td>STOCKWELL VICTORIA NB</td>
<td>716</td>
<td>Stockwell</td>
<td>I</td>
<td>6246</td>
<td>899</td>
</tr>
</tbody>
</table>

This methodology uses a different access/egress/interchange classification than that presented in the RODS results. As RODS has until recently been limited to London Underground stations, this data treats the London Underground as a separate system from Overground, DLR, and National Rail lines. Thus, a movement consisting of a transfer from the Overground to an Underground line is considered an access, not an interchange (and an egress in the reverse case). Movements that do not involve the Underground (such as a transfer from the DLR to National Rail, or a passenger exiting from the Overground) are categorized as interchanges. However, for the purpose of this methodology, such movements are considered from the perspective of an Oyster user traveling on the full public transport system. Thus, an interchange between LU and another system is an interchange, as it will not produce an entry or exit transaction on an Oyster card.

**Step 2.** For access (egress) movements, the RODS value for each movement is divided by the total of all access (egress) movements at that station in the time period. This results in the proportion of entering (or exiting) passengers that use each line in each direction.

RODS proportions are calculated using the counts from the full time period. At the 15-minute level, the figures for each movement are inherently less reliable because of the much smaller sample size. The percentage of passengers entering at Pimlico that board the southbound Victoria Line may vary greatly from one 15-minute interval to the next, and from one day to another. However, one can reasonably expect the same percentage to be fairly consistent over the entire AM peak. Consequently, the period-level proportions are considered more reliable.

**Step 3.** Interchange-on movements are those where passengers transfer from another line onto the line of interest. For interchange-on (interchange-off) movements, the RODS count of each movement is divided by the sum of RODS egress (access) movements at all subsequent (preceding) stations along the line-direction path. That is, the interchange from the Jubilee southbound to the Victoria Line southbound at Green Park would be divided by the sum of all egress movements at Victoria Station, Pimlico, Vauxhall, Stockwell, and Brixton. (The denominator includes all egress movements, not only those coming from the Victoria Line southbound.)
As in the previous step, the numerator is calculated for the full time period. For each element in the denominator, values are summed across a time window equal in length to the entire time period, but offset by the amount shown in Table 3-2. For time windows that don’t fall exactly on the quarter hour, the first and last time interval values are interpolated. Thus, for a 6:40 to 9:40 time window, values are summed from the various time intervals as follows:

\[
\frac{5}{15} (6:30 - 6:45) + (6:45 - 7:00) + \cdots + (9:15 - 9:30) + \frac{10}{15} (9:30 - 9:45)
\]

*Multiply RODS proportions by gateline counts*

**Step 4.** For each station, retrieve all gateline entries and exits for each 15-minute interval within the time period. In addition, retrieve entry and exit counts for every offset window for each station as determined in Step 1 (see Table 3-2). This involves interpolating the values at the boundaries of the window (see Step 3).

**Step 5.** For access (egress) movements, each RODS proportion (defined by a unique combination of station, line, and direction) is multiplied by the total gateline entries (exits) at that station for each 15-minute interval. This is the actual number of passengers entering a station and boarding a line-direction path.

**Step 6.** For interchange-on (interchange-off) movements, each RODS proportion (which is defined by a unique start line-direction and end line-direction combination) is multiplied by the sum of all gate exits (entries) at subsequent (preceding) stations along the end line-direction path. For each subsequent (preceding) station and time interval, the exits (entries) are interpolated for the offset time interval determined in determined in Step 1 (see the example in Table 3-2).

**Aggregate movements**

**Step 7.** The change in passengers at each station during each time interval is then calculated by subtracting the sum of the egress and interchange-off movements from the sum of the access and interchange-on movements. For the starting terminus in a single direction (for example, Walthamstow Central on the Victoria Line southbound), the flow on the first link (Walthamstow Central to Blackhorse Road) is merely the sum of the access movements.

**Step 8.** For each 15-minute interval, the flow on each direction-specific link is then the sum of the change in passengers at each station between the origin terminus and the link’s start station. Every element in the sum – that is, the change in passengers at each preceding station – is calculated for the offset time interval determined in Step 1.

For the first several intervals in the period, the offsets for upstream stations fall before the start of the current period. (For instance, the flow on Oxford Circus-Green Park from 7:00-7:15 is dependent upon the change in passengers at Blackhorse Road from 6:44 – 6:59.) To calculate these figures, the current period RODS proportions are used, and the gate counts are interpolated for the actual offset time window. (Thus, the Blackhorse Road calculation would use AM peak period RODS proportions, but gate counts interpolated for 6:44 – 6:59.) Using the already-calculated current period RODS proportions greatly simplifies the computational process, and the proportions are assumed to be relatively comparable between the end of one period and the beginning of the next.
Step 9. Sum the flow on each link over multiple 15-minute intervals to get the flow for longer time windows. In this example, summing the link flows over all twelve intervals produces the flow for the whole AM peak period (7:00 – 9:00 a.m.).

3.5.3 Example flows on the Victoria Line

The figures below show the results of using the methodology described in Section 3.5.2 to determine flow on the Victoria Line during the AM peak period (7:00 – 10:00 a.m.). This profile shows that the most congested southbound link during the morning peak is Euston to Warren St. This is true for both the full time period and the peak 15-minute interval (8:30 – 8:45). In the northbound direction, the highest flows are seen on the link from Victoria to Green Park. The shapes of the profiles are very similar between the full period and peak interval for both directions.

For comparison, link loads from the general RODS results alone are also shown. The profiles generated from RODS are alike in both shape and magnitude, though the flow values from this methodology tend to be slightly larger than those from RODS. The most noticeable difference is the higher estimates on the last few northbound links during the peak 15 minutes (see Figure 3-10), though as these volumes are much lower than maximum load point, capacity is not an issue.

![Graph showing link flows on the Victoria Line](image1)

![Graph showing link flows on the Victoria Line](image2)

Figure 3-7. Link flows on the Victoria Line (southbound 07:00 - 10:00)  
Figure 3-8. Link flows on the Victoria Line (southbound 08:30 - 08:45)
Table 3-4 examines the differences between the results of this method and RODS in more detail. There is a general pattern of increasing discrepancies between the two sources moving down the line, as slight calculation differences at the link level accrue. Calculated flow values in the first part
of the time period tend to be larger than the corresponding RODS values, whereas the reverse is true in the last few 15-minute intervals. This is primarily due to differences between the RODS and gateline exit counts. As mentioned in Section 3.3.4, RODS does not reconcile exit counts with gateline data. Gateline exits tend to be slightly larger than RODS exits in the first half of the period, and smaller than RODS in the second half. Thus, in calculating flow, fewer exits are subtracted in the first half of the period, resulting in larger flow values. The difference between RODS and gateline exits also affects the number of passengers interchanging on, but this influence is small compared with the large numbers of exits.

Figure 3-11 and Figure 3-12 display peak-interval flows against the capacity on each link. For the most part, the scheduled capacity is sufficient to handle these passenger volumes. However, the peak load point in the southbound direction is above the capacity. Furthermore, the scheduled capacity (shown in solid black) represents the maximum supply provided under ideal conditions. Trains are occasionally cancelled or withdrawn from service for a variety of reasons, including asset malfunctions, passenger actions, or a lack of staff availability. In addition, even if trains are merely delayed, and not withdrawn, it reduces the effective capacity for some period of time. As shown here, a lack of even one train trip in a 15-minute interval can reduce capacity below demand. (Once a delayed train begins moving again, the effective capacity may rise above the scheduled level temporarily if several trains are running close together. Depending on the number of passengers that have accrued during the delay, this extra capacity may or may not be able to quickly serve all unmet demand.)

Figure 3-11. Link flows and capacity on the Victoria Line (southbound 08:30 - 08:45)
The issue of capacity is a critical one because it has such a significant effect on the quality of service received by passengers. Insufficient supply results in denied boardings, over-crowded trains, and longer journey times as passengers are forced to wait for the second train in order to board. In the worst of the cases shown in the figures above, the unmet demand of the southbound link from Euston to Warren Street with two cancelled trains is approximately 2,200 passengers. It is important to note that unmet demand is not necessarily equivalent to denied boardings as subsequent trains may arrive at shorter headways to make up for one that is delayed or withdrawn from service.

Additionally, passengers may board at higher densities during delayed conditions (effectively increasing the train’s capacity). True capacity limits are not well known; LU calculates capacities at densities ranging from 4 to 7 passengers per square meter (for standing passengers). TfL most often uses capacities based on 5 passengers per square meter, as does this thesis. While planning staff at TfL believe the 7 pass./m² “crush” capacity to be unrealistic in most scenarios, the hard evidence to support this is thin. Neither is there a good understanding of how much variation exists between passengers’ personal space preferences; anecdotally, some riders are willing to board extremely crowded trains while others elect to wait for the next one. The subject of actual capacity and its variation is an area that merits further research.

Table 3-5 demonstrates how actual supply can differ considerably from the schedule when disruptions occur. This table shows the number of trains and the corresponding passenger capacity in each direction (the number of southbound-running trains is measured at Euston station, and northbound at Victoria; the respective peak load points). The 15-minute numbers, in particular,
suggest that high numbers of passengers were affected by the loss of supply on this day (as the period-level scheduled capacities have a large buffer compared to the expected demand).

Table 3-5. Scheduled and actual service on the Victoria Line (14 Nov 2013)

<table>
<thead>
<tr>
<th></th>
<th>Southbound 07:00 - 10:00</th>
<th>Southbound 08:30 - 08:45</th>
<th>Northbound 07:00 - 10:00</th>
<th>Northbound 08:30 - 08:45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled trains</td>
<td>91</td>
<td>8</td>
<td>91</td>
<td>8</td>
</tr>
<tr>
<td>Actual trains</td>
<td>67</td>
<td>6</td>
<td>63</td>
<td>5</td>
</tr>
<tr>
<td>Scheduled capacity</td>
<td>90,909</td>
<td>7,992</td>
<td>90,909</td>
<td>7,992</td>
</tr>
<tr>
<td>Actual capacity</td>
<td>66,933</td>
<td>5,994</td>
<td>62,937</td>
<td>4,995</td>
</tr>
</tbody>
</table>

These results show that this method of calculating flow is an extremely useful tool. The generated flows can be compared to scheduled and actual supply data in order to identify the times of day and parts of the line at risk for crowding. The flows produced by this method are similar overall to those derived directly from RODS. This is to be expected, since RODS is a primary data source for this methodology, and these results were calculated using average gate counts from 25 weekdays, much as RODS gives general results for the month in which it is conducted. However, this methodology can be adapted to estimate flow values on a specific day, either by using single-day gate counts, or by using Oyster values from a single day and expansion factors.
Examining trends of supply and demand each in its own right can provide valuable insight for a transit agency. However, it is essential to remember that these dimensions do not exist in isolation, but as interrelated components of transit system performance. Analyzing all three dimensions jointly potentially offers agencies even greater benefits than any study that focuses on a single dimension alone.

This chapter defines a framework for exploring how supply and demand interact to produce service performance. Service performance, by definition, is the degree to which the services supplied meets passenger demand. This relationship is not a straightforward link between inputs and outputs; supply and demand can influence each other as well as service performance. While the general direction of these relationships is usually inherently obvious, the details of when, where, and how much these dimensions interact are not well understood. A better understanding of how these dimensions impact service performance can inform both long-term strategies, such as travel demand management schemes and service planning, and short-term operational tactics, such as gate closures and passenger notifications.

This thesis develops a framework for analyzing how supply, demand and service performance interact. The primary focus is the effects of the two input dimensions (supply and demand) on performance, but the interdependencies between supply and demand are also considered. The objective of this thesis is not to comprehensively describe the relationships themselves, but to provide a means of analyzing the available data in order to better understand these relationships.

The first section of this chapter discusses the factors that can affect service performance. Section 4.2 sets out how each cause can be identified using the available data. The complete framework for analyzing the three dimensions collectively is defined in Section 4.3. Section 4.4 presents a limited example of this analysis and discusses the implications of the results. The final section discusses the limitations of the framework, and conclusions regarding its viability.

**4.1 Drivers of service performance**

In striving to improve the understanding of service performance, it is natural to focus on the conditions that result in poor service performance. Causes of poor service performance essentially fall into two categories: low supply and high demand. This section describes the scenarios in which poor service performance is primarily due to low supply, and when it is due mainly to high demand.
These categories can be broken down further; low supply may be due either to incident-related delays, or to headway variability. High demand can take the form of crowding on trains, or crowding on platforms.

Incident-related delays are those that have a clearly identifiable cause, such as a mechanical failure of the track or signal infrastructure, or the train itself, or a customer falling ill or being injured while in transit. These causes may be either endogenous (a driver late for his or her shift) or exogenous (a severe storm affecting operations). TfL’s standard operating procedure is to record and track all such incidents that last two minutes or more, including any effect on service (whether trains were delayed or cancelled, and the duration of each). The other, less concrete, form of reduced supply is uneven headways. Long headways have an obvious negative effect on passengers, who experience both longer waiting times and fewer, and hence more crowded, trains. However, a situation in which some headways are longer than scheduled and others shorter also reduces capacity, and thus, the total supply. Even if the average headway remains as scheduled, the average wait time increases (see Section 3.1), because more passengers experience the longer headways. Uneven headways may or may not be the result of a specific event. Train delays certainly cause headways to be more variable, but headways can also vary due to differences in driving styles between train operators and variations in dwell times.

Crowding due to high demand can take several forms, each of which affects service performance in multiple ways. Boarding and alighting is more difficult when trains are crowded, and when trains near capacity, denied boardings occur. Crowding on trains also contributes qualitatively to passengers’ perceptions of service. Crowding on the platform, meanwhile, can also impact the boarding and alighting process. In extreme cases, platform crowding leads station staff to close the faregates until passenger numbers return to safe levels.

While supply-side and demand-side causes can contribute to poor service performance concurrently, in many cases one source dominates. Naturally, supply-driven causes are the principal driver of service performance in contexts where demand is low. Along the spatial spectrum, supply drives service performance in the outer fare zones of the Underground service area. Temporally, off-peak periods are (except in unusual cases) affected only by supply-side contributors, though supply issues impair peak period performance as well. Conversely, demand-driven causes emerge in Central London and during peak periods. This thesis explores this relationship and attempts to quantitatively capture this paradigm by identifying to what extent each cause contributes to service performance across different spatial and temporal contexts.

