Quantifying Passenger Impact of Disruptions on Metro Lines

by

Mark Perelmuter

Bachelor of Engineering in Civil Engineering
The Cooper Union for the Advancement of Science and Art (2018)

Submitted to the Department of Urban Studies and Planning
in partial fulfillment of the requirements for the degree of

Master of Science in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2020

© Massachusetts Institute of Technology 2020. All rights reserved.
Quantifying Passenger Impact of Disruptions on Metro Lines

by

Mark Perelmuter

Submitted to the Department of Urban Studies and Planning on May 20, 2020
in partial fulfillment of the requirements for the degree of Master of Science in Transportation

Abstract

Disruptions occur frequently in urban rail transit systems. Whether due to asset failure, passenger action, weather, or other causes, disruptions often force passengers to change their preferred route or mode, defer their travel to a later time, or avoid making the trip altogether. Researchers and transit network operators have devoted significant time to understanding how passengers respond to disruptions, and modeling the impact that these responses have on the network state. They have also explored many avenues for mitigating the impact of disruptions once they occur.

The goal of this research is to understand the effect that infrastructure investments, specifically in track layout, have on mitigating the impact of major disruptions, and to model the effect of these disruptions by developing a simplified passenger assignment model, which aims to accurately represent the impact of a major disruption, such as a partial line suspension, on a transit network while having a sufficiently short computation time to be useful for sketch planning and similar first-order alternatives analysis. Both parts of the work are applied to the London Underground, specifically the Piccadilly line, as a case study.

The track layout analysis framework outlines a method to determine optimal locations for new track crossovers and compute the benefit that they have on reducing the impacts of unplanned partial line suspensions, or planned closures, by allowing trains to operate over a greater portion of the affected line. This benefit is then used as an input to a business case, which finds that, for the categories of benefit considered, investment in track layout enhancements on the Piccadilly Line is not justified.

The simplified assignment model strikes a balance between accurately representing the behavior of passengers during unplanned disruptions, specifically the difference between expected and experienced network state that is characteristic of these disruptions, and keeping the scope of the model sufficiently small to ensure quick computation time. Despite the model’s simplifications, its results show promise in accurately representing the state of the network during a disruption, and identifying the most overcrowded links on a network, while having a computation time shorter than that of existing models.

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

Thesis Supervisor: Haris N. Koutsopoulos
Title: Professor of Civil and Environmental Engineering, Northeastern University
Acknowledgements

This thesis would not have been possible without the mentorship, advice, and support of my advisors, Nigel Wilson and Haris Koutsopoulos. I am grateful to them for their guidance throughout this process, which has definitely made me a better student and researcher. Others at MIT, among them Jinhua Zhao, Fred Salvucci, and Mei-Chun Yiu, were also important in helping this work come to fruition.

I would like to thank Transport for London for sponsoring this research, and the many people there with whom I worked. Kelvin Blackie, Menno Yap, Liam McGrath, Chris Baitup, Sandra Weddell, Noori Sharma, Chris Locke, and others welcomed me to TfL, offered valuable insight, answered my many questions in great detail, and taught me a lot about the operations of the Underground. My experience would have been significantly diminished without their assistance and advice.

I cannot forget to thank my parents, who have offered me love and support throughout this entire process. Special thanks to my friends at the Transit Lab, at Sidney-Pacific, and at MIT as a whole have made these two years in Boston outstanding and unforgettable. Although our time together ended somewhat abruptly, I will always look back upon it fondly.
# Table of Contents

1 Introduction................................................................................................................................................. 8  
1.1 Motivation.................................................................................................................................................. 8  
1.2 Research Goals.......................................................................................................................................... 10  
1.3 Research Framework.................................................................................................................................. 11  
1.4 London Underground.................................................................................................................................. 14  
1.4.1 The Piccadilly Line.................................................................................................................................. 16  
1.4.2 Overview of Data Sources...................................................................................................................... 18  
1.4.3 Disruptions on the Underground.......................................................................................................... 18  
1.5 Thesis Organization.................................................................................................................................... 20  

2 Disruption Modeling in the Literature and in Practice.................................................................................... 22  
2.1 Literature Review.......................................................................................................................................... 22  
2.1.1 Schedule vs. Frequency Based Modeling .............................................................................................. 22  
2.1.2 Graph Theory Approaches.................................................................................................................... 24  
2.1.3 Static Assignment Models..................................................................................................................... 25  
2.1.4 Dynamic Assignment Models............................................................................................................... 26  
2.1.5 Simulation............................................................................................................................................... 27  
2.2 Current Practice at London Underground.................................................................................................. 27  
2.2.1 Underground Ridership data: RODS, Oyster, and Wifi.......................................................................... 28  
2.2.2 Journey Time Metric.................................................................................................................................. 29  
2.2.3 Pre-modeling (LCH) Approach.............................................................................................................. 30  
2.2.4 Retrospective (EJT) Approach................................................................................................................ 38  

3 Track Layout Analysis....................................................................................................................................... 40  
3.1 Literature Review.......................................................................................................................................... 40  
3.2 Current Practice............................................................................................................................................ 43  
3.3 Analysis Procedure and Results.................................................................................................................. 45  
3.3.1 Assessment of Existing Track Layout....................................................................................................... 45  
3.3.2 Terminal Capacity Analysis...................................................................................................................... 49  
3.3.3 Selection of Potential Crossover Locations............................................................................................. 55
List of Figures

Figure 1-1. Tube map.......................................................................................................................... 15
Figure 1-2. The Piccadilly line – layout, service patterns, and infrastructure .............................. 17
Figure 1-3. The Piccadilly line – schematic....................................................................................... 17
Figure 1-4. Incident occurrences and impacts by length of initial delay, 2010-2019 ...................... 19
Figure 1-5. Delays on the Piccadilly Line, 2018 ............................................................................... 20
Figure 2-1. Railplan network hierarchy ............................................................................................. 33
Figure 2-2. Railplan scenario comparison .......................................................................................... 35
Figure 2-3. Northern line diagram ..................................................................................................... 37
Figure 3-1. Example terminal configuration ....................................................................................... 41
Figure 3-2. Piccadilly line track layout ............................................................................................... 46
Figure 3-3. Frequency of unscheduled short turns on the Piccadilly line ...................................... 48
Figure 3-4. Geometric conflicts at relay (left) and stub-end (right) terminals ............................ 49
Figure 3-5. Waterfall diagram illustrating blocking back ................................................................. 50
Figure 3-6. The eastern section of the Piccadilly line .................................................................... 51
Figure 3-7. Average run time: Piccadilly Line eastbound (by section) .......................................... 52
Figure 3-8. Distribution of eastbound dwell times at Arnos Grove ................................................. 53
Figure 3-9. Example of facing and trailing point crossovers ............................................................ 56
Figure 3-10. Proposed track layouts ................................................................................................ 57
Figure 3-11. Proposed Crossrail 2 route ............................................................................................ 61
Figure 3-12. LCH by year and time of day for Covent Garden crossover ........................................ 62
Figure 3-13. LCH by year and time of day for Turnham Green crossover .......................................... 62
Figure 3-14. LCH by year and time of day for Alperton crossover ................................................... 63
Figure 3-15. Percent reduction in LCH by year and time of day for Covent Garden crossover .. 63
Figure 3-16. Percent reduction in LCH by year and time of day for Turnham Green crossover . 63
Figure 3-17. Percent reduction in LCH by year and time of day for Alperton crossover ............ 64
Figure 3-18. Percent reduction in LCH resulting from Crossrail 2 ................................................. 67
Figure 4-1. Simplified assignment procedure ................................................................................. 77
Figure 4-2. Network representation as a graph ............................................................................... 79
Figure 4-3. Railplan demand profile ................................................................................................. 87
Figure 4-4. Example Network for Route Choice .............................................................................. 89
List of Tables

Table 2-1. Journey Time Metric weighting factors ................................................................. 30
Table 2-2. Trips per hour modeled in Railplan ........................................................................ 32
Table 2-3. Assumed changes to the network between 2021 and 2041 ........................................ 33
Table 3-1. Disruption Service Patterns ..................................................................................... 60
Table 3-2. Top causes of relevant incidents ............................................................................ 70
Table 3-3. Benefit of crossover installation for rush hour unplanned disruptions ....................... 70
Table 3-4. Benefit of crossover installation for planned weekend closures ............................... 72
Table 3-5. Aggregate benefit of crossover ............................................................................... 72
Table 4-1. Penalties associated with walk link elements .......................................................... 81
Table 4-2. Service status guidelines ......................................................................................... 91
Table 4-3. Total passenger hours by category, simplified model ............................................... 96
Table 4-4. Total passenger hours for selected categories, base Railplan approach .................... 97
Table 4-5. Comparison of flows on selected critical links ......................................................... 99
Table 4-6. Overflow link usage inbound .................................................................................... 102
Table 4-7. Line by line OTT for selected lines, in passenger hours (unweighted) ....................... 103
Table 4-8. Line by line PWT for selected lines, in passenger hours (unweighted) ..................... 104
Table 4-9. Relevant links with most denied boardings ............................................................. 105
Introduction

The primary focus of this thesis is the development of a simplified transit passenger assignment model to represent network performance in the event of unplanned disruptions on the London Underground (LU) in an accurate, yet efficient way. This involves consideration of two issues critical to any assignment model: passenger behavior and model scope. This model is then validated and compared with the existing modeling methodology used by LU staff. In addition, a chapter of this thesis is dedicated to analyzing the degree to which track layout modifications can mitigate the passenger impact of disruptions on a metro line, in particular the Piccadilly line of the Underground. It was this work, and the modeling limitations encountered therein, that inspired the development of the simplified assignment model.

1.1 Motivation

The London Underground, which opened in 1863 and today has 270 stations and 250 miles of track, is the world’s twelfth busiest metro system with a daily ridership of approximately five million. Like the metro systems of most metropoles in the Western world, the Underground was integral to London’s development throughout the twentieth century. Unsurprisingly for a system with such a long lifespan, its assets require periodic maintenance, and incidents occur frequently. These incidents, whether caused by infrastructure failure, passenger behavior, or other reasons, cause disruptions that inconvenience and delay passengers. (The remainder of this thesis will use the words incident and disruption as defined by Freemark (2013): an incident is the initial event that causes train operation to be affected, and a disruption is the cumulative effect of the incident on train service). Incidents take a wide range of forms, from a minor door problem on a train causing it to be delayed by a few seconds to a power failure which may require a line suspension for several hours. Broadly speaking, disruptions can be subdivided into three categories by length of initial delay. “Minor” disruptions, with an initial delay duration of ten minutes or less, “moderate” disruptions, with an initial delay between ten and thirty minutes long, and “major” disruptions, with an initial delay of more than thirty minutes. The primary distinction between these three categories is the preferred type of response by Underground service controllers and, consequently, the behavior of passengers in the network.

According to CuPID, LU’s internal incident tracking system (Transport for London, 2020f), the Underground experienced 30,791 incidents\(^1\) in the 2018-19 fiscal year, resulting in 30,077,159 Lost Customer Hours (LCH, a metric discussed in detail in Section 2.2.3) for passengers. 89% of these incidents, causing 76% of the total LCH, were minor disruptions.

---

\(^1\) Excluding industrial action
Moderate disruptions accounted for 9% of incidents and 14% of the LCH, and major disruptions were 2% of incidents and 9% of the LCH.

Passenger perception of journey delay varies by disruption magnitude. In particular, it is non-linear with respect to delay duration (Bates et al, 2001): passengers generally perceive a lengthier delay to be more detrimental than two delays, each of half the duration. This is, in part, due to the time buffer that passengers leave themselves, starting their journeys early to ensure arriving on-time (Uniman et al, 2010). If a disruption causes a delay that is less than this buffer, the passenger will still arrive at his/her destination on-time, so he/she is unlikely to find this overly problematic. Once the delay exceeds the buffer and the passenger arrives at his/her destination after the desired time, the perceived disbenefit increases rapidly.

The effects of a major disruption on passenger flow through the Underground network as a whole are more complex than those of minor or moderate disruptions. A minor disruption typically causes no change in passenger behavior, both because information about such incidents is typically not communicated to passengers who are not on the affected train(s) and because the small resulting increase in travel time is unlikely to tip the passenger’s preference in favor of a different route or mode (Freemark, 2013). A moderate disruption is likely to cause some passengers who can shift to other paths without a significant increase in travel time to do so, but route “stickiness” (Yap, 2019), or the propensity to remain on one’s habitual route during a disruption even if another route is slightly faster, means that most passengers will not change their travel behavior. A major disruption, though, is usually well advertised to passengers, and the significant increase in travel time on the disrupted segment causes many affected passengers to reroute themselves through the network. This means that the effects of a major disruption on the full network cannot be estimated through simple calculation, and must be modeled to achieve accurate results.

The above two reasons, as well as the significant passenger impact of each major incident once it does occur, motivate the author to focus the research work on major disruptions. The modeling work related to track layout in Chapter 3, as well as the simplified assignment model developed in Chapter 4, focus on major disruptions.

Many previous researchers have looked at a wide variety of measures, such as service control strategies and passenger information protocols, to mitigate the impact of disruptions of various magnitudes. One area that has not been extensively explored, however, is the degree to which the track infrastructure affects the passenger impact of a disruption. In particular, the addition of a crossover at a critical point on a line can lessen the impact of an incident by allowing a greater portion of the line to remain operational. This phenomenon can be seen readily with the Bank blockade, a multi-week 24/7 closure of the Northern line between Kennington and Moorgate scheduled for 2021 as of this writing (Transport for London, 2020a). The reason for this closure is to enable large-scale station expansion work at Bank station and rerouting of the adjacent tunnels. If the crossover at Moorgate were not present, southbound trains would not be able to turn around anywhere south of Euston, so the entire Bank branch of the line would have to be shut down. Therefore, the presence of the crossover at Moorgate allows more of the Northern line to remain
open, transporting more passengers to the affected stations and reducing the passenger impact of the closure. Although this closure is planned, the same principle applies to unplanned disruptions.

Traditionally, the modeling done to explore this question was detailed and precise. However, it was also labor-intensive and made assumptions generally intended for steady-state network conditions, rather than for unplanned disruptions. This motivated the main focus of this thesis: the development of a simplified passenger assignment model that would be suitable for modeling unplanned disruptions on the Underground. The principal goals of this model are to be accurate, run efficiently, and generate easily understandable and actionable results.

1.2 Research Goals

The objectives of this research are threefold:

1. Develop a framework to estimate the benefit of track layout enhancements on line performance

When Underground trains get off schedule (for whatever reason), it is generally the responsibility of controllers to restore timetabled operation as quickly and efficiently as possible. For them to be able to do this, the track layout must have features - typically crossovers or sidings - that give controllers flexibility and allow trains to be diverted from their planned paths to get back on schedule. Such diversions usually arise in two situations (Carrel, 2009). First, when a minor disruption occurs, it usually causes several trains to fall behind schedule, but by only a few minutes. The controller action in such cases is typically to short-turn the trains affected a few stations before their scheduled terminal and send them back in the other direction at the scheduled time. The location of crossovers near the ends of a metro line thus determines how efficiently and quickly controllers can get the affected trains back on schedule: if there is no crossover at an appropriate location, the controller will be unable to reverse the train there, and will be forced to either extend the train to the next available reversing location, which might keep it behind schedule, or instruct it to reverse even earlier, which would put it ahead of schedule. Second, in the case of a major disruption, there is often an impassable segment due to an incident blocking the track, so trains are reversed near the incident location because they are unable to proceed beyond it. Typically, the closer to the incident the trains are able to get before reversing direction, the more stations can be served and the lower the passenger impact of the disruption. Therefore, the location of crossovers and sidings along a metro line directly affects the magnitude of disruption that incidents cause.

As discussed in Section 3.1, determining the track layout on urban metro systems has traditionally been more an art than a science. Even when scientific principles are applied, site-specific factors tend to override the application of general rules. The aim of this research is to formalize a systematic decision-making process for this problem. The factors affecting the use of crossovers and sidings are analyzed for disruptions of all magnitudes; a framework is developed that uses data on incident frequency, incident severity, and operating characteristics to recommend crossover locations that should best mitigate the impact of major disruptions and planned closures.
This method examines the Piccadilly line of the London Underground and uses Railplan, a Transport for London (TfL) modeling tool discussed further in Section 2.2.3.

2. Develop a simplified passenger assignment model to estimate the impact of unplanned disruptions on a metro line

There are many public transport modeling tools available in the literature. In general, these lie on a spectrum, with simplicity and computation speed on one end, and accuracy and detail on the other. Simplified approaches, such as those based in graph theory, have the potential to work well for a transit agency’s need to get model results quickly, but are infrequently used by transit agencies because they often make strong simplifying assumptions about passenger behavior and network performance. On the other hand, more sophisticated models, such as those involving dynamic assignment, can provide unique insight into the scenario being analyzed and what actions transit managers can take; however, long computation times often mean that practitioners find it infeasible to use them. In addition, there is often a “black box” effect, where there is such a vast array of inputs that the relationships between input and output are unclear.

This research aims to strike a balance between these two extremes. The simplified transit assignment model developed here aims to be nimble enough for agency stakeholders to use it for their modeling tasks. At the same time, it must also be accurate, to develop users’ trust in the model and enable actions to be taken based on its results. This accuracy must be reflected both in its calibration and in the assumptions about passenger behavior which are built into the model.

3. Validate the simplified assignment model by comparing it with the results of modeling work by TfL staff, and apply the model to the London network

As discussed in Section 2.2.3, TfL modelers have sought to improve on their existing Railplan modeling software, and therefore have developed an “imperfect knowledge” Railplan approach that uses the same modeling software but revises the procedure to more accurately reflect plausible passenger behavior during unplanned disruptions. Preliminary results by TfL show significant differences between the results from the two versions of the Railplan model (Yap, 2019), suggesting that passenger behavior assumptions play a large role in modeling disruptions. The results from the modeling tool, including both aggregate estimates of passenger impact and link-by-link volumes and crowding factors, are compared to those from both Railplan models, and the model is applied to several disruption scenarios on the Underground network.

1.3 Research Framework

The idea that spurred this research was the longstanding belief among London Underground staff that a large part of the Piccadilly line’s many performance issues (Freemark, 2013) were caused, or made worse, by its heavily constrained track layout. Incidents can happen frequently and anywhere on the system, but once an incident occurs (and especially after it concludes) it is up to the line controllers to get trains, and train operators, back on schedule so normal service can continue for the rest of the day. Although there is some slack in the timetable
to enable late trains to get back on schedule without intervention, the operational instability and extreme crowding associated with disruptions often mean that this is insufficient and timetabled operation must be restored through controller actions. These usually consist of some combination of the following actions: holding, expressing, canceling, or short-turning (Carrel, 2009). Holding a train is the simplest strategy, but only makes a train later and must be applied carefully on a high-frequency rail line where the following train is usually close behind. Expressing is generally avoided in principle on the Underground, with the exception of the Metropolitan line that operates fast and slow services as part of the timetable. Canceling is the easiest strategy to manage, since controllers can remove trains from the network until their scheduled slot passes the siding where those trains are stored, but is limited by the siding capacity, as well as the need to provide service during incident recovery (Venancio, 2016). The final strategy, short-turning, is therefore commonly employed for recovery from both routine delays and major disruptions. Its inherent dependence on available track infrastructure means that its applicability is often limited. Rahbee (2001), Carrel (2009), Venancio (2016), and others have examined controller behavior during disruptions, but most of this work considered the existing network, and not possible modifications. This research takes a different approach, considering possible modifications to the track layout and assuming a set of controller actions to determine the reduction in passenger impact from major disruptions.

This approach is implemented using TfL tools, datasets, and staff knowledge. The first step, understanding the causes and frequencies of delays on the Piccadilly line, uses CuPID, a log that records all incidents with an impact of over two minutes. Criteria are devised to identify the incidents which are relevant to this study. Because this study focuses on major disruptions that lead to partial line suspensions, the great majority of disruptions are excluded because of their minimal impact. For example, a ten-minute train delay that results from a passenger illness is too short for controllers to reverse trains near the incident, so such an incident is not deemed relevant. In addition, incidents must be near the proposed crossover to be deemed relevant; because most crossovers connect the two running lines and do not give trains room to wait out of traffic, reversing is not done using such crossovers if through service is running.

Visits to the line control center and discussions with other operational staff gave the author valuable insight. Line controllers in particular, whose daily work is limited by the available track layout, had many recommendations for additional crossovers and sidings. Many of these were most useful for smaller-scale service interventions, such as short-turns of individual trains an hour or two after an incident to get them back on schedule. As discussed previously, this type of service intervention is not the primary focus of this research. Their other recommendations, in combination with knowledge of structural constraints such as tunnel walls and other infrastructure, inform the short list of proposed track layout modifications. A set of service patterns is then created for each combination of disruption location and presence, or absence, of a nearby crossover. This involves analyzing operational practices, track geometry, and train characteristics to determine how quickly a train can turn around at a given crossover (extant or proposed). This turnaround time then informs the service frequency that can be operated. Service patterns operated during past closures and
disruptions are used as a point of comparison to determine the operability of the proposed service patterns.

Finally, the proposed service patterns are modeled using Railplan, a static equilibrium-based assignment model built on Emme software (INRO, 2020). (Railplan is described in detail in Section 2.2.3). The output of this model is used to estimate the degree to which the installation of a given crossover can reduce the impact of a nearby disruption; this figure is then multiplied by the expected frequency of unplanned major disruptions, and planned weekend closures, to determine the annual benefit of the crossover. By comparing this with the cost of installation, the business case is prepared.

Analyzing this modeling process reveals several shortcomings of the current approach using Railplan. The primary one, from a practical perspective, is computation time - each run takes one to two hours to complete. The computation of the several dozen scenarios involved in this study thus took several weekends; the author spent the intervening weekdays configuring the next batch of runs, which is quite labor intensive. There are also several secondary drawbacks to Railplan. The first is the set of assumptions surrounding passenger behavior. Railplan is designed to model steady-state networks, where passengers have a good idea of the state of the network before they choose their routes. This is not reasonable when modeling unplanned disruptions - when a link on the network fails, passengers do not know how other passengers will re-route. This is critical on a network such as the Underground because of crowding impacts - during disruptions passengers are frequently unable to board several trains in a row due to crowding. This means what may appear to be the “quickest route” may not actually be the quickest once denied boardings are considered. Representing this lack of passenger knowledge is critical to understanding the impact of a disruption. Another secondary drawback of Railplan is its “black box” nature as a result of its large scope and complexity; a model scope that spans all transport modes and the entirety of Great Britain makes assessing why a certain input led to a certain output challenging, at best.

This inspired the author to pursue the development of a simplified assignment model that would be fast enough for efficient use yet accurate for modeling unplanned disruptions. Unlike Railplan, this model (which is written in Python) focuses exclusively on the London Underground network, with some ‘overflow links’ and walk links between stations added, but no buses, ferries, etc. modeled explicitly. The network and demand are imported from Railplan, with the demand converted to station-to-station matrices, divided into 15-minute timebands, and assigned to the undisrupted network in accordance with existing travel patterns. Then, a disruption is introduced. Passengers on routes not directly affected by the disruption do not change their routes. Passengers using the affected segments, meanwhile, choose another route on the basis of conditions (i.e. crowding) on the remainder of the network as they were prior to the disruption, and the knowledge of the disruption that they are assumed to have. After assignment is complete for each time period, crowding discomfort penalties are recalculated, and any passengers denied boarding are reassigned to the next time period. No iteration is performed; passengers make their initial routing decisions based on the network as they understand it and then proceed over their chosen route regardless of crowding. The results are then compared with the results from both Railplan models: one which
performs equilibrium-based assignment and the other which considers passenger behavior under unplanned disruptions in a way somewhat similar to the simplified model, but still models the entirety of Great Britain and all transit modes.

### 1.4 London Underground

The London Underground (also referred to as the Tube) is the world’s oldest metro network, having opened in 1863. Since that time it has grown to include 270 stations and 250 miles of track, and serves roughly five million daily riders. The system is vital to the urban fabric of London, transporting commuters to their jobs, tourists to the sights, and children to their schools. The system does not operate in isolation, of course, but interacts with other transport operators and with the wider social, economic, and political context.

The Underground is a subsidiary organization of Transport for London (TfL) (Transport for London, 2020g), which is a branch of the London government. Transport for London controls nearly all transport modes in London: urban rail (Underground, Overground, Tramlink, Docklands Light Railway, TfL Rail), surface (bus, paratransit, taxi, and cycle hire, as well as street design), river, and aerial cable car, though, except for the Underground, most services are operated under contract by private parties. The only major transport mode in London outside TfL’s purview is longer distance rail, which is controlled by National Rail.

Figure 1-1 shows London Underground services, commonly referred to as the Tube map. It depicts all TfL rail-based modes as well as the cable car and piers for the river services.

Over the last several years, TfL has faced significant financial challenges, for several reasons (Topham, 2018a). A principal factor was the 2015 decision by the UK government to eliminate the operating subsidy, which had provided TfL with £700 million annually. A second cause for reduced revenue was the Fares Freeze, a mayoral initiative to keep fares for TfL services constant since 2016. Third, a decline in ridership (2% year-on-year in 2018) has meant that fare revenue is farther below projections. Fourth, budget projections had assumed that Crossrail would open on schedule in 2018, and would attract more riders and thus increase fare revenue; the project’s multi-year delay in opening has postponed this financial boost. Fifth, the financial uncertainty associated with Brexit has increased certain costs for TfL, most directly the cost of borrowing money (Ianvisits, 2019). Finally, the COVID-19 pandemic ongoing as of this writing has dramatically reduced fare revenue, putting additional strain on the budget (BBC News, 2020).

The response to these pressures has included a significant cost-cutting effort with the postponement of many planned investment programs, including the re-signaling of the Piccadilly line (Topham, 2018b). Given the aging infrastructure on the Piccadilly line, this delay has the potential to cause decreased signal reliability and consequently line performance in coming years. Thus, Underground managers are interested in ways to improve performance with relatively low investment, such as enhancing the track layout to improve incident recovery.
Figure 1-1. Tube map (Transport for London, 2020c)
Another consequence of these budget cuts has been the increasing centralization of functions across operating groups. For example, a single transit planning and modeling team is now responsible for all modes of transport. Accordingly, in this thesis the modeling staff who analyze the Underground will be referred to as TfL staff. TfL has very well-developed data infrastructure and analysis capability, with a plethora of datasets describing most aspects of its operation. Three datasets related to the Underground that are used for this research are summarized here, and described in detail later in the thesis.

1.4.1 The Piccadilly Line

The first part of the Piccadilly line core opened in 1906, stretching from Finsbury Park in the east to Hammersmith in the west, as part of a set of three lines built by Charles Tyson Yerkes’s Underground Electric Railways Company of London, which are now the Bakerloo, Northern, and Piccadilly lines. Save for a one-station branch line in Central London, from Holborn to Aldwych (formerly Strand) station, which operated from 1907 to 1994, this section of the line remains in operation, unchanged, today. In the 1930s the line was extended on both ends, via new tracks to Cockfosters in the east and via existing track (replacing District Railway services) to Uxbridge and Hounslow West in the west. The line remained in this configuration until the 1970s, when the extension to Heathrow Terminals 2 and 3 was opened. The loop to Terminal 4 followed in 1986, and the branch to Terminal 5 in 2008.

The line today is 71 km long, with 53 stations, and has two western branches and a single trunk through central London. The vast majority of the line consists of two tracks. The stations at Uxbridge, Arnos Grove, and Cockfosters have three tracks, South Ealing and Northfields have four tracks, and Heathrow Terminal 4 has one track in a unidirectional loop. The section between Barons Court and Acton Town has four tracks, with the District line on the outer tracks stopping at all stations and the Piccadilly line on the inner two skipping several stations. Piccadilly and District line trains share tracks between Acton Town and Ealing Common, but otherwise their trains operate independently. Piccadilly and Metropolitan line trains share tracks between Rayners Lane and Uxbridge.

During the rush hour, the Piccadilly line operates at 2.5 minute headways (24 trains per hour) on the central trunk. Because of scheduled short turns at Rayners Lane, Northfields, and Arnos Grove, and the presence of multiple branches on the western end of the line, the remainder of the line sees less service. Figure 1-2 shows the Piccadilly line’s layout, service patterns, and infrastructure, and Figure 1-3 shows a schematic diagram of the line.

Today, the Piccadilly line carries over 210 million passengers annually. It is a critical link across London, passing through major transit hubs including Kings Cross, St Pancras and Heathrow Airport, tourist destinations such as Piccadilly Circus, and large residential areas. It has not seen a comprehensive investment and infrastructure upgrade program for over forty years, the longest such period for any tube line today except the Bakerloo. The Piccadilly line operates a fleet of 87 trains of 1973 stock, which last had a major refurbishment in the early 2000s (Squarewheels, 2010). A new fleet of trains is on order from Siemens, and expected to enter revenue service in
2023 (Reynolds, 2018). The signaling infrastructure is similarly antiquated. Although the Piccadilly Interim Control Upgrade (PICU) project, completed in 2019, replaced the line control center and the punchcard-era signal control technology previously used, the back-end signal infrastructure remains unchanged, and financial difficulties have caused a postponement of long-discussed re-signaling plans (London Reconnections, 2018). It is against this challenging background that London Underground managers try to maximize the line’s performance. For these reasons, the Piccadilly line was chosen as the case study for this research.

Figure 1-2. The Piccadilly line – layout, service patterns, and infrastructure. (Freemark, 2013)

Figure 1-3. The Piccadilly line – schematic. (Transport for London, 2020e)
1.4.2 Overview of Data Sources

CuPID (Contract Performance Information Database) is the internal database that London Underground uses to keep a record of all disruptions affecting passengers’ travel through the network. This log includes all incidents with a disruption duration of more than two minutes. This includes many incidents affecting train service - asset failures, passengers taken ill, etc. - as well as non-incident events (cancellations caused by operator shortages, for example) and in-station events that do not affect trains, such as lift failures. For each disruption, a wide range of information is recorded, including time, location, duration of initial incident, duration of effect on service, and cause code. This cause code is determined through an allocation process based on the root cause of the disruption. For this analysis, the most important data field in CuPID is the LCH (Lost Customer Hours) – the estimated impact of the disruption. This field is populated by matching the real-world disruption against an extensive library of incidents, for each of which an LCH figure has been pre-calculated. This process is described in more detail in section 2.2.3.

NetMIS is the database that contains historical train location data for the Underground. It draws its data on actual train locations from Trackernet, which is the real time model board used by line controllers, which in turn obtains data from signal and train control computers, remote telemetric units connected to legacy signals, and the Connect train-based radio system. This is then matched to timetable train movements drawn from iCart, the Underground’s internal electronic timetabling system. NetMIS provides data including location, train number, timestamp (for both wheel-start and wheel-stop at stations), and signal information. Wheel-start and wheel-stop times are calculated using an offset from the track circuit activation times that comprise the raw input data. This data makes possible a wide range of analyses of headway, dwell time, running time, and incident impact.

There are several data sources that aim to describe ridership patterns on the Underground. The two principal ones are RODS, which is survey based, and Oyster, which is based on farecard data. RODS (Rolling Origin-Destination Survey) periodically surveys passengers to determine their travel patterns, and their choice of route. The survey data (necessarily limited by sample size) is then extrapolated to estimate total passenger flows. Data based on Oyster, on the other hand, is much more comprehensive because it is drawn from all fare-paying passengers. However, it only records when passengers tap in and out of the system; therefore, it cannot be used to determine the route a passenger takes through the network. Because of the routing uncertainty, and other historical reasons related to model calibration, many TfL models remain dependent on RODS data, despite the existence of Oyster data since its introduction in 2003. TfL is currently making efforts to gather data about passenger routing through aggregating Wifi data from cell phones (Transport for London, 2017a), but analysis of collected data is still in the preliminary stages.

1.4.3 Disruptions on the Underground

As defined in section 1.1, disruptions on the Underground can be divided into three categories, minor, moderate, and major, depending on the length of the initial incident. The primary reason for drawing this distinction between disruption types is the variation in service
controller response depending on the initial delay. Although service controllers cannot predict how long an incident will last, and this uncertainty often causes difficulties when deciding how to respond, their experience and judgement usually let them decide on a response strategy as the incident occurs (Carrel, 2009).

Minor disruptions warrant little response by controllers. Incidents that last less than ten minutes cause little train lateness and this delay can typically be recovered at the termini.

Moderate disruptions mainly see controllers mitigating their effects after the disruption has cleared and the only problem is late-running trains. Some trains at the ends of the line may be sent to the depot to prevent train congestion near the incident location, but trains en-route will generally not be diverted so as to keep crew and train assignments as close as possible to the schedule. Assuming that the incident is resolved in a short period of time, this ensures that the affected trains and operators will be only slightly off schedule, as opposed to the hour (or more) offset from schedule that would result if trains were reversed mid-trip.

Major disruptions cause controllers to focus on maintaining service while the incident is underway, in addition to mitigating its impacts after the incident is resolved. In the case of an incident that blocks a section of track, this is usually done by short-turning trains on each side of the affected section to keep service operating on the rest of the line.

![Figure 1-4. Incident occurrences and impacts by length of initial delay, 2010-2019](image)

Figure 1-4 shows the distribution of delays, by length of initial delay, for all incidents recorded in CuPID (note the logarithmic scale on each axis). Note that in CuPID, “initial delay” denotes the length of the incident itself, while “duration” refers to the time until service has returned to schedule. Incidents with an initial delay of 0 minutes are excluded from this plot; the vast majority of these are station-based incidents such as lift failures. Industrial action (strike) incidents are also excluded.

As expected, the most common incidents are short: 89% of incidents last less than ten minutes. For incidents that have an initial delay of less than (approximately) 100 minutes, the
relationship between initial delay and average LCH, and initial delay and frequency of occurrence, is a power relationship. For delay durations greater than 100 minutes, there is a greater range in incident occurrence and impact. This is often caused by the common practice of installing a temporary fix for problems such as signal failures, which allows train service to resume and thus reduces the impact of the disruption without ending the disruption itself.

The location of an incident can be as critical to its passenger impact as the time it takes to resolve. This is especially true on a line such as the Piccadilly, where the presence of several branches and many locations on the outskirts where short-turning is possible means that incidents on the central core have a much greater impact than those closer to the ends of the line, where scheduled frequencies are lower and capacity exists on nearby lines.

Figure 1-5 shows the spatial distribution of incidents on the Piccadilly line, separated into sections corresponding to crossover and siding locations. As before, industrial action and delays with a duration of zero minutes are excluded, as are station-based incidents (i.e. elevator failures) which have no impact on train service. Clearly, the outer sections of the line (the western branches, and Arnos Grove to Cockfosters) are comparatively delay-free; the principal locations for delay are the express-running section between Acton Town and Barons Court, as well as the segment east of Kings Cross. This delay distribution helps determine which disruptions justify modeling to understand and investment to mitigate.

Figure 1-5. Delays on the Piccadilly Line, 2018 (adapted from Transport for London, 2020e)

1.5 Thesis Organization

- Chapter 2 describes the current state of disruption modeling. This chapter begins with a literature review covering several approaches to the subject. This is followed by an extensive description of the several complementary approaches that the London Underground currently uses to understand the impact of disruptions on its network.
• Chapter 3 discusses the track layout analysis. It begins with a literature review and discusses the ways in which crossovers and sidings are used on the London Underground, and the benefits of each. Next, it describes the analysis procedure that London Underground staff currently use to determine the benefit of a given track layout element. Finally, it establishes a procedure for estimating this benefit, and models this for several proposed crossovers on the Piccadilly line. It concludes with the business case for the installation of each crossover considered.

• Chapter 4 describes the simplified passenger assignment model. It begins with a description of the model inputs. Then, the assumptions underlying the model and their rationale are discussed. Next, the procedure behind the model is described and results presented. These results are validated by comparing against the results from the two Railplan approaches used by London Underground modeling staff, and by looking at specific link flows.

• Chapter 5 summarizes the thesis contributions, results, and conclusions. It gives several recommendations to London Underground and outlines several potential directions for future research.
2 Disruption Modeling in the Literature and in Practice

Understanding the impact of disruptions is a common area of focus in public transport modeling. A wide range of modeling tools have been developed to help estimate this impact. These models have been fine-tuned to match their applications - quick estimates used by practitioners for preliminary alternatives analysis, fine-detail data used to examine passenger-specific impacts, and everything in between. This chapter begins with a review of the main modeling approaches. Then, the current state of disruption modeling on the London Underground is summarized.

2.1 Literature Review

This section addresses two principal areas of difference between many of the public transport models in the literature. The first - the difference between schedule- and frequency-based models - refers to how passengers are assumed to behave on the network. The second is the tradeoff between level of detail and computational effort, with four types of models discussed, in increasing order of complexity: graph-theoretic, static, dynamic, and simulation-based models.

2.1.1 Schedule vs. Frequency Based Modeling

There are two primary approaches to modeling transit networks, whether in the context of steady-state operations or disruptions: schedule- and frequency-based. Schedule-based models represent each vehicle departure on a transit line explicitly, while frequency-based models approximate the line based on its headway and capacity, without considering individual vehicles. The modeler’s choice between a schedule- or frequency-based model should be based on the way passengers are expected to perceive the transit service and how this affects their behavior. User perceptions and behavior are governed by numerous system parameters; in this case, the most important one being service frequency.

Schedule-based models are best when service is infrequent, because then passengers typically consider each vehicle arrival/departure. Before beginning their trip, they will often consult a timetable and time their arrival at the station to minimize their expected wait time (Gentile et al, 2016). In their paper that lays out a deterministic assignment model and describes its implementation in the EMME/2 software package, Constantin et al (2002) express this behavior by introducing the concepts of “implicit earliness and lateness,” which includes the difference between a passenger’s arrival time at a station and their departure in a vehicle, and “explicit earliness and lateness” which describe the difference between a passenger’s preferred departure time and that which is delivered by a vehicle. The model they introduce outlines penalties for each. Similarly, Carraresi et al (1996) define a “local disutility” at the origin, destination, and each transfer location, which is a function of the difference between a passenger’s ideal and actual
arrival (departure) times. This idea is combined with alternative passenger assignment algorithms to derive an optimization model that minimizes total passenger waiting time by modifying vehicle departure times while satisfying driver and fleet availability constraints.

One critical aspect of understanding the performance of networks such as the Underground, particularly during disruptions, is vehicle capacity. Often, a large portion of a disruption’s impact comes from passengers being unable to board trains because of the capacity constraint, so it is important that a schedule-based model that considers each individual vehicle trip takes this into account. For example, Nuzzolo et al (2012) develop a joint departure time and route choice model that explicitly considers passengers denied boarding and iterates over multiple days to reflect passengers’ learning about network conditions. This paper analyzes the impact on passenger assignment of several parameters relevant to passenger behavior, such as FIFO (first-in-first-out) vs RIFO (random-in-first-out) behavior when boarding a train, and the impact of being denied boarding on future passenger choices, and models on-vehicle crowding and failure-to-board. The paper concludes, based on application to a simple network and preliminary application to the Naples transit network, that such a model converges to a unique solution and can be useful to examine network crowding in detail.

As service frequency increases, passenger perceptions of the network shifts: fewer passengers rely on a timetable when planning their departure time, which means that although most users will have a sense of typical headways and running times, they will regard specific train arrivals as random events, without a precise arrival time (Gentile et al, 2016). As with the London Underground (Transport for London, 2020b), while such systems usually maintain detailed timetables that are used for operator and rolling stock scheduling and may even be available to the public, passengers generally do not take such information into account. Transport for London’s practice is to consider services with headways of 15 minutes (or less) as “turn up and go,” i.e. where passengers do not consider the timetable when planning their departure time (Stubbs, n.d.). In such cases, the modeling emphasis shifts from understanding passengers’ preference for particular departures to understanding their behavior under crowded conditions (Wilson and Nuzzolo, 2008).

Many researchers apply schedule-based modeling techniques to these high-frequency scenarios. However, frequency-based models allow the network to be considered at an aggregate level without considering each vehicle explicitly, which reduces computation time and model complexity while still providing useful insight. These models are most useful to determine average load, or congestion, on a line across a timespan (Gentile et al, 2016). One of the fundamental papers in this area (Spiess and Florian, 1989) introduces the concept of “strategy,” which describes how passengers use available information to choose their path through the network. This paper outlines a linear optimization algorithm that can be used to solve the equilibrium problem. Though the concept of a strategy is not exclusive to frequency-based models, it is particularly useful in analyzing the route choice problem in such models. Cepeda et al (2006) construct a frequency-based, equilibrium model and formulate it as an optimization problem which minimizes a function of the difference between the network state and perfect equilibrium. They then apply this model
to the transit networks of Stockholm, Winnipeg, and Santiago (Chile), and demonstrate that the algorithm converges quickly and produces realistic results.

Schmocker et al (2008) develop a model that addresses what they perceive to be the main shortcoming of frequency based assignment models - namely, the typical lack of “peakiness” in such models due to the aggregated nature of the demand matrix. They do this through the development of a “fail-to-board probability” that considers those passengers who do not fit on a train in a given timeband as being accommodated by trains in the next timeband. They apply this model in a case study of the London Underground and find that several links on both the Victoria and District lines do not have sufficient capacity to meet demand between 8:30 and 9:00AM, though there is sufficient capacity in other periods. This matches well with TfL data (Greater London Authority, 2019), supporting the validity of their approach.

For reasons of computational efficiency as well as the high frequency nature of Underground service, as described in Chapter 4, the model developed here is frequency-based. As a result, the remainder of this literature review will focus on frequency-based models.

2.1.2 Graph Theory Approaches

Of the four approaches to disruption modeling described in the rest of this section - graph theory, static, dynamic, and simulation models - those based on graph theory are the simplest and most abstract. These approaches describe networks primarily in terms of their topological properties, although some also include passenger flow information (King and Shalaby, 2016). Topological approaches, at their base, consider the network in the abstract sense as a collection of links and nodes, and estimates of the impact of a disruption are based on the importance (to be defined) of the disrupted link(s) or node(s). A typical metric used to assess the importance of a given part of the network is betweenness centrality, which is a measure of the percentage of shortest paths between all origin-destination (OD) pairs that include a given network element. Derrible (2012) uses this metric to gain an understanding of transfer behavior, which is critical in metro systems, characterizes the topological properties of 28 metro systems around the world and identifies these networks’ most important stations. Although he does not explicitly state it, he implies that the stations with the highest centrality are those on which a disruption would cause the greatest passenger impact.

Other, network-wide metrics can also be used to compute disruption impacts, by measuring the change in the metric in the event of a disruption. Such metrics include directness - the number of transfers needed to make a journey - and connectivity - which describes the number of transfer possibilities in a network. These have been shown to correlate well with common real-world performance indicators such as system ridership (Derrible and Kennedy, 2009).

Latora and Marchiori (2001) define a metric they call efficiency, which is the inverse of the distance (or travel time) between any two nodes; by extension, the average efficiency of a graph is the average efficiency over all node pairs. They then used this metric to estimate the impact of disruptions (the removal of link(s) or node(s)) on the Boston (MBTA) network, and identified the links and nodes that would cause the greatest degradation in performance if they were removed.
In their study, where performance was defined as the total travel time between each pair of nodes, without consideration of actual passenger demand, they determine that the Boston subway network is generally not fault-tolerant. A follow-up paper (Latora and Marchiori, 2005) expands on this concept to identify those additional links which would most increase the robustness of the system.

Jenelius et al (2005), in their study of the Swedish road network, propose several metrics that consider traffic demand, including importance, which quantifies the impact of the loss of a link or node on the total travel times of all users of the network, and exposure, which measures the degree to which a certain group of passengers is vulnerable to disruptions. They then quantify the vulnerability of the network to both random and targeted attack, or the degree to which the removal of a link or node would impact total travel times. King and Shalaby (2016) use Latora and Marchiori’s efficiency metric to consider the impact of random node removals on the Toronto transit network, and correlate Jenelius’s importance metric with the computed impact of disruptions, and find that the links with the highest importance are those with the fewest nearby alternatives, not necessarily those in the center of the network.

However, these approaches all make two critical assumptions - lack of capacity constraints on the network, and independence of passenger volume and travel time - that limit their applicability to modeling disruptions in particular and suggest more in-depth analysis, such as that by King and Shalaby (2016) using an EMME-based static assignment model. In a system where capacity exceeds demand, such an assumption might well be valid, however in a network such as the Underground, where demand in a short time period can exceed supply, and the network experiences denied boardings even without disruptions, the impact of any disruption will be complex, because the shifting of passengers to alternative routes will cause a sharp increase in crowding and denied boardings on those routes.

2.1.3 Static Assignment Models

This approach is similar to that used in many traffic assignment applications. The fundamental principle is one of “stationarity;” that flows and demands are constant throughout the study period. This, by definition, implies that demand cannot exceed supply, because otherwise ever-increasing queues would form and the network state would not be stationary (Gentile et al, 2016). Hunt (1990) constructs a static assignment model of the transit network of Edmonton, Canada, and introduces a logit model to determine passenger route choice. His findings, mainly focused on network design, suggest that passenger outcomes will improve if routes are designed primarily to reduce the number of transfers passengers must make, even if this increases in-vehicle time.

De Cea and Fernandez (1993) construct a static equilibrium based assignment model based on these principles, and suggest several non-linear optimization techniques to solve it. Their model includes crowding, with vehicle capacity constraints represented explicitly. Crowding is included as a node-based penalty representing denied boarding time (as opposed to a link-based penalty representing travel discomfort), and the utility of a given route is defined using a metric known as
“effective frequency,” which approximates the impact on passenger waiting time of being unable to board vehicles due to capacity constraints.

King and Shalaby (2016) use EMME, which is the software underpinning Railplan, to analyze the impact of station removals - cessation of services through a station - on the Toronto subway system. Interestingly, they find that the largest disruption results from removing Lawrence station (which is neither a transfer station nor in the CBD). This is because Lawrence lies on a heavily used branch line with no nearby alternatives that are not already at capacity. Though this model gives useful results that are borne out in reality (according to the TTC, delays at Lawrence alone accounted for 11.2% of total system delay-minutes), it is important to recognize that it is still a static model. In other words, the supply and demand do not change during the three-hour period (6:00 to 9:00 AM) being studied. This fails to consider ‘peakiness’ - the fluctuations in demand over short periods within the morning rush period - which leads to underestimation of the true delay impact.

### 2.1.4 Dynamic Assignment Models

Unlike static assignment models, which assume that a network is in a steady state and unchanging, dynamic assignment models add the element of time to the analysis. These models consider several time-dependent elements of a transit network, including passenger propagation, supply and demand fluctuations, and queueing (Gentile et al, 2016); this allows a more detailed look at several aspects of transit network operation and can support a deeper understanding of the network state than is possible with the assumptions underlying a static model.

Schmocker (2006), for example, notes that it is possible to consider passengers who choose a new route if they are unable to board the first train at their current station. This behavior, seemingly irrational when examined with a static model, is a good representation of actual behavior during a disruption. He develops an equilibrium-based model that dynamically considers congestion on the London Underground by explicitly considering capacity constraints and removing any passengers denied boarding from the downstream links that they would otherwise have occupied. Its route choice algorithm incorporates overcrowding and denied boarding information, as well as vehicle arrivals for those passengers with multiple paths from origin to destination. He demonstrates that this approach accurately reflects the crowding patterns on the Underground.

Meschini et al (2007) construct a dynamic equilibrium model that aims to capture congestion effects on both road and transit modes using a frequency-based model. Their approach dynamically loads passengers onto network links and updates the performance of each link as appropriate, and incorporates an average line access time model to estimate denied boardings. They then apply this model to the commonly used Sioux Falls network and find that the model properly considers congestion effects and has reasonable computation time. Tong and Wong (1999) develop a stochastic dynamic model that incorporates a time-dependent route choice approach using a branch and bound algorithm. Monte Carlo simulation is applied to solve the resulting model, which is applied to the Hong Kong MTR network as an example to show the
model’s ability to identify the most overcrowded links and explore alternatives to resolve this overcrowding.

2.1.5 Simulation

Simulation is the most micro-scale, and therefore the most detailed and computationally intensive, of the four approaches to modeling outlined here. It allows in-depth analysis of the behavior of each passenger (or vehicle) at any point on the network. Sumalee et al (2009), for example, devise an assignment model that explicitly considers the probability that passengers will be able to sit at some point on their journey, based on parameters, such as the desire to sit, that vary over time. They assume that seated passengers do not experience crowding-based discomfort. Thus, passengers who assume that they will have a chance to sit will expect a lower generalized journey time than passengers who do not. This, in turn, means that considering seat allocation will cause passengers to choose to travel closer to their ideal departure time, because the possibility of sitting (and thus reducing perceived crowding-based discomfort) outweighs the discomfort penalty from traveling during the most crowded time during the peak period. Removing the seat allocation process, on the other hand, will make passengers more likely to consider the crowding discomfort along their entire journey as opposed to only the portion where they expect to stand, thus encouraging a shift to less crowded times. Teklu (2008) constructs a Monte Carlo simulation model that includes capacity constraints and explicitly models passenger arrivals, and applies it to a small network. He then compares his results to those of de Cea and Fernandez (1993), who analyzed the same network using an optimization approach, and finds that the simulation is more accurate because it explicitly considers the impact that capacity constraints have on passenger route choice. Cats (2014) uses BusMezzo, an agent-based simulation model, to examine the impact of real-time information systems on passenger decisions. With the ongoing expansion of media conveying service information, both general and personalized, it is important to understand the ways in which tactical information provision may change passenger routes and departure time choices on the network; Cats’s work is designed to help transit operators assess the impact of real time information provision.

2.2 Current Practice at London Underground

This section describes the current practice at the London Underground for modeling the impact of disruptions. It begins by discussing the various sources of ridership data with the advantages and disadvantages of each. Then, the Journey Time Metric (JTM), the metric that LU uses to quantify passenger journey times, is introduced. Finally, the two main approaches used to quantify delay impact are discussed: Excess Journey Time (EJT) is a retrospective calculation based on service as operated, and Lost Customer Hours (LCH) is a modeling approach that uses Railplan, a static assignment model based on EMME software, to generate an incident library that actual incidents are then matched against. A new approach to incident modeling that is currently
under development by Transport for London planners, the “imperfect-knowledge” Railplan approach, is also described.

2.2.1 Underground Ridership data: RODS, Oyster, and Wifi

London Underground has several sources of data that it uses to estimate ridership: RODS, Oyster (smartcard), and Wifi. The oldest of these is RODS, the Rolling Origin-Destination Survey, which dates to 1998 though similar origin-destination surveys have been performed on the Underground since 1970 (Transport for London, 2017c). Surveys are performed in the autumn, between 07:00 and 24:00 on weekdays, and are considered on a rolling 15-year basis. These surveys aim to capture a wide variety of data on passenger journeys, including:

- Origin and destination
- Time of travel
- Route
- Access and egress modes
- Demographic data
- Journey frequency
- Journey purpose
- Ticket type

As with any survey-based approach on a network such as the Underground, its principal limitation is that it has a small sample size – approximately 20 to 30 thousand surveys annually (Gordillo, 2006). Even with fifteen years of data considered, the sheer volume of ridership on the Underground means that there are many types of users that the survey does not reach, and the data thus cannot capture. In addition, journey patterns change over time, and the inclusion of dated survey results to increase the sample size may skew the results by considering journeys that are no longer made, at least in the same way.