4.2 Identifying contributing causes of delay

Each of the higher-level causes described in the previous section manifests in different ways. In order to ascertain which of the underlying causes are at work, an array of supply and demand measures are examined. Each possible cause potentially affects each of these measures differently, and by examining which measures are higher or lower than typical levels, it is possible to attribute service performance to a primary cause. These measures are summarized in Table 4-1.
4.2.1 Indicators of high demand

Crowding on the platform can be identified through two measures. It is most easily observed through an increase in entries at the station in question. If a station becomes crowded due to high demand (that is, in absence of supply-side problems), then the number of entries must necessarily be higher than normal. In addition, in-station crowding will generally result in higher dwell times as both the boarding and alighting processes are likely to take longer on congested platforms.

One measure indicative of crowding on trains is the flow along the train’s path of travel. Flow can be estimated using the approach detailed in Section 3.5 for either a single link or a series of links that are used to travel on a specific origin-destination (OD) pair. High demand on a section of line can also be directly represented by the number of passengers traveling on select OD pairs. Though a single segment is used on dozens of OD pairs, identifying a selection of those that are certain to use the segment provides a small sample for comparison with relative flow levels. This approach suggests measuring the number of journeys taken on a collection of similar OD pairs that use the portion of the line that is under investigation. As a passenger’s exact path cannot be known from AFC data alone, only pairs with origins and destinations on the same line are used. Same-line OD pairs can almost certainly be assumed to have taken place on the line they share, rather than on another path through the system requiring an interchange. The number of passengers on specific OD pairs should be used as a supplement to measures of flow, as these journeys are only a fraction of those included in the flow. The advantage of OD demand is that it can be measured directly, whereas flow is an estimate based on many assumptions. Flow is a more complete measure, but OD demand is more accurate. OD journeys can also be a useful proxy in conditions that invalidate the method used to determine flow, such as severe disruptions or special events. The last metric used to identify on-train crowding is dwell time. Long dwell times can be a result of either crowded platforms or crowded trains, as both make boarding and alighting more difficult, and as such it can be an indicator of either cause.

4.2.2 Indicators of low supply

When a train is delayed, longer running times are a direct result. In some cases, however, a train may make up for lost time following a delay, particularly on longer trips or after very short incidents. The approach taken in this chapter avoids evaluating running times on journeys with scheduled running times over an hour, and recognizes that short delays cannot be identified from running time alone. Frequency is also used to detect disruptions, as delays and cancellations will be reflected in fewer trains per hour. Frequency is an extremely useful measure for gauging train supply, as it is descriptive of service over some time interval, yet is sensitive to the impact of just a single delayed train.

Headway variability can be inferred by observing headways directly, but also by quantifying the dispersion of headways. Here, dispersion is measured as the standard deviation of all headways in the past half-hour. The variation in headway is also observed through frequency, its inverse. Although using both frequency and headways as indicators may seem redundant, frequency is included merely to present a different view of the supply conditions. In some cases, frequency is simpler to interpret, as the impact of a single high headway can be easy to overlook when examining individual headways, but is clearly reflected in a rolling measure of frequency.
Finally, two measures are used to evaluate service performance. Journey time is the primary metric, as it captures delays at every stage of a trip (except for the small portion of access and egress time that occurs outside the faregates). Journey time is also a customer-focused metric, as it is measured at the individual passenger level (rather than by train or station), and is extremely important to passengers. Additionally, journey time can be measured at very disaggregate levels (for the OD pair), and a reasonable sample can be obtained over a short time span. Service performance is secondarily measured by wait time. An advantage of wait time is that it is consistent with current TfL practices, as the agency already uses it as a performance metric. However, wait time is an aggregate measure, and therefore can be less informative for shorter periods of time.

4.3 Framework for analyzing system dimensions

This framework is designed to examine how supply, demand, and service performance characteristics change across different contexts, in order to understand how supply and demand affect service performance. There are several types of variability that are potentially of interest: spatial variability, or how the dimensions change from the low-demand outer zones of the TfL network to the high-demand setting in Central London; intra-day variability, or variation across a time period or across hours of the same day, and longitudinal variability, or the variation at a given time of day from one day to another. This framework assumes that a sample of 40 weekdays is available for all datasets.

Temporal scope

Out of the possible causes of poor service performance discussed in Section 4.1, the demand-side causes are the least understood. As previously established, demand-induced performance problems emerge during the peak of the peak period, and in or near Central London. As the when of demand-side causes is already known with some precision – the peak of the peak is a fairly narrow temporal scope – this framework focuses on where capacity-constrained conditions exist. Accordingly, this framework concentrates on spatial variability over the two types of temporal variability. The temporal scope is therefore limited to the peak 30 minutes in the AM peak.

Spatial scope

This framework is discussed in the context of the westbound Piccadilly Line during the AM peak (see Figure 4-1). However, the concepts outlined here can be applied to other high-frequency lines that serve both low- and high-demand areas, and that have reasonably closely-spaced stations to lend flexibility in choosing and aggregating OD pairs.
Unit of analysis

The line is analyzed using OD collections, rather than individual OD pairs. An OD collection is a group of origin-destination pairs in which all origins are similar (both in their geographic location and in their supply and demand characteristics) and all destinations are similar. In the example of the Piccadilly Line during the AM peak, westbound journeys (into Central London) are analyzed using multiple OD collections. The first OD collection is comprised of origin stations near the eastern terminus, and destination stations in Central London. This method compares OD collections with successively shorter trip lengths, and a consistent group of destination stations. An example of the OD collections to be used is shown in Table 4-2, and Figure 4-1 shows their locations along the line.

Measurement of delay indicators

Measures of supply, demand, and service performance are normalized and aggregated across each collection. This method focuses on several specific indicators out of those shown in Table 4-1 which are normalized and aggregated across each collection. Each individual journey time is divided by the median over the 40 days for that OD pair. The resulting normalized journey times can then be examined both in aggregate (the average of all normalized journey times across an OD collection) or as the distribution over the collection. Although the median is used to normalize, journey times are aggregated across an OD collection using the average of the normalized journey times. This is because when examining service performance for a specific sample, it is important to include the effects of any unusually high (or low) journey times, as they may be indicative of interactions between dimensions.
Table 4-2. Origin-Destination collections including peak load point

<table>
<thead>
<tr>
<th>Collection</th>
<th>Origin stations</th>
<th>Destination stations</th>
<th>Min journey time</th>
<th>Max journey time</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFS-CVG</td>
<td>Cockfosters</td>
<td>Covent Garden</td>
<td>30</td>
<td>37.5</td>
</tr>
<tr>
<td></td>
<td>Oakwood</td>
<td>Leicester Square</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Southgate</td>
<td>Piccadilly Circus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGR-CVG</td>
<td>Arnos Grove</td>
<td>Covent Garden</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Bounds Green</td>
<td>Leicester Square</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wood Green</td>
<td>Piccadilly Circus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPL-CVG</td>
<td>Turnpike Lane</td>
<td>Covent Garden</td>
<td>16</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>Manor House</td>
<td>Leicester Square</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Piccadilly Circus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARL-CVG</td>
<td>Arsenal</td>
<td>Covent Garden</td>
<td>9</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>Holloway Road</td>
<td>Leicester Square</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caledonian Road</td>
<td>Piccadilly Circus</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Demand is represented by flow on each link. Flow, as discussed in Chapter 3, is the number of passengers that use a link in a specified interval of time. If that interval of time is large relative to time it takes to travel a link, a passenger may use more than one link in the given time interval. That is, within a 15-minute time interval, a passenger may be counted on three different links, if each link has a running time of five minutes. Consequently, if summing flow across links, some passengers will be double-counted. This effect can be avoided by examining flow only at the link level. Therefore, for each OD collection, demand is presented as the flow over the peak 30 minutes on each link. For each link, the flow is normalized to the median flow (on that link, in the peak 30 minutes) across the 40-day sample. Because each OD collection overlaps with other collections, only the flow on links that are not shared are of interest when comparing two collections. By looking at the normalized flows on each link that appears in the scope of analysis, it is possible to identify whether a jump in demand (relative to expected levels) occurred on a particular link or links.

Running time can be manipulated similarly to demand. The running time of each train on each OD pair is normalized by the scheduled (or by the median) running time for that OD pair. Each train’s normalized running time can then either be examined as a distribution, or averaged to obtain a single value for the collection as a whole.

Supply is also analyzed using headways. Headways can be measured from the first train that departs before the 30-minute window to the last train that departs within the window (to avoid calculating an unrealistic number if a train passes just inside or outside of the window boundary). Headways are measured at every origin station in a collection, and each headway is normalized by the scheduled value for that station. Headways provide several useful metrics: first, the average normalized headway for a collection, the standard deviation of normalized headways, and capacity. Capacity is determined by taking the inverse of headways. Capacity can be expressed as either percentage of scheduled capacity, by taking the inverse of the normalized headway values, or as actual capacity, by taking the inverse of the non-normalized average headway and multiplying by train capacity. Comparing normalized demand and capacity measures is useful for drawing general inferences, while comparing actual demand and capacity numbers allows the number of denied boardings at a station to be estimated.
Comparing some collections with origins well outside Central London with collections that have origins close to the most congested sections of the line can highlight the effects of crowding. For the collection CFS-CVG, only a small part of the journey is affected by crowding; whereas for the ARL-CVG collection, the majority of the trip is congested. As the location of the origin stations – and thus the demand on the line – changes incrementally, it may be possible to identify how quickly service performance changes, and where it becomes severely impacted.

Table 4-2 demonstrates some of the criteria to consider when grouping OD pairs into collections. The journey times shown in the right columns are derived from scheduled running times. Each collection has about the same number of origin and destination stations. Journey times for all OD pairs in a collection should be within the same range; therefore, the number of origin and destination stations is limited to three each. For the same reason, journey times between stations in a group of origins (or destinations) should be similar and relatively short. Selecting origin (or destination) stations that are temporally close together ensures that the overall journey times on all OD pairs within a collection are comparable in length. For example, the scheduled running time between Southgate and Arnos Grove is 4.5 minutes; somewhat longer than any of the other link running times on this segment of the Piccadilly Line. Therefore, these two stations should not be in the same origin group.

Additionally, supply levels should be kept comparable between collections. 18 trains per hour serve the origin stations in the first collection, whereas at all the other stations in Table 4-2 trains are scheduled at the full peak frequency of 24 trains per hour. On the demand side, Finsbury Park is excluded from any of the collections, as it has notably different demand patterns. Due to the interchange with National Rail, Finsbury Park sees much higher boardings than any other station east of King’s Cross. Furthermore, the Victoria Line is a viable alternative for many Piccadilly Line destinations from Finsbury Park. Thus, these OD pairs should be excluded, as these passengers cannot all be assumed to use the Piccadilly Line.

<table>
<thead>
<tr>
<th>Collection</th>
<th>Origin stations</th>
<th>Destination stations</th>
<th>Min journey time</th>
<th>Max journey time</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFS-KXX</td>
<td>Cockfosters</td>
<td>King’s Cross</td>
<td>24.5</td>
<td>29.5</td>
</tr>
<tr>
<td></td>
<td>Oakwood</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Southgate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGF-KXX</td>
<td>Arnos Grove</td>
<td>King’s Cross</td>
<td>15.5</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Bounds Green</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wood Green</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPL-KXX</td>
<td>Turnpike Lane</td>
<td>King’s Cross</td>
<td>10.5</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>Manor House</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARL</td>
<td>Arsenal</td>
<td>King’s Cross</td>
<td>3.5</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>Holloway Road</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caledonian Road</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The measures described above can be analyzed across different longitudinal scopes in order to answer different questions. If the question asked is, “How did the level of demand and supply impact service performance on day x?” then it is most useful to examine these measures (after
normalization and any aggregation) for a single day. By contrast, if the question is about general patterns, such as, “What locations are most at risk for denied boardings, and to what extent?” then these measures can be examined together over all days in the sample. While many of these measures can simply be averaged over all 40 days (e.g., the 40-day average of normalized journey times on one OD collection), boiling a large sample down to a single value may not provide sufficient insight into the inter-dimensional relationships. A more comprehensive method is to analyze the suite of measures on each day across all days. This is a more in-depth task, but if the question to be answered is specific enough, it can be limited to a manageable scope. For instance, if the days with the greatest discrepancies in journey time and running times are selected, then each OD collection’s values can be compared with its relative demand measures to identify if specific stations or links have recurring capacity issues.

One possible sub-analysis in this framework involves comparing OD collection’s journey times to running times. In a scenario where demand is below capacity, the distribution of journey times should be similar to the distribution of running times. However, if there are denied boardings, there will be some individual journey times that are greater than the corresponding running times (relative to the majority of journey times). In a case where large numbers of denied boardings occur, journey times will have a wider distribution. (In theory, denied boardings should incur a bimodal distribution, with the two local maxima representing passengers who boarded the first train and those who had to wait for the second. However, at peak-hour headways, the variation in access and egress times washes out the distinction between the two modes.)

This method of analysis takes advantage of Oyster data. Oyster journey time is a useful service performance measure for several reasons. Oyster journeys provide a large sample of data which is particularly useful in an analysis where the variation in service performance is of as much interest as the actual service performance values. Secondly, Oyster journey times are more representative of actual passenger journeys than estimates derived from AVL data, as they are measured at the individual passenger level. Thirdly, Oyster journey times include both wait time and in-vehicle time – two important aspects of quality from the passenger’s perspective – in a single measure.

### 4.4 Example Analyses

This section presents a limited analysis of two days using some of the techniques described above. This approach compares supply metrics (standard deviation of headways and frequency) with service performance (journey time) and demand (flow). In lieu of full OD collections, the supply metrics are examined at four stations along the line, and journey times are generated for two OD pairs. From east to west, Bounds Green, Manor House, Holloway Road, and King’s Cross are roughly evenly spaced along the Piccadilly Line from its eastern end to Central London (see Figure 4-1). All of these stations are on the portion of the line where service is scheduled at 24 trains per hour during the peak periods; the highest capacity provided on the Piccadilly. The first three of these stations each represent a different OD collection as shown in Table 4-2, while King’s Cross represents a collection that spans the peak load point. This analysis uses data from Oct. 20 – Dec 14, 2013 (as described in Section 3.4).

This approach begins by examining service performance on a day that is known to have had supply-side problems. On Dec. 9, 2013, the Piccadilly Line experienced five separate disruptions during the
morning peak period, four in the westbound direction and one eastbound at Turnpike Lane, on the easternmost segment of the line. There was also a disruption at Arnos Grove just before the peak period. The information logged in CuPID for these disruptions is summarized in Table 4-4. There were several instances of train drivers not being available to work; while spare drivers were assigned to the first two, the third train lacking a driver (at 7:13) was cancelled. Thus, the duration is the length of time from when the train should have entered service until it actually entered service. Although the long duration would seem to imply that this is a major disruption, a single train cancellation has no effect on other trains, and therefore is not by itself considered serious. The signal failure at 8:41 was by far the worst incident of the morning; as a result, multiple trains were short-turned and reformed (renumbered to match the schedule of a later train).

Table 4-4. Disruptions logged in CuPID (9 Dec 2013)

<table>
<thead>
<tr>
<th>Start Time</th>
<th>Location</th>
<th>Direction</th>
<th>Duration</th>
<th>Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>6:44</td>
<td>ARNOS GROVE</td>
<td>EB</td>
<td>7</td>
<td>Train operator not available</td>
</tr>
<tr>
<td>7:00</td>
<td>ARNOS GROVE</td>
<td>WB</td>
<td>3</td>
<td>Train operator not available</td>
</tr>
<tr>
<td>7:13</td>
<td>COCKFOSTERS DEPOT</td>
<td>WB</td>
<td>201</td>
<td>Train operator not available</td>
</tr>
<tr>
<td>7:31</td>
<td>LEICESTER SQUARE</td>
<td>WB</td>
<td>3</td>
<td>Customer tripped while boarding</td>
</tr>
<tr>
<td>7:45</td>
<td>TURNPIKE LANE</td>
<td>EB</td>
<td>5</td>
<td>Customer reported burning smell</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(spurious)</td>
</tr>
<tr>
<td>8:41</td>
<td>LEICESTER SQUARE</td>
<td>WB</td>
<td>48</td>
<td>Signal failure</td>
</tr>
</tbody>
</table>

Figure 4-2 shows the frequency of trains in the preceding half hour at the four stations. The scheduled frequency is also shown (though the exact scheduled frequency does vary slightly between 11 and 13 trains per 30 minutes, these fluctuations are smoothed over wider intervals). The effects of the signal failure can be clearly seen beginning at 8:45. Service at King's Cross still begins to recover more quickly than at the other stations, but only very slightly. Frequency is also quite sensitive to small disruptions, as is shown by the slight dip below the scheduled level in the first half hour of the period. This is likely the result of the earlier delays caused by unavailable drivers.
In Figure 4-3, the standard deviation of all headways in the past half hour is shown over the AM peak period at each of the four stations. As in the previous figure, King’s Cross seems the least impacted by disruptions. Holloway Road has the most variation out of the four, and the two easternmost stations are very similar. From this view, it is very clear how the supply at each station was affected over time. By this metric, the effects of the disruption are felt very suddenly, and dissipate just as rapidly at all stations but King’s Cross. However, there appear to be no ill effects at the start of the period from the more minor incidents. This suggests that standard deviation of headways is not as useful a metric as is frequency.

Both Figure 4-2 and Figure 4-3 show that King’s Cross is not as affected by the signal failure incident as the other stations. This is slightly surprising, given that King’s Cross was the closest to the incident. Another way to examine the effects of this incident is shown in Figure 4-4, which displays the distribution of headways from 8:30 to 9:00. The signal failure occurred in the middle of this window, which is also when demand peaks.

Under normal circumstances, one would expect the headway distribution at Bounds Green (the station furthest from Central London) to have the lowest mode and the narrowest spread. Trains are easier to keep on schedule at outer stations where demand is lower, and even in the case of disruptions, the lack of extreme demand such as seen in Zones 1 and 2 means that it is easier to recover from delays. Meanwhile, stations closer to the inner zones would have wider distributions that peak at slightly higher headways.

Figure 4-4 does seem to suggest that service was more affected at the inner stations on Dec. 9, but this could merely be due to the proximity to the signal failure. Bounds Green, the furthest from both high demand and the incident, does have the lowest mode. While most of the headways at King’s Cross were low, it also experienced one very long headway of 18 minutes. (With only a few exceptions, all scheduled headways at these stations are 2.5 minutes during the peak.) King’s Cross
is shown here to be more affected than other stations, but in the previous two figures recovered earlier. A likely explanation is that while King’s Cross felt the impacts of the signal failure first, it also was the first to experience relief when trains began moving normally again. The effects at Manor House and Holloway Road are similar; both have a few headways moderately above the schedule but nothing resembling the extreme headway at King’s Cross. This suggests that delays resulting from the signal failure propagated back upstream. At the outset of the incident, King’s Cross was the most affected, and stations further up the line were only slightly affected (Manor House and Holloway Road), or not affected at all if far enough away (Bounds Green).

The next step examines the effects of the documented disruptions on journey time. Journey time is measured on two OD pairs: Bounds Green to Holborn and Caledonian Road to Holborn. These OD pairs are two of the highest demand OD pairs on the Piccadilly Line during the AM peak period. Both OD pairs are affected by the characteristically high demand in Central London, but journeys that begin at Bounds Green are also affected by the system conditions in the less-used outer zones. Journey times on these OD pairs have distributions given in Figure 4-5, which shows the 90th, 50th, and 10th percentile travel times of all journeys with exit times during the morning peak. (Journeys are grouped into 15-minute intervals to determine the percentiles.) Though the trip from Bounds Green takes roughly twice as long, the two journeys have comparable distributions. The Caledonian Road-Holborn distribution ought to have a narrower range of journey times, given that it is shorter, and therefore has less variation in on-train time. Although the spread of the Caledonian Road-Holborn distribution is smaller than that of Bounds Green-Holborn at the beginning and end of the period, near the middle the two have similar spreads. Around 8:45, where journey times peak (as do passenger loads), both distributions have a range of about 8 minutes between the 10th and 90th percentile.
Figure 4-5. Distribution of journey times (Piccadilly Line, westbound)

Figure 4-6 and Figure 4-7 display typical journey times on the two origin-destination pairs as well as the journey times on this particular day. Each figure shows the 90th, 50th, and 10th percentile journey time, similar to Figure 4-5. Each of the colored points on these graphs represents an individual passenger's journey time on Dec. 9, plotted by exit time. The impacts of the disruptions to the line are clear as both OD pairs experienced severe delays around the time of the signal failure, with increased travel times of up to 50 minutes above the range of normal values. The absolute increases in journey time are similar between the two OD pairs. However, travel times on the second OD pair, originating at Caledonian Road, seem to recover more quickly, as by 9:30 most trips are under 20 minutes (though many of these are still above the 90th percentile). Interestingly, the minor delays that occurred at the beginning of the period do not appear to have affected service performance.

Figure 4-6. Typical and individual journey times (Piccadilly, BGR-HOL, 9 Dec 2013)
To understand the demand on these days, flow is determined for each of the links in the scope of analysis. Flow is computed as described in Section 3.5, but instead of gate entries and exits, this analysis uses Oyster entry and exits. Expansion factors are computed for each station and hour of the day to scale the Oyster transactions up, so that the total number of transactions includes passengers with magnetic tickets as well as those using smart cards. This allows flow to be computed on a specific day (as gateline counts are only available as a weekday average, rather than by individual day). Estimating flow by individual day allows the distribution of flow within a given spatial and temporal scope to be created. As an example, Figure 4-8 shows the median flow (out of the 40 weekdays in the dataset) on the westbound Piccadilly Line both for the full AM peak (light blue) and the peak 30 minutes (dark blue). The peak load point is on the link from King’s Cross to Russell Square (a result consistent with the general line loadings generated from RODS data).

An examination of the flow on Dec. 9 implies that service performance affects demand, rather than the reverse. Figure 4-9 shows the westbound flow on each link from Cockfosters, the northeastern terminus, to Covent Garden, in Central London during the peak 30 minutes (8:30 – 9:00). Figure 4-10 shows how the flow changes over the AM peak period on the peak load link, King’s Cross to Russell Square. Both views show that the flow is slightly lower than typical. The flow on the peak load link during the peak 30 minutes is about 8,000 passengers, compared to the median of 8,700 passengers as shown in Figure 4-8 (about an 8% decrease). Figure 4-10 also shows the 10th and 90th percentile flows on this link. The flow on Dec. 9 drops below the 10th percentile just after 9:00, and stays below this threshold for 45 minutes. Given the knowledge that this part of the line experienced disruptions, the reduction in flow may be the effect of supply-side problems on demand. It is almost certain that during the 18-minute headway at King’s Cross shown in Figure 4-4, station staff made announcements to waiting passengers, and many other channels communicate the state of service to passengers both inside and out of the system. Passengers aware of the delays may have sought other routes (or modes), thereby lowering demand. This is all the more feasible in Central London, where many alternative routes and modes are available.
Figure 4-8. Median link flows (Piccadilly Line, westbound)
In contrast to Dec. 9, Nov 27 was a relatively uneventful day. No notable disruptions occurred during the AM peak period. The supply-side metrics show that service was run quite close to schedule on this day as shown in Figure 4-12 and Figure 4-11. Headway standard deviation is consistently low across all four lines (Figure 4-12). Frequency is close to the scheduled value, although there is slightly more variation from 8:30 on. There is no apparent cause for the slight dip in frequency at 8:45 (and again at 9:30 at Bounds Green); it is apparently just normal variation in service.
Nov. 27 was one of the highest-demand days in the dataset of 40 weekdays. Figure 4-13 shows the flow across each link during the peak 30 minutes, and Figure 4-14 shows the flow on the highest-demand link across the morning period. The flow from King’s Cross to Russell Square between 8:30 and 9:00 was over 9,100 passengers. Overall, the flow on this link was at the 90th percentile nearly the whole morning period.

In spite of the high demand, journey times were relatively normal from Bounds Green to Holborn and from Caledonian Road to Holborn, as shown in Figure 4-15 and Figure 4-16. Like Figure 4-6 and Figure 4-7, the two figures below show the 10th, 50th, and 90th percentile journey times out of the 40-weekday sample (identical to the distributions shown in Figure 4-5). Each colored point represents an individual passenger’s journey time on Nov. 27 plotted at the time they tapped out. Though both OD pairs had a few passengers (one at 7:30 and another at 9:40 in Figure 4-16, as well as others
right on the 90th percentile line in both figures) with high journey times, the majority of journeys were well within the distribution, and a few even had abnormally low journey times.

**Figure 4-15. Typical and individual journey times (Piccadilly Line, BGR-HOL, 27 Nov 2013)**

The precise flow on the link from King's Cross to Russell Square is shown in Figure 4-17. (As flow is measured in passengers per some interval of time, it is technically a step function as shown here. Other depictions of flow over time in this chapter are smoothed for clarity.) The theoretical capacity is also displayed, assuming 6 trains arrive every 15 minutes at perfectly even headways. The actual capacity is also shown, based on the 15-minute frequency at King's Cross. Around 8:45, this link exceeds the theoretical capacity, but just barely. However, from 8:20 to 8:40 there was a slight dip in the frequency – a difference of just a few long headways in a row – and suddenly this link is over capacity by more than 1,000 passengers. In fact, all links from Turnpike Lane to Leicester Square
had flows of over 3,000 passengers in the interval from 8:30 to 8:45. This temporary reduction in
capacity was not due to any train cancellations or delays, but merely a few trains running slightly
behind schedule, resulting in three headways in a row above 3.5 minutes.

In theory, the discrepancy between actual capacity and demand during the peak 30 minutes on Nov.
27 should have resulted in denied boardings. However, it is impossible to confirm this suspicion from
the journey times shown in Figure 4-11 and Figure 4-12 above. The distribution of journey times is
quite wide (see Figure 4-5), particularly during this part of the time period. A number of these
journeys are above the median travel time, but almost none are above the 90th percentile. If a
customer is left behind, but successfully boards the next train, his or her journey time increases by
only a few minutes – hardly enough to stand out. Such small differences in travel time could also be
due to slower access and egress times than the average passenger. Verifying denied boardings from
journey time requires both that the increased journey times be noticeably higher than other times on
the same day, and that a number of passengers exhibit these increased journey times. In the case of
Nov. 27, neither of these criteria are met.

![Figure 4-17. Comparing flow to capacity (Piccadilly Line, westbound, KXX-RSQ, 27 Nov 2013)](image)

Based on the demand measures, it can be surmised that Nov 27 experienced both crowding on
trains and crowding on platforms. However, because the journey times are not consistently high
(Figure 4-15 and Figure 4-16), it cannot be concluded that customers experienced poor service
performance as defined by increased journey time, but rather that they likely experienced poor
service performance in terms of crowding. In addition, this day exhibits signs of headway variability
between 8:15 and 9:00, which was a major cause of the apparent denied boardings in the busiest
part of the line segment.

4.5 Summary

While the framework presented in Section 4.3 offers a potentially valuable tool for analyzing demand
and supply-side impacts on service performance, it does have several limitations. This framework
involves a great deal of data that can be analyzed at different temporal and spatial scopes and a
variety of levels of granularity. Several types of OD collections are suggested, and the various measures can be analyzed both as an individual day against a distribution, or as multiple days in series. A significant difficulty is identifying the appropriate spatial and longitudinal scope of analysis.

An additional limitation is that the required data manipulation results in units that are not intuitive. The need for normalization, and in some cases aggregation or statistical functions such as standard deviation transforms well-understood values into units that are meaningless to the average observer. In applying this methodology, care should be taken to present the results in a format that the intended audience can easily interpret.