Oyster data, available since the introduction of the so-named farecard in 2003 and now including records from contactless bank cards, phones, and other payment media, represents the next generation of ridership data. It records each Oyster card (or other payment medium) ID, and the time and location of each tap-in and tap-out. Although it cannot determine trip purpose or demographic data as does RODS, its vastly greater sample size means that it has become the preferred dataset for many types of Underground ridership analysis. Because London is a closed fare system, where passengers are required to tap in and out, passenger destinations do not need to be deduced, as must be done in cities with open fare systems. Route choice must be inferred, however, because most passengers only use their Oyster cards to enter and exit the Underground, but not when transferring between lines. Modeling, calibration with RODS, and verification through other data sources are used to infer passenger route choice.

Wifi data is the most recent method to infer ridership, with its main purpose being to capture passenger routes through the network, which it does by tracking passengers’ phones.
Because almost every Underground station is equipped with Wifi, a Wifi-enabled phone can be tracked on its journey through the system as it tries to connect to Wifi routers at successive stations. Although variation in smartphone ownership among Underground ridership may introduce bias, such data collection can provide much greater coverage of Underground ridership than RODS and can therefore provide better information, specifically on route choice. For instance, a data collection pilot in 2016 revealed thirteen different paths that were used by at least 1% of the passengers traveling between Kings Cross and Waterloo (Transport for London, 2017a). This data can also be used to understand passenger behavior during disruptions, as those passengers who usually use a disrupted link can be tracked over their alternate routes through the network. This is particularly valuable since conventional data collection means such as Oyster, are often waived when there are serious disruptions on the network. Wifi data is not a replacement for Oyster, because it does not directly record system entry or exit time and only represents a sample of Underground passengers, but the combination of these two datasets has the potential to provide a deeper understanding of passenger behavior on the Underground.

Current practice reflects a gradual shift by TfL staff from RODS to Oyster. Newer analyses, such as those used to derive station-specific crowding information with fifteen-minute precision, are based primarily on Oyster data. However, there are many legacy models, including those used for EJT and LCH calculations, that are based on RODS data for historical reasons and to facilitate trend analysis. Work is underway to update EJT calculations to better incorporate Oyster data, and to integrate Wifi data (which is now being collected system-wide) into TfL’s models.

2.2.2 Journey Time Metric

London Underground uses two primary metrics to measure daily performance - Lost Customer Hours (LCH) and Excess Journey Time (EJT). Both metrics are based on the Journey Time Metric (JTM), which aims to assign a time value to each passenger’s journey that is weighted to reflect the discomfort the passenger experiences, whether due to long waiting times, using stairs, or increased crowding on trains and in stations.

The Journey Time Metric consists of five components (London Transport, 1999):

- **Access, Egress, and Interchange (AEI):** this takes into account turnstile entry, movement from faregate to platform, and lift and escalator failure information to create a model of pedestrian flow within the station.
- **Ticket Purchase Time (TPT):** this metric uses survey data on average queueing and interaction times with ticket machines and average percentage of people using a ticket machine to infer how much time passengers spend on ticket purchase.
- **Platform Wait Time (PWT):** this metric uses train movement data to determine the average wait time on the platform. Crowding levels are taken into account, both in the context of platform crowding (which causes discomfort) and denied boardings, where the passenger waits for more than one train before being able to board.
- **On Train Time (OTT):** this metric uses train movement information and inferred crowding data to compute the time a passenger spends on a train, weighted by a crowding factor.
• Closures: this metric considers the impact of disruptions of more than 30 minutes, including both unplanned incidents and planned closures.

This metric does not consider all time equally. Instead, to reflect the inconvenience of traveling in a crowded train or climbing stairs, the following weights are applied:

Table 2-1. Journey Time Metric weighting factors (London Transport, 1999)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Time weighting factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking (through ticket halls/passageways)</td>
<td>2</td>
</tr>
<tr>
<td>Walking up stairs</td>
<td>4</td>
</tr>
<tr>
<td>Combination of walking up and down stairs</td>
<td>3.25</td>
</tr>
<tr>
<td>Riding escalators/lifts</td>
<td>1.5</td>
</tr>
<tr>
<td>Walking down stairs</td>
<td>2.5</td>
</tr>
<tr>
<td>Ticket queuing time</td>
<td>3</td>
</tr>
<tr>
<td>Ticket purchase time</td>
<td>2</td>
</tr>
<tr>
<td>Waiting on platform</td>
<td>2</td>
</tr>
<tr>
<td>Travelling on train</td>
<td>variable (between 1 and 2.48 depending on the level of crowding)</td>
</tr>
</tbody>
</table>

To evaluate the performance of the network at any time, the actual values for each of these metrics are compared with the ‘scheduled’ values to compute the excess time. For the portions of the calculation that do not depend directly on train movements - AEI and TPT - scheduled and actual values are derived from surveys, ticket machine logs, and pedestrian models. Planned engineering works are classified as scheduled closures, while unplanned disruptions account for actual closures.

2.2.3 Pre-modeling (LCH) Approach

In 2003, the London Underground contracted out maintenance and asset refurbishment responsibility to the private sector under a Public Private Partnership (PPP) in response to years of declining performance and asset condition (Schaefer, 2018). As part of this agreement, the private partners were responsible for ensuring service availability, which was measured using a metric called Lost Customer Hours (LCH). This metric was used to identify the degree to which each incident on the network affected service delivered; the root causes of incidents were recorded and the private partners were penalized for an excessive number of incidents caused by factors within
their purview. Although the functions provided by the private partners under the PPP were brought back under TfL control by 2010, LCH is still used today: the assignment of a root cause and magnitude of disruption to each incident allows detailed performance reports to be created for each of the Underground’s performance areas - customer service, assets, operations, etc. LCH figures are also used to inform business cases, prioritize maintenance and investment projects, and analyze planned closures (Transport for London, 2019b).

The LCH metric is an attempt to capture all the impacts of a disruption. This has several important implications. First, it attempts to measure the impact of not just the incident phase of a disruption, but also the service recovery phase. Although the implementation of the service recovery process is inherently unpredictable, depending on line controllers’ experience, the position of trains and crews, and passenger behavior, the LCH calculation assumes a typical service recovery process. Second, the specific type of incident is taken into account. For example, a signal failure that causes reduced speed and capacity over a section of line is modeled differently than a signal failure that prevents trains from operating over a given section of track. Third, ridership profiles at the time of the disruption are taken into account, because a suspension late on a Sunday evening will have a significantly lower impact on passengers than a similar suspension at the height of a weekday morning peak. Finally, it is a measure of not just the direct travel time impact of a given disruption, but of the total social disbenefit. Thus, it uses the Journey Time Metric, instead of a measure of actual passenger time spent on the network, to quantify the impact of a disruption.

LCH and the Incident Library

Because of the large number of incidents occurring on the Underground, many of which are in some sense unique, it is infeasible to model each incident individually to determine its LCH impact. Thus, a set of approximately 6,600 incidents across all lines, locations, and impacts on service has been modeled. As an example, a signal failure on the Victoria line that lets service run throughout the line, but requires a frequency reduction, is modeled as nine separate incidents: 10, 20, …, 90% reduction in frequency. For each incident, the resulting service pattern is modeled (using a demand forecasting model, pedestrian flow model, train service simulation tool, and Railplan), and the LCH value computed.

These results are used to create an incident library. Meta-models combine and interpolate the results from the above models to create the library, which represents each combination of 17 disruption types, all Underground lines, all incident locations and times, for planned and unplanned incidents - 22 million combinations in all (Transport for London, 2019b). This is stored in the form of a lookup table that is then used by CuPID (discussed in Section 1.4) to determine the LCH value of an incident.

Since this methodology was developed in 1996, ridership patterns have changed significantly, as has the transit network itself. As a result, the incident library has been updated several times. The most recent update, performed in 2018, assumes projected 2022 ridership and network state and is expected to be used until the mid-2020s.
**Railplan**

Railplan is a static, iterative, equilibrium-based assignment model, based on Emme software (Transport for London, 2019a). First developed in the late 1980s, Railplan models London and the Southeast of England, and includes an aggregate representation of the rest of Great Britain. As a London-focused model, the level of network detail decreases with distance from London. For example, bus services outside London are not modeled, many services (of any mode) outside the Southeast are omitted, and rail stations outside the M25 (London’s orbital motorway) are modeled as having only a single platform, as opposed to those within the M25 which are modeled separately by direction, platform group, or by individual platform. Railplan includes all rail-based modes (Underground, National Rail, Docklands Light Railway, and Croydon Tramlink) as well as London buses and walking, but does not include modes such as the Emirates Air Line (a cross-river cable car in east London), riverboat services, or cycling.

Railplan carries out the last step - passenger assignment - of the four-step urban transport modeling process. The first three steps - trip generation, distribution, and modal split - are performed by the London Transportation Studies (LTS) model, which generates a demand matrix that is an input to Railplan (Transport for London, n.d.). This demand matrix divides Great Britain into 4106 zones, with zone size increasing with distance from London. While most of these zones are geographic areas, some represent demand originating outside Great Britain and arriving by air or Eurostar. Railplan is used for modeling the network state at five-year intervals (2021, 2026, etc.); Table 2-2 shows the number of trips per hour modeled by year and time period.

<table>
<thead>
<tr>
<th>Year</th>
<th>AM Peak</th>
<th>Off Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021</td>
<td>1,023,031</td>
<td>693,341</td>
</tr>
<tr>
<td>2031</td>
<td>1,138,761</td>
<td>779,109</td>
</tr>
<tr>
<td>2041</td>
<td>1,222,154</td>
<td>849,378</td>
</tr>
</tbody>
</table>

Like demand, supply is assumed to change over time, with investment in existing infrastructure allowing increased frequencies, and construction of new infrastructure expanding the network. Table 2-3 shows the changes in the network that Railplan assumes between 2021 and 2041.

Like nearly all transit models, Railplan models the network as a collection of links and nodes over which lines operate. Figure 2-1 shows the network hierarchy in a Railplan model.
Table 2-3. Assumed changes to the network between 2021 and 2041 (Transport for London, 2019a)

<table>
<thead>
<tr>
<th>Line</th>
<th>Year</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Speed 2</td>
<td>2026</td>
<td>Phase 1 (10 tph) in line with HS2 Business Case Assumptions, associated changes to WCML services; all GWR Paddington and Elizabeth Line services call at Old Oak Common</td>
</tr>
<tr>
<td>Circle, H&amp;C, District, Metropolitan</td>
<td>2026</td>
<td>Full implementation of Four Lines Modernisation (4LM), increased frequency</td>
</tr>
<tr>
<td>Piccadilly</td>
<td>2026</td>
<td>New rolling stock with higher capacity, increased frequency</td>
</tr>
<tr>
<td>Docklands Light Railway</td>
<td>2026</td>
<td>New rolling stock with higher capacity</td>
</tr>
</tbody>
</table>

Figure 2-1. Railplan network hierarchy (Transport for London, 2019a)

In the context of disruption modeling, Railplan makes several critical assumptions. First, as a static equilibrium-based model, it assumes that the network operates in a steady state. This means that any disruption being considered is assumed to last for the entirety of the period being analyzed (7:00 to 10:00 for morning peak, 10:00 to 16:00 for inter-peak, or 16:00 to 19:00 for evening peak are the three typical study periods). Although this is an obvious simplification of
reality, it allows like-for-like comparisons between disruptions without the complications of modeling service recovery after a disruption.

Second, the steady-state assumption implies that demand never exceeds capacity since, if it did, queues would build up and the network would no longer be in a steady state. This is remedied by considering denied boardings implicitly, as part of the crowding discomfort factor calculations, instead of explicitly, by estimating the number of passengers denied boarding. Although this approximation impacts the results of any analysis because it does not consider the fact that passengers denied boarding are delayed before boarding links farther downstream, with a well-calibrated crowding penalty function, this effect can be approximated. Railplan’s crowding function is discussed in more detail in Section 4.3.6.

Third, as with any iterative, equilibrium-based model, Railplan assumes that passengers consider the weighted travel time of each link when selecting a path through the network. This means that passengers are assumed to be fully aware of the weighted travel time - and therefore the crowding - on each link. This may be a realistic assumption for a strategic model that is designed to model a typical day with no disruptions, because over time passengers will likely have explored possible paths and learned about the crowding on each. These assumptions are flawed when modeling disruptions, however; this motivated the development of the imperfect knowledge Railplan model, described below.

Railplan generates a wide range of outputs. The highest-level of these is a total number of (generalized) passenger hours spent on the network by all passengers, weighted to include discomfort factors associated with crowding, walking, etc. as specified in the JTM. For disruption analysis, this is useful as a bottom-line figure that, when compared to the equivalent figure for an undisrupted network, represents the impact of a disruption on passengers. This is then broken down into real hours (unweighted by crowding) and crowding penalty by transport mode. Further subdivisions are at the link and line level, which allow for many comparisons, either numerical or graphical. Figure 2-2 depicts the Railplan results interface comparing the passenger volumes for the Underground network between the 2021 reference case (no disruptions) and a scenario modeling the suspension of the Piccadilly line service from Kings Cross to Hyde Park Corner on a link-by-link basis.

Green links are those on which the undisrupted case has a greater passenger volume, and red links are those on which the disrupted case has a greater volume; the thickness of the bar corresponds to the magnitude of the change in passenger flow. The line with large green bars is the Piccadilly line; as expected, during a disruption when its frequency is greatly reduced, its passenger volume decreases. The Victoria line, which roughly parallels the Piccadilly through central London, sees large increases in volume. Other lines see minor increases or decreases in volume because some passengers change their route due to crowding shifts on the network.
Imperfect Knowledge Railplan Approach

The base Railplan approach described above makes several assumptions that make it imperfect for modeling the impact of unplanned disruptions.

First, as discussed above, passengers are assumed to have perfect knowledge of the disrupted service characteristics. In the modern day, the prevalence of smartphones and other media for real-time provision of service information means that this assumption is not as critical as it once was. Nevertheless, there are still some passengers - those who do not check the service status before beginning their journey, and those who are en route when a disruption occurs and do not hear the relevant announcements - who only change their behavior when they directly encounter a disruption, either by being unable to travel over a suspended section or by encountering extreme crowding when attempting to board a train. In the literature, the fraction of those passengers assumed to be aware of the disruption is typically taken to be 53% (Yap, 2019).

Second, passengers are assumed to have perfect knowledge of changes in service patterns on the portion of the line/network not affected by the disruption. Often, an unplanned disruption will cause frequencies to change on segments of the line that are not directly affected, whether by increasing (as on the Northern line when one of the two branches through central London is disrupted, and trains are rerouted via the other branch) or decreasing (during most partial line suspensions controllers increase headways to avoid overloading a temporary terminal). Because such frequency changes are usually ad-hoc, aside from a high-level message such as “Severe Delays” there is usually no information about them provided to passengers.
Third, passengers are assumed to have perfect knowledge of real-time and future crowding levels, which will be the result of other passengers’ behavior. For a steady-state network, a passenger’s daily experience will develop knowledge regarding the network’s typical state. In an assignment model such as Railplan, this is typically replicated through iteration, where passengers are able to take the crowding encountered by previously assigned passengers into account when planning their route. In an unplanned disruption, however, aside from frequent exposure to similar disruptions, there is no way for passengers to understand how other passengers will behave on other parts of the network. While real-time train crowding data is available internally and could potentially be made available to passengers, in the near future it is unrealistic to expect output from an algorithm that could predict crowding fifteen to thirty minutes in advance, which would be necessary for passengers to reliably re-plan their trips across the network, to be regularly provided to passengers.

Fourth, the system is assumed to be at equilibrium, with all passengers choosing the route that is best for them. During an unplanned disruption, passengers have no way of knowing which route is optimal, because they do not know the crowding conditions elsewhere on the network. Finally, the effects of many severe disruptions on the Underground violate one of the primary assumptions behind an equilibrium-based model - that supply exceeds demand.

These assumptions reflect the nature of Railplan’s primary intended use - strategic modeling of steady-state conditions with no disruptions or, in other words, what the network will look like and how passengers will behave “on a typical day.” However, these assumptions are not appropriate for modeling unplanned disruptions. To remedy this mismatch of tool and application, Transport for London staff have developed a methodology they deemed “imperfect-knowledge” which is intended to represent passenger behavior under disruptions more accurately while still using the Railplan toolkit. The methodology involves taking the results of a base-case scenario and modifying it as follows (Yap, 2019):

1. A disruption, such as a partial line suspension, is introduced. All links on which the frequency of service is changed are denoted as ‘affected.’
2. All passengers traveling over affected links are denoted as ‘affected passengers’ and removed from the network.
3. Affected passengers who are aware of the disruption (assumed to be 53% of all affected passengers) are assumed to be aware of headways, running times, (transfer) waiting time and crowding levels as in the undisrupted case, but not how these parameters change during the disruption. These passengers are reassigned to the network.
4. Affected passengers who are unaware of the disruption (the remaining 47% of all affected passengers) are assumed to be aware of headways, running times, and (transfer) waiting time as in the undisrupted case, but assume no crowding on the affected line. These passengers are reassigned to the network. This lack of crowding awareness simulates “route stickiness,” or the tendency of passengers to prefer their originally planned route where possible; in particular, passengers unaware that a disruption is ongoing will proceed on their usual route until they are unable to continue.
Application of this procedure to the “Bank blockade” - the planned closure of the Bank branch of the Northern line (the upper branch through Zone 1 in Figure 2-3) for several months - produced interesting insight. Although this is a planned closure, modeling it as an unplanned one gives some insight into which routes passengers initially prefer, as compared with those that they prefer once they understand the network’s new levels of crowding. This allows Transport for London to develop passenger messaging to discourage use of the most overcrowded links.

This case study predicted significant changes in demand in three main areas of the network. First, the Morden branch of the Northern line sees a decrease in demand, instead of an increase. This is to be expected, considering the elimination of Northern line trains via the Bank branch, which increases journey times to stations on those branches and encourages passengers to pursue other routes. The base Railplan model, however, assumes that passengers are aware of the 1 tph increase in Morden branch service that will be operated during the closure, and accordingly assigns more passengers to the branch, which is unrealistic. The imperfect knowledge model’s finding suggests that the Morden branch is unlikely to become significantly more crowded as a result of the disruption, which allows TfL staff to focus their passenger messaging on other sections of the network.

Second, the Northern line services to/from Moorgate see more demand in the imperfect knowledge model than they do in the base model. Because of the track configuration at Moorgate, its throughput as a terminus is restricted to approximately 16 tph, which is insufficient to accommodate the prior demand on that section. If passengers are unaware of the frequency reduction, as is the case in the imperfect knowledge model, they will all attempt to use these services and overload them. The base Railplan model, on the other hand, assumes that passengers have had experience with overcrowded trains on this section of the line and will pursue alternative routes if available. These findings suggest that passenger messaging is necessary for services to/from Moorgate, advising passengers of alternate routes to reduce crowding.
Third, the Jubilee line is significantly busier in the imperfect knowledge model results than in the base model. The Jubilee line provides a fast connection between the two branches of the Northern line (at Waterloo and London Bridge) and is thus ideal for many passengers affected by the closure. Even today, without a disruption, the Jubilee line is over capacity on this section, so the addition of Northern line passengers will result in significant overcrowding and denied boardings. In the imperfect knowledge model passengers do not take this into account, while in the base Railplan model Jubilee line passengers (even those who do not normally use the Northern line) choose to change their route to avoid the crowding on this section. As with the services to/from Moorgate, this suggests that passenger messaging along the Jubilee line is advisable to discourage passengers from using it.

As of this writing, full-scale applications of this approach to disruptions on the Underground are still being explored, and because of the way it is configured the model cannot generate bottom-line aggregate LCH numbers as does the base Railplan approach. Therefore, for creation of the incident library and for modeling most incidents, TfL staff continue to use the base Railplan model. The track layout analysis in Chapter 3 also uses the base Railplan approach. The passenger behavior assumptions outlined here have similarities with those of the simplified assignment model proposed in Chapter 4; a comparison of results between the two Railplan methodologies and the simplified model will also be presented in Section 4.4.

2.2.4 Retrospective (EJT) Approach

Excess Journey Time (EJT) is a newer metric that aims to provide a more holistic understanding of the network’s performance. Unlike LCH, which accounts for delays on an incident-by-incident basis, EJT uses actual train movement data to compute aggregate journey times on a daily basis. This corresponds more closely to the service that passengers experience, because it considers all deviations from the schedule (not just those caused by incidents). However, the measure’s aggregate nature means that it is difficult to determine the root cause of any particular deviation from the schedule. This is particularly true when multiple incidents occur in the same period, or when a disruption at one end of the line affects crew changes at the other end of the line hours later, because service recovery has left drivers off schedule. There are simply too many variables in the operation of a metro - passenger volumes and actions, driver door-closing behavior affecting station dwell time, driver behavior affecting run times (on manually driven lines), reversing times at termini, and controller decisions during service recovery - for strictly causal relationships between incidents and delivered service to be determined.

Excess Journey Time consists of the following components:

On Train Time: this measures the time that passengers spend on the train, not weighted by crowding discomfort factors. The train movement component of this calculation is drawn from NETMIS. The passenger volumes are computed using an assignment process that computes ‘line-
loads’ on line sections, each made up of several stations, using historical RODS survey data at the 15-minute timeband level.

Platform Wait Time: this measures the time that passengers spend waiting on the platform, not weighted by crowding discomfort factors. This is determined using the same assignment process as in OTT.

On Train Crowding: this is a crowding discomfort factor based on the ratio of how many people are on the train to the seated and crush capacities of the train. It is calculated using the same assignment process as in OTT. This function is quadratic in the number of passengers, which reflects the increasing discomfort passengers experience as their train becomes more crowded.

Left Behind: this is a penalty associated with not being able to board the first train to arrive, due to crowding. It is calculated using the assignment process described above.

On Train Delay: this is an additional penalty associated with on-train time being worse than scheduled.

Short Tripping: this is a penalty for passengers forced to alight from their trains and wait when their initial train is reversed prior to its scheduled destination. Because this calculation assumes that passengers board the next train (if there is space), this is used for single-train short turns and not partial line suspensions, where all trains passing through a given location are reversed.

Platform Crowding: this is a crowding discomfort factor based on passenger congestion on the platform. It is calculated using the passenger assignment process described above.

Planned Closures: this takes into account the passenger disbenefit of closures for construction work.

Unplanned Closures: this considers unplanned line suspensions, station closures, etc.

Today, EJT is used as a way to assess overall network performance. Because it captures the service as delivered and does not rely on accurate incident recording or similar modeling assumptions, it is considered to provide a more holistic view of passenger experience on the Underground. Efforts are underway to resolve its main drawbacks - its aggregate nature and lack of attribution of delay to specific incidents - by breaking the data down into 15-minute timebands and to station-to-station segments; however, even this is not likely to lead to causal relationships between incidents and delays in the way that LCH attempts to do.
3 Track Layout Analysis

This chapter describes a methodological approach to analyzing the track layout of the Piccadilly line and estimating the benefit of proposed improvements to the track layout. It begins with a literature review, outlining past research on track layout design and analysis. Then, London Underground’s current approach to analyzing modifications to its track layout is described. This consists of two components: an estimate of how often each track layout component (e.g. crossover or siding) is used, and a business case analysis that converts this usage frequency into an economic benefit figure and compares it to the cost of maintenance and upkeep. Next, the analysis procedure for the addition of a new crossover is summarized, and applied as a case study to three proposed crossovers on the Piccadilly line, at Covent Garden, Turnham Green, and Alperton. This analysis begins with a review of how controllers use the existing track layout and the factors that influence their choice of how to route trains during a disruption. Next, the relationship between track layout and terminal capacity is discussed, and several factors that increase track occupancy time, thus decreasing capacity, are outlined. A strategy for remedying this issue, thus increasing capacity on the Piccadilly line, is proposed. Then, the proposed modeling framework for quantifying the benefit of a new crossover is presented, and applied to the above three crossovers in the context of unplanned disruptions and planned weekend closures. Results and business case outcomes are presented. Limitations in the methodology are discussed and extensions are proposed to more fully quantify the potential benefits of a given crossover.

3.1 Literature Review

The relationship between track layout and metro line performance is underexplored in the literature. This relationship takes two main forms: micro- and macro-scale. Micro-scale analysis assesses the capacity of a terminal, junction, or other interlocking (a set of signals and switches at a location) and then relates that to the line’s overall capacity and reliability. For example, the following equation, adapted from the Transit Capacity and Quality of Service Manual (Transit Cooperative Research Program, 2003), relates the parameters involved in the operation of a two-track terminal with the crossover located before the station:

\[
\frac{t_j}{2} \leq \left( H - t_s - \sqrt{\frac{2(P + T + CS)}{a_s + a_d}} - \sqrt{\frac{(P + T + CS)}{2a_s}} \right)
\]  

(3-1)
where:
\[ a_s = \text{initial service acceleration rate} \]
\[ C = \text{switch angle factor (5.77 for #6 switch, 6.41 for #8 switch, and 9.62 for #10 switch)} \]
\[ d_s = \text{service deceleration rate} \]
\[ H = \text{train headway} \]
\[ P = \text{platform length} \]
\[ S = \text{track separation} = \text{platform width} + 1.6 \text{ m (5.25 feet)} \]
\[ T = \text{distance from cross-over to platform} \]
\[ t_l = \text{terminal layover time (in seconds)} \]
\[ t_s = \text{switch throw and lock time} \]

Figure 3-1 shows this configuration.