The small-scale exploration of supply, demand, and service performance measures in the preceding section supports the assertions made in Section 4.3 that this type of analysis can identify interactions between dimensions. Comparing flow with supply-side metrics can be used to identify situations in which reduced supply limits demand in turn. This concept could be further investigated by comparing flow to records of when passenger notifications were circulated. Customers' responsiveness to notifications can be evaluated by determining whether passengers begin to shift away from a line before or after notifications are disseminated.

Journey times, in addition to supply-side measures, are another means of describing the rate of service recovery at different points along the line. This chapter also shows that frequency is a particularly useful tool for capturing the nuances of supply-side problems, while the standard deviation of headways can show how these problems vary from one part of the line to another.

By examining flow in conjunction with actual frequencies, areas that are likely subject to capacity constraint issues can be identified. Oyster data may be used to further confirm the existence of denied boardings by comparing journey times to running times in cases where denied boardings are suspected. This type of analysis is best done as an intra-day comparison, where individual normalized journey times on a particular day are compared against each other. High-demand OD pairs, in particular, have wide journey time distributions, and journeys that experience denied boardings are included in these distributions. This makes comparing journey times to their distributions less useful in the context of crowding (though it continues to be useful in analyzing the effects of incident-related delays). An additional means of identifying denied boardings using Oyster journey time is to compare actual (non-normalized) journey times to running times. While journey times will always be slightly higher than running times due to the inclusion of access and egress time, journey times that exhibit a bimodal distribution over a small temporal scope (such as the peak 30 minutes) are indicative of denied boardings.
As extracting data in real time from automatically collected data systems has become possible in public transportation systems, research on the benefits of instantaneous information has markedly expanded. Real-time AVL data has yielded applications ranging from dynamic control strategies (Eberlein, Wilson, & Bernstein, 2001) to improved passenger information systems (Tang & Thakuriah, 2012; Watkins, Ferris, Borning, Rutherford, & Layton, 2011). This chapter explores the development of a metric that describes service in real time using Oyster data.

Conditions that are trending towards atypical may be a precursor to events on either the demand or supply side that could benefit from mitigation. Improved real-time indication can help operational staff anticipate and respond to emerging problems. For example, high levels of passenger demand may prompt staff to close gates in order to keep platform crowding below critical levels. A real-time measure could also help controllers assess the progress of service recovery after a disruption; currently no systematic information about demand is available in the control room. Such a metric could also be used to improve customer communications. Using aggregate Oyster data, this approach offers a different view of the system than that furnished solely by AVL-based sources.

This effort has two parts: the development of the metric, and an assessment of its potential value. In both parts, the intent is not to create a fully functional tool, but to examine whether such a measure is worth further research and development for eventual application across a range of spatial and temporal scenarios, at various levels of granularity. The first section begins by reviewing the concept of passenger accumulation as it was initially developed, and its limitations. Section 5.2 proposes improvements to the initial measure. The tools and processes currently available to TfL for indicating system state are summarized in Section 5.3. Finally, this chapter concludes by assessing the potential for passenger accumulation to act as an alternate or supplementary measure to TfL’s existing tools.

5.1 Previous research on real-time applications of AFC data

The concept of passenger accumulation was originally introduced by Freemark (2013) as a means of identifying disruptions and quantifying their impacts on passengers. Freemark theorizes that if service disruptions slow journeys but do not affect the rate at which passengers enter the system, more passengers will accumulate in the system than is normal. While disruptions specifically are not the subject of this thesis, any condition that deviates from the norm is of interest (whether caused by a disruption or otherwise).

Freemark defines passenger accumulation at time $T$ as the number of passengers who have tapped in by time $T$ but have not yet tapped out. Using Oyster data, the total number of exits can be subtracted from the total entries to track the number of users in the system at a given time. Thus, this measure describes system state using the tap-ins of currently active passengers to assess how
the system is performing. By contrast, measuring journey time requires waiting for a passenger to exit in order to characterize their journey. A metric that relies only on active, not past, passengers could be used in real time with less of a time lag than journey time metrics based on completed trips.

Passenger accumulation is determined as shown below:

\[
p_{i,g}(T) = \sum_{n \in m} \left( \int_0^T e_n(t) - x_n(t) \, dt \right)
\]

Equation 5-1

Where

- \( p_{i,g}(T) \) = passenger accumulation on day \( i \) and line segment \( g \) at time \( T \)
- \( e_n(t) \) = cumulative entries at time \( t \) at station \( n \)
- \( x_n(t) \) = cumulative exits at time \( t \) by passengers who entered at station \( n \)
- \( m \) = the set of all stations in the line segment \( g \)

In applying the passenger accumulation concept to the Piccadilly Line, Freemark splits the line into five segments and determines accumulation on each segment independently, as in Equation 5-1. In order to ensure that the passengers counted actually used the Piccadilly line, entries from all interchange stations are excluded. The difference between entries and exits is recalculated every 10 minutes. An example of Freemark’s results, displaying the Trunk segment of the Piccadilly Line (Acton Town – King’s Cross), is shown in Figure 5-1 for five consecutive weekdays in September 2011.

Figure 5-1. Absolute passenger accumulation on the Piccadilly Line between Acton Town and King’s Cross (Freemark, 2013)\(^{12}\)

This measure can then be compared to conditions on a typical day. Passenger accumulation’s potential value stems from the ability to compare a specific day (or time period) to the expected

\(^{12}\) This chart shows a spike in accumulation during the PM peak period, but no such rise in the morning. This can be explained by the isolation of passengers to their original line segments. During the AM peak, most riders enter the system outside Central London, rather than in the Trunk segment, and are counted on the section of line where they board.
conditions in order to describe how well the service is performing. One way to accomplish this comparison is by defining the Excess Passenger Accumulation as the difference between the actual and expected passenger accumulation, as shown below.

$$Excess \ Passenger\ Accumulation = p_{i,g}(t) - p_{s,g}(t)$$

Equation 5-2

Where $p_{i,g}(t)$ = passenger accumulation on day $i$ and line segment $g$ at time $t$

$p_{s,g}(t)$ = passenger accumulation on a typical day $s$, on line segment $g$ at time $t$

Continuing the example shown in Figure 5-1, Freemark plots excess passenger accumulation on a single day (Sep. 22, 2011) and uses another day from the same week (Sep. 21) as the typical day $s$ (see Figure 5-2). A number of delays occurred in the Eastbound direction between 17:00 and 20:00 on September 22. This chart clearly shows larger values of passenger accumulation on the study day within the Trunk and Heathrow segments.

While passenger accumulation has the potential to provide valuable insight into the system state as both a real-time and passenger-centric measure, the original methodology suffers from several inadequacies that limit its utility. These weaknesses are summarized below, and will be addressed in Section 5.2.

![Figure 5-2. Excess passenger accumulation for the previous example (Freemark, 2013)](image)

The limitations of this method are:

- **Omission of interchange passengers.** The exclusion of interchange journeys severely limits this metric. It excludes two types of journeys: those on the Piccadilly Line that began at an interchange station, and those that transfer onto (or off) the Piccadilly from (or to) another line. The omission of transfer stations is a significant issue, particularly for the trunk segment of the line through Central London. Here, only 4 of 13 stations (excluding the boundary stations, King’s Cross and Acton Town; and the rarely-served Turnham Green) do not have

---

13 A complete list of all stations used in each segment is provided in Appendix D.
interchanges. This results in too small a sample of stations on the line’s most crucial segment to yield consistently reliable and usable results. Furthermore, Freemark notes that because stations served by multiple lines cannot be used, the method as proposed cannot be applied to segments with shared track, such as the joint section of the Piccadilly and Metropolitan lines, or any part of the Circle line.

- **Unknown path of travel.** While a passenger is known to have tapped in at a particular station at a specific time, this method makes no inferences about where they may go within the system. After entering, a user could still be in the designated segment, in a different segment of the same line, or on a different line entirely. Thus, the passenger accumulation of a line segment may be very different from the actual number of riders on that segment at that time. This produces two problems. First, the implicit assumption that riders are still on the line segment where they initially boarded may produce spurious passenger accumulation results. Secondly, passenger accumulation is directionally ambiguous. High accumulation could indicate an excess of passengers in either (or both) direction(s).

- **No further spatial disaggregation.** The original method breaks the Piccadilly, one of the longest Underground lines, into just five segments. While passenger accumulation at the line segment level can be valuable as a general descriptor of performance, it may be too broad to be useful in the context of spurring real-time operational interventions.

- **Small sample size for defining typical conditions.** Only a single day was used to define $p_{s,g}$ rather than a series of days. This day $s$ was assumed to be “normal” based on a comparison of its absolute passenger accumulation to other days in the same week. Furthermore, day $s$ was the day before study day $i$, and consequently the comparison may be compromised by systematic day of week variation.

### 5.2 Expanding the passenger accumulation metric

This thesis aims to develop the concept of passenger accumulation further by addressing some of the limitations identified in the previous section. In doing so, this approach relies upon a modified version of the flow methodology presented in Section 3.5, altering the assumptions in order to estimate the system state in real time. The proposed method allows passenger accumulation to be used at a variety of spatial resolutions, providing greater flexibility for its potential applications.

#### 5.2.1 Developing an improved methodology

Of the previously identified limitations, the greatest weakness is the exclusion of interchange stations. This research incorporates interchange stations by using RODS data to assign entering passengers to lines. This will greatly increase the number of customers included, as it allows for consideration of all passengers arriving on (and leaving) the line at interchange stations, whether they are entering (or exiting) via the station entrance or from another line. Using RODS data also addresses the second limitation, by inferring passengers’ direction of travel and interchange movements.

This method further augments the original accumulation metric by adding flexibility in the spatial scope of analysis. Increasing the number of line segments used provides information about
accumulation at a more disaggregate level. As with the method to calculate flow, this approach provides a means to determine flow at the link level.

Another critical enhancement is based on improving the small sample size used to define “normal” in the original accumulation metric. Here, typical passenger accumulation is represented as the distribution of accumulation over 40 weekdays, thus providing a more systematic baseline against which an individual day can be compared.

At the most disaggregate level, passenger accumulation is closely related to flow. Flow, as described in Section 3.5.1, is the number of passengers that use the line between two adjacent stations during a specified time interval. Thus, the flow on the link from Euston to Warren Street from 8:05 to 8:10 is the number of passengers that enter the link between 8:05 and 8:10. Depending on the length of the interval, and the journey times between stations, one passenger may travel on multiple links. If a passenger enters Euston at 8:05, he or she is counted on the Euston-Warren Street link, the Warren Street-Oxford Circus link (a minute and a half later, at 8:06:30) and on the Oxford Circus-Green Park link (entering at 8:08:56). Each passenger is counted only once on each link in the time interval, but may be counted multiple times at the segment or line level.

Accumulation, by contrast, is an instantaneous measure. In an ideal implementation – that is, without the limitations mentioned earlier – the passenger accumulation between Euston and Warren Street at 8:10 is the number of passengers active between the two stations at that exact point in time. When measuring accumulation, therefore, each passenger is only counted once, regardless of the spatial level of aggregation.

As the time interval used to calculate flow shrinks, the number of passengers that appear on more than one segment diminish. Across a one-minute interval, very few passengers are counted twice (the scheduled running time between Covent Garden and Leicester Square, the shortest link in the network, is one minute exactly). If flow is calculated across a one-second interval, each passenger is counted only on the link where they are in that second – a definition nearly identical to that of accumulation. In essence, flow across an interval of infinitesimal length is equal to passenger accumulation.

However, it is not possible to determine flow at one-second (or smaller) intervals. Because Oyster data is only recorded to the minute in this case, flow at its most precise level can only be determined for a one-minute interval. While the number of passengers traveling on a link during one minute is greater than the flow across a one-second interval, the shape of the curve produced by measuring flow across sequential one-second intervals is identical to that of flow across one-minute intervals. Furthermore, this relationship holds as the interval widens (for instance, the shape of the flow graph is identical between flow across one-minute and five-minute intervals, as shown in Figure 5-3).

The analysis presented here deviates from the original definition of passenger accumulation previously presented Section 5.1, though it remains true to the concept (and thus will continue to be termed accumulation). This thesis computes flow, rather than instantaneous accumulation, across very short intervals. Five-minute intervals are used for practicality.
Figure 5-3 shows that regardless of whether five-minute or one-minute intervals are used, the shape of the curve is identical (five-minute intervals are plotted on the left axis, and one-minute intervals on the right). The five-minute values are exactly five-times the one-minute interval values. As will be shown later in the chapter, passenger accumulation is a relative measure. The actual accumulation unit values are not themselves meaningful, but rather are used to compare passenger accumulation on a specific day to its distribution over a number of days. As long as both the single-day values and the distribution over multiple days are determined using the same time interval, the individual day's relative performance remains the same. Applying five-minute intervals also reduces the noise associated with using one-minute intervals.

The approach used here follows the flow methodology described in Section 3.5.2, with three exceptions.

- Five-minute intervals are used instead of 15-minute intervals.

- This method uses Oyster tap-ins and tap-outs instead of gateline entries and exits. As explained in 3.3.1, TfL's Oyster data is currently available in near-real time. Furthermore, in a real-time application, the temporal detail – transactions to the minute rather than in 15-minute intervals – provided by Oyster is essential. Although this approach excludes customers paying by magnetic tickets, this is considered a minor limitation, as they are a small minority of all customers, and as ticket customers are not likely to impact the overall travel patterns observed in the line-level Oyster and RODS data.

- The number of interchange passengers is assumed to be proportional to the number of passengers entering at upstream stations. In Section 3.5, interchange-ons are expressed as a proportion of the passengers that exit at downstream stations at a later time. Because this method uses exits that occur after the time interval for which flow is being determined (or "future" exits), it cannot be applied in real-time. Interchanging-on passengers are therefore determined by multiplying earlier upstream entries from Oyster by the number of passengers.
interchanging-on in RODS as a proportion of the earlier upstream entries in RODS. This methodology assumes that the number of interchange-on passengers at a particular station is proportional to the number of passengers entering the line (from outside the system) upstream. For example, the number of passengers interchanging on to the southbound Victoria Line at Finsbury Park is calculated as a proportion of the passengers who earlier entered the system at Seven Sisters, Tottenham Hale, Blackhorse Road, and Walthamstow Central.