![Figure 3-1. Example terminal configuration](adapted from Transit Cooperative Research Program, 2003)

Though this equation, as written, determines the layover time necessary to support a given headway, it can be rearranged to calculate any variable. For example, if the minimum possible layover time is controlled by immutable factors such as the time needed to charge a train’s brakes\(^2\), the minimum possible headway can be calculated by rearranging Equation 3-1 as follows:

\[
H \geq \frac{t_l}{2} + t_s + \frac{2(P + T + CS)}{a_s + a_d} + \sqrt[2]{\frac{(P + T + CS)}{2a_d}}
\]  
(3-2)

Researchers such as Gill (2000) and Hong (2017) relax the assumption of track configuration and discuss the capacity implications of various switch configurations for metro line termini, and illustrate how a simulation model can aid such analysis. Van Oort and van Nes (2010) take this process one step further and apply such a capacity analysis to comparing different types of termini. They find that, of the three common types of two-track termini, a terminal with a crossover and two tail tracks located behind the platform can best handle high train frequencies.

---

\(^2\) When a train is dwelling at the terminal and there is no operator in the cab, the train’s brakes (which operate on compressed air) are in the emergency position, which means that the train’s brake pipe is not pressurized. When an operator enters to begin the return trip, the train’s air compressors must refill the brake pipe with compressed air so that the brakes can be released.
There are fewer studies related to macro-scale design of the track layout of an entire metro line, however. Most practitioners who are tasked with adding or removing interlockings on existing metro systems, or designing new lines, use a combination of site-specific constraints, comparisons with other systems, and their experience, instead of any formal scientific approach.

Historically, track layouts on new lines were governed by a set of very general principles. When the four oldest subway systems in the United States (Boston, New York, Chicago, and Philadelphia) were built in the early twentieth century, for example, the recommended guidance was to locate crossovers approximately one mile apart “to make operation of a part of the line possible in case of accident” (Glaser, 1918). The same article recommended the addition of sidings “at convenient points along the line” for storage of disabled trains and extra trains for rush hour service, but left the question of their location unanswered. Another journal article from the same period merely noted that “engineers have tried to avoid… switches, crossings, etc. in places other than close to stations where trains must stop, and to locate them on the farther side rather than on the near side” of stations (Lavis, 1914). This enables minimizing delay, both from waiting for switches to be aligned and from reducing speed because of switch geometry, on interstation segments where trains should be moving quickly. Even today, comprehensive texts such as Pyrgidis (2016) can be very brief when discussing crossover placement, recommending only that it “account for maintenance activities [and] emergency cases.”

In the modern day, some approaches have become more scientifically rigorous. Abramovici (1986) developed a detailed formulation relating desired frequency, crossover spacing, and operating pattern (running trains express in the non-peak direction, or fleeting, for example). However, his main objective was placing crossovers to permit single-track operation in emergencies. This is only practical on low-frequency rapid transit systems; Washington DC’s WMATA, for example, uses such service patterns to accommodate maintenance work, but suffers long headways as a result, and so restricts such operation to weekends unless absolutely necessary (WMATA, 2018). Abramovici notes that in cases of very high traffic, any single track operation will need to be accompanied by route shortening, frequency reduction, or a decrease in the number of trains on the line. On a system such as the London Underground, with scheduled headways as low as 100 seconds, such strategies are already implemented regularly after even a minor incident: the Victoria line, for example, has more trains operating during the rush hour than platforms to accommodate them, so even the slightest perturbation will quickly cause backups with trains held in tunnels (Trendall, 2017). Given the intensity of service, single-tracking is clearly infeasible for unplanned or planned disruptions on most Underground lines, and the primary use of a crossover (or siding) in an emergency will be to reverse trains.

The increase in complexity of modern-day construction, which results from denser cities and more stringent safety regulations than existed decades ago, means that constructability and safety considerations often prevail when determining crossover locations. In the 1990s, during the construction of the Jubilee Line Extension, locations of crossovers and reversing sidings were initially proposed at high-traffic locations, both for emergencies and for regular short-turns of some trains. However, tunneling considerations (particularly ground settlement) and cost concerns
ultimately proved decisive in selecting locations for reversing facilities (Mitchell, 2003). During
the design of Crossrail 1 and 2 in the 2000s and 2010s, three main operational objectives were
pursued for proposed reversing locations - maintaining service over the unaffected part of the line
in case of an incident, access to and around the worksite for engineering trains and other
maintenance vehicles, and removal of a disabled train in an emergency (Richardson, 2019).
However, the primary determinants of crossover locations to meet these objectives are still
engineering decisions including settlement and other constructability issues, as with the Jubilee
line, as well as factors involving ventilation and fire safety (Richardson, 2019).

3.2 Current Practice

In the modern day, when government agencies large and small must contend with limited
budgets and competing priorities, a simple demonstration of net benefit is insufficient to justify
investment in any infrastructure project. A comprehensive assessment of the costs and benefits of
a project - economic and otherwise - must be developed. The resulting business case, outlines the
various alternatives available for a given project, analyzes the benefits and drawbacks of each, and
recommends a course of action. Business cases are often thought of as being purely economic, but
it is also important to consider any strategic implications.

The strategic case for a project is, broadly, the degree to which it helps an organization
achieve its broader goals. For example, Policy 14 of the London Mayor’s Transport Strategy seeks
to make “the transport system navigable and accessible to all and [reduce] the additional journey
time that disabled and older users can experience” (Greater London Authority, 2018). A project
proposing the installation of elevators at a Tube station, for example, would make it quicker and
easier for disabled people to get around London, thus aligning itself with the Mayor’s Transport
Strategy and TfL’s strategic goals. This alignment could form part of the business case, or allow
the project to receive funding even if the expected usage of the proposed elevator is not high
enough to justify the project on purely economic grounds. Strategic factors, in combination with
legal, political, and other constraints, account for the real-world needs of organizations and the
people they serve in a way that a purely economic calculation omits.

Once the strategic justification for a project is understood, the economic case must be
made. This is an examination of the two sides of any project - benefit and cost - and calculation of
a benefit-cost ratio (BCR). Because TfL is a public entity charged with delivering a social good,
both benefit and cost are computed not only in monetary terms, but also in terms of impact on
society. These impacts on society can be broadly divided into three levels by scope. The first
includes directly quantifiable metrics that impact a passenger’s journey. This includes factors
related to journey time - actual in-vehicle time, waiting time, access time, discomfort penalties due
to crowding, etc. This is the primary factor behind most investment schemes for minor projects,
and is used by models such as Railplan. The second level includes broader impacts on society,
including improvements to reliability, pollution, noise, ride quality, appearance, etc. These metrics
still impact the journey of a given passenger, but also include externalities affecting society as a
whole. The third level is the broadest and attempts to quantify the impact of a project on large-
scale changes in land use, commuting patterns, and economic activity. This type of calculation is generally found in the business cases for megaprojects such as Crossrail 2 (Haylen, 2019).

Costs are the other side of the equation in any business case, and include a wide variety of factors. First, any disbenefits to passengers or society as a whole must be quantified. For example, if the construction of a new above-ground line causes an increase in noise for surrounding residents, this is taken into account as a negative externality. Most costs, however, are more directly economic. The most obvious is up-front cost, which includes all expenses associated with the initial delivery of a project - design, construction, project management, administrative overhead, etc. Ongoing expenses must also be considered, which can include additional fuel required for operation of an asset, extra staff costs, and maintenance on the new infrastructure. Finally, life-cycle costs must also be considered. Because even the best-maintained assets eventually fail and require overhaul or replacement, a full business case will consider all these costs.

TfL’s approach to analyzing the benefit of crossovers and sidings follows the business case methodology (Transport for London, 2017b) and considers several different categories of costs and benefits. The analysis behind the 2016 removal of the siding between Holborn and Tottenham Court Road on the Central line (Powell and McInulty, 2016), for example, considered benefit with regard to planned closures, unplanned closures, the ‘bolt-hole effect’ (explained below), and speed increases for all through trains; aside from the additional fare revenue expected as a result of providing a faster and more reliable service, all these are benefits to passengers, which are converted to an economic benefit to society using value-of-time calculations. The frequency of planned closures was estimated based on historical and expected maintenance needs, and unplanned closure rates were estimated using CuPID data. The ‘bolt hole effect’ refers to the ability to store a failed train in a siding, out of traffic, until the end of the service day. By getting this disabled train out of the way of normal service faster, the impact of a disruption is reduced. However, the analysis found that this use of a siding occurs very infrequently in practice, given the comparatively low number of severe fleet-based failures and the desire to immediately move a failed train to the depot to avoid disrupting service later in the day. Speed increases for all through trains are, aside from capital renewal costs, the dominant factor in the business case for the Central line siding. The presence of a switch will sometimes require lower speeds due to altered signal block lengths or to reduce wear and tear on switch components. Because they apply to every train every day, even a slight increase in maximum speeds that saves a few seconds can create a significant total passenger benefit. For the analysis of crossovers in this chapter, it is assumed that speeds would not be decreased as a result of the installation of a given crossover. This is of course a question of engineering, which is beyond the scope of this thesis, but a 2006 analysis of reinstalling the crossover at Covent Garden (Tube Lines, 2006) found that speeds for through trains would not be reduced.
3.3 Analysis Procedure and Results

This section examines the benefits that modifying track layout can have on Piccadilly line performance. After a summary of the existing track layout and factors that influence line controllers’ decisions, two types of analysis are presented. The first explores the parameters that affect the capacity of a terminal, and how these factors cause congestion at Arnos Grove. Suggestions for mitigating this congestion are presented. The second involves the use of Railplan to calculate the benefit that a crossover has on mitigating the impact of unplanned partial line suspensions and planned weekend closures.

3.3.1 Assessment of Existing Track Layout

The first step in enhancing the track layout of a metro line is understanding the existing track infrastructure and how it is used. The Piccadilly line’s track layout is presented in Figure 3-2.

Conversations with line controllers gave insight into the criteria they use when deciding how to use the track layout to get service back on schedule after an incident, given the control strategies described in Section 1.4.3. One critical factor considered is passenger management. When a train is reversed, all passengers on board exit, either because the train will be turning around mid-tunnel or in a siding (where, for safety reasons, passengers are not permitted) or simply because it will be heading in the opposite direction and thus not continuing in the direction those passengers are traveling. Once passengers have alighted from the train, they must either wait for the next train continuing in the same direction (if there is one), in which case the platform must have space to accommodate them in addition to the passengers already there, or use the station passageways to transfer to another line or to leave the Underground entirely, in which case these passageways must have sufficient capacity. This is a particular concern at stations such as Covent Garden that do not have escalators and rely solely on elevators for vertical circulation, with very limited capacity. The presence of such stations usually require the train to be cleared of passengers one station before it otherwise would have been. Although passenger crowding is unlikely to prevent the use of a given crossover to turn trains, it can have major implications on the way service is operated during a disruption, and is thus important to consider when planning the services to operate during a disruption.

A second factor is reversing time. This includes three main components: train movement time, crew movement time, and passenger movement time. (Other components, such as the time for signals to clear and switches to be realigned after a train reverses, are minor in comparison). Train movement time is the easiest to determine, using either simulation or direct calculation similar to Equation 3-1. The only inputs necessary are properties of the rolling stock - train length and acceleration/deceleration curves - and track information - geometry and speed limits.

---

3 This scenario, where the train reverses beyond the station platform, is referred to as a “relay terminus” (Lee, 2002). The other case, where the train reverses while dwelling in the station, is referred to as a “stub-end terminus.”
Crew movement time is more difficult to predict because of variability across train operators and, when trains are reversing at a platform, the impact of platform crowding on operator walking time. According to Underground staff, it is this component that causes the greatest uncertainty in capacity calculations, and encourages the use of rules of thumb instead of precise calculation. The third component, passenger movement time, is even more difficult to predict because it depends on factors as wide-ranging as quality of information provision, crowding, and speed of passenger alighting. However, passenger movement time is generally less of a constraint on overall reversing time than the other two factors. Whether the reversing point is stub-end or relay, the time needed for the operator to walk to the other end of the train generally exceeds that needed for passengers to alight. These three components together give reversing time, of which the inverse is maximum possible frequency. If this frequency is judged to be too low, controllers will sometimes opt not to operate a service because trains and stations would quickly become unsafe due to overcrowding, and will adapt their contingency service pattern to ensure passenger safety.

A third factor considered by controllers is asset reliability. Controllers are unlikely to order a train to use a crossover if they believe it may malfunction. The rarity of use of most crossovers and the rarity of failure of each asset mean that such judgements are usually made based on a controller’s personal experience, rather than historical asset performance data. Generally, the crossovers that are believed to be less reliable are those which are inspected and maintained infrequently. The ones that receive less frequent inspection and maintenance are usually those that are used less frequently (because they are believed to be least critical to service). This (circular) reasoning leads to the conclusion that crossovers which are seldom useful fall into a cycle of disuse and under-maintenance and are not trusted to operate even when they could help maintain as much train service as possible. As a result, effort should be taken during major projects such as line signal upgrades to ensure that each crossover is useful in a variety of disruptions, planned and unplanned, so that it will be used regularly. For a given track layout, it is important to remember that the mere presence of a crossover does not guarantee that controllers will use it.

A useful complement to understanding controllers’ rationale for their decisions is data which shows the frequency with which each reversing location is used for unscheduled short-turns. Because there is no digitized, central log of controller actions, obtaining this data requires inferring these actions from NETMIS logs of train location. Venancio (2016) and TfL staff created two algorithms to perform this inference. The results of applying the TfL algorithm to NETMIS data from the last six months of 2018 from the Piccadilly line are presented in Figure 3-3. As Venancio noted in her work, NETMIS data is frequently incomplete, missing data such as train ID or location for any given entry. Although her algorithm inferred some of this missing data, and LU has made efforts to improve the quality of the data, gaps remain. Therefore, the data in Figure 3-3 is best interpreted as a measure of where trains are reversed most often, as opposed to exactly how often they are reversed there.
The four most commonly used short-turn locations are on the outer ends of the line, which suggests that the dominant purpose of unscheduled short turns is to return late trains to schedule. Arnos Grove, on the eastern end of the line, is the most common reversing location because it is well configured to operate as one. The presence of a third platform, located between the two main platforms, means that trains can reverse without impeding through traffic and without needing to detrain\(^4\). In addition, a major crew depot is located at Arnos Grove, making it convenient for controllers to reverse and reform (assign to another scheduled slot) trains there because a new train operator is likely to be boarding anyway. Finally, Arnos Grove is located nine minutes from the Cockfosters terminal, enabling a short-turning train to recover fifteen to twenty minutes of late running, which is just above the lateness threshold at which controllers begin short-turning. These factors combined create an incentive for a large number of short turns. However, as discussed in the following section on terminal capacity, this frequent use of Arnos Grove for short-turns can cause operational problems.

The second most common reversing location, Northfields, is convenient for many of the same reasons as Arnos Grove. Its two extra platforms mean that trains can be reversed and detrained without interfering with through traffic. It also serves as both a crew and train depot, meaning that reformations or withdrawals from service are easy to implement.

Wood Green, two stations away from Arnos Grove, is usually used to short-turn trains scheduled to turn at Arnos Grove. This station has a siding located to the north of the platforms. The two main constraints that preclude controllers from using it more frequently are detraining

\(^4\) London Underground requires trains that are reversing between stations to be emptied of passengers. Detraining is the process of making sure that all passengers have exited the train.
time and crew reliefs. Because trains at Wood Green reverse using the siding, the train operator and station staff (if any are present on the platform) must make sure that the train is empty before the train enters the siding. This takes time and has the potential to cause knock-on delays on following trains. Crew reliefs are a constraint because train operators who are scheduled to be relieved (for a meal break, or at the end of their shift) at nearby Arnos Grove cannot be short- turned and sent on another round trip.

South Harrow is generally used to short-turn trains terminating at Rayners Lane. It does not have a reversing siding or a third platform, but the relatively low train frequency on the Rayners Lane branch means that reversing in the station generally does not delay other trains. The line’s other reversing locations are used only sporadically. Hammersmith, Acton Town, Green Park, and Kings Cross are generally used during service suspensions, for getting disabled trains off the line or to the depot, or for schedule recovery after a major disruption. Ruislip, Heathrow Central, and Hatton Cross are usually used for recovery from routine train lateness.

3.3.2 Terminal Capacity Analysis

Before considering the relationship between a line’s track layout and its performance during disruptions, it is worth examining the line’s operation on a normal day without disruptions, focusing particularly on termini. In high-frequency metro systems, termini are often the binding constraint on headway (Lee, 2002). This is primarily due to two factors: conflicting train movements and high platform occupancy times. Conflicting train movements are caused by both geometric and directional conflicts. Geometric conflicts are inherent in any stub-end or relay terminal that has more than one track, as is shown in Figure 3-4, where for both terminal layouts the path of the blue train (which is ending its run) conflicts with the path of the red train (which is starting its run).

![Figure 3-4. Geometric conflicts at relay (left) and stub-end (right) terminals.](image)

Whether before or after a station, there must be a crossover to allow trains to change tracks; inevitably, some trains crossing from one track to the other as they arrive at the terminal will conflict with a departing train. Solutions involving grade separation to avoid this issue are generally impractical in a rapid transit systems due to space and budget constraints. Directional conflicts, also inherent to all non-loop termini, are caused by the simple fact that trains reverse direction at a terminal. Because the train’s entry to and exit from a terminal platform (or siding, in the case of relay termini) are at the same point, a train must wait for its predecessor to clear the platform and switch (plus a safety buffer) before beginning to enter. Whereas at a mid-line station a train can approach or enter a platform as the train ahead is leaving, which increases capacity, at a terminal this is not possible.
Unlike conflicting train movements, high platform occupancy times are influenced almost entirely by factors under the agency’s control. They are still critical to overall terminal capacity, of course: by definition, a terminal cannot accommodate more trains at a time than it has platforms, so for steady-state operation one train must depart for each one that arrives. Therefore, the platform occupancy time of each train at the terminal (which includes dwell time as well as entry and exit time) cannot exceed the headway multiplied by the number of platforms. Because entry and exit times are (relatively) constant, the critical variable in terminal capacity is dwell time. This same reasoning also applies to mid-line stations: the lower bound on headway is platform occupancy time within a station, which is easiest to modify through lowering dwell time.

As mentioned in the previous section, Arnos Grove on the Piccadilly line is used frequently for unscheduled short-turns. In addition, several trains per hour are scheduled to reverse there. Because a train reversing has higher platform occupancy time than a through train (as shown below), both of these factors have the potential to cause delays due to capacity constraints. Also, since Arnos Grove is the location of a major crew depot, many operators board and alight there. Even when everything is running on schedule and all operators are where they should be, this inherently causes some additional dwell time. During disruptions, when operators are frequently out of place and the ad-hoc nature of controller action makes the risk of miscommunication greater, this problem is compounded.

Figure 3-5. Waterfall diagram illustrating blocking back  (Transport for London, 2020d)

Congestion entering Arnos Grove is commonly cited by LU staff as one of the critical constraints facing the Piccadilly line. Despite its peak service level of 24 tph, which is comparatively low by Tube standards, regular “blocking back” - train congestion as trains await an available platform - causes delays to eastbound trains approaching Arnos Grove (and the passengers on them) on a daily basis. Figure 3-5, a fragment of a “waterfall diagram” (also called a Marey chart, or time-space diagram) of the eastbound Piccadilly line on February 21, 2019, illustrates this phenomenon. No incident occurred near Arnos Grove in this period, but train 243-9 took ten minutes to travel from Bounds Green (BGR) to Arnos Grove (AGR), likely due to interference from train 322-8, which was itself delayed by a westbound train leaving Arnos Grove. This delay to train 243-9, in turn, affected at least three subsequent trains.
A two-part analysis is conducted to examine capacity at Arnos Grove. The first part seeks to quantify this problem by analyzing eastbound running times in the area. The second part aims to determine the degree to which two factors - reversing and crew changes - affect the platform occupancy of a given train. For the purposes of this analysis, NETMIS wheel-stop to wheel-start data are used; therefore, entry and exit times are not considered. These data are calculated by applying offsets, representing entry and exit times, to the track circuit occupancy times. Although this analysis does not prove a causal relationship exists between high platform occupancy and blocking back, which would require simulation, it does establish the scope of the problem and examines some of the parameters that could be addressed to mitigate its impact.

For quantifying eastbound running times, the eastern section of the line (from Kings Cross to Cockfosters) was divided into four segments: Kings Cross to Finsbury Park (segment 1), Finsbury Park to Wood Green (2), Wood Green to Arnos Grove (3), and Arnos Grove to Cockfosters (4). An overview of this section of the line is shown in Figure 3-6.

Figure 3-6. The eastern section of the Piccadilly line (adapted from Transport for London, 2020e)

Segment 1 is used as a control from a terminal congestion standpoint, because it is far away from any regularly used terminal and thus can be considered to have typical mid-line free-flow conditions. It is a control to determine the effect that other factors, most notably increased dwell times during rush hour, may have on travel time. Segment 2 is approaching Wood Green, where some trains are reversed if they are behind schedule. In addition, if terminal congestion approaching Arnos Grove is severe, the queue of trains can reach back into this section. Segment 3 is the primary section that suffers from terminal congestion on a routine basis. Segment 4 is not constrained by the Arnos Grove capacity; any delays on this section are caused by capacity constraints at Cockfosters or by trains entering and exiting the depot at Oakwood.
Figures 3-7 (a)-(d) show average running times by hour as derived from NETMIS data for weekdays in November 2018. For each section, travel times of greater than 60 minutes are excluded, as these are assumed to be caused by either data errors or significant disruptions in the area, which are outside the scope of this terminal capacity analysis.

(a) Kings Cross to Finsbury Park (Segment 1)  (b) Finsbury Park to Wood Green (Segment 2)

(c) Wood Green to Arnos Grove (Segment 3)  (d) Arnos Grove to Cockfosters (Segment 4)

Figure 3-7. Average run time: Piccadilly Line eastbound (by section)

Segment 1 shows the expected distribution for a mid-line section. The early afternoon run time, which is used as a baseline due to the lack of crowding or crew reliefs, is approximately 8.5 minutes. The highest run time during the evening peak is 9.6 minutes, which is an increase of 12.9% over the early afternoon. Considering that this is a heavily used section of the Piccadilly line, this is likely due to higher dwell times associated with higher peak ridership.

Segment 2 shows a higher variance of run time. The early afternoon run time on this section is 8.0 minutes, while the evening peak run time is as high as 9.4 minutes. This is a difference of 17.5%. Because many passengers transfer from the northbound Victoria line to the eastbound Piccadilly line at Finsbury Park, this section of the line also has high ridership. The variation in run time reflects the influence of high ridership on dwell time, as well as some effect of terminal operations. When terminal congestion at Arnos Grove is severe it can impact this section; in addition, any train reversing at Wood Green will likely cause a brief backup on this section as it detrains and is sent into the siding.
Segment 3 clearly shows the impact of the terminal capacity constraint at Arnos Grove. The early afternoon shows a run time of 6.7 minutes, while the evening peak run time is as high as 8.4 minutes (a difference of 25.3%). In addition, the morning peak run times on this section are even worse than in the evening peak: at 10:00 the average run time is 8.6 minutes (28.3% higher than the baseline). Passenger volumes are lower on this segment than on Segments 1 and 2, so high dwell times likely do not account for this increase. Rather, it is caused by a combination of several factors - train crew changes, reversing, and stabling of trains into the Arnos Grove siding - that take place concurrently. Determining which of these factors impacts terminal congestion most significantly, as discussed below, is critical to mitigating delays.

Segment 4 shows that the train service before the late evening largely returns to normal once trains pass Arnos Grove. With trains on this section of the line relatively uncrowded, run times are fairly consistent for much of the day. The main spike in run times occurs around 22:00, when trains are being withdrawn from service at Oakwood and Cockfosters and sent to the depot. Just as at Wood Green, trains being withdrawn from service must be verified by staff to be empty. This takes time and delays other trains on the line. However, because such withdrawals usually take place during the late evening hours and do not impact the rush hour, they are not considered further in this study.

The next step in the analysis is determining which factors affect eastbound dwell times (which contribute significantly to congestion) at Arnos Grove. Two factors are considered - reversing and crew changes. These data are drawn from three days (February 19-21, 2019) which had minimal disruptions. All trains with a recorded dwell time of zero seconds or less (which constitute 2% of all trains) are removed from the sample, since these are assumed to be data errors or out-of-service trains. Figure 3-8 shows the distribution of eastbound dwell times for four categories of trains: terminating (reversing) trains with crew changes, terminating trains without crew changes, through trains with crew changes, and through trains without crew changes.

![Distribution of Eastbound Dwell Times at Arnos Grove](image)

Figure 3-8. Distribution of eastbound dwell times at Arnos Grove
crew changes, through (non-reversing) trains with crew changes, and through trains without crew changes. NETMIS calculates these dwell (wheel-stop to wheel-start) times by applying an offset to track circuit occupancy times, which are derived from signal data.

These results clearly demonstrate that, as expected, trains which reverse at Arnos Grove have significantly higher dwells than those that do not. A large share of this difference is due to scheduled layover (terminal dwell) time, which is approximately six minutes (360 seconds) for most trains during the peak periods. The median dwell times for terminating and through trains differ by just under six minutes, which suggests that most trains that reverse spend approximately six minutes doing so. However, even those terminating trains with dwell times below the first quartile, which are most likely behind schedule and sent on their return journey as soon as possible, have dwell times significantly higher than the third quartile of dwell time for through trains. This large difference is likely caused by two factors. The first is cab set up time, the time it takes to set up the operator’s cab at the front of the train to proceed in the opposite direction. The second is conflict with westbound through trains. If a westbound train is ready to depart from the middle platform at Arnos Grove but another westbound train is leaving from the side platform at the same time, the train in the middle platform will wait at the platform. Because the eastbound through trains are not merging with any other trains as they pass Arnos Grove, they do not experience this delay.

The presence of a crew change also has a significant impact on dwell time. For reversing trains, the addition of a crew change lowers the median dwell time from 395 seconds to 385 seconds (2.5%). Crew changes on terminating trains step back\(^5\), which is employed as part of the schedule during peak periods at Arnos Grove. Stepping back allows the dwell time to be less than would be necessary if the operator had to walk from one end of the train to the other. However, the use of stepping back during peak periods - when layover times are generally shorter to permit higher frequencies - means that it is difficult to separate correlative and causal relationships between stepping back and lower dwells. For through trains, on the other hand, the presence of a crew change can only increase the dwell time, because otherwise the operator would remain in the cab as at a normal station. Crew changes also do not affect the timetable for through trains. Because the timetable is created before the crew duties are created, according to TfL staff a dwell time of 60 seconds is assumed for all through trains at crew change locations. For through trains, the addition of a crew change increases the median dwell time from 26 seconds to 59 seconds (127%).

\(^5\) Typically at London Underground, a train operator arriving at a terminal will walk to the other end of the train and take the same train back down the line. This constrains reversing time to be no less than the time it takes the train operator to do this. At termini that see a particularly intensive service, such as Arnos Grove, stepping back is employed. This means that operators, after arriving at a terminus, do not take the same train in the other direction, but instead “step back” and allow another operator to take over, while they wait for the next (or second, or third) subsequent departure which they will drive on the return trip. This means that reversing time is now constrained by the time it takes the operator to board the train, as opposed to walking down the platform. This can dramatically reduce platform occupancy time and thus increases reversing capacity.

Stepping back requires additional train operators, because there is always one (or more) waiting on the platform instead of operating a train, so for organizational reasons it is usually only done as part of the timetable, for planned weekend closures, or for particularly protracted disruptions where there is time to dispatch the additional operators.
Although any causal relationship between dwell time reduction and capacity increase would need to be confirmed through simulation, these findings suggest several courses of action for TfL decision makers. One option that has been considered by LU in the past has been the introduction of an additional reversing platform at Oakwood to relieve pressure at Arnos Grove. This platform would accommodate trains extended from Arnos Grove, which would decrease total platform occupancy time at Arnos Grove and likely also increase capacity. Other possibilities include adding station staff at Wood Green to decrease the detraining time, or introducing stepping back at Cockfosters to increase its capacity as a terminus, both of which could reduce the number of trains reversing at Arnos Grove.

3.3.3 Selection of Potential Crossover Locations

The primary focus of this chapter is modeling the impact of potential new crossovers on major disruptions and planned closures. The first step in this process, the selection of crossover locations to be analyzed, is strongly constrained by the tunnel infrastructure in the central core and by the locations of existing crossovers on the surface sections. In the context of mitigating major disruptions, it does not make sense to propose an additional crossover if an existing one is located close by.

On the surface section between Arnos Grove and Cockfosters at the eastern end of the line, the presence of a third track at Arnos Grove, a deep tunnel section at Southgate, and a crossover at Oakwood means that there is no scope to introduce an additional interlocking. The aforementioned third platform at Oakwood would increase terminal capacity on the eastern end of the Piccadilly line. However, its proximity to reversing facilities at Arnos Grove and Cockfosters means that it would not significantly mitigate the impact of service suspensions. Therefore, it is not considered further in this research.

On the central deep-tunnel section of the line, from Arnos Grove to Barons Court, the only economical modification options are those where the tunnel infrastructure allows a crossover to be placed. In practice, this means those locations where crossovers existed historically, but have since been removed. There are two such locations: York Road, which is an abandoned station between Kings Cross and Caledonian Road, and Covent Garden. York Road is located immediately adjacent to Kings Cross which has a crossover (in fact, the latter crossover replaced the former), so the restoration of a crossover there makes no sense. Covent Garden, however, is more promising due to the relatively long distance (approximately two kilometers on either side) from the adjacent crossovers. The absence of a crossover at Covent Garden was noted in the 1990s, when frequent security alerts at Kings Cross caused large sections of the Piccadilly line to be suspended, and in 2005, when the 7/7 terrorist attack caused the line to be suspended for several weeks between Wood Green and Hyde Park Corner, eliminating Piccadilly line service to all of the major central London transfer stations (Benson, 2005). A crossover at Covent Garden would have allowed service from the west to reverse there, preserving connections with the Bakerloo, Jubilee, Northern, and Victoria lines.
On the western surface section, the lack of tunnels means fewer constraints on crossover placement. The primary criteria for selecting potential crossover locations are thus the frequency of delays and the location of existing crossovers and sidings. The long, delay-heavy nature of the section from Hammersmith to Acton Town motivated the proposal of a crossover at Turnham Green, which is near the middle of this section and has Piccadilly line trains stop during early morning and late evening hours. Farther west, the Heathrow branch’s history of incremental extensions means that there are crossovers every two or three stations, so there is little advantage to installing additional ones on that section. The Rayners Lane branch, however, is plain-line between the District line’s divergence point near Ealing Common and South Harrow, a distance of nearly eight kilometers. This prompted the analysis of a crossover at Alperton, which is near the middle of this section.

For a full-scale study of crossover benefit, once the proposed locations for the crossover are selected the specific track layout must be chosen. The type of crossover must be chosen - facing point, trailing point, or both. Facing point crossovers are those that ‘face’ the train - in other words, when a train is traveling in the usual direction for a given track, it can proceed along one of two routes. Trailing point crossovers are the opposite, where a train traveling in the usual direction would reverse to use the crossover. In Figure 3-9 below, assuming left-hand running as on the Underground, crossovers 1 and 3 are trailing point and crossover 2 is facing point.

![Figure 3-9. Example of facing and trailing point crossovers](image)

Though facing point and trailing point crossovers are both single crossovers, they have slightly different effects on signaling and on the speed of trains using the crossover. The exact location must also be selected, dependent on train speeds in the area, track superelevation (cant), and structural and other constraints. More detailed track engineering and train service simulation would be necessary to determine the impact on service of each of these crossover options. For this study, trailing point crossovers are assumed to be located just east of the station being analyzed at all three proposed locations, because trailing point crossovers are generally preferred for safety reasons, because a switch malfunction will not cause the train to derail as often, and proximity to a station allows that station to be efficiently used as a terminus. Figure 3-10 shows the proposed
modifications at each of the three proposed locations, with current layout on the left and proposed layout on the right.

Figure 3-10. Proposed track layouts

3.3.4 Generation of Service Patterns

Once the crossovers to be analyzed are determined, the service patterns that controllers would implement when using them are generated. As mentioned previously, the primary focus of this research is on partial line suspensions after major incidents. For each crossover, an incident is assumed to occur on one side of the crossover, and service is therefore suspended between the new crossover and the next crossover down the line. With current infrastructure, the service suspension will extend in both directions to the nearest adjacent crossovers. For example, if an incident requiring a service suspension occurs at Piccadilly Circus, with current track infrastructure, service would be suspended between Hyde Park Corner and Kings Cross. With the addition of a crossover at Covent Garden, the service would only need to be suspended between Hyde Park Corner and Covent Garden.

The exact service patterns assumed are governed by several factors. The first is the timetable. Particularly on the Rayners Lane branch, which has a low scheduled frequency, frequencies actually operated should not exceed those timetabled. On the trunk section, of course, this does not control as other constraints usually preclude achieving the scheduled frequency. The second factor is reversing capacity. To avoid major delays, controllers do not run more service on a given section than the terminal can handle. Therefore, on a line (section) without branches the
terminal’s reversing capacity determines the maximum service levels on the entire line. Additional mid-line termini are generally not used unless they have a third (siding) track, because the time needed for a train to reverse on the main line would delay following trains which are not reversing. The third factor is the nature of the disruption - planned or unplanned. If a closure is planned, such as for weekend engineering work, the ability to get train operators to step back increases capacity. Even when stepping back is not implemented, planned closures are communicated in advance to passengers and operators, reducing confusion. Reduced confusion, in turn, decreases the time that passengers spend boarding and alighting, and operators spend communicating with controllers to understand their assignments, thereby reducing total reversing time. Finally, controller behavior is also important. During an unplanned disruption, controllers have to deal with the incident itself, as well as ensuring that crews and trains are in the right place at the right time. Therefore, controllers generally prefer simple service patterns with self-contained services where possible.

As mentioned previously, reversing time is the critical determinant of reversing capacity, and crew movement time is the critical determinant of reversing time. Over the years, London Underground has developed the following rules of thumb for reversing capacity:

- **Unplanned reversing**: 5-6 tph (trains per hour) per berth
- **Planned reversing**: 7.5-8 tph per berth
- **Planned reversing with stepping back**: 10-12 tph per berth

A berth is a track where a train changes direction. The number of berths that can be used in parallel determines capacity. Of course, at some termini these guidelines are exceeded on a daily basis, particularly on the Victoria line, where two-track termini at each end support a 36 tph service. However, this requires careful optimization and near-perfect operation and relies heavily on the topologically simple nature of the Victoria line, where there are no branches and all trains travel the full length of the line.

With these considerations in mind, service patterns are generated, several based on the following assumptions:

- **Train service on other lines is not affected by the disruption on the Piccadilly line.** In particular, Metropolitan line service to Uxbridge and District line service to Ealing Broadway are maintained, but not increased. For unplanned disruptions, there is typically no time to arrange such service increases, and for planned (weekend) disruptions, the timetabled service is usually considered sufficient.
- **Rail replacement bus service is not provided in the central section, where there are many Tube alternatives.** However, on the western end of the line, rail replacement buses are provided for planned disruptions, at a five-minute headway.
- **For planned disruptions, reversing capacities of up to 12 tph are assumed, as stepping back is generally provided where necessary during planned outages.** For unplanned disruptions, reversing capacities of 6 tph are assumed.

For each crossover, three scenarios are examined: the “base,” “west side,” and “east side” suspensions. For each proposed crossover on the Piccadilly line, each relevant incident occurs
either to its west or to its east. If the crossover does not exist, then regardless of the incident location, the service suspension will extend to the nearest extant crossovers in both directions. For instance, if the proposed crossover at Covent Garden does not exist then incidents at Green Park or Russell Square will both require a suspension from Kings Cross to Hyde Park Corner. In the tables and figures below, this scenario is denoted as having “base” infrastructure. If the crossover is installed, the location of the incident affects the required service suspension. For the Covent Garden crossover, an incident at Green Park would require a suspension from the crossover west to Hyde Park Corner, and an incident at Russell Square would require a suspension from the crossover east to Kings Cross. These scenarios are denoted as “west side” and “east side” respectively. The benefit of a crossover for west side incidents is determined by comparing the passenger impact values of the “base” and “west side” cases, and for east side incidents by comparing the “base” and “east side” cases. Table 3-1 shows the service patterns that are assumed to operate under each suspension scenario.

3.3.5 Modeling

The next step is to model these service patterns in Railplan. Each disruption is modeled for two time periods - the weekday morning peak, from 7:00 to 10:00, and the off peak from 10:00 to 16:00. The disruptions, and their corresponding service patterns, are assumed to remain in effect for the entire period being studied, with the passenger benefit being computed as the percentage reduction in LCH with the addition of the crossover. The scenarios are created by modifying four different base case scenarios, representing different years: 2021, 2031, and 2041, as well as 2041 with the addition of Crossrail 2 to the network. “2041 with Crossrail 2” scenarios are created by modifying 2041 scenarios to add Crossrail 2 and remove the National Rail branch services, mostly operating out of Waterloo, that it will replace. However, land use and demand changes as a result of Crossrail 2 are not considered. Although land regeneration and the ability to construct additional housing are key goals of the Crossrail 2 project, to avoid confounding variables, the same demand matrices are used as inputs to Railplan for all 2041 scenarios.

As currently planned, Crossrail 2 will follow a southwest-to-northeast trajectory through central London and will intersect the Piccadilly line at two stations and broadly parallel it north of Holborn. This makes it a particularly interesting case study for determining the degree to which the addition of a new line can, by providing an additional alternative route, mitigate the impact of disruptions on an existing line, and reduce the benefit that any track layout modifications on the existing line will have. The proposed route for Crossrail 2 is shown in Figure 3-11.
### Table 3-1. Disruption Service Patterns.

<table>
<thead>
<tr>
<th>Proposed Crossover</th>
<th>Scenario type</th>
<th>Suspended section</th>
<th>Unplanned - AM peak</th>
<th>Planned - Off peak</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Covent Garden</strong></td>
<td>Base</td>
<td>Kings Cross - Hyde Park Corner</td>
<td>T5-HPC 6tp, T4-NFD 6tp, RLN-ACT 6tp, KXX-CFS 6tp</td>
<td>T5-HPC 6tp, T4-NFD 6tp, UXB-HPC 3tp, RLN-HPC 3tp, KXX-CFS 12tp</td>
</tr>
<tr>
<td></td>
<td>West side</td>
<td>Kings Cross - Leicester Square</td>
<td>T5-LSQ 6tp, T4-NFD 6tp, RLN-ACT 6tp, KXX-CFS 6tp</td>
<td>T5-LSQ 6tp, T4-NFD 6tp, UXB-LSQ 3tp, RLN-LSQ 3tp, KXX-CFS 12tp</td>
</tr>
<tr>
<td></td>
<td>East side</td>
<td>Holborn - Hyde Park Corner</td>
<td>T5-HPC 6tp, T4-NFD 6tp, RLN-ACT 6tp, HOL-CFS 6tp</td>
<td>T5-HPC 6tp, T4-NFD 6tp, UXB-HPC 3tp, RLN-HPC 3tp, HOL-CFS 12tp</td>
</tr>
<tr>
<td><strong>Turnham Green</strong></td>
<td>Base</td>
<td>Hammersmith - Acton Town</td>
<td>T5-ACT 3tp, T4-ACT 3tp, RLN-ACT 3tp, GPK-CFS 6tp, HMD-CFS 12tp</td>
<td>T5-ACT 6tp, T4-NFD 6tp, RLN-ACT 6tp, GPK-CFS 6tp, HMD-CFS 12tp</td>
</tr>
<tr>
<td></td>
<td>West side</td>
<td>Hammersmith - Turnham Green</td>
<td>T5-TGN 6tp, T4-NFD 6tp, RLN-ACT 6tp, GPK-CFS 6tp, HMD-CFS 12tp</td>
<td>T5-TGN 6tp, T4-NFD 6tp, RLN-TGN 6tp, GPK-CFS 6tp, HMD-CFS 12tp</td>
</tr>
<tr>
<td></td>
<td>East side</td>
<td>Turnham Green - Acton Town</td>
<td>T5-ACT 3tp, T4-ACT 3tp, RLN-ACT 3tp, TGN-CFS 6tp, GPK-CFS 6tp, BCT-CFS 6tp</td>
<td>T5-ACT 6tp, T4-NFD 6tp, RLN-ACT 6tp, TGN-CFS 12tp, BCT-CFS 6tp</td>
</tr>
<tr>
<td><strong>Alperton</strong></td>
<td>Base</td>
<td>Acton Town - Rayners Lane</td>
<td>Regular service, with Rayners Lane and Uxbridge trains diverted to Northfields. No Piccadilly line service Rayners Lane - Uxbridge</td>
<td>Regular service, with Rayners Lane and Uxbridge trains diverted to Northfields. No Piccadilly line service Rayners Lane - Uxbridge. Bus service provided on the suspended section, 12bph</td>
</tr>
<tr>
<td></td>
<td>West side</td>
<td>Acton Town - Alperton</td>
<td>Same as above, with the addition of 6tp RLN-ALP</td>
<td>Same as above, with the addition of 6tp RLN-ALP</td>
</tr>
<tr>
<td></td>
<td>East side</td>
<td>Alperton - Rayners Lane</td>
<td>Regular service, except only 6tp operate from the trunk to Alperton. All other trains to Rayners Lane or Uxbridge are diverted to Northfields.</td>
<td>Regular service, except Rayners Lane and Uxbridge trains are truncated to Alperton. Bus service provided on the suspended section, 12bph</td>
</tr>
</tbody>
</table>

Station abbreviations: ACT: Acton Town; ALP: Alperton; BCT: Barons Court; CFS: Cockfosters; GPK: Green Park; HMD: Hammersmith; HOL: Holborn; HPC: Hyde Park Corner; KXX: Kings Cross St Pancras; LSQ: Leicester Square; NFD: Northfields; RLN: Rayners Lane; TGN: Turnham Green; T4: Heathrow Terminal 4; T5: Heathrow Terminal 5
3.4 Results

Railplan generates detailed results for each scenario, including the passenger flows and crowding on each link in the network. These can be compared in aggregate - as a bottom-line LCH figure for a given disruption - or in more detail, looking at flows on each link to determine which (other) lines are most impacted by a disruption. For this study, although the detailed link flow based analysis described above is used for qualitative validation, here we focus on the aggregate figures. The primary focus of the analysis is understanding the degree to which a given crossover can mitigate the impact of an incident, and how this mitigation changes with time of day, and with the year being analyzed. Figures 3-12, 3-13, and 3-14 show the passenger impact of a disruption near each of the three proposed track layout modifications in 2021, 2031, 2041, and 2041 with Crossrail 2. Figures 3-15, 3-16, and 3-17 show the percentage reduction in disruption impact from installing each crossover.
For this analysis, for each year and proposed crossover there are three scenarios considered: “base,” “west side,” and “east side,” as described in Section 3.3.4. In these figures, “AM” refers to scenarios during the morning peak period, while “OP” refers to off peak scenarios. All figures incorporate weighted travel times in accordance with the Journey Time Metric.

Figure 3-12. LCH by year and time of day for Covent Garden crossover

Figure 3-13. LCH by year and time of day for Turnham Green crossover
Figure 3-14. LCH by year and time of day for Alpertont crossover

Figure 3-15. Percent reduction in LCH by year and time of day for Covent Garden crossover

Figure 3-16. Percent reduction in LCH by year and time of day for Turnham Green crossover
There are two counteracting principles at work here. First, having a crossover near an incident allows train service to operate over a greater portion of the line. Passengers who gain more direct trips as a result, benefit because they need to make fewer transfers and have shorter overall travel time. Passengers who still change their route also benefit: because there are fewer displaced passengers, so alternative routes are less crowded, causing less discomfort and fewer denied boardings. This means that adding a crossover will substantially lower the passenger impact of a disruption. This will be referred to as the service extension benefit.

On the other hand, the presence of a disruption, in particular a service suspension (of any size), typically causes a dramatic reduction in service frequency throughout the line. Ad-hoc terminals like those generally used to turn trains during partial line suspensions usually have a much lower reversing capacity than the line’s scheduled frequency, as described above. As a result, to avoid train congestion it is necessary for controllers to greatly reduce frequency during a partial line suspension. This causes overcrowding and displaces passengers - even those who do not use the suspended section of the line - to other routes. Adding a crossover and extending the line’s residual service by two or three stations will not mitigate this. This means that, for more crowded conditions, a new crossover will not have significant benefit to passengers. This will be referred to as the disruption capacity constraint. It is the relation between these two principles that governs the net passenger impact of a disruption.

As expected given its location in the trunk section of the Piccadilly line, the impact of the crossover at Covent Garden is mainly governed by the disruption capacity constraint during the morning peak period. The reduction in service from the normal schedule of 24 tph to the 6 tph operated under disrupted conditions means that most passengers can no longer be accommodated by the line and must find alternative routes. The remaining passengers will experience increased crowding and related discomfort. This impact overshadows the benefit of any service extension that a new crossover may provide. This is true for all years being modeled, regardless of changes in train capacity, frequency on neighboring lines, etc. Higher demand in 2031 and 2041 causes the passenger impact of a service suspension to increase, but the number of people accommodated by
a 6 tph service remains essentially constant. Therefore, as expected, the disruption capacity constraint dominates and the model results do not show a noticeable change in morning peak passenger benefit of the Covent Garden crossover for the three years analyzed. In each case, the reduction in Lost Customer Hours caused by the disruption is 3-4%.

For disruptions during the off peak period, the total impact of a given disruption is smaller than during the peak due to lower passenger flows. For these disruptions, though, the crossover is proportionally more beneficial. When the disruption is to the west of the crossover (between Hyde Park Corner and Covent Garden), trains from the east reverse at Covent Garden, but because of station access limitations they are assumed to discharge passengers at Holborn. Without the crossover being present, they would have reversed at Kings Cross, one mile away, which has transfers to five other lines including the Victoria line, which runs roughly parallel to the Piccadilly through central London. The extension to Holborn only provides service to two additional stations and a transfer to one other line, the Central. Therefore, the service extension benefit of the extension is limited. In 2021, the crossover reduces the impact of such a disruption by 6.3%. In 2031, when capacity on other lines has been increased, this benefit decreases to 5.1% because a larger number of passengers are able to take alternative routes, which do not depend on how far Piccadilly line trains operate. Between 2031 and 2041, demand is assumed to increase but the network is unchanged. Thus, alternative routes get more crowded and more passengers must use the Piccadilly line and can therefore benefit from the service extension; thus, the benefit of the crossover increases to 6.9%.

When the disruption is to the east of the crossover (between Kings Cross and Covent Garden), however, trains from the west reverse at Covent Garden and discharge passengers at Leicester Square, instead of Hyde Park Corner. This presents a significant advantage to passengers, adding direct service to the heart of the West End as well as transfers to the Bakerloo, Northern, Jubilee, and Victoria lines. Between 2021 and 2031, the increase in service levels on the District line (which parallels the Piccadilly through much of western London) decreases the passenger benefit of the crossover from 19.8% to 17.8%. Between 2031 and 2041 the percentage reduction remains roughly constant because the capacity of parallel links remains unchanged.

During the morning peak, the benefits of the crossover at Turnham Green depend strongly on the incident location. For an incident east of the crossover, the installation of the crossover allows the residual service from the west to take passengers as far as Turnham Green, where they can transfer to the District line (Richmond or Ealing Broadway branches) to continue their trip with minimal additional travel time. Without the crossover, the residual service only operates as far as Acton Town, where there is half as much District line service available (only the Ealing Broadway branch serves Acton Town). Thus, extending residual service to Turnham Green doubles the capacity of the closest alternative route, dramatically reducing the overcrowding that passengers experience, and reducing the overall impact of the disruption by 15.3% in 2021. This percentage is smaller in 2031 and 2041 as the Four Lines Modernization project, currently underway, increases train frequency on the District line.
For an incident west of Turnham Green, the crossover allows trains from the trunk section of the line to operate to Turnham Green, instead of Hammersmith. However, for passengers to access trains on the remainder of the line west of Acton Town, they are still required to use District line services heading to Ealing Broadway; the only change for these passengers is that they transfer to the District line at Turnham Green instead of Hammersmith, which has negligible benefit, as reflected in Figure 3-16.

During the off peak period, the crossover benefit during a disruption on the east side is, percentagewise, lower than during the morning peak period: 10.4% in 2021, and varying only slightly thereafter. This is likely due to the relatively low off-peak ridership on the western branches of the line, especially because Heathrow, the major demand generator on the western end of the Piccadilly line, is also served by Crossrail. The disruption capacity constraint is not as strong a factor during the off peak, so extending service to Turnham Green to give passengers access to additional District line trains is not as beneficial as during the peak.

A disruption on the west side causes the crossover to be more beneficial during the off peak. This is because of the service patterns assumed to be operating. Extending trains to Turnham Green allows the use of the siding at Hammersmith for additional trains along the trunk, which discharge passengers at Barons Court and use the siding to reverse. This adds service to the entire line, and also allows passengers on those trains to access the District line. For these reasons, the benefit of the crossover for such disruptions is estimated to be 14.7%. As during the peak, the benefits decrease slightly by 2031 as the parallel District line gets a frequency upgrade, then increase in 2041 as demand increases but supply does not.

As at Turnham Green, the benefits of the crossover at Alperton are heavily influenced by the location of the incident, though they are lower in magnitude because of the relatively low ridership on the Rayners Lane branch. In case of an incident to the west of the crossover, the residual service operates from Alperton east through the trunk to Cockfosters, which gives passengers east of Alperton an unaffected journey except for any additional platform wait time. This cuts the length of the suspended section by half. This means that the benefit of the crossover for such disruptions is between 24% and 26% for each year analyzed. The reason it is not higher is related to through riders, who board at stations between Rayners Lane and Uxbridge and use the Piccadilly line. Although the section with no service is shortened by half, the assumed absence of Piccadilly line service west of Rayners Lane means that passengers on that section divert to the Metropolitan line into Central London.

When the incident is located to the east of the crossover, service is suspended between Acton Town and Alperton. This leaves a residual shuttle service operating between Alperton and Rayners Lane. This keeps this section connected to Central London, though indirectly: passengers are required to transfer at Rayners Lane to the Metropolitan line. It is this indirect nature of the alternative route that causes the benefit of the crossover for this type of disruption to be only (approximately) 5% for each year analyzed.

During the off peak, the benefits of the Alperton crossover in 2021 are approximately equal, as a percentage of total disruption impact, to those during the peak period. In 2031, however,
the benefit of the crossover declines sharply, from 23.9% to 17.1%, for west-side suspensions. This is likely because of the increase in frequencies on the Ealing Broadway branch of the District line scheduled for 2026. After this increase, passengers at stations close to the Ealing Broadway branch such as North Ealing, and (in particular) Ealing Common, will be able to complete their trips more quickly using the District line, and will therefore experience less benefit from the Piccadilly line service.

The degree to which the introduction of Crossrail 2 affects the impact of a given disruption depends greatly on the disruption location. For example, the impact of a disruption on the Rayners Lane branch in western London, which has a low frequency and is far from the proposed route of Crossrail 2, will not be mitigated by the new line. On the other hand, Crossrail 2 will decrease the impact of a disruption in the central trunk section significantly because it offers a nearby parallel route for passengers to use, without introducing additional demand to the network (as discussed above). Figure 3-18 shows the degree to which the addition of Crossrail 2 mitigates the impact of each suspension analyzed above.