Strictly speaking, the third point above is a potentially weak assumption – passengers interchanging on are affected by service on the lines they are coming from, while entries on another part of the line are not. Similarly, disruptions at outlying stations on the line may decrease the number of upstream entries, but have no effect on passengers transferring from other lines at stations in Central London. The primary weakness of this assumption therefore is its inability to accurately represent situations in which demand is high (or low) in one part of the network, but not another.

However, this assumption is sensitive to network-wide variation in demand. If a major event causes demand to increase (or decrease) throughout the Underground system, it is reasonable to assume that both interchange-ons and direct boardings increase (or decrease) proportionately. A further advantage of this approach is that it is symmetric with the way interchange-off passengers are calculated. Most importantly, this assumption allows interchange-on passengers to be calculated in real time, by using tap-ins that have already occurred.

5.2.2 Applications of the proposed methodology

Figure 5-4 and Figure 5-5 show an example of passenger accumulation on a day with relatively good service. Though this example uses past data, the accumulation is recalculated every five minutes, as if it was in real time. As explained in the previous section, accumulation here is calculated as five-minute flows, so the accumulation shown at 8:00 is actually the flow between 7:55 and 8:00. This

![Graph showing passenger accumulation by segment](image)

Figure 5-4. Passenger accumulation by segment (Victoria Line, southbound, 1 Nov 2013)
example shows passenger accumulation on the Victoria Line on a single day. The line is broken into four segments, a similar geographical unit to that shown in Figure 5-1 and Figure 5-2, although these segments are decidedly smaller than those used by Freemark. As explained in Section 5.2.1, passengers may be counted more than once at the segment level. Therefore, one must be cautious when directly comparing segments. The segments used here have roughly similar travel times during the peak (as taken from the schedule, see Table 5-1), and therefore, comparable rates of passengers who are counted twice. However, accumulation can still not be quantitatively compared from one segment to another. It is possible to make broad generalizations if two segments have similar travel times, but vastly different magnitudes of accumulation; however, if the accumulation values are close, it cannot be determined which is actually higher. The view of passenger accumulation presented in Figure 5-4 and Figure 5-5 is most useful for comparing accumulation on a single segment to itself at different times of day (or in the opposite direction).

**Table 5-1. Running time by segment**

<table>
<thead>
<tr>
<th>Segment</th>
<th>Num. Links</th>
<th>Running Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walthamstow (WAL-SVS)</td>
<td>3</td>
<td>7.25</td>
</tr>
<tr>
<td>Islington (KXX-SVS)</td>
<td>3</td>
<td>9.25</td>
</tr>
<tr>
<td>Central (KXX-VIC)</td>
<td>5</td>
<td>9.5</td>
</tr>
<tr>
<td>Brixton (BRX-VIC)</td>
<td>4</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Despite this limitation, there are several aspects worth noting when examining passenger accumulation at the segment level. Not unexpectedly, the Central segment has the highest overall accumulation southbound, and in the northbound direction during the PM peak (by nearly a factor of two; this is a large enough difference to conclude that the exact number of passengers on the Central segment is larger than the other segments at these times). The outer three segments also display directional peaking patterns. The two northern segments, Walthamstow and Islington, have higher southbound accumulation (of riders traveling into the city) in the morning and higher
northbound accumulation during the PM peak, while the Brixton segment displays the opposite trend.

On the Central segment in both directions, as well as the southbound Islington and northbound Brixton segments, it can be seen that the morning period has a sharper peak than the afternoon. In addition, the accumulation on the northbound Brixton segment during the morning peak appears comparable to that on the Central segment in the same direction, though one would expect the accumulation at the end of the line to be much lower than in Central London (especially given that the Central segment is slightly longer as shown in Table 5-1). The northbound accumulation in the AM peak is by far the greatest accumulation seen on the Brixton segment at any time in either direction.

As discussed in the preceding section, one advantage of using RODS to estimate passenger accumulation is the added spatial flexibility. Figure 5-6 shows the Central segment from Figure 5-4 broken down to the link level (two of the five links are omitted for clarity). Even within this small span of the line, directional patterns emerge. The northernmost link, King’s Cross-Euston, has high accumulation during the morning and low accumulation in the afternoon relative to the others. During the PM peak, the three links are neatly arranged from south to north in order of decreasing passenger accumulation. The relative magnitudes of accumulation on these links in the afternoon show the build-up of accumulation along the line as commuters are carried home out of Central London.

Figure 5-6. Passenger accumulation by link (Victoria Line, southbound, 1 Nov 2013)

In general, the segment or line level promises the most useful view of passenger accumulation. On long lines, such as the Central or Piccadilly, service can vary from one part of the line to another. Areas with satisfactory performance may dampen out problems, rendering them invisible. However, on shorter lines such as Victoria (which has just 16 stations to the District Line’s 60), the line level may offer a reasonable perspective. If the line level runs the risk of providing insufficient detail, the link level, by contrast, often provides too much. In order to be useful, real-time information must be
easy to absorb. Examining accumulation for every link, even on a line as short as Victoria, is more information than can be easily processed. However, the link level may add value as a supplementary view of the system. In particular, looking at specific links – such as those around the peak load point – could be useful, especially as the magnitudes of accumulation can be compared directly at the link level, unlike across segments. An ideal implementation of passenger accumulation would allow the user to drill down to the link level as necessary, perhaps to ascertain whether all links in the segment are performing equally, or whether unusual levels of passenger accumulation are unevenly distributed.

Passenger accumulation is particularly useful in the context of comparing days. Figure 5-7 shows the accumulation on the southbound Central segment (which had the highest accumulation out of all directional segments in Figure 5-4 and Figure 5-5) and compares it against two other days in the same week. While all three days have similar accumulation during the middle of the day and in the afternoon, they are quite different during the morning peak period. The middle of the week has the highest accumulation, as is generally true for demand across most modes of travel. (Though not shown here, the accumulation on Tuesday and Thursday of the same week fell between the Wednesday and Friday levels.) However, the difference between Monday and Friday is considerably larger than the difference between Wednesday and Friday, though one would expect the beginning and end of the week to have comparable numbers of passengers. At the peak of the peak, around 8:45, the difference between Oct. 28 and Nov. 1 is about 3000 passengers, or a 33% difference. This discrepancy suggests that Oct. 28 was an unusual day – perhaps it merely had very low demand, but the low accumulation could also be the result of supply side problems.\textsuperscript{14}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure5-7.png}
\caption{Passenger accumulation across days (Victoria Line, southbound, Central segment)}
\end{figure}

\textsuperscript{14} In fact, Oct. 28 had numerous delays and cancellations, so the low accumulation is almost certainly the result of these supply-side issues. However, this is only known by an ex-post examination of the recorded disruptions. Causality cannot be determined from passenger accumulation alone, either in real time or in retrospect.
This approach can be taken a step further by examining a single day against every other day, providing a systematic baseline for passenger accumulation. Figure 5-8 and Figure 5-9 display the distribution of passenger accumulation for 40 weekdays on the Victoria Line, in the northbound direction during the PM peak. The distribution is shown by percentiles. For example, at 18:30, an accumulation value of 14000 for the whole line would be between the 20th and 30th percentile. This historical data provides the backdrop against which real-time passenger accumulation can be used.

![Figure 5-8. Distribution of passenger accumulation (Victoria Line, northbound)](image1)

![Figure 5-9. Distribution of passenger accumulation (Victoria Line, northbound, Central segment)](image2)
A comparison of Figure 5-8 and Figure 5-9 reveal that the two distributions are quite similar. The segment-level distribution (Figure 5-9) has three distinct peaks at approximately 17:25, 17:50, and 18:20, which are present but less noticeable in the line-level distribution. (Due to the high frequency of arrivals, the lower frequency of arrivals on the Victoria Line, northbound, 31 Oct 2013)
of trains during the PM peak period, these crests cannot be explained by supply, but rather are interpreted as commuters consistently leaving work at these times over others.) Given the similarity between the two distributions, this thesis uses the smoother and more broadly applicable line-level distribution as the basis for examining passenger accumulation on the Victoria Line. Figure 5-10 displays passenger accumulation plotted on a single day with minor delays, October 31, 2013, against the distribution shown in Figure 5-8.

This figure shows how the day's accumulation progresses over the PM peak period. Each graph represents a snapshot of real-time passenger accumulation at a different point over the course of the afternoon. Passenger accumulation starts off on the low side during this period, staying below the 50th percentile until about 17:00. Accumulation stays in the low to middle range of the distribution until about 18:00, when it starts rising, reaching the 85th percentile at about 18:25. It then falls, returning to about the 10th percentile at 18:40, and staying low through the end of the period.

This day presents an interesting case. The beginning and end of the period suggest that, all else equal, Oct. 31 had relatively low demand. As it turns out, there was a train delay and subsequent withdrawal from service that began at 18:00, when the accumulation in Figure 5-10 starts to climb. This suggests that, due to the delay, passengers exited the system later than they otherwise would have, resulting in higher accumulation. It subsequently returned to its previously low levels as the service returned to normal, and the delayed passengers reached their destinations and exited the system.

While the rise at 18:00 is too slow to be valuable in real-time – by the time it becomes noticeable, the defective train has already reentered service – the tail end of this peak is a potentially useful source of information. Passenger accumulation suggests that the effect of the delay continues to impact passengers well after the trains have returned to the schedule. Accumulation doesn’t return to the median until after 18:30 (a second slight delay occurred at 18:28 that could have contributed to a slower recovery, although it was at Brixton, where northbound demand is particularly low during the afternoon peak period). Thus, this graph could be an additional source of information to controllers about the effectiveness of their strategies in terms of passengers, rather than in terms of trains.

5.3 Current TfL tools for indicating system state

TfL's existing real-time measures are based on data from the AVL system (see Section 3.3.2). AVL data can be used to evaluate any supply-side measures by comparing their actual and scheduled values. Any metric based on AVL data can potentially be obtained in real time.

A wealth of automatically calculated supply-side measures can be found on Heartbeat, TfL's internal web-based collection of analytics and performance reporting resources. Heartbeat provides numerous tools for examining supply and service performance retrospectively, including frequency graphs (a rolling count of the trains per hour), headways, dwell times, lateness (the difference between actual and scheduled departure time) and wait time plotted at select stations over the course of the day.
Figure 5-11 and Figure 5-12 are examples of two such graphs for Oct. 31. Figure 5-11 shows the number of trains that served Oxford Circus in each half hour window (for example, 15 trains passed from 18:00 to 18:30), the scheduled hourly frequency, and the actual hourly frequency using a centered rolling average. The frequency dips somewhat below the scheduled rate around 17:30, and recovers around 18:30. From 19:00 to 20:00, the frequency is slightly above schedule as the small backup clears, and trains catch up to their scheduled arrival and departure times. Figure 5-12 paints a similar picture. From 18:00 to 19:45, most trains are about 5 minutes late.

Figure 5-11. Heartbeat frequency graph (Victoria Line, northbound at Oxford Circus, 31 Oct 2013)

Figure 5-12. Heartbeat lateness graph (Victoria Line, northbound at Oxford Circus, 31 Oct 2013)

This frequency graph shows service problems earlier than they occur, due to the centered rolling average, but if this were a lagged rolling average – as would be used in real time – they would
appear at 18:00, the same time as they appear on the lateness chart. The frequency graph also
doesn’t indicate service problems from 19:00 to 20:00. This is because although each train left the
station a few minutes behind its scheduled departure time, the number of trains serving each station
was as scheduled. Thus, one could conclude that frequency is a better metric than lateness for
describing the impact on passengers of delayed trains. As long as trains are maintaining scheduled
headways, passengers aren’t affected by late-running trains – unless the preceding delays are
serious enough to cause crowding and denied boardings.

Figure 5-13 shows a waterfall diagram for the southbound Victoria Line over one hour. Each line
represents a single train trajectory (as it moves from the bottom to the top of the chart) and its
progress along the line (moving from left to right). Evenly spaced lines at consistent slopes indicate
good service, as can be seen from 17:30 to 17:55. The gap that appears around 18:00 signifies an
unusually long headway between trains, reaffirming the perturbation in service shown at the same
time in Figure 5-11 and Figure 5-12.

Figure 5-13. Heartbeat waterfall diagram (Victoria Line, northbound, 31 Oct 2013)

Analysis by TfL staff is not limited by the automatically generated Heartbeat charts; and is often
based on using NetMIS data directly. However, while any AVL-based supply or service performance
measures generally can be made available in real time, the data and analytics available through
Heartbeat are designed to be ex-post tools. Operations staff – the group to whom real-time data is
most valuable – use a limited set of information. This is partly to simplify their job, which requires
fast decision-making in response to rapidly changing situations.

In the control context, service is primarily measured using TfL’s Centralised Train Following System
(CTFS). CTFS is a real-time program developed by the London Underground that measures train
lateness by comparing actual and scheduled arrival times at select stations. Trains are color-coded
according to lateness; trains that are 6 (or fewer) minutes late are shown in green, trains that are 7
to 14 minutes late are yellow, and trains that are 15 (or more) minutes late are red. CTFS is the
primary information source used by controllers to determine what interventions (such as short-
turning, reforming, or cancelling trains) are necessary to maintain service quality.
The other key tool at service controllers’ disposal is TrackerNet, a depiction of the track and signal system that displays train locations in real time. TrackerNet is strictly a visual illustration of the system; it provides no analytical information. While controllers do not have access to real-time headway metrics, they can determine approximate headways by inspection from TrackerNet.

TrackerNet layers live trains on top of a representation of the track infrastructure which includes stations, signals (not shown above) reversing points, and sidings. Each train is color-coded by its destination and displays its train number (or destination if the system is unable to determine a train’s identifier).

A possible improvement to these two systems employed by controllers is to combine the information shown in CTFS with TrackerNet. Train lateness, the most important detail provided by CTFS, could be added to the web-based implementation of TrackerNet as an additional indicator appearing next to each train icon (or alternating with the train number). This minor change could simplify controllers’ intake of information, by providing the critical data in one view.