![Figure 3-18. Percent reduction in LCH resulting from Crossrail 2](image)

Clearly, the line’s greatest benefit is for suspensions in the central area. The addition of Crossrail 2 is estimated to reduce the passenger impact of morning peak suspensions in the central area by approximately 15%, regardless of which of the three suspension scenarios in the Covent Garden area - base-case infrastructure, crossover installation and west-side suspension, or crossover installation and east-side suspension - is considered. During the off peak, this benefit decreases to between 8% and 10%, depending on the scenario. This is due to higher available capacity during the off peak, caused by increased reversing capability for planned disruptions and decreased demand, which means that the new capacity introduced by Crossrail 2 is not as beneficial.

The effect of disruptions in the Turnham Green area is also reduced by the introduction of Crossrail 2. Although its route is not close to Turnham Green, the reduction in frequency on the
Piccadilly line trunk that is inherent to any suspension means that Crossrail 2 can accommodate passengers who would otherwise be denied boarding by Piccadilly line trains. This is particularly true on the eastern section of the line, where the two lines run parallel and share two transfer stations. During the morning peak this translates to a benefit of 10% to 11%. During the off peak, the increased supply and decreased demand cause this benefit to decrease to approximately 5%.

For suspensions near Alperton, on the Rayners Lane branch, the addition of Crossrail 2 has minimal impact because frequency on the Piccadilly line trunk is not affected by the disruption and the Crossrail 2 route does not go anywhere near the western section of the Piccadilly line.

3.4.1 Business Case Analysis

Clearly, to quantify the net benefit of the installation of any crossover, a full analysis of the factors discussed in Section 3.2 is necessary, including both costs and benefits.

A full examination of the costs of the installation of a crossover is an exercise in project management and electrical, mechanical, and structural engineering that is beyond the scope of this work. Therefore, a cost estimate is taken from a previous study of a crossover at Covent Garden and adjusted for inflation to reflect an expected figure today. In 2005, the installation and maintenance of this crossover was estimated to cost approximately £10 million (Benson, 2005) which, when adjusted for inflation (Bank of England, 2020), would be approximately £15 million now. This figure includes an optimism bias, which accounts for the systematic tendency for project promoters to underestimate such costs, of 66% - the TfL recommended value for atypical capital improvements such as a new crossover (Transport for London, 2017b). The costs of the other two crossovers considered in this chapter are likely lower, because their surface location makes installation easier and may require fewer custom parts than the deep-tunnel location of the Covent Garden crossover, but as an order-of-magnitude estimate this £15M cost is assumed to apply to each of the three crossovers.

The benefits of a crossover include a wide range of factors that are beyond the scope of this work. This study will compute the benefit of each crossover only in the cases of major disruptions and planned weekend closures. However, it will not consider the benefits that result from allowing more opportunities for short-turns, more efficient engineering train operations, removing disabled trains more quickly, or other factors.

In the context of major unplanned disruptions, the total benefit of a crossover can be estimated by considering the annual benefit as an annualized cash flow for some number of years and applying a discount rate to determine the net present value. For the purposes of this study, a period of fifty years is assumed, as this is the approximate interval between line signal upgrade projects, which are generally when track layout changes are considered.

The annual benefit, in turn, can be computed by multiplying the total impact of relevant disruptions by the percentage reduction in disruption impact that the crossover allows. The data on incident impact and frequency is drawn from CuPID. Notably, it is assumed here that asset failure rates (and thus disruption frequencies) remain constant over the study period. This is not strictly accurate; as assets continue to age, their reliability generally decreases and failure rates
increase. On the other hand, when a line upgrade project is completed and most of a line’s assets are in good condition, failure rates are usually low. Given TfL’s present uncertainty regarding financing, scheduling, and scope of any future Piccadilly line upgrade, it is difficult to determine the impact that it will have on asset failure rates and thus on line performance. Therefore, failure rates are assumed to remain constant.

As mentioned above, only relevant disruptions should be considered when quantifying the disruption-related benefit. In short, a ‘relevant’ disruption is one that would cause a controller to order trains to use the crossover being analyzed. There are two main factors that determine an incident’s relevance.

First, obviously, is the incident location. Here, controller behavior assumptions play a large role. For this analysis, as above, it is assumed that service operates as far as it can, and service that short turns at other locations to increase frequencies is not operated as it would interfere with through trains. For instance, a disruption at Arsenal which causes a partial line suspension from Kings Cross to Wood Green will see trains from the western end of the line operate as far as Kings Cross. It is assumed that no additional trains from the west reverse at Hyde Park Corner (the next crossover west of Kings Cross), as their reversing time would impede the operation of through services. Thus, relevant incident locations are those that occur between two adjacent crossovers, but not at the location of either. For the proposed crossover at Covent Garden, for instance, where the nearest existing crossovers are at Kings Cross and Hyde Park Corner, relevant incidents are those that take place at Green Park, Piccadilly Circus, Leicester Square, Holborn, and Russell Square. Incidents at Covent Garden itself are not considered because they would interfere with using that crossover to reverse trains, and thus produce no benefit. Incidents at Kings Cross and Hyde Park Corner would see the crossover at Covent Garden being used; however, they would require a greater length of suspended service than was modeled in this analysis, thus they are not considered. This is, of course, a conservative estimate; a more fully-fledged analysis would also consider such extended closures.

Second is the incident type. This involves determining which incidents cause partial line suspensions, which draws on several data fields in the CuPID database. First, the “Service Disruption category” field contains entries for partial line suspensions; all incidents thus marked are deemed relevant. However, many suspensions, particularly those that are shorter-term, are denoted using other categories, so the “cause factor” field is analyzed next. Whenever a passenger ends up under a train, service is usually suspended, so all incidents where the cause factor is a PUT (person under train) are marked as relevant. Finally, there are also other incident types, such as signal failures, that sometimes cause suspensions. To isolate these, the incident description field is searched for the words “suspend” or “reverse,” which are generally included for incidents where crossovers are relevant. Table 3-2 shows the ten cause factors with the highest number of Lost Customer Hours caused by relevant incidents at all locations on the Underground over a nine-year period from June 2010 to May 2019.
Table 3-2. Top causes of relevant incidents

<table>
<thead>
<tr>
<th>Cause Factor</th>
<th>Total LCH</th>
<th>Number of Incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer - Suicide</td>
<td>4,155,468</td>
<td>559</td>
</tr>
<tr>
<td>Signals - Train Detection</td>
<td>2,652,197</td>
<td>680</td>
</tr>
<tr>
<td>Customer - Illness / Accident</td>
<td>1,842,476</td>
<td>571</td>
</tr>
<tr>
<td>Track &amp; Civils - Plain Line</td>
<td>1,680,524</td>
<td>257</td>
</tr>
<tr>
<td>Fleet - DIS - ATC</td>
<td>1,499,447</td>
<td>568</td>
</tr>
<tr>
<td>Fleet - DIS - Traction</td>
<td>1,418,371</td>
<td>776</td>
</tr>
<tr>
<td>Signals - Relays</td>
<td>1,406,969</td>
<td>451</td>
</tr>
<tr>
<td>Signals - Points Indication</td>
<td>1,316,907</td>
<td>499</td>
</tr>
<tr>
<td>National Rail - Network Rail</td>
<td>1,274,515</td>
<td>590</td>
</tr>
<tr>
<td>Fleet - Staff Error</td>
<td>1,269,697</td>
<td>868</td>
</tr>
</tbody>
</table>

Once all relevant incidents are selected and their frequencies determined, the total reduction in Lost Customer Hours is computed based on the percentage decrease in disruption impact computed using Railplan modeling, as described above. Percentage reductions in LCH compared with the base case are used, and not the exact values computed previously, to take into account varying disruption durations. This is then converted to monetary terms, using TfL’s recommended figure for Value of Time of £9.39 per hour (Transport for London, 2017b), with the results shown in Table 3-3.

Table 3-3. Benefit of crossover installation for rush hour unplanned disruptions

<table>
<thead>
<tr>
<th>Crossover</th>
<th>Average LCH/year</th>
<th>% LCH reduction</th>
<th>Total LCH reduction/year</th>
<th>Annual benefit (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covent Garden – incident on west side</td>
<td>11,875</td>
<td>3.7%</td>
<td>1,003</td>
<td>9,415</td>
</tr>
<tr>
<td>Covent Garden – incident on east side</td>
<td>14,444</td>
<td>3.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turnham Green - incident on west side</td>
<td>0</td>
<td>0.0%</td>
<td>1,389</td>
<td>13,046</td>
</tr>
<tr>
<td>Turnham Green - incident on east side</td>
<td>9,080</td>
<td>15.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alperton - incident on west side</td>
<td>9,384</td>
<td>24.7%</td>
<td>2,639</td>
<td>24,784</td>
</tr>
<tr>
<td>Alperton - incident on east side</td>
<td>6,698</td>
<td>4.8%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note that the suspension on the western side of Turnham Green does not have any LCH benefit because no incidents on the Piccadilly line are recorded as having taken place at Chiswick Park, which is the only station between Turnham Green and Acton Town. Although the LCH per year is highest for incidents in the central section near Covent Garden, the percent LCH reduction for incidents near Alperton is higher due to the service extension benefit prevailing over the disruption capacity constraint there. Thus, the crossover at Covent Garden is the least beneficial of the three for unplanned disruptions.

Planned off peak closures are significantly harder to predict in the future because of their dependence on TfL’s financial situation, asset condition data, and priorities. In the last decade, several lines (the Victoria, Jubilee, and Northern, as well as the Sub-Surface lines) have benefitted from extensive investment in the form of line upgrade programs. These projects have replaced signaling systems, as well as upgraded track, stations, and other systems, and in several cases replaced the rolling stock. Once such upgrades have been completed, weekend closures are rare, typically occurring no more than annually on any given line section. However, during these upgrade projects there are many weekend closures, because the Tube’s usual closure hours (approximately thirty-five hours per week on lines without Night Tube services, and twenty-five on lines with Night Tube services) are insufficient for such a large amount of work. The number of these closures varies greatly by line; for the Northern line upgrade, which incorporated many lessons from the Jubilee line upgrade, and was executed without significant schedule overruns, eight full-line closures were required along with assorted partial-line closures; the Jubilee line, in contrast, required dozens of closures (Harvie, n.d.). For the purposes of this assessment, the upgrade is assumed to take place in 2030, with five weekend closures on each side of each crossover.

Because the Railplan model considers a 6-hour period for off-peak scenarios, an internal TfL tool is used to derive LCH values for full-weekend closures. These are then scaled by the percentages calculated using Railplan to determine the benefit for full-weekend planned closures, with the results shown in Table 3-4.

Once the annual benefit (for unplanned disruptions) and per-weekend benefit (for planned disruptions) are computed, the total benefit is computed by assuming an asset life and number of closures per year. A constant 1.5% discount rate for benefits is assumed, as per the recommendation of the Office of Rail and Road (Transport for London, 2017b). A fifty-year period is assumed, with no changes in incident frequency and one planned weekend closure on each section every other year. Five closures on each section are assumed to be necessary for the line upgrade, which is assumed to take place in 2030. 2021 LCH reduction values are assumed between 2020 and 2030, 2031 values between 2031 and 2040, and 2041 values (without Crossrail 2) beyond that. The total aggregated crossover benefits are shown in Table 3-5.
Table 3-4. Benefit of crossover installation for planned weekend closures

<table>
<thead>
<tr>
<th>Crossover</th>
<th>Base case LCH</th>
<th>% LCH reduction</th>
<th>LCH reduction / weekend</th>
<th>Benefit / weekend (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covent Garden – suspension on west side</td>
<td>80,516</td>
<td>6.3%</td>
<td>5,098</td>
<td>47,877</td>
</tr>
<tr>
<td>Covent Garden – suspension on east side</td>
<td></td>
<td>19.8%</td>
<td>15,916</td>
<td>149,453</td>
</tr>
<tr>
<td>Turnham Green – suspension on west side</td>
<td>13,771</td>
<td>14.7%</td>
<td>2,025</td>
<td>19,023</td>
</tr>
<tr>
<td>Turnham Green – suspension on east side</td>
<td></td>
<td>10.4%</td>
<td>1,434</td>
<td>13,469</td>
</tr>
<tr>
<td>Alperton – suspension on west side</td>
<td>40,085</td>
<td>23.9%</td>
<td>9,561</td>
<td>89,781</td>
</tr>
<tr>
<td>Alperton – suspension on east side</td>
<td></td>
<td>7.7%</td>
<td>3,101</td>
<td>29,118</td>
</tr>
</tbody>
</table>

Table 3-5. Aggregate benefit of crossover

<table>
<thead>
<tr>
<th>Crossover</th>
<th>Annual benefit unplanned, 2021, (£)</th>
<th>Annual benefit planned, 2021, (£)</th>
<th>Total benefit over 50 years (10^6 £)</th>
<th>Benefit for five closures in 2030, discounted (£)</th>
<th>Total benefit (10^6 £)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covent Garden</td>
<td>9,415</td>
<td>98,666</td>
<td>3.568</td>
<td>850,171</td>
<td>4.418</td>
</tr>
<tr>
<td>Turnham Green</td>
<td>13,046</td>
<td>16,247</td>
<td>0.939</td>
<td>139,991</td>
<td>1.079</td>
</tr>
<tr>
<td>Alperton</td>
<td>24,784</td>
<td>59,450</td>
<td>2.643</td>
<td>512,262</td>
<td>3.156</td>
</tr>
</tbody>
</table>

Notably, the annual benefit for planned weekend closures is substantially higher than that for unplanned disruptions. This is largely caused by the comparatively brief duration of major unplanned disruptions – typically two or three hours at most, compared with almost fifty service hours for a weekend closure – and by the relationship between the service extension benefit and disruption capacity constraint, whereby peak hour disruptions in central London cause overcrowding that overshadows the benefit of extending service to several additional stations.

Using these estimated benefits, the BCR for the Covent Garden crossover is 0.37, for Turnham Green 0.09, and for Alperton 0.26. Clearly, on their face these figures are not sufficient to justify any of these investments, however these figures are just one component of a business
case, and they incorporate several conservative assumptions and simplifications that, if considered in greater detail, might help justify the installation of a given crossover, as described below.

First, this analysis only considers the benefits of any given crossover that most directly relate to journey time, and does not consider any higher-level benefits such as perceived reliability (which may encourage more use of the network, generating additional fare revenue). For unplanned disruptions, the Piccadilly line is particularly vulnerable to incidents in the central area, where the distance between crossovers means that any incident has the potential to take a large section of the line out of service. In particular, the lack of crossovers between Hyde Park Corner and Kings Cross means that any incident in that section deprives the western half of the line of its main central-area destinations and transfers. Extending this service to Covent Garden would allow passengers to make critical transfers in the central area, thus increasing their perception of the network’s reliability as a whole in addition to shortening their journey times.

Second, only major disruptions and planned closures are considered here. As discussed in Section 2.3.1, the most crossover use is at the outer edges of the line, which suggests that crossovers are primarily used to reverse trains after minor disruptions to get them back on schedule. In addition, there are other uses for crossovers - allowing engineering trains faster access to (and around) worksites, getting disabled trains back to the depot faster, etc. - which are not considered here. A full business case would need to quantify the frequency of use of each crossover for each of these cases, and estimate an economic benefit for each type of use.

Third, as discussed previously, only incidents that take place between interlockings are considered. This results in a conservative estimate of total crossover benefit, because many incidents, particularly signal failures, take place at interlockings. Considering these incidents is feasible, but would require a wider range of service patterns to be modeled and would perhaps warrant a more in-depth study of controller responses to disruptions.

Fourth, off peak, unplanned closures are not explicitly modeled. This means that the percentage benefit determined from peak closures is applied to all unplanned disruptions, whether they occur during the peak or off peak. This is a conservative assumption because, as discussed previously, service extensions during the off peak are proportionally more beneficial than during peak periods. The lower crowding levels during the off peak mean that when capacity is reduced as a result of a disruption, a greater share of the line’s total passenger flow is able to take advantage of any service extension. This means that accounting for off peak unplanned disruptions separately should yield a larger benefit for each crossover.

3.5 Conclusions

The study of track layout on metro lines is a largely underexplored field. There are a wide variety of factors that influence the provision of crossovers and sidings during line construction, and many more that affect how they are used by line controllers if they exist. There are many use cases for a crossover or siding: mitigation of major disruptions by maintaining residual service, mitigation of minor disruptions by short-turning trains to get them back on schedule, as well as others involving engineering trains, getting failed trains out of the way of service, etc. This work
discusses the different use cases for each, and outlines a methodology for computing the benefit of crossover installation to mitigate the impact of major disruptions and planned closures.

Three possible crossover locations on the Piccadilly line are analyzed in the case of major disruptions. These crossovers, the placement of which is determined by site constraints, proximity to other crossovers, and line controller preferences for using crossovers during disruptions, are well-distributed throughout the line, with one in the central section, one on the inner-western section of the line, and one on the Rayners Lane branch in the west of London.

For all three crossovers, a preliminary business case analysis shows that major incidents and planned closures occur too rarely for additional crossovers to be economically justifiable. For controllers to begin using a crossover to reverse trains during a major disruption, it must be serious enough to cause a service suspension on the affected segment, and be expected to last long enough that controllers will not simply wait out the disruption. Aside from line upgrades, which occur only every few decades and are subject to political priorities and financial constraints, planned closures on a given section of line typically occur less than once a year. With these restrictions, there are at most a few cases a year where a crossover would be useful, so within the context analyzed here there is no economic justification for installation of any new crossover. However, there are additional uses of a crossover that should be studied further before any investment decision is made. One of these, terminal capacity, is analyzed in this chapter. Because insufficient terminal capacity affects every rush-hour train operating every day, even a minor reduction in terminal delay can produce significant passenger benefit. A second use of a crossover, which is the most commonly used one on outlying sections of the line, is recovery from minor delays. To analyze the benefit of a crossover from this perspective, it would be necessary to perform more in-depth modeling to determine the benefits (and disbenefits) to passengers, crews, and overall on-time performance of short-turning a given train. This is particularly useful on crossovers on the outer edges of the line, such as the proposed one at Alperton analyzed here.
This chapter describes the development and application of a simplified assignment model, including its inputs, modeling and assignment procedure, results, and comparison with approaches based on Railplan.

First, the motivations behind the development of the model are discussed. Although the approaches described in Chapter 3 - both, those found in the literature and those used in practice at TfL - are robust and provide valuable insight, there is a gap between over-simplified approaches that do not provide sufficiently detailed output to support decisions on investment, passenger information, or disruption mitigation strategies, and more complex approaches that provide the desired level of detail but are too computationally demanding to use for first-order analysis when many scenarios are being considered. This model aims to fill that gap.

Second, the simplified model’s inputs are described. For reasons of implementation and comparability, the model primarily draws its inputs from the Railplan network specification, consisting of nodes, the links that connect nodes, and the services that operate over these links. The demand that is loaded onto the network is also drawn from Railplan’s demand matrices.

Third, the model procedure is outlined. This section describes the model’s scope (restricted largely to the Underground), structure, route choice assumptions, assignment procedure, and calculation of various performance metrics.

Fourth, the model is applied to the following partial line suspension scenarios for the AM peak (07:00 to 10:00), 2021 network, that were introduced in Chapter 3:

- Piccadilly line suspension Kings Cross to Hyde Park Corner
- Piccadilly line suspension Kings Cross to Leicester Square
- Piccadilly line suspension Holborn to Hyde Park Corner

These are the “base,” “east side,” and “west side” scenarios respectively, as defined in Section 3.3.4, for the proposed crossover at Covent Garden. In addition to the model being useful for analyzing the impacts of real world disruptions, these applications suggest it can also be another way to estimate the benefit of installation of new crossovers.

Finally, the proposed model results are compared with the base and imperfect knowledge Railplan methodology results for the same scenarios.

4.1 Goals

The methodological and computational shortcomings of the Railplan approach to modeling disruptions outlined in Chapter 3 suggest the need for a simpler, faster tool for first-order analysis. This tool should allow rapid evaluation of alternative scenarios. It could also be used to produce a library of incidents for later matching with real-world disruptions, as is currently done using
Railplan, or to gain a high-level understanding of the benefit of certain incident mitigation strategies or investments, with more in-depth analysis reserved for those options found to be most promising. For example, a sensitivity analysis could be conducted to determine the relationship between the frequency operated on the portion of a line unaffected by an incident and the impact of that incident; the results of this could then inform decisions on real-world operational procedures that could increase the frequency of such residual service. Decisions on exact location of infrastructure investments, whether small scale such as crossovers or megaprojects such as Crossrail 2, could also be informed by using this tool to narrow the possibilities before more in-depth analysis.

These motivations inspire the development of a simplified assignment model. The model proposed here is a quasi-dynamic, shortest-path, London Underground-focused, fixed-demand model designed to analyze the impact of major service disruptions. The principal goal of this model is to provide a tool for rapid first-order analysis of the impact of disruptions on the Underground network, the results of which can then be used to analyze the impact of disruption mitigation strategies or the benefits of infrastructure investment. As a “sketch-planning” tool, this model can relax certain assumptions about passenger behavior, and omit certain detailed outputs, to decrease computation time, but must still produce credible forecasts. Aside from being a sketch-planning tool for first-order evaluation, this tool also aims to give decision-makers insight into the impacts of a disruption that may not be accurately represented with the existing Railplan approach. In addition, the flexibility present in a simplified model can allow TfL planners to more easily conduct specialized analyses than with a standard assignment model.

4.2 Inputs

The model requires several inputs to represent both supply and demand on the transport network. In this work, most of these inputs are drawn from TfL’s Railplan model to simplify model validation and comparison with existing modeling results. However, there are other sources that can be considered for many of these inputs; for example, demand can be drawn from Oyster data, rather than the RODS data used in Railplan.

The inputs, which are discussed in more detail in section 4.3, are as follows:

- Network structure (nodes and links). For this model, each entrance and line-direction pair at each station is assumed to be a separate node, as discussed in section 4.3.2.
- Services
  - List of routes (services) on the network
  - Frequency, and list of links served, for each route
  - Vehicle capacities
- Demand
  - Demand between each OD pair over the entirety of the study period
  - Demand profile over time
4.3 Procedure

This section describes the model procedure. A summary of the procedure is presented graphically in Figure 4-1.

Figure 4-1. Simplified assignment procedure

4.3.1 Scope

In striking a balance between accuracy and computational speed, neither a purely static nor a fully dynamic model is desirable. A static model does not consider the fluctuations of network supply and demand over time. Even Railplan, a static model which attempts to simulate this effect by varying supply and demand using assumed profiles (as described below), does not explicitly
consider any interdependence between time periods. This neglects one of the primary impacts of congestion - that passengers who are denied boarding during one time period, but continue waiting on the platform, increase the demand in the following time period. A fully dynamic model would take this into account, but would be computationally demanding and more difficult to calibrate, verify, and apply. In addition, aside from Schmocker et al (2008) there are very few dynamic frequency-based models in the literature; most are schedule-based, which introduces additional complexity and does not match well with passenger behavior on a high-frequency transit network such as the Underground (Gentile et al, 2016). To balance the advantages and disadvantages of these two approaches, a quasi-dynamic approach is selected, where the time period being considered (the three-hour morning peak, in the case study) is divided into fifteen-minute timebands, each of which is considered independently except when passengers are denied boarding due to overcrowding. In other words, within a given timeband, the network is analyzed only considering passengers specified by the departure time demand profile (discussed in Section 4.3.3) to be within that timeband. These passengers are assumed to complete their journeys entirely within the given timeband, except if they are denied boarding due to overcrowding; if this occurs, they are moved to the next timeband at the station where they are denied boarding.

The fact that passengers who are not denied boarding remain within their initial timeband regardless of journey time is, of course, an approximation. If the demand profile is assumed to represent passenger entries into the system, it is not perfectly accurate to consider someone traveling at 08:30 from an outlying station, such as Oakwood, to central London together with someone traveling at 08:30 from a closer-in station, such as Manor House, to central London, because the travel time between the two origins means that these two passengers will not be on the same train. However, the relatively small changes in demand between any two adjacent timebands mean that the impact of such a mismatch will not have a significant impact on the model results. In other words, considering passengers originating at 08:30 at Manor House together with passengers originating at 08:15 at Oakwood, which would be more realistic, would not significantly impact the model results due to the relatively small changes in passenger flow between 08:15 and 08:30. In addition, because more than 90% of Tube journeys during the morning peak begin in Zones 1-4 (Transport for London, 2018), the journey times of these passengers are sufficiently short relative to a fifteen-minute timeband that this approximation does not result in significant errors.

Of course, the choice of timeband length is important. Shorter timebands require greater granularity in demand data and make the assumption that passengers remain within one timeband for the entirety of their trip less plausible. Longer timebands, on the other hand, may fail to capture peakiness - the demand spikes over ten or fifteen minutes that cause the greatest congestion on the network - because it averages loads over a longer time period. For example, while the ten minutes at the peak of the peak may see denied boardings at a particular station, the ten minutes directly before or after may have enough unused capacity that using a thirty-minute timeband will not capture these denied boardings. The logical conclusion of extending timeband length is to consider the entire study period as a single timeband, which is essentially a static model with the drawbacks
discussed previously. To balance these factors, fifteen-minute timebands are selected for this study.

Unlike Railplan, which focuses on the entire public transport network in South East England but models some services across all of the United Kingdom, this model is focused on the London Underground. Consistent with the goal of developing a ‘sketch-planning’ tool that gives a high-level understanding of the impact of disruptions, all non-Underground modes and services are excluded. The Greater London transport network is vast and diverse, incorporating everything from short riverbus routes to high speed train services. While modeling it all, as Railplan does, is of course useful in the context of strategic planning, the computational effort required to do so may not be necessary here, given our objectives. Commuters from outside London are unlikely to alter their travel paths into central London as a result of a disruption on a Tube line, and even within London the majority of bus or regional rail routes will not be affected by a single Tube incident. Therefore, the scope of this model is restricted to the Underground, with two minor exceptions: auxiliary walk links between stations in central London and “overflow links” connecting suburban stations to central London. These links are specified based on the disruption being considered, as will be discussed below.