5.4 Evaluation of passenger accumulation as a real-time indicator

In this section, passenger accumulation is compared to existing tools (both real-time and ex-post). By analyzing, in retrospect, passenger accumulation as if in real time, and comparing it to what is known about system conditions after the fact, it is possible to assess the value it adds to TfL’s current tools. This evaluation is based on whether passenger accumulation can provide better information than that currently available according to three criteria: currency (real-time versus retrospective), accuracy, and the level of granularity at which it can describe the system.

Figure 5-15 is an example of passenger accumulation on a day with relatively high afternoon peak period demand and some service deficiencies. Up to 17:10, demand hovers in the middle of the distribution, ranging from the 26th to the 62nd percentile. It then climbs, rising above the 90th percentile at 17:45. Though the accumulation stays high the rest of the period, it peaks at 18:45 and then drops off slightly. Clearly, something caused demand to be so high on Nov. 5; from 18:30 on, this day has the highest accumulation of any in the 40-day sample. Additionally, the spike and then drop at 18:45 could indicate an unusual occurrence in the system. A surge in accumulation can be caused by an increase in demand, delays that prevent users from tapping out, or some combination thereof. If minor service problems are to blame, then accumulation should gradually return to its previous level as excess passengers trickle out of the system once good service is restored. In a more serious disruption, accumulation may drop below typical levels as passengers choose to take alternate modes or routes.
In the case of Nov. 5 (Guy Fawkes Day), several disruptions did in fact occur between 18:15 and 20:00 (see Figure 5-16). Minor delays arose at both Seven Sisters and Pimlico, but the more notable incident was severe crowding at Brixton Station. Demand was unusually high due to the traditional Guy Fawkes fireworks display, and a malfunctioning escalator slowed the rate at which passengers exited the station, exacerbating the crowding. As a result, station staff closed Brixton to entering passengers and the station functioned as exit-only until after 20:00.

These events are consistent with the narrative that can be inferred from passenger accumulation. The high accumulation from 17:45 on was likely due to people traveling to the fireworks display. The subsequent drop at 18:45 reflects the lack of entries at Brixton Station, and the fact that it remains high at the end of the period suggests that it took some time for the crowding at the station to disperse after the partial closure.

The increased demand from the fireworks show highlights a limitation of this approach. Special events (such as the festivities for Guy Fawkes Night in this example) that drive high demand in a particular direction or part of the network are a departure from the general demand patterns represented in RODS. This undermines the validity of using RODS data to infer customers' paths through the system.

Given that three of the observed disruptions happened on the southern end of the line, further information can be gleaned from passenger accumulation at the segment level. Although the
accumulation in Figure 5-17 has much more variation than at the line level (Figure 5-15), there are some similarities. Both views show that passenger accumulation stays within the distribution until 17:00 (although the range is wider at the segment level), and that accumulation then climbs above the 90th percentile. Both show accumulation peaking at 18:45 and then decreasing. However, the segment level shows accumulation as a series of increasingly sharp spikes from 17:45 to the end of the period, appearing much more volatile than the line level accumulation. The large dip at 18:35 is particularly worrisome. This drop begins at 18:25, about the time the station entry gates were closed, however, given the peak that soon follows, one cannot conclude that this drop is the result of the closure. In fact, the series of crests and troughs that appear at the Brixton segment level cannot be attributed to any cause with certainty.

Examining passenger accumulation side by side with some of the previously described metrics offers insight into the value accumulation can add to TfL's existing tools. Three supply-side measures derived from Netmis data – headways, frequency, and train lateness – are shown in Figure 5-18 through Figure 5-20. Each of these metrics is calculated for a single station (Oxford Circus). The Victoria Line is quite short relative to other London Underground lines, and does not have complicated service patterns (fewer trains run on the segment between Seven Sisters and Walthamstow Central, but all trains serve the rest of the line). Supply-side metrics tend to be fairly consistent along the length of the line. Consequently, using one station to represent the whole line is a reasonable approximation in this case.

Figure 5-18 shows the preceding headway of each departing train. Headways are reasonably close to their scheduled values until about 18:30, and then only one train has a high enough headway to indicate that service was very poor. Compared to passenger accumulation, headways give very little indication of the quality of service. The single outlying headway suggests that delays were brief and resolved quickly. Headways also give no information as to the number of passengers affected. Finally, it should be noted that headway metrics cannot account for events that have no bearing on trains, such as a partial station closure, but that have a considerable impact on passengers.
Frequency, as seen above, performs slightly better as a descriptor of system state. Figure 5-19 displays the number of trains that passed in the preceding half hour. A lagged frequency metric can be used in real time (unlike the centered rolling average in Figure 5-11), and limiting the window to only 30 minutes instead of an hour results in a measure that adapts more quickly to the current situation. Here, frequency gives a much better idea of the length of the delay than the headways; frequencies are low from 18:15 through the end of the period. The frequency graph indicates that the supply of trains starts deteriorating at 18:30, the same as the headways chart. However, this is four minutes after Brixton is closed to entries, and 12 minutes after the first recorded delays. Accumulation, in contrast, hints at problems earlier, but with less certainty. The spike at 18:15 does suggest problems, but is less straightforward to interpret, and doesn’t specify whether the cause is on the demand side or the supply side (or both).

Lateness, as shown in Figure 5-20, is less consistent with the other supply metrics. Trains begin to fall behind schedule at 17:35, though there is no clear reason from the disruptions identified in Figure 5-16. A possible explanation is that the high demand is lengthening dwell times, which would...
Figure 5-20. Lateness at Oxford Circus (Victoria Line, southbound, 5 Nov 2013)

be consistent with the line-level passenger accumulation values. While the jump at 18:25 is expected, lateness does not return to normal levels by the end of the period, unlike passenger accumulation and the other two measures.

Although not initially within the scope of this analysis, one possible improvement to the use of AVL data suggests itself from comparing Figures 5-18 through Figure 5-20. Out of the three AVL measures shown here, frequency is the most effective at communicating the severity and timing of delays. Lateness can overstate the magnitude of delays in cases where train arrivals are offset but headways are meeting scheduled durations. Meanwhile, charting headways doesn’t indicate how long the effects of a delay last, and can understate the impact if only one or two headways are high. Real-time frequency data, therefore, could prove a valuable addition to the control environment, as it has the potential to help controllers both assess a disruption as it evolves, and evaluate the recovery of the service.

Headways, frequency, lateness, and other AVL-based supply-side metrics can provide precise data on the location of trains and their performance to schedule, and are easy to interpret. However, they provide little indication of the extent to which passengers are affected. Consequently, problems may be exaggerated if trains are falling off schedule when demand is low, or if even headways are maintained. By contrast, if trains are quickly restored to schedule but high demand has caused crowding, supply-side metrics may underrepresent the severity of the situation.

Passenger accumulation can provide information about system state by describing when and where customers are most affected. High accumulation can foreshadow crowding in the system, and thus alert staff to a situation in which small problems may be exacerbated by unusually high demand. In addition, passenger accumulation can indicate how long poor service lasts from the passenger’s perspective. This information is potentially valuable to control staff as a real-time, passenger-centric source of feedback regarding control interventions.

However, it can be difficult to explain what causes fluctuations in accumulation based on this measure alone. Thus, the best use of accumulation is in conjunction with other measures. Though AVL data is available in real time, only the train lateness measure is currently used in real time.
At the segment level, passenger accumulation is significantly more variable, and proves to be difficult to interpret reliably. Examining accumulation on a small segment presents the risk of reading too much into variation that is nothing more than noise. In this regard, AVL-based metrics have an advantage. Because AVL data can be calculated for any station, supply-side metrics offer a greater level of spatial detail than accumulation. However, some AVL metrics can be difficult to aggregate, as normal service at one station can dampen the effects of poor service at another.

For longer Underground lines and those with complex service patterns, segments are likely a necessary means of breaking up the information into meaningful pieces. In addition, the segment level can act as a source of supplementary information when more detail is desired.

Finally, there is one key limitation to passenger accumulation. Because of the dependence upon RODS data to describe relative travel patterns, the implementation of passenger accumulation as described in this thesis is not valid in cases of severe disruptions such as partial line suspensions, or large events that change the “normal” directional patterns on any line. However, a future version of passenger accumulation that better accounts for passenger interchanges may overcome this obstacle. One possible means of achieving this improvement may be provided with the advent of data from mobile phones.

At this point, passenger accumulation cannot be interpreted with enough consistency for this thesis to definitively conclude that it provides value. Additional exploration is necessary to test the full potential of the accumulation metric.
This thesis explores the dimensions of transit system performance in the hope that a better understanding of these dimensions can lead to improved operational decisions that will benefit both transit agencies and their passengers. This chapter reviews the research conducted to this end. Section 6.1 summarizes the contents of the preceding chapters and the key findings of each. Recommendations to Transport for London are presented in 6.2. Finally, Section 6.3 suggests areas that could benefit from future research.

6.1 Research summary

Chapter 3 introduces the three dimensions of transit system performance on which this thesis centers. System performance is comprised of demand from passengers, the service supplied by the transit agency, and service performance. The characteristics of each of these dimensions are discussed, noting that supply includes both speed and capacity. Service performance is defined as the product of supply and demand, and is chiefly characterized by regularity, although speed can also be considered an attribute of performance. The primary relationships of interest are those of supply and demand on service performance, but supply and demand also influence one another.

The chapter then presents a number of measures that can be used to quantify each of these dimensions using automatic fare collection and automatic vehicle location data. Several of the specific performance metrics that TfL uses are discussed, and it is noted that TfL’s measures tend to focus on the supply dimension. In particular, the measures that are used in operational contexts focus almost exclusively on the impact of disruptions to service supplied. Therefore, there is potential for greater use of metrics that reflect the impacts of poor service upon passengers. In addition, this chapter covers TfL’s current baselining practices, which often involve the use of small, and therefore potentially unrepresentative, sample sizes.

Focusing on the topic of demand measurement, this thesis develops a methodology to calculate flow, or the passenger loads on a link between two adjacent stations. This approach calculates flow from gate (or smart card) entry and exit counts using data from the Rolling Origin Destination Survey to infer passenger routes. This method is then applied to the Victoria Line, and used to identify locations where demand is close to exceeding estimated capacity. The results show that during the busiest 15 minutes of the morning period, demand already exceeds capacity at the peak load point in the southbound direction. The number of line segments above capacity jumps to four southbound, and one northbound, with just one fewer train than scheduled. When two trains are removed from service, demand exceeds supply on five southbound and four northbound links.

Chapter 4 aims to provide insight into the relationships between the three dimensions by examining them concurrently. The factors that affect service performance are discussed and broken down into
supply-side and demand-side causes. The measures of each dimension introduced in Chapter 3 are revisited to investigate how each may help determine the possible causes of poor service performance. Flow, as outlined in the previous chapter, is identified as a key metric for defining line demand.

This chapter develops a framework to analyze the impacts of supply and demand on service performance. The spatial scope is defined as a segment of a line from the low-demand outer zones of the London Underground network to the high-demand areas of Central London. The peak 30 minutes of the peak is the recommended temporal scope. The framework introduces the concept of OD collections, or groups of OD pairs with similar origin and destination stations. The basic premise of the analysis is that comparing supply, demand, and service performance measures, normalized to expected levels and aggregated across collections, can provide an improved understanding of how these dimensions interact. Particular emphasis is placed on identifying conditions where demand exceeds capacity.

A small-scale analysis is then conducted using the eastern part of the Piccadilly line during the AM peak as an example. This analysis focuses on two specific days; one with numerous disruptions affecting service, and another with no disruptions, but high demand. The first case revealed that comparing flow with supply-side metrics can be used to identify conditions in which a lack of supply incentivizes passengers to change routes (or modes), resulting in a corresponding reduction in demand. The second case showed that actual capacity as calculated from supply-side metrics, together with demand measures, can suggest when and where denied boardings occurred.

Chapter 5 introduces the concept of passenger accumulation as a metric that can indicate system state in real time. First, a prior implementation of passenger accumulation at TfL is described, along with its limitations; namely that it omits passengers using interchange stations. The previous definition of accumulation is restricted by its inability to infer passengers’ direction of travel or determine whether a passenger has left the line via an interchange, its lack of spatial flexibility, and the small sample size used to determine normal accumulation levels.

This research then offers a new method to estimate passenger accumulation. The proposed approach builds on the methodology used to estimate flow in Chapter 3, but is modified to determine flow in real time. This method is applied to the Victoria Line, and the results are examined from different perspectives. It is concluded that examining passenger accumulation is not particularly useful at the segment level on the Victoria Line, as there is too much variation in the data for it to be interpreted definitively. However, segments may be useful for viewing passenger accumulation on other lines, as one segment on a longer line may be equivalent to the full extent of the Victoria Line.

This example is followed by a description of TfL’s existing tools, both retrospective and real-time, for indicating network conditions. Finally, the chapter compares passenger accumulation with TfL’s measures to assess whether accumulation can provide value as a real-time indicator of system state. Passenger accumulation has an advantage over AVL metrics in that it is customer-centric; however, it is difficult to clearly identify underlying causes to changes in accumulation. Furthermore, accumulation is unreliable when there are major disruptions (such as a partial line suspension) or special events, as these circumstances invalidate the underlying assumptions regarding interchange
activity and path choice. Thus, the chapter concludes that accumulation as it is currently proposed cannot be interpreted consistently enough to be valuable in a real-time context.

6.2 Recommendations

The objectives of this thesis were intended to be valuable specifically to the London Underground. This research identifies several opportunities for TfL to make improvements by applying the findings of this thesis and enhancing the use of their available data.

- **Use a larger sample size for baselines.** Many of TfL’s existing analyses compare transit system performance on an individual day to a single day in the past. This practice can distort the resulting picture of performance if the baseline day is not representative. This is particularly critical in the context of demand, which has no single value that can function as a standard. This thesis uses a sample of 40 weekdays, and compares demand and service performance on individual days to their respective distributions over the 40-day period.

- **Use distributions when more precise classifications of performance are needed.** Building upon the recommendation above, if large sample sizes are available, standards can be represented as distributions rather than a single target value. Distributions are flexible in that they allow the user to select the appropriate level of acceptability (such as the 75th or 95th percentile), or to define a range of values as the target. The analysis in this thesis reveals that many measures have wide distributions, and consequently, a range of values may be more realistic for defining standards. Though this approach is not practical for all types of analyses, it can add value when finer characterization of system performance is necessary and when a larger sample size is available.

- **Include frequency as a supply-side measure in operational contexts.** The analyses of supply-side measures in this thesis show that frequency better represents supply conditions than many other AVL metrics. It is sensitive to small changes in train operations, and captures the passenger perspective better than does train lateness. Lateness is currently TfL’s most commonly used measure of supply, yet passengers are not affected by whether or not a specific train arrives at its stipulated time, but rather whether the next train arrives as soon after the preceding one as is scheduled. Including frequency as a key decision factor may lead service controllers to implement service recovery strategies that better consider passenger impacts. This could be achieved by adding a frequency metric to either CTFS or the web-based version of TrackerNet.

- **Use journey time to evaluate service recovery.** Chapter 4 shows that journey times can be used to determine when passengers no longer feel the effects of a disruption. This is an improvement over current measures, such as train lateness and wait time, because using AFC data directly represents the passenger experience. Oyster journey times can be used to supplement these existing measures.

- **Jointly analyze AVL and AFC measures to identify denied boardings.** Two types of analyses are discussed in Chapter 4 for exploring capacity constraints. The framework suggests comparing running times with Oyster journey times, as denied boardings are indicated by
journey times that are notably higher than their corresponding running times, and that exhibit a wider distribution. A second method, applied in the example analysis, compares actual capacity (derived from frequency) to demand (measured as flow, the number of passengers on a link in a given time interval), and demonstrates that the two can be used together at the link level to identify conditions where capacity is exceeded.

- **Improve data quality and consistency.** TfL has a plethora of available data sources that far exceed those mentioned in this thesis. However, performing analysis that combines data from multiple sources can be hindered by inconsistencies between datasets. For example, the agency has numerous different fields to identify stations; including the station name, three-letter code (termed sutor code), station code, station number, and national location codes. Different data sources may use one or several of these. Furthermore, even sources that use the same field may be in conflict. For instance, a few stations have different national location codes in Oyster data than in RODS. While making these changes are relatively simple in theory, they may be difficult to implement due to the institutional barriers that exist in any large organization.

### 6.3 Future research

Several topics addressed in this thesis would benefit from further exploration. Additionally, some of the analysis conducted in this work raised other questions that were not included in the original scope, but that would be both interesting and potentially valuable to transit agencies to study. These areas are outlined below.

- **Explore alternate methods of inferring passenger route choice within the London Underground.** The methods of estimating both flow and the accumulation metric are notably limited by their dependence upon RODS to infer path choice. Using another source of passenger route information as an input to these processes may provide more accurate results. In the case of flow, in which destinations are already known, a possible approach is to use a discrete choice model to assign Oyster journeys to possible routes through the system. One example of this is the research by Zhan Guo (2008). TfL already has possible paths by OD pair and generalized journey times for each path from the Passenger Route Finder tool, but the results are not consistent with the limited route choice information provided by RODS, and the tool's underlying model is not well understood. Other alternate sources of passenger route choice may become available in the future, as technology makes possible new data streams such as cell phone location information.

- **Further study of the passenger accumulation metric.** The analysis in this thesis suggests that accumulation may have more value in evaluating service recovery from the passenger perspective than in anticipating changes to system state, which was the focus of this analysis. However, the greatest obstacle to using passenger accumulation is in identifying the underlying cause of changes in accumulation. Further work should be done on the process of identifying what factors most affect accumulation under various circumstances. The accumulation metric's value is dependent upon the ability to understand changes in accumulation consistently and with confidence.
Application and refinement of the framework to analyze the relationships between dimensions. This thesis introduces a framework that can be used to jointly analyze supply, demand, and service performance metrics, but applies only a limited version of the analysis. As this preliminary investigation does suggest that this type of analysis has merit, a full-scale application of the framework is a potential next step. This framework is designed to explore the effects of capacity-constrained conditions in particular, and the results could prove particularly valuable to a transit agency as this is a topic that is often not well understood. Given that this framework is a new construct introduced in this thesis, future work should also assess its usefulness and make revisions to its methods as necessary.

Examine the impacts of passenger notifications on travel behavior. This research discusses the interdependence of supply and demand, and suggests that jointly analyzing measures of both can identify situations in which one exerts a strong influence on the other. Passenger notifications are one possible context in which to further analyze this relationship. Examining records of when notifications were disseminated versus demand measures can improve the understanding of the extent to which passengers change paths or modes on their own initiative, rather than as a result of notifications. Using journey times could further support this analysis, by identifying when passengers took a shorter (or much longer) route through the network. A better understanding of the impacts of passenger notifications could either identify possible improvements to the current system, or reinforce the effectiveness of existing practices.

Investigate supply-side problems that lack a clear cause. Supply-side deficiencies can often be attributed to a particular disruption as recorded by operational staff. However, in some cases, a line may experience issues, yet no corresponding incident is recorded. Further analysis could reveal factors that lead to these types of problems, and inform the development of strategies to prevent or mitigate them.
## APPENDIX A: UNDERGROUND STATION CODES