4.3.2 Structure

The primary component of the model is the network structure, modeled in Python’s NetworkX package as a directed graph. There are two graphs, one for the base case and one for the disrupted case, which are considered separately. The graphs do not have a time dimension; instead, the algorithm resets all relevant parameters in the graph (such as link volume) after considering each timeband. The properties of the graph’s links and nodes are described in this section.

Figure 4-2 schematically represents the types of links and nodes found in the graph. Except for the purposes of allocating PWT and OTT by line, the distinction between different lines (services) is not considered in the graphical representation.

Figure 4-2. Network representation as a graph
Nodes

In the directed graph there are three types of nodes: station entrance nodes, platform nodes, and dummy nodes. Each node is assigned a unique five (or six) digit number that represents its borough, station, and type (platform, station entrance, etc.), and has two properties: east and north coordinates expressed in meters using the British National Grid. These coordinates are only used to define auxiliary walk links; for all other walk links, survey data is used for an accurate representation of walking distances, and for non-walk links, timetable data is used for travel times.

All stations are modeled with separate nodes representing platforms and station entrances/exits, with walk links connecting them. Each station has one or more station entrances that are connected to (at least) two platform nodes. At most stations, platforms are defined separately for each line-direction combination. This means that most stations have (at least) three nodes: one for the northbound (or westbound) platform, one for the southbound (or eastbound) platform, and one for the station entrance/exit. Termini, even those such as Mill Hill East which have only a single platform, are modeled with two platform nodes, one for arriving passengers and the other for departing passengers. A complex station such as Baker Street has eight platform nodes: two (by direction) each for the Bakerloo, Hammersmith & City, Jubilee, and Metropolitan lines. Because the Circle line shares tracks and platforms with the Hammersmith & City line at Baker Street, it is not modeled using separate nodes.

There are several edge cases where modeling limitations required the creation of “dummy nodes” in the network specification. One of these represents international travel through Heathrow Airport. This is modeled as an additional station through which all Heathrow branch trains pass. The other type of node, also present at Heathrow, is a node with no transfer penalty that is modeled as the terminal of Piccadilly line services via Heathrow Terminal 4, which run via a one-directional loop and thus do not have a single, well-defined terminus. Although Piccadilly line links are modeled as running via these nodes, because dwell times are included in node-to-node travel times instead of being modeled at the node level, their presence does not significantly impact the model results.

Links and Services

In the simplified assignment model there are four types of links: in-station walk links, transit links, auxiliary walk links between stations, and overflow links. All links are unidirectional, having the following attributes: origin node, destination node, mode, actual travel time (in minutes), expected travel time (in minutes), and volume (which is an output of the assignment process). (Expected travel times are used when computing route choice, while actual travel times are used when computing perceived journey times, as described below). Transit links also have the following attributes: capacity, seated capacity, frequency, lines served, and number of denied boardings (which is another assignment output). Note that in the remainder of this section, “actual” travel time refers to unweighted time, “perceived” refers to weighted time used for performance metric calculation, and “expected” refers to weighted time used for route choice.
In-station Walk Links

In-station walk links connect platform and station entrance nodes. For each, the actual distance is derived from survey data, converted to time based on a walk speed of 5 km/h (3.1 mph), and then increased by the factors listed in Table 4-1, which are the same ones used by Railplan. Though these are not the same penalties as specified for walk links in the JTM, they are adopted here to ensure consistency with Railplan results. These penalties, combined with the factor of 2 applied to all walk links as discussed in Section 4.3.6, give a good approximation of the penalties for traversing different station elements specified in the JTM. These are used for both route choice and performance metric calculation.

Table 4-1. Penalties associated with walk link elements  (Transport for London, 2019a)

<table>
<thead>
<tr>
<th>Type of Delay</th>
<th>Time Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevator (up or down)</td>
<td>1 sec per vertical meter + 75 sec</td>
</tr>
<tr>
<td>Escalator (up or down)</td>
<td>1.5 sec per vertical meter</td>
</tr>
<tr>
<td>Stairs (up or down)</td>
<td>2.5 sec per vertical meter</td>
</tr>
<tr>
<td>Ticket gate</td>
<td>6 sec per gate crossed</td>
</tr>
</tbody>
</table>

Platform wait time (which is defined as the time until the first train arrives, not including denied boarding time which is calculated separately), is assumed to be half the scheduled headway at that platform and is included in the expected walking time for links that are used to access platforms. (For clarity of results, perceived platform wait times are computed and recorded separately from walk times.) Note that this definition of PWT assumes that all passengers are interested in boarding the first train to arrive; this neglects cases where multiple lines stop at the same platform but serve distinct destinations. This underestimates platform wait time at those platforms served by multiple lines, which include nearly all stations on the Circle line, or by multiple branches of a single line, as on the Central and District lines. In addition, platform wait time for same-platform transfers is not considered, although these are relatively infrequent on the Underground. It should be noted, however, that these simplifications are unlikely to have a significant impact on quantifying the impact of a major disruption such as a partial line suspension, particularly during the morning rush hour, for two reasons. First, the majority of passengers (64%) during the morning peak end their journeys in Zone 1 (Transport for London, 2018a), the center of London. Because their destinations are almost always on the trunk section of any branched line, or on the Circle line section of the sub-surface network where lines share track, the great majority of passengers are likely to want to board the first train rather than waiting for a specific line. (The exception to this is on the Northern line, which has two northern branches and two parallel trunk portions. For that line, this issue could be remedied by doubling all southbound platform wait times on the northern branches, since alternate trains from each branch serve each trunk). Second, most
of the impact of a disruption such as a partial line suspension is due to sections with no service, or with dramatically reduced frequency and hence capacity. This does not depend on the wait time for a train to a particular branch. Therefore, this underestimation takes place on both the base-case and disrupted-case scenarios, and thus does not affect the estimated impact of the disruption.

**Transit Links**

Transit links are those over which Underground trains operate. They connect platform nodes at different stations. Specific service patterns are not associated with the links that those services travel over. In other words, while the lines operating over a given link are recorded for allocation of platform wait time and on-train time by line, the aggregate travel time figures are only influenced by a link’s total frequency and capacity. Actual travel times are obtained from the network specification, or timetable information, and include dwell time as part of the travel time. Expected travel times also incorporate crowding discomfort penalties, which are used in the route choice model. The crowding penalties are calculated based on a typical day with no disruptions using Equation 4-1 in Section 4.3.6, the crowding penalty formula. In the case of a disruption, some passengers expect higher travel times on the disrupted line(s), as described in Section 4.3.4.

To determine the capacity of each transit link, the capacity of each vehicle must be known. While the seated capacity of an Underground carriage is well defined, the standing (and thus total) capacity depends on the assumed density of standing passengers. Assumptions on crush-load capacity vary: Railplan, for example, assumes a maximum density of 7 ppsm (standing passengers per square meter of standing room) for its crowding calculations, while most operators - including Transport for London in other applications - consider 4.5-5 ppsm the maximum practically achievable standing density (Hirsch and Thompson, 2011). In this model, a maximum density of 4.5 ppsm is assumed for all non-disrupted lines, and 5 ppsm for disrupted lines. The choice of standing density to use is, of course, up to the judgement of the modeler, but there are several factors to consider. The theoretical crush capacity of a vehicle, which is typically calculated at 7 ppsm, is achievable only with very high loading times, if it is achievable at all. Passengers are unlikely to willingly subject themselves to such crowded conditions in anything but the most dire emergencies, particularly on a high-frequency network such as the Underground where trains typically arrive every few minutes. Furthermore, the high dwell times required to achieve such loads impede the operation of the service; thus, a higher total throughput can be achieved by loading fewer people onto a given train, allowing the next train to arrive more quickly. Passengers on networks such as the Underground are generally aware of this tradeoff; thus, when there are no disruptions, they do not tend to overcrowd trains. When there are disruptions, however, they are typically more anxious to board the train, meaning that the number of people transported will increase. The exact numbers to use for both undisrupted and disrupted maximum crowding density are subject to the modeler’s judgement, or can be validated using automated passenger counters or other data sources.

There are several locations on the Underground network where representing real-world service patterns causes unrealistic results. This is primarily caused by the shortest path assumption
At locations on the network with infrequent direct services where a transfer would otherwise be necessary. For example, every fifteen minutes one Northern line train from Morden is scheduled to operate via the Charing Cross branch after Kennington instead of continuing via the Bank branch as do all others. Because the network specification models each platform as a separate node, and Kennington has separate platforms for each branch, this service is modeled as a link between Oval (the station south of Kennington) and the Charing Cross branch platform at Kennington. If this service did not exist, all passengers going between the Morden and Charing Cross branches would be modeled as transferring at Kennington from the Bank branch platform to the Charing Cross branch platform. Instead, because the shortest path between the two is on the direct service, all passengers are assigned to use the direct service and none are assumed to transfer at Kennington; this significantly overcounts the direct service and causes unrealistically high levels of denied boardings at Oval. Thus, it is necessary to model the shortest path as transferring at Kennington (which most passengers do anyway if no direct service is imminent); this is done by increasing the expected travel time on the link from Oval to Kennington (Charing Cross branch) to a very high number to discourage passengers from using it, and increasing the capacity of the link from Oval to Kennington (Bank branch) to include that of trains that will operate via Charing Cross. Similar adjustments are made at other locations on the network with similar service patterns. This overestimates platform wait time and walking time for passengers who would have taken a direct journey, but the effect on overall network-wide results is minor.

There are also several sections of the Underground, notably on the Piccadilly and Metropolitan lines, on which some trains operate express, bypassing some stations. This behavior is represented by having separate platform nodes for express trains at each bypassed station. These platform nodes are not connected by walk links to the remainder of the station, so passengers do not board or alight there and (because dwell times are not explicitly considered, as described above) the addition of these stations has no impact on model results.

One critical assumption made here is that service patterns are assumed constant throughout the study period. Of course, this is a simplification, because on most Underground lines service frequencies vary (somewhat) throughout the day, and even during a peak period. However, Transport for London’s efforts to lessen peak crowding by operating frequent service during the “shoulders” of the peak period, as well as practical limitations on run times and rolling stock and crew availability that make it impractical to operate maximum service frequency only during the 15 or 30 minutes when it is most needed, mean that service levels can be assumed to be roughly constant during the AM peak. A more accurate distribution of service frequency could be added to the model by specifying different service frequencies for different timebands, but the accuracy gain would need to be balanced against the increased model setup and computation time.

One important implication of a constant service pattern over the study period is that the disruption being analyzed is assumed to last for the entire study period - three hours for AM peak period scenarios. This is of course a simplification, but the reduction in the number of scenario parameters makes the model easier to verify and to apply. Service recovery strategies for short disruptions are usually the result of an ad-hoc process that varies from case to case, while service
during a prolonged disruption such as a partial line suspension is usually more consistent as controllers implement an emergency service pattern (Carrel, 2009). Therefore, for a sketch-planning tool it is useful to consider the network during a disruption to be in a steady state. This is also easier computationally - modeling a steady state disruption and then scaling its impacts based on incident duration, as London Underground does for its incident library, requires modeling fewer discrete incidents than would be necessary modeling various incident durations individually.

**Auxiliary Walk Links**

Auxiliary walk links are used to connect stations which are close together in Central London, on the theory that passengers are likely to walk to their destination if they can use the Underground to get to a station close to their destination during a disruption. For example, if the Piccadilly line is suspended between Kings Cross and Hyde Park Corner, a passenger going from Cockfosters to Russell Square will likely take the Tube to Kings Cross and walk from there to their final destination. Even when the destination station is still served by the Underground - Holborn, for example - it is often faster to walk above ground than to make an additional transfer underground. Which walk links to add depends on the scenario being studied, as well as the modeler’s judgement. The goal of adding these links is to represent those paths that passengers may walk during a disruption but not otherwise (if passengers would use these links without a disruption, this would be reflected in the initial demand matrix). In our case study, walk links are added in the disrupted case between all Piccadilly line station entrances from Kings Cross to Hyde Park Corner, inclusive. This strikes a balance between giving regular Piccadilly line passengers alternative paths through central London and not adding extraneous links that would allow passengers to use a non-Underground mode to complete a trip that the demand matrix specifies they take entirely on the Underground.

**Overflow links**

Overflow links represent mode shift away from the Underground for trips from suburban origins into central London. This is critical for scenarios where, as in the case study, there is a significant reduction in capacity on a section of the line, particularly a dead-end branch. Without the introduction of overflow links or some other way of modeling demand diversion, passengers attempting to board on a branch with severely reduced capacity will be faced with unrealistically long waiting times due to being denied boarding and be unable to change their route because no other Underground route to central London exists. Instead, these overflow links are meant to represent modal shift to parallel bus, National Rail, or other modes, and are generally used by passengers who arrive at a station and find a significant number of people waiting for a train (as described in the following section on route choice).

As with auxiliary walk links, the exact number and location of overflow links to include depends on the specific disruption and on the modeler’s judgement. There are typically at least two parallel rail alternatives for any given Tube line, as well as numerous alternatives involving buses, cycling, driving, etc. The goal of adding overflow links is to approximate those alternatives
which, as with auxiliary walk links, may be used by habitual Underground passengers during disruptions. A balance must also be struck between modeling simplicity - keeping the model focused on the Underground - and accuracy - reflecting the fact that, for example, many National Rail trains serve intermediate Tube stations on their way to their Zone 1 termini.

For the case study, these links are modeled in the disrupted case from each station between Manor House and Cockfosters to Kings Cross, and for each station between South Ealing and Hounslow West to Paddington and/or Waterloo (as appropriate), depending on proximity to TfL Rail or South Western Railway services respectively. Additional links are modeled to Gunnersbury and Ealing Broadway, for those cases where the detour route passes through those stations, for passengers not heading all the way to Central London who are still unable to board the Piccadilly line at their origin station in a reasonable time. Travel times on these links are calculated using Google Maps travel times for buses and National Rail services; the difference in travel time between an overflow link and its corresponding Underground route can vary significantly depending on which bus and rail routes are available nearby. For example, the overflow link time from Turnpike Lane to Kings Cross is 38 minutes, based on the Google Maps recommendation of using the 41 and 390 buses, as compared to the Piccadilly line run time between these stations of 19 minutes. The overflow link time from Cockfosters to Kings Cross, meanwhile, is 43 minutes, and assumes the use of the 298 bus and a Thameslink train. The Piccadilly line journey between these stations takes 35 minutes. Because these links are simply intended to approximate passenger behavior, and there are always several options - various bus routes, multiple parallel rail lines, or even other modes such as driving - these links are assumed to have infinite capacity, and crowding effects are ignored. The lack of crowding impact on overflow links represents an underestimation of the disruption impact, but given the links’ function as an approximation of actual passenger behavior it may be an acceptable assumption. In addition, discomfort penalties associated with walking, waiting, etc. on overflow links are modeled as a 20% increase in expected, and perceived, travel time. Though this parameter is obviously an approximation and is subject to calibration based on observed mode shift during actual disruptions, the additional walking, transferring, etc. that are assumed in the use of an overflow link make 1.2 a plausible estimate for this coefficient.

Another possible approach to the issue of mode shift away from the Underground would be to assume that some demand is suppressed, and either not consider these passengers within the model at all or provide an overflow link directly from their origin to their destination. Given the directional nature of passenger flows on the Underground during the AM peak period, this is likely a less reasonable approach. Within central London the Underground is practically the only last-mile distribution network: surface roads, including buses, do not have sufficient capacity for the passenger volumes that take the Underground; aside from Thameslink and soon-to-open Crossrail, National Rail trains do not serve destinations inside the Circle line, which is where many jobs are located; and walking is impractical for many trips in an area the size of central London, which is several square miles. Therefore, most passengers going to central London who are unable to use their preferred Tube line due to a disruption will complete their trip on a different Tube line.
Considering overflow links directly from origin to destination, or removing some passengers from the demand matrix altogether, would not take these rerouted passengers into account.

4.3.3 Demand

Transport for London has many sources of ridership data; each having advantages and disadvantages that are important to consider when choosing which to use in a model. As discussed in Section 2.2.1, the most accurate and comprehensive source is smartcard (bankcard and Oyster farecard) data. It provides information on journey origin, destination, and tap-in and tap-out time, resulting in a robust dataset for any assignment model. Despite these advantages, for historical reasons, ease of calibration, comparison with prior year results, and to incorporate route choice, Railplan’s demand matrices are calibrated using RODS data. For the sake of comparison with the Railplan model results, the proposed model also uses Railplan demand data as its source. Adapting the model to use Oyster data instead would be feasible, but the differences between Oyster data and RODS data would make the model’s results for properties like link-level loads more difficult to validate.

To avoid considering the entire UK public transport network, the simplified model uses demand figures that are restricted to just the Underground network. In other words, as extracted from Railplan, the origin for each passenger is the first Underground platform encountered on his/her trip, and the destination the last. This produces an origin-destination matrix with 645 unique origins (destinations) and approximately 68,000 pairs with some ridership. This matrix is modified so that all passengers begin their trips at a station entrance rather than at a platform. With the exception of passengers who enter the Underground without using a station entrance - chiefly those transferring from National Rail services that stop within the paid area of Underground stations - this accurately reflects passenger behavior. Further, all origin-destination pairs with less than one passenger over the AM peak period are removed. This change, done for computational reasons, reduces the number of OD pairs by 46% but total passenger volumes by only 0.5%. In the real world, of course, it is unrealistic for any origin-destination pair to have fractional demand; such figures are an artifact of the Railplan assignment procedure. For a study concerned with congestion, where passenger volumes on a single train can approach 1,000, such fractional demands are insignificant - especially when subdivided further into timebands - and thus it is reasonable to restrict modeling to those OD pairs with more than one passenger during the peak period. After this filtering, 36,224 OD pairs remain, with a total passenger flow of approximately 1.5 million passengers for the AM peak.

Next, this demand matrix is divided into timebands. As discussed below, this model is quasi-dynamic, with the study period (for the case study scenarios, the three-hour AM peak) divided into fifteen-minute timebands with demand assigned to each timeband. To represent fluctuations in demand, the demand allocation across timebands cannot be uniform; some demand profile must be assumed to avoid underestimating crowding during the peak of the peak. The demand profile can be defined based on a range of sources - passenger destination arrival times, origin departure times, time of crossing a cordon such as the Zone 1 boundary, or some other
metric. In a quasi-dynamic model such as this, the choice of demand profile will affect the results, because it will determine which passengers are considered together on which sections of the network. For consistency with Railplan results, the demand profile shown in Figure 4-3 is assumed to be the departure time distribution for all OD pairs.

![Figure 4-3. Railplan demand profile (adapted from Transport for London, 2019a)](image)

4.3.4 Route Choice

Passenger route choice is an extensively researched topic in the literature, with a broad range of approaches to model the heterogeneous behavior of passengers (Gentile et al, 2016). Some passengers prefer fewer transfers, or taking a slower path if that provides a greater chance of getting a seat. Other passengers, such as tourists, might prefer a shorter travel time but choose a suboptimal route due to lack of familiarity with the network. On the Underground, it is well known that passengers often take a range of routes between many OD pairs. According to TfL data, passengers heading from Kings Cross to Waterloo, for example, use more than a dozen different routes (Transport for London, 2017a).

As with any passenger assignment model, a route choice algorithm must be selected, which involves a tradeoff between accurately representing the behavior of each passenger and improving computational performance while focusing on getting credible aggregate results. As a sketch-planning tool, this model takes the latter approach, and assumes that all passengers choose their shortest paths in terms of expected travel time from origin to destination, weighted in accordance with the Journey Time Metric (with several caveats, as described below). This is a suitable approximation for modeling disruptions during the peak period, when the majority of passengers are commuters who prioritize travel time minimization more than other passengers (Borgesson and Eliasson, 2019). Infrequent riders are likely to use trip-planning applications to plan their trips,
which also generally recommend the fastest route. Thus, it is reasonable to assume that each passenger will try to get to their destination as quickly as possible, and concerns such as comfort are less likely to encourage passengers to take longer paths. Dijkstra’s algorithm (Dijkstra, 1959) is used to compute shortest paths.

These shortest paths are assumed to be independent of timeband. In other words, the shortest path between two points at 7:00 is assumed to remain the shortest path between those points at 8:30. This reflects the primary type of passengers during the morning peak - commuters - who are most sensitive to travel time and thus are not significantly influenced by factors such as crowding, which change over the course of the peak period. In addition, even those passengers who would be interested in changing their route to a less crowded one would be hard-pressed to do so because of the crowding levels present on all the Underground’s lines at the height of the AM peak period. According to Underground data (Greater London Authority, 2019), between 8:00 and 9:00 the least crowded line is the District, which sees 85% capacity utilization at its peak load point, and the most crowded is the Northern, where the equivalent figure is 130%. Finally, most passengers, particularly daily commuters, usually travel at a consistent time. This means that they are unlikely to know how the crowding on each link they are considering changes over the course of the peak period. By force of habit, they are likely to consider their usual route as the shortest if they travel at a different time, even if the changes in network crowding mean that a different route may, in fact, be better. A more detailed model could calculate the changes in shortest paths across timebands; doing so might provide more accurate results as passengers adapt to timeband-specific crowding levels. However, in the interest of computation time, such an approach is not pursued in the simplified assignment model described here.

Without disruptions, all passengers consider the expected travel time, which includes crowding based discomfort, on each link when determining the shortest path to their destination. For the case study, this is derived from the Railplan model, but in principle it could be derived from any equilibrium-based model or other representation of user preferences.

When there is a disruption, user expectations of the network are more complex. Passengers are typically made aware of the disruption through public address announcements, real-time apps, or some other means of communication. However, passengers do not know the crowding (and thus the perceived travel time) that results as passengers reroute themselves over the network. This gap between passenger expectation and passenger experience is critical to understanding the disruption impact.

When a disruption occurs on a line, passengers still expect service on other lines to be as on a normal (undisrupted) day. Even though passengers from the disrupted line may change their route, thus making other lines more crowded, passengers do not take the crowding impacts of this rerouting into account when selecting a route. While experienced passengers might be broadly aware of overall crowding trends in case of a disruption, it is unreasonable to expect them to take second-order crowding impacts into account when planning their route. This also means that passengers whose routes are not directly affected by a disruption do not reconsider their routes. For example, during a Piccadilly line disruption, a passenger choosing between the Northern and
Victoria lines to travel between Warren Street and Stockwell will be unaware that the Victoria line is significantly more crowded because many Piccadilly line passengers will have diverted to it. This passenger will not choose a new route, because their route is not affected by the disruption. The question of which passengers’ routes are “affected” is a subjective one, but according to TfL recommendations (Freemark, 2013) a service status of “Severe Delays” - which is typically posted during major disruptions including partial line suspensions - is intended to encourage passengers to change their routes. Therefore, for the scenarios in the case study, all links on the Piccadilly line are denoted as “affected” - this includes the suspended section itself, around which passengers obviously need to reroute, and the rest of the line, which sees a significant frequency reduction - when passengers are advised of “Severe Delays.”

For the disrupted line, on the other hand, passengers are assumed to be aware of the disruption. If there is a partial line suspension, they seek to avoid the sections that are suspended. For sections that have significantly reduced service, passengers are assumed to be aware of this reduction through service status apps or display boards. Although passengers are generally not informed of specific service frequencies during disruptions, a service status of “Severe Delays” is typically posted, and real-time departure information remains available. This means that most commuters will have a general understanding of the impact of a disruption on the service. This is translated into route choice by assuming that all passengers, except those boarding at the line’s most outlying stations, who typically get seats and thus would not experience discomfort from denied boarding or additional onboard crowding, expect line crowding to increase by the same factor by which frequency is reduced. For example, if the Piccadilly line, which typically operates with a frequency of 24 tph, is operating at 6 tph, for the purposes of route choice, passengers will assume that it is four times as crowded as usual. Because the crowding penalty function is quadratic, as discussed below, this creates a large disincentive to use the line. (The crowding penalty multiplier is capped at 1.5, however, to ensure that passengers are not overly discouraged from using the line. This cap is a parameter that can be calibrated upon further study of passenger behavior during disruptions).

![Figure 4-4. Example Network for Route Choice](image)

As an example to illustrate the route choice assumptions, consider the network shown in Figure 4-4, with the unweighted expected travel times on each link measured in minutes (as shown), and the headway on each line assumed to be two minutes. For a passenger traveling from...
SE1 (Station Entrance 1) to SE7, the shortest path is line A for the entirety of the trip, SE1-1-2-3-5-7-SE7, with an expected travel time of 19 minutes (the platform wait time at 1 is included in the walk link from SE1 to 1). A passenger traveling from SE2 to SE7 will, likewise, prefer to travel SE2-2-3-5-7-SE7 with an expected travel time of 15 minutes. Now, consider a disruption on line A, which forces a suspension of service between 5 and 7 and a 50% reduction in service on the rest of the line. Because the passenger traveling from SE1 to SE7 boards at the line’s origin, he/she will not expect any additional journey time as a result of the frequency reduction, aside from the increase in platform wait time. As a result, this passenger’s preferred route will be SE1-1-2-3-5-6-7-SE7, with a total expected journey time of 24 minutes (three minutes walking from SE1 to 1 and waiting at 1, twelve minutes travel time on line A, two minutes walking and waiting at 5, six minutes travel time on line B, and one minute walking from 7 to SE7). The passenger traveling from SE2 to SE7 will, however, expect an increase in travel time on line A because of the crowding caused by the frequency reduction. Supposing that the frequency reduction (crowding increase) causes the expected travel time for each line A link to increase by three minutes, then this passenger’s preferred route will be SE2-2-3-4-5-6-7-SE7, with an expected travel time of 25 minutes.