<table>
<thead>
<tr>
<th>Sutor Code</th>
<th>Station Name</th>
<th>Sutor Code</th>
<th>Station Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT</td>
<td>Acton Town</td>
<td>CPS</td>
<td>Clapham South</td>
</tr>
<tr>
<td>AGR</td>
<td>Arnos Grove</td>
<td>CRD</td>
<td>Caledonian Road</td>
</tr>
<tr>
<td>ALD</td>
<td>Aldgate</td>
<td>CRX</td>
<td>Croxley</td>
</tr>
<tr>
<td>ALE</td>
<td>Aldgate East</td>
<td>CST</td>
<td>Cannon Street</td>
</tr>
<tr>
<td>ALP</td>
<td>Alperton</td>
<td>CTN</td>
<td>Camden Town</td>
</tr>
<tr>
<td>AME</td>
<td>Amersham</td>
<td>CWD</td>
<td>Chorleywood</td>
</tr>
<tr>
<td>ANG</td>
<td>Angel</td>
<td>CWF</td>
<td>Canary Wharf</td>
</tr>
<tr>
<td>ARC</td>
<td>Archway</td>
<td>CWR</td>
<td>Canada Water</td>
</tr>
<tr>
<td>ARL</td>
<td>Arsenal</td>
<td>CYL</td>
<td>Chancery Lane</td>
</tr>
<tr>
<td>BAL</td>
<td>Balham</td>
<td>DEB</td>
<td>Debden</td>
</tr>
<tr>
<td>BAR</td>
<td>Barbican</td>
<td>DGE</td>
<td>Dagenham East</td>
</tr>
<tr>
<td>BAY</td>
<td>Bayswater</td>
<td>DGH</td>
<td>Dagenham Heathway</td>
</tr>
<tr>
<td>BBB</td>
<td>Bromley-by-Bow</td>
<td>DHL</td>
<td>Dollis Hill</td>
</tr>
<tr>
<td>BCT</td>
<td>Barons Court</td>
<td>EAC</td>
<td>East Acton</td>
</tr>
<tr>
<td>BDE</td>
<td>Barkingside</td>
<td>EBY</td>
<td>Ealing Broadway</td>
</tr>
<tr>
<td>BDS</td>
<td>Bond Street</td>
<td>ECM</td>
<td>Ealing Common</td>
</tr>
<tr>
<td>BEC</td>
<td>Becontree</td>
<td>ECT</td>
<td>Earl's Court</td>
</tr>
<tr>
<td>BER</td>
<td>Bermondsey</td>
<td>EDG</td>
<td>Edgware</td>
</tr>
<tr>
<td>BGR</td>
<td>Bounds Green</td>
<td>EFY</td>
<td>East Finchley</td>
</tr>
<tr>
<td>BHL</td>
<td>Buckhurst Hill</td>
<td>EHM</td>
<td>East Ham</td>
</tr>
<tr>
<td>BHR</td>
<td>Blackhorse Road</td>
<td>ELE</td>
<td>Elephant &amp; Castle</td>
</tr>
<tr>
<td>BKG</td>
<td>Barking</td>
<td>EMB</td>
<td>Embankment</td>
</tr>
<tr>
<td>BLF</td>
<td>Blackfriars</td>
<td>EPK</td>
<td>Elm Park</td>
</tr>
<tr>
<td>BNG</td>
<td>Bethnal Green</td>
<td>EPP</td>
<td>Epping</td>
</tr>
<tr>
<td>BNK</td>
<td>Bank</td>
<td>ERY</td>
<td>East Putney</td>
</tr>
<tr>
<td>BOR</td>
<td>Borough</td>
<td>ERB</td>
<td>Edgware Road (B)</td>
</tr>
<tr>
<td>BOS</td>
<td>Boston Manor</td>
<td>ERD</td>
<td>Edgware Road (C/D/H)</td>
</tr>
<tr>
<td>BPK</td>
<td>Belsize Park</td>
<td>ESQ</td>
<td>Euston Square</td>
</tr>
<tr>
<td>BRX</td>
<td>Brixton</td>
<td>ETE</td>
<td>Eastcote</td>
</tr>
<tr>
<td>BST</td>
<td>Baker Street</td>
<td>EUS</td>
<td>Euston</td>
</tr>
<tr>
<td>BTX</td>
<td>Brent Cross</td>
<td>FAR</td>
<td>Farringdon</td>
</tr>
<tr>
<td>BUR</td>
<td>Burnt Oak</td>
<td>FBY</td>
<td>Fulham Broadway</td>
</tr>
<tr>
<td>BWR</td>
<td>Bow Road</td>
<td>FLP</td>
<td>Fairlop</td>
</tr>
<tr>
<td>CFS</td>
<td>Cockfosters</td>
<td>FPK</td>
<td>Finsbury Park</td>
</tr>
<tr>
<td>CHF</td>
<td>Chalk Farm</td>
<td>FRD</td>
<td>Finchley Road</td>
</tr>
<tr>
<td>CHG</td>
<td>Chigwell</td>
<td>FYC</td>
<td>Finchley Central</td>
</tr>
<tr>
<td>CHM</td>
<td>Chesham</td>
<td>GFD</td>
<td>Greenford</td>
</tr>
<tr>
<td>CHP</td>
<td>Chiswick Park</td>
<td>GGR</td>
<td>Golders Green</td>
</tr>
<tr>
<td>CHX</td>
<td>Charing Cross</td>
<td>GHL</td>
<td>Gants Hill</td>
</tr>
<tr>
<td>CLF</td>
<td>Chalfont &amp; Latimer</td>
<td>GHR</td>
<td>Goldhawk Road</td>
</tr>
<tr>
<td>CLW</td>
<td>Colliers Wood</td>
<td>GPK</td>
<td>Green Park</td>
</tr>
<tr>
<td>CNT</td>
<td>Canning Town</td>
<td>GPS</td>
<td>Great Portland Street</td>
</tr>
<tr>
<td>COL</td>
<td>Colindale</td>
<td>GRD</td>
<td>Gloucester Road</td>
</tr>
<tr>
<td>COV</td>
<td>Covent Garden</td>
<td>GRH</td>
<td>Grange Hill</td>
</tr>
<tr>
<td>CPC</td>
<td>Clapham Common</td>
<td>GST</td>
<td>Goodege Street</td>
</tr>
<tr>
<td>CPK</td>
<td>Canons Park</td>
<td>GUN</td>
<td>Gunnersbury</td>
</tr>
<tr>
<td>CPN</td>
<td>Clapham North</td>
<td>HAI</td>
<td>Hainault</td>
</tr>
<tr>
<td>Sutor Code</td>
<td>Station Name</td>
<td>Sutor Code</td>
<td>Station Name</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------</td>
<td>------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>HAR</td>
<td>Harlesden</td>
<td>MGT</td>
<td>Moorgate</td>
</tr>
<tr>
<td>HAW</td>
<td>Harrow &amp; Wealdstone</td>
<td>MHE</td>
<td>Mill Hill East</td>
</tr>
<tr>
<td>HBT</td>
<td>High Barnet</td>
<td>MLE</td>
<td>Mile End</td>
</tr>
<tr>
<td>HBY</td>
<td>Highbury &amp; Islington</td>
<td>MNR</td>
<td>Manor House</td>
</tr>
<tr>
<td>HCH</td>
<td>Hornchurch</td>
<td>MON</td>
<td>Monument</td>
</tr>
<tr>
<td>HDN</td>
<td>Hillingdon</td>
<td>MOR</td>
<td>Morden</td>
</tr>
<tr>
<td>HIG</td>
<td>Highgate</td>
<td>MPK</td>
<td>Moor Park</td>
</tr>
<tr>
<td>HLN</td>
<td>Hanger Lane</td>
<td>MYB</td>
<td>Marylebone</td>
</tr>
<tr>
<td>HMD</td>
<td>Hammersmith (D/P)</td>
<td>NAC</td>
<td>North Acton</td>
</tr>
<tr>
<td>HMP</td>
<td>Hampstead</td>
<td>NEA</td>
<td>Neasden</td>
</tr>
<tr>
<td>HMS</td>
<td>Hammersmith (C/H)</td>
<td>NEL</td>
<td>North Ealing</td>
</tr>
<tr>
<td>HNC</td>
<td>Hounslow Central</td>
<td>NEP</td>
<td>Newbury Park</td>
</tr>
<tr>
<td>HND</td>
<td>Hendon Central</td>
<td>NFD</td>
<td>Northfields</td>
</tr>
<tr>
<td>HNE</td>
<td>Hounslow East</td>
<td>NHG</td>
<td>Notting Hill Gate</td>
</tr>
<tr>
<td>HNW</td>
<td>Hounslow West</td>
<td>NHR</td>
<td>North Harrow</td>
</tr>
<tr>
<td>HOH</td>
<td>Harrow-on-the-Hill</td>
<td>NHT</td>
<td>Northolt</td>
</tr>
<tr>
<td>HOL</td>
<td>Holborn</td>
<td>NOG</td>
<td>North Greenwich</td>
</tr>
<tr>
<td>HPC</td>
<td>Hyde Park Corner</td>
<td>NWB</td>
<td>North Wemble</td>
</tr>
<tr>
<td>HPK</td>
<td>Holland Park</td>
<td>NWD</td>
<td>Northwood</td>
</tr>
<tr>
<td>HRC</td>
<td>Heathrow Terminals 123</td>
<td>NWH</td>
<td>Northwood Hills</td>
</tr>
<tr>
<td>HRD</td>
<td>Holloway Road</td>
<td>NWP</td>
<td>Northwick Park</td>
</tr>
<tr>
<td>HRV</td>
<td>Heathrow Terminal 5</td>
<td>OAK</td>
<td>Oakwood</td>
</tr>
<tr>
<td>HST</td>
<td>High Street Kensington</td>
<td>OLD</td>
<td>Old Street</td>
</tr>
<tr>
<td>HTF</td>
<td>Heathrow Terminal 4</td>
<td>OLY</td>
<td>Kensington (Olympia)</td>
</tr>
<tr>
<td>HTX</td>
<td>Hatton Cross</td>
<td>OST</td>
<td>Osterley</td>
</tr>
<tr>
<td>ICK</td>
<td>Ickenham</td>
<td>OVL</td>
<td>Oval</td>
</tr>
<tr>
<td>KBY</td>
<td>Kingsbury</td>
<td>OXC</td>
<td>Oxford Circus</td>
</tr>
<tr>
<td>KEN</td>
<td>Kennington</td>
<td>PAD</td>
<td>Paddington</td>
</tr>
<tr>
<td>KET</td>
<td>Kenton</td>
<td>PER</td>
<td>Perivale</td>
</tr>
<tr>
<td>KEW</td>
<td>Kew Gardens</td>
<td>PGR</td>
<td>Parsons Green</td>
</tr>
<tr>
<td>KGN</td>
<td>Kensal Green</td>
<td>PIC</td>
<td>Piccadilly Circus</td>
</tr>
<tr>
<td>KIL</td>
<td>Kilburn</td>
<td>PIM</td>
<td>Pimlico</td>
</tr>
<tr>
<td>KNB</td>
<td>Knightsbridge</td>
<td>PIN</td>
<td>Pinner</td>
</tr>
<tr>
<td>KPK</td>
<td>Kilburn Park</td>
<td>PLW</td>
<td>Plaistow</td>
</tr>
<tr>
<td>KTN</td>
<td>Kentish Town</td>
<td>PRD</td>
<td>Preston Road</td>
</tr>
<tr>
<td>KXX</td>
<td>King's Cross St. Pancras</td>
<td>PRY</td>
<td>Park Royal</td>
</tr>
<tr>
<td>LAM</td>
<td>Lambeth North</td>
<td>PUT</td>
<td>Putney Bridge</td>
</tr>
<tr>
<td>LAN</td>
<td>Lancaster Gate</td>
<td>QBY</td>
<td>Queensbury</td>
</tr>
<tr>
<td>LAT</td>
<td>Latimer Road</td>
<td>QPK</td>
<td>Queen's Park</td>
</tr>
<tr>
<td>LEY</td>
<td>Leyton</td>
<td>QWY</td>
<td>Queensway</td>
</tr>
<tr>
<td>LGR</td>
<td>Ladbroke Grove</td>
<td>RAY</td>
<td>Rayners Lane</td>
</tr>
<tr>
<td>LON</td>
<td>London Bridge</td>
<td>RED</td>
<td>Redbridge</td>
</tr>
<tr>
<td>LSQ</td>
<td>Leicester Square</td>
<td>RKY</td>
<td>Rickmansworth</td>
</tr>
<tr>
<td>LST</td>
<td>Liverpool Street</td>
<td>RMD</td>
<td>Richmond</td>
</tr>
<tr>
<td>LTN</td>
<td>Loughton</td>
<td>ROD</td>
<td>Roding Valley</td>
</tr>
<tr>
<td>LYS</td>
<td>Leytonstone</td>
<td>ROY</td>
<td>Royal Oak</td>
</tr>
<tr>
<td>MAN</td>
<td>Mansion House</td>
<td>RPK</td>
<td>Ravenscourt Park</td>
</tr>
<tr>
<td>MAR</td>
<td>Marble Arch</td>
<td>RPK</td>
<td>Regent's Park</td>
</tr>
<tr>
<td>MCR</td>
<td>Mornington Crescent</td>
<td>RSQ</td>
<td>Russell Square</td>
</tr>
<tr>
<td>MDV</td>
<td>Maida Vale</td>
<td>RUG</td>
<td>Ruislip Gardens</td>
</tr>
<tr>
<td>Sutor Code</td>
<td>Station Name</td>
<td>Sutor Code</td>
<td>Station Name</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------</td>
<td>------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>RUI</td>
<td>Ruislip</td>
<td>TPL</td>
<td>Turnpike Lane</td>
</tr>
<tr>
<td>RUM</td>
<td>Ruislip Manor</td>
<td>TTH</td>
<td>Tottenham Hale</td>
</tr>
<tr>
<td>SBC</td>
<td>Shepherd's Bush</td>
<td>UPB</td>
<td>Upminster Bridge</td>
</tr>
<tr>
<td>SBM</td>
<td>Shepherd's Bush Market</td>
<td>UPK</td>
<td>Upton Park</td>
</tr>
<tr>
<td>SEL</td>
<td>South Ealing</td>
<td>UPM</td>
<td>Upminster</td>
</tr>
<tr>
<td>SFD</td>
<td>Stratford</td>
<td>UPY</td>
<td>Upney</td>
</tr>
<tr>
<td>SFS</td>
<td>Southfields</td>
<td>UXB</td>
<td>Uxbridge</td>
</tr>
<tr>
<td>SGT</td>
<td>Southgate</td>
<td>VIC</td>
<td>Victoria</td>
</tr>
<tr>
<td>SHL</td>
<td>Sudbury Hill</td>
<td>VUX</td>
<td>Vauxhall</td>
</tr>
<tr>
<td>SHR</td>
<td>South Harrow</td>
<td>WAC</td>
<td>West Acton</td>
</tr>
<tr>
<td>SJP</td>
<td>St. James's Park</td>
<td>WAL</td>
<td>Walthamstow Central</td>
</tr>
<tr>
<td>SJW</td>
<td>St. John's Wood</td>
<td>WAN</td>
<td>Wanshead</td>
</tr>
<tr>
<td>SKN</td>
<td>South Kensington</td>
<td>WAR</td>
<td>Warwick Avenue</td>
</tr>
<tr>
<td>SKT</td>
<td>South Kenton</td>
<td>WAT</td>
<td>Watford</td>
</tr>
<tr>
<td>SNB</td>
<td>Snaresbrook</td>
<td>WBP</td>
<td>Westbourne Park</td>
</tr>
<tr>
<td>SPK</td>
<td>Stonebridge Park</td>
<td>WBT</td>
<td>West Brompton</td>
</tr>
<tr>
<td>SRP</td>
<td>South Ruislip</td>
<td>WCL</td>
<td>Whitechapel</td>
</tr>
<tr>
<td>SSQ</td>
<td>Sloane Square</td>
<td>WCT</td>
<td>White City</td>
</tr>
<tr>
<td>STA</td>
<td>Stanmore</td>
<td>WDL</td>
<td>Wood Lane</td>
</tr>
<tr>
<td>STB</td>
<td>Stamford Brook</td>
<td>WEM</td>
<td>Wemble Central</td>
</tr>
<tr>
<td>STG</td>
<td>Stepney Green</td>
<td>WFD</td>
<td>Woodford</td>
</tr>
<tr>
<td>STK</td>
<td>Stockwell</td>
<td>WFG</td>
<td>West Finchley</td>
</tr>
<tr>
<td>STN</td>
<td>Sudbury Town</td>
<td>WGN</td>
<td>Wood Green</td>
</tr>
<tr>
<td>STP</td>
<td>St. Paul's</td>
<td>WHD</td>
<td>West Hampstead</td>
</tr>
<tr>
<td>SVS</td>
<td>Seven Sisters</td>
<td>WHM</td>
<td>West Ham</td>
</tr>
<tr>
<td>SWC</td>
<td>Swiss Cottage</td>
<td>WHR</td>
<td>West Harrow</td>
</tr>
<tr>
<td>SWF</td>
<td>South Woodford</td>
<td>WIM</td>
<td>Wimbledon</td>
</tr>
<tr>
<td>SWK</td>
<td>Southwark</td>
<td>WJN</td>
<td>Willesden Junction</td>
</tr>
<tr>
<td>SWM</td>
<td>South Wimbledon</td>
<td>WKN</td>
<td>West Kensington</td>
</tr>
<tr>
<td>TBE</td>
<td>Tooting Bec</td>
<td>WLG</td>
<td>Willesden Green</td>
</tr>
<tr>
<td>TBY</td>
<td>Tooting Broadway</td>
<td>WLO</td>
<td>Waterloo</td>
</tr>
<tr>
<td>TCR</td>
<td>Tottenham Court Road</td>
<td>WMP</td>
<td>Wimbledon Park</td>
</tr>
<tr>
<td>TEM</td>
<td>Temple</td>
<td>WMS</td>
<td>Westminster</td>
</tr>
<tr>
<td>TGR</td>
<td>Turnham Green</td>
<td>WPK</td>
<td>Wembley Park</td>
</tr>
<tr>
<td>THB</td>
<td>Theydon Bois</td>
<td>WRP</td>
<td>West Ruislip</td>
</tr>
<tr>
<td>THL</td>
<td>Tower Hill</td>
<td>WSP</td>
<td>Woodside Park</td>
</tr>
<tr>
<td>TOT</td>
<td>Totteridge &amp; Whetstone</td>
<td>WST</td>
<td>Warren Street</td>
</tr>
<tr>
<td>TPK</td>
<td>Tufnell Park</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B: TERMS USED IN THIS THESIS

**Link**
A section of line between two adjacent stations.

**Segment**
A section of a line between any two stations. A segment is made up of multiple links.

**Service performance**
A specific dimension of system performance. Service performance is a function of supply and demand. Elsewhere may also be referred to as service quality.

**System performance**
A general term referring to how a transit system functions as a whole. Not a synonym for service performance.

**Time period**
TfL defines 13 time periods per week for scheduling and analysis. There are six weekday periods (Early, AM peak, Interpeak, PM peak, Evening, and Late) and seven weekend periods. TfL also refers to these internally as timebands.

**Operational period**
TfL divides the year into periods of four weeks for operational and analysis purposes, rather than using months. The first of the 13 operational periods falls in April. Many performance measures are assessed at the operational period level.
APPENDIX C: TFL PERFORMANCE METRICS

Platform Wait Time (PWT)
TFL uses the specific term PWT to refer to wait time on rail services. PWT is calculated using the standard formula shown in Equation 2.1.

Excess Platform Wait Time (Excess PWT)
Excess PWT is defined as the difference between actual and scheduled platform wait time. Actual platform wait time is calculated as shown above. More information on PWT and Excess PWT is available in Tfl’s Platform Waiting Time Guide (Transport for London, 2010).

Journey Time Metric (JTM)
JTM is a composite performance metric that combines five components: ticket purchase time; access, egress, and interchange time; PWT; in-vehicle time; and closures. The closures category incorporates both planned and unplanned closures, defined as disruptions that last more than 30 minutes plus the scheduled headway. Each element of journey time is weighted based on passenger perception, and measured using a combination of sources such as train movement data (NetMIS), models, disruption data (CuPID), and surveys.

Excess Journey Time (EJT)
EJT is defined as the difference between the actual JTM, and the “ideal” journey time. Each element of journey time has an associated “perfect” or scheduled time, which are then combined to produce the full “ideal” journey time. More information is available from Tfl (Transport for London, 2012b).

Reliability Index
The Reliability Index is a combination of two measures: Train Service Reliability and Train Service Reliability Stations. Train reliability is determined by measuring operators’ response times against the latest of the following criteria: signal clear time, minimum dwell or layover, or scheduled departure time. Station reliability measures dwell time, as wheel stop to wheel start, against established target values at specific stations (with exceptions for signal impedance as necessary). The composite measure is calculated as the percentage of days that the daily average meets the target score.

Lost Customer Hours
Lost Customer Hours is a measure of passenger inconvenience that is used to estimate the impacts of a disruption. Rather than a direct measure of travel time; LCH represents a generalized cost of a disruption. It includes not only increased journey time, but the additional cost imposed by conditions such as crowding levels and taking the stairs instead of an escalator. LCH accounts for both the direct and knock-on effects of a disruption. LCH is calculated using data from CuPID, NetMIS, and a number of models.
APPENDIX D: PICCADILLY SEGMENTS (FREEMARK)

Non-interchange stations used for passenger accumulation by Freemark

<table>
<thead>
<tr>
<th>Heathrow</th>
<th>Rayners Lane</th>
<th>Trunk</th>
<th>North</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heathrow T 5</td>
<td>South Harrow</td>
<td>Knightsbridge</td>
<td>Caledonian Road</td>
</tr>
<tr>
<td>Heathrow T 4</td>
<td>Sudbury Hill</td>
<td>Hyde Park Corner</td>
<td>Holloway Road</td>
</tr>
<tr>
<td>Heathrow T 1,2,3</td>
<td>Sudbury Town</td>
<td>Covent Garden</td>
<td>Arsenal</td>
</tr>
<tr>
<td>Hatton Cross</td>
<td>Alperton</td>
<td>Russell Square</td>
<td>Manor House</td>
</tr>
<tr>
<td>Hounslow West</td>
<td>Park Royal</td>
<td></td>
<td>Turnpike Lane</td>
</tr>
<tr>
<td>Hounslow Central</td>
<td>North Ealing</td>
<td></td>
<td>Wood Green</td>
</tr>
<tr>
<td>Hounslow East</td>
<td></td>
<td></td>
<td>Bounds Green</td>
</tr>
<tr>
<td>Osterley</td>
<td></td>
<td></td>
<td>Arnos Grove</td>
</tr>
<tr>
<td>Boston Manor</td>
<td></td>
<td></td>
<td>Southgate</td>
</tr>
<tr>
<td>Northfields</td>
<td></td>
<td></td>
<td>Oakwood</td>
</tr>
<tr>
<td>South Ealing</td>
<td></td>
<td></td>
<td>Cockfosters</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