The crowding function used here, as discussed below, effectively tries to model both crowding discomfort and denied boarding time jointly. Denied boarding time is inherently node-based, which a given passenger only experiences once (when boarding), and not on every link traversed. However, estimating it as a link-based value is a useful approximation for several reasons. First, passengers who see “Severe Delays” on the status board generally perceive this as making their trip longer by some percentage, not by a fixed value (Freemark, 2013). As seen in Table 4-2 below, which shows the guidance given to network controllers on which service status to post during a disruption, such a status is often used when journey times are increased (due to a signal failure, track condition, or other reason), as well as for frequency reductions. Second, particularly on dead-end branches of lines with severe capacity reductions, it is nearly impossible for passengers to predict true denied boarding times. If there is no shift to overflow links or other lines, trains on these branches will be filled at the first few stations with passengers who are mostly traveling to central London, and passengers at downstream stations will be unable to board any arriving trains. Any estimate of denied boarding time based on this principle alone would be on the order of hours, as passengers at downstream stations wait for the peak period to end. Any passenger shift to overflow (or other) links, which occurs in reality and enables passengers at these downstream stations to board trains, is impossible for a passenger to estimate in advance when choosing a route. Thus, any passenger estimate of denied boarding time at a particular station that takes this mode shift into account would be arbitrary and inaccurate. Third, a node-based calculation of denied boarding time would not take into account passenger willingness to board an already overcrowded train that may have been deemed full by passengers at previous stations. As discussed previously, the standing capacity of a train is not a constant, but depends on passenger behavior and willingness to board, which itself depends on time already spent waiting, among other factors. Although a node-based estimate of denied boarding time could take these three
factors into account by making additional assumptions about passenger behavior, a link-based approximation such as the one used here is a useful approximation of the effects of denied boardings on passenger route choice.

Table 4-2. Service status guidelines. (Transport for London, via Freemark, 2013)

A. Minor Delays.

<table>
<thead>
<tr>
<th>Station</th>
<th>Trains between Heathrow Airport-Arnos Grove</th>
<th>Trains between Acton Town-Uxbridge and Arnos Grove-Cockfosters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Times</td>
<td>Other Times</td>
<td>Core Times</td>
</tr>
<tr>
<td>Headways</td>
<td>3x normal lasting &gt; 10 mins for more than one train</td>
<td>4x normal lasting &gt; 15 mins for more than one train</td>
</tr>
<tr>
<td>Trains moving slowly</td>
<td>&gt; 10 mins of blocking back with 3x normal headway for more than one train</td>
<td>&gt; 10 mins of blocking back with 3x normal headway for more than one train</td>
</tr>
<tr>
<td>Stoppage/Sit down</td>
<td>Between 10-15 mins</td>
<td>Between 15-20 mins</td>
</tr>
<tr>
<td>% of Scheduled Trains in Service</td>
<td>Between 75-85%</td>
<td>Between 70-85%</td>
</tr>
</tbody>
</table>

B. Severe Delays.

<table>
<thead>
<tr>
<th>Station</th>
<th>Trains between Heathrow Airport-Arnos Grove</th>
<th>Trains between Acton Town-Uxbridge and Arnos Grove-Cockfosters</th>
<th>All Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Times</td>
<td>Other Times</td>
<td></td>
<td>All Times</td>
</tr>
<tr>
<td>Headways</td>
<td>4x normal lasting &gt; 15 mins for more than one train</td>
<td>5x normal lasting &gt; 20 mins for more than one train</td>
<td>3x normal or there are &gt; 2 consecutive cancellations</td>
</tr>
<tr>
<td>Trains moving slowly</td>
<td>&gt; 15 mins of blocking back and/or trains being terminated early</td>
<td>&gt; 20 mins of blocking back and/or trains being terminated early</td>
<td>&gt; 20 mins of blocking back and/or trains being terminated early and/or trains working platform-to-platform</td>
</tr>
<tr>
<td>Stoppage/Sit down</td>
<td>&gt; 15 mins gap for more than one station</td>
<td>&gt; 20 mins gap for more than one station</td>
<td>&gt; 15 mins</td>
</tr>
<tr>
<td>% of Scheduled Trains in Service</td>
<td>&lt; 75%</td>
<td>&lt; 70%</td>
<td>&lt; 70%</td>
</tr>
</tbody>
</table>
4.3.5 Assignment

When demand on a transit line exceeds capacity, and some passengers are unable to board the first train to arrive, it is sometimes important to consider which passengers are able to board first and which passengers have to wait for another train. On branches of a line with severe reductions in capacity - in the case study, the sections of the Piccadilly line between Acton Town and Heathrow Airport, Acton Town and Uxbridge, and Finsbury Park and Cockfosters - the simplified model reflects this principle by assigning passengers to the network sequentially by origin station (first all passengers originating at Cockfosters, then those at Oakwood, etc.). This reflects the real-life behavior of passengers on the network, where those boarding further out fill up trains and prevent those further downstream from boarding. However, if the only desired model output is the aggregate system-wide total travel time, and the fraction of passengers heading to a given destination is the same for all origins, then model results are largely independent of passenger assignment sequence. Even a basic principle of real-world transit operations - that people already onboard a train get priority to continue on that train over people trying to board - can be neglected with no impact on overall travel time, because the aggregate figure is not affected if one passenger alights to wait for the following train while another passenger boards. Therefore, to speed computation time and simplify the model by making assignment order independent of route choice, once the passengers originating on the branches of a line with severe capacity reductions are assigned, the simplified model assigns all other passengers in numerical order by origin node number - that is, essentially arbitrarily. This does not change the link flows or denied boardings, even though it makes computing travel times and other metrics by passenger or OD pair much more difficult.

During the passenger assignment process, some links will be assigned more passengers in a fifteen-minute timeband than the available capacity. When this occurs, any passengers who cannot be accommodated, and do not choose to reroute due to overcrowding as discussed below, are added to the next timeband’s demand matrix, starting at the node that they have already reached and ending at their destination. This is not a realistic representation of actual passenger behavior, because passengers will not alight at an intermediate station to allow a waiting passenger to board. However, as with the OD pair order of passenger assignment discussed above, assuming that this behavior occurs does not impact the overall model results because the switching of two passengers does not change the aggregate system travel time.

This explicit treatment of passengers denied boarding differs from the Railplan approach, which considers denied boardings implicitly as part of the crowding function. The Railplan approach is appropriate for a strategic model representing steady-state behavior on an undisrupted network, where passengers are aware of congestion on the network and passengers generally do not choose routes on which they will be denied boarding for more than one or two trains. The model developed here, however, aims to analyze passenger flows during major disruptions. During such disruptions, there may be significant numbers of passengers denied boarding, whether due to insufficient capacity or insufficient passenger knowledge of the system state, and their inability to
pass through a given link significantly affects crowding downstream. Thus, it is reasonable to consider denied boardings explicitly.

Explicitly considering passengers being denied boarding can cause some passengers to still be waiting at the end of the AM peak period. Omitting these passengers from consideration would underestimate their travel time through the network; to avoid this, the model is extended to run past the end of the study period, but with zero new demand being introduced, to allow all passengers to propagate through the network. Although this underestimates the crowding impact that these passengers would experience from midday ridership, for the case study such impact is likely to be minor given the low crowding levels after the peak period ends at 10:00 (Transport for London, 2018).

One final real-world passenger behavior that must be modeled to obtain accurate results is the reconsideration of a route when significant overcrowding is encountered. When passengers come across such an overcrowded link that they do not expect to be able to board a train in a reasonable time, they will look for another route, if one with a reasonable journey time exists. The simplified model assumes similar behavior: if passengers arrive at a platform and find that the number of denied boarding passengers exceeds a quarter of the line’s total capacity per fifteen-minute timeband, they will find the next-shortest path that does not use the given link. If its expected travel time is less than thirty minutes longer than that of the current path, and they have not already re-routed themselves on their journey, they will switch to the new path. The exact figures specified here - 25% of line capacity and 30 minutes extra travel time - can be calibrated based on further study of revealed passenger preference, but are chosen here as reasonable estimates based on the author’s understanding of Underground station crowding and passenger behavior during disruptions. For a line operating on a two minute headway, for example, 25% of line capacity represents slightly less than two trainloads of people waiting on the platform. This threshold, representing a full platform with some spillover, is when newly arriving passengers are assumed to choose another route.

4.3.6 Performance Metrics

As per the Journey Time Metric, discussed in Chapter 2, travel on crowded trains is perceived as taking longer than the true travel time, because of the discomfort associated with crowding. In both Railplan and the simplified assignment model, each link is assigned a penalty that is a function of the crowding on that link over the three-hour peak period. The perceived travel time is then calculated by increasing the true travel time by this crowding penalty factor. As described in the Railplan User Manual (Transport for London, 2019a), the crowding penalty for a given link is calculated as follows (for the Underground; the coefficients differ for other modes):

\[ P = 0.0272 + 1.9337x + 0.9037x^2 \]  

(4-1)

Where \( x = (\text{Volume - Seated Capacity})/(\text{Crush Capacity - Seated Capacity}) \) and \( P = \) crowding penalty factor
This crowding function is calibrated to give reasonable results with crush capacities based on a standing density of 7 ppsm. Although the simplified assignment model assumes a maximum standing density of 4.5 ppsm or 5 ppsm for capacity and denied boarding calculations, for computation of crowding penalty the simplified model uses capacity based on 7 ppsm. This allows for fairer comparison with Railplan results.

Once passenger assignment is complete, this formula is used to compute perceived crowding, which is then converted to a crowding penalty factor by which true journey times are multiplied to give total perceived journey time. As discussed previously, this crowding formula implicitly includes denied boarding time, as well as pure crowding-based discomfort. Therefore, the simplified model’s explicitly calculated denied boarding times are considered in the final results as a component of the crowding penalty, and not in addition to it.

Passengers denied boarding clearly experience a delay that is not accounted for by simply adding them to the next timeband’s demand matrix; this delay must be calculated. This is done based on an estimate of the next time period’s demand on a given link, the number of passengers denied boarding on the link, and the link capacity. Notably, FIFO behavior on platforms is not assumed, because of heterogeneity in passenger behavior on crowded platforms. As with passenger assignment order, as discussed in the previous section, a more in-depth consideration of passenger boarding order would make it easier to calculate per-passenger travel time but would not change aggregate system-wide metrics.

If there is insufficient excess capacity in timeband \((n+1)\) for all passengers denied boarding in timeband \(n\) to be accommodated, then those passengers who are not denied boarding further (from timeband \((n+1)\) to timeband \((n+2)\)) will depart on the desired link at a uniform rate over timeband \((n+1)\). Thus, given that passengers are assumed to arrive at a uniform rate within timeband \(n\), these passengers experience an average denied boarding time of 15 minutes (i.e. the length of each timeband). Passengers who are denied boarding from timeband \(n\) to \((n+1)\), and then to \((n+2)\), will experience this 15 minute delay in timeband \(n\) in addition to the delay incurred in timeband \((n+1)\).

On the other hand, if there is sufficient excess capacity in timeband \((n+1)\) for all passengers denied boarding in timeband \(n\) as well as all passengers who arrive in timeband \((n+1)\), then the queue of passengers will dissipate before the end of timeband \((n+1)\). These passengers will experience a 7.5 minute denied boarding time from the middle to the end of timeband \(n\), and then an additional denied boarding time proportional to the proportion between the denied boarding volume and the excess capacity on the link. The following equations capture this behavior:

\[
t = 7.5 \times \left( 1 + \min\left( 1, \frac{L_n + V_{n+1}}{C} \right) \right)
\]

(4-2)

\[
V_{n+1} = V_n + \frac{f(b_{n+1})}{f(b_n)}
\]

(4-3)
Where:
\[ t = \text{denied boarding time penalty} \]
\[ L_n = \text{total number of passengers denied boarding in timeband } n \]
\[ V_n = \text{accommodated volume, excluding passengers denied boarding, in timeband } n \]
\[ C = \text{capacity of the current link} \]
\[ f(b_n) = \text{fraction of total system demand assigned to timeband } n \]

The model’s final step is to compute total passenger hours (H) from its components: for the Underground, platform wait time (PWT) until the first arriving train, unweighted on train time (OTT), crowding penalty (CP), denied boarding time (DB), and walk time on in-station links (WK), as well as walk time on auxiliary walk links (AW) and time on overflow links (OF). Denied boarding time is considered as a component of the crowding penalty for consistency with Railplan, so it is not added separately. Walking time on in-station links is taken to include the penalties associated with stairs, escalators, etc. outlined in Table 4-1. A factor of 2 is then applied to walk and platform wait/denied boarding times; this factor of 2 for all walk links, when combined with the factors outlined in Table 4-1 for station features such as stairs and escalators, produce penalties approximately consistent with the JTM. A factor of 1.2 is applied to overflow links as discussed in Section 4.3.2. This is summarized in Equations 4-4 and 4-5.

\[
H = \sum_i (w_i x_i) = w_{\text{pwt}} x_{\text{pwt}} + w_{\text{ott}} x_{\text{ott}} + w_{\text{cp}} x_{\text{cp}} + w_{\text{wk}} x_{\text{wk}} + w_{\text{aw}} x_{\text{aw}} + w_{\text{of}} x_{\text{of}}
\]

(4-4)

\[
H = \sum_i (w_i x_i) = 2 x_{\text{pwt}} + x_{\text{ott}} + x_{\text{cp}} + 2 x_{\text{wk}} + 2 x_{\text{aw}} + 1.2 x_{\text{of}}
\]

(4-5)

Where:
\[ w_i = \text{weight of component } i \]
\[ x_i = \text{unweighted value of component } i \]

4.4 Results

This section describes the application of the simplified model described above to three suspension scenarios on the Piccadilly line. First, the model is validated by comparison with the Railplan approaches. Next, several types of results are presented. First, the LCH impact of each suspension is computed, and total passenger travel time is divided by type (platform wait time (PWT), on train time (OTT), etc.). Next, PWT and OTT are broken down by line to provide a more in-depth understanding of the impact of each disruption on different parts of the network. The denied boarding volumes on the most overcrowded links are then presented; such data is useful to help Underground staff predict station overcrowding. Finally, model computation time is discussed.

The following scenarios are considered here for the 2021 network:
• Undisrupted scenario: Normal network, with no disruption
• Scenario 1: Piccadilly line suspension Kings Cross to Hyde Park Corner
• Scenario 2: Piccadilly line suspension Kings Cross to Leicester Square
• Scenario 3: Piccadilly line suspension Holborn to Hyde Park Corner

These are the same scenarios used in Chapter 3 to estimate the benefit of the Covent Garden crossover. Scenario 1 represents the suspension that would result if an incident occurred in the area given the present track infrastructure; Scenarios 2 and 3 represent the service patterns that would result if the crossover were installed and an incident occurred to its east or west, respectively. Each scenario assumes a 6 tph residual service frequency from Kings Cross (or Holborn) to Cockfosters and from Hyde Park Corner (or Leicester Square) to Heathrow Terminal 5, with additional shuttle services from Rayners Lane to Acton Town and from Heathrow Terminal 4 to Northfields on the western branches.

4.4.1 Validation

This section describes the validation of the model, which involves comparison with the two Railplan approaches. The comparison involves both aggregate system-wide figures, to determine the overall impact of the disruption on the network, and link-specific volumes, to understand the passenger behavior which comprises that impact.

In the sketch-planning context, it is most useful to compare high-level figures, such as passenger hour totals by category and disruption total LCH values. These are presented for the simplified model in Table 4-3, and are separated into categories: for the Underground, platform wait time (PWT), on train time (OTT), crowding penalty (CP), denied boarding time (DB, which is a subset of CP as discussed previously), and walk time (on in-station walk links), as well as walk time on auxiliary walk links and overflow link time. (This considers approximately 1.5 million passengers during the AM peak period, of whom approximately 11% use the Piccadilly line when there is no disruption, assuming the simplified model’s route choice algorithm.)

Table 4-3. Total passenger hours by category, simplified model

<table>
<thead>
<tr>
<th>Scenario</th>
<th>PWT</th>
<th>OTT</th>
<th>CP (incl. DB)</th>
<th>DB</th>
<th>Walk</th>
<th>Aux walk</th>
<th>Overflow</th>
<th>Total (weighted)</th>
<th>LCH</th>
<th>Lost min / pax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undisr.</td>
<td>85,863</td>
<td>398,792</td>
<td>202,832</td>
<td>10,411</td>
<td>128,855</td>
<td>0</td>
<td>0</td>
<td>1,031,058</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Scn. 1</td>
<td>95,185</td>
<td>391,743</td>
<td>227,510</td>
<td>26,108</td>
<td>129,547</td>
<td>3,682</td>
<td>13,845</td>
<td>1,092,692</td>
<td>61,634</td>
<td>2.47</td>
</tr>
<tr>
<td>Scn. 2</td>
<td>95,054</td>
<td>391,832</td>
<td>225,728</td>
<td>26,639</td>
<td>129,349</td>
<td>3,596</td>
<td>12,814</td>
<td>1,088,935</td>
<td>57,877</td>
<td>2.31</td>
</tr>
<tr>
<td>Scn. 3</td>
<td>95,745</td>
<td>392,741</td>
<td>227,245</td>
<td>29,841</td>
<td>129,353</td>
<td>2,554</td>
<td>12,063</td>
<td>1,089,764</td>
<td>58,706</td>
<td>2.35</td>
</tr>
</tbody>
</table>
Table 4-4. Total passenger hours for selected categories, base Railplan approach

<table>
<thead>
<tr>
<th>Scenario</th>
<th>OTT (Tube)</th>
<th>CP (Tube)</th>
<th>Walk (in-station, all modes)</th>
<th>Walk (other)</th>
<th>PWT (all modes)</th>
<th>LCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undisr.</td>
<td>403,800</td>
<td>202,819</td>
<td>183,343</td>
<td>752,685</td>
<td>238,754</td>
<td>--</td>
</tr>
<tr>
<td>Scn. 1</td>
<td>385,668</td>
<td>213,303</td>
<td>181,652</td>
<td>755,678</td>
<td>244,082</td>
<td>43,632</td>
</tr>
<tr>
<td>Scn. 2</td>
<td>386,647</td>
<td>212,857</td>
<td>181,677</td>
<td>755,525</td>
<td>244,203</td>
<td>41,911</td>
</tr>
<tr>
<td>Scn. 3</td>
<td>386,092</td>
<td>213,201</td>
<td>181,638</td>
<td>755,433</td>
<td>244,159</td>
<td>42,020</td>
</tr>
</tbody>
</table>

As a comparison, the base Railplan approach’s results for selected equivalent categories are presented in Table 4-4. The structure of the Railplan results makes comparisons for some categories infeasible – for example, PWT values are presented in aggregate for all modes – but where feasible these figures are presented only for the Underground. All component figures are unweighted (except CP, which are weighted figures by definition); LCH figures presented here are weighted sums of the components.

For the undisrupted case, the simplified model slightly underestimates OTT when compared to Railplan, but calculates the total CP to be virtually identical. For the disrupted cases, OTT is slightly higher in the simplified model, but CP increases more noticeably. Broadly, the simplified model assumes that fewer passengers will shift modes than Railplan does, and therefore that Underground links will be more crowded. In an equilibrium model such as Railplan, passengers have full knowledge of crowding when planning their trip. Thus, they are likely to avoid the links that are most crowded as a result of the disruption, such as those on the Piccadilly line, and shift to other lines or modes. In the simplified model, passengers do not have this knowledge and thus do not avoid such links, causing additional crowding there.

For a sketch-planning tool, the most important output is the total number of lost passenger hours. The overall LCH figures are consistently approximately 40% higher for the simplified model than they are in the base Railplan model. There are several reasons for this. First, the LCH figure represents the difference between passenger-hours values for two scenarios, and is thus inherently smaller than the passenger-hours values themselves; therefore, a relatively small change in total passenger time may cause a large change in LCH. For example, in Scenario 1, the simplified model estimates the crowding penalty as being roughly 14,200 passenger-hours higher than Railplan does, a 6.7% increase; meanwhile the simplified model’s LCH value for Scenario 1 – which includes crowding penalty as well as several other categories of travel time – is 18,000 passenger-hours higher than Railplan’s.

Second, the simplified model represents passenger behavior during an unplanned disruption that does not attempt to achieve equilibrium. This means that all passengers choose an optimal route given the behavior of all other passengers, which is clearly unrealistic. One of the main reasons why disruptions have such a large impact is because passengers do not know which route will be optimal when they begin their journeys. In addition, passengers who are not directly affected by the disruption will often prefer to stick to their habitual route (Yap, 2019), even though the increase in crowding that results from passengers who reroute themselves as a result of a disruption might mean that it would have been faster to choose another route. In short, passenger
route choice during disruptions tends to be suboptimal, both from the user and system perspective, which increases disruption impact; an equilibrium-based model does not take this into account.

Third, the simplified model focuses on the Underground. To some degree such a limitation on passenger behavior is unrealistic, as passengers will often consider regional rail or other modes if the Underground is disrupted and rerouting to those modes will provide a faster trip. However, there may be good reasons for passengers’ reluctance to change modes, including factors such as the Tube Map. For example, Guo (2011) notes that passengers prefer routes that appear shorter on a map even if they take longer in reality. Topham (2015) notes that according to Mike Brown, then Managing Director of the Underground, the addition of the London Overground services to the Tube Map played a large part in dramatically increasing their ridership. Therefore, it may reasonably be expected that not all passengers will consider modes that do not appear on the Tube map if they are not familiar with them. For Piccadilly line disruptions, of the three primary rail alternatives considered for overflow links only the TfL Rail service that operates from Paddington to the west is shown on the Tube map. Further research is warranted to explore mode shift to modes on the Tube map (such as other Underground lines or the Overground) as opposed to those not on the map (such as most National Rail lines).

Fourth, the simplified model assumes that all passengers take the shortest path from origin to destination. While commuters - most passengers during the morning peak period - prioritize travel time reduction more than other passenger groups do (Borgesson and Eliasson, 2019), there is always heterogeneity in passenger route choice. In particular, some passengers may prefer to take a slightly slower route if they think it will be less crowded. In a non-equilibrium model such as this one, where passengers choose a route before they know how crowded it will actually be, assuming that passengers take a broader range of routes (thus dissipating crowding) is likely to decrease total passenger travel time.

Despite these caveats, the simplified assignment model’s results follow the same trends as Railplan’s. Based on the simplified model’s LCH results in Table 4-3, the installation of the crossover at Covent Garden will reduce the impact of incidents to its west by 4.8%, and of incidents to its east by 6.0%. As was calculated with the base Railplan model in Chapter 3, the crossover is more beneficial during incidents to its east, because its service pattern enables the western end of the Piccadilly line to have direct connections to the central transfer stations, than during incidents to its west, which only extends Piccadilly line service to one additional transfer station.

The approach underlying the simplified assignment model shares many passenger behavior assumptions with the imperfect knowledge Railplan approach, so it is useful to compare the results from the two Railplan models with those from the simplified model. The particular implementation of the Railplan imperfect knowledge procedure means that it does not generate a bottom-line aggregate number of passenger hours or LCH as do the other two models. Thus, comparison must be made on a link-by-link basis. This analysis compares the two Railplan approaches to the simplified model for the Kings Cross to Hyde Park Corner suspension scenario (Scenario 1).

For an in-depth understanding of the differences in the results, the link loads for the AM peak period are compared for selected critical links between the three approaches – simplified
model, base Railplan, and imperfect knowledge Railplan – and with the total capacity over the peak period in Table 4-5.

Table 4-5. Comparison of flows on selected critical links

<table>
<thead>
<tr>
<th>Line</th>
<th>Simplified</th>
<th>Base RP</th>
<th>Imp. knowl. RP</th>
<th>Capacity (simplified model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounds Green - Wood Green</td>
<td>Piccadilly</td>
<td>8,736</td>
<td>7,833</td>
<td>13,116</td>
</tr>
<tr>
<td>Manor House - Finsbury Park</td>
<td>Piccadilly</td>
<td>17,457</td>
<td>12,462</td>
<td>26,113</td>
</tr>
<tr>
<td>Finsbury Park - Highbury &amp; Islington</td>
<td>Victoria</td>
<td>82,880</td>
<td>78,844</td>
<td>86,497</td>
</tr>
<tr>
<td>Osterley - Boston Manor</td>
<td>Piccadilly</td>
<td>10,438</td>
<td>10,537</td>
<td>11,434</td>
</tr>
<tr>
<td>South Ealing - Acton Town</td>
<td>Piccadilly</td>
<td>11,097</td>
<td>10,644</td>
<td>13,829</td>
</tr>
<tr>
<td>Sudbury Town - Alperton</td>
<td>Piccadilly</td>
<td>8,070</td>
<td>4,054</td>
<td>5,184</td>
</tr>
<tr>
<td>North Ealing - Ealing Common</td>
<td>Piccadilly</td>
<td>10,886</td>
<td>4,816</td>
<td>6,560</td>
</tr>
<tr>
<td>South Kensington - Sloane Square</td>
<td>Circle/District</td>
<td>59,108</td>
<td>54,005</td>
<td>58,511</td>
</tr>
<tr>
<td>Knightsbridge - Hyde Park Corner</td>
<td>Piccadilly</td>
<td>1,100</td>
<td>1,014</td>
<td>1,497</td>
</tr>
</tbody>
</table>

Of the three models, the base Railplan model assumes that passengers have the most knowledge; the imperfect-knowledge Railplan model assumes that they have the least knowledge. Therefore, because passenger knowledge of crowding and residual service directly influences route choice and thus link crowding, if all else is equal the link flows as calculated by the simplified model should lie between the base and imperfect knowledge results. The base Railplan approach, unlike the simplified model, assumes that passengers are aware of the behavior of other passengers, which means that passengers will consider crowding on their desired links when selecting a route, and are therefore more likely to avoid these links. Therefore, the base Railplan approach can be generally expected to provide a low estimate of the volumes on critical links - those that are disrupted or primary alternatives to disrupted links. The imperfect knowledge approach, on the other hand, assumes that 47% of passengers are unaware of the disruption, which means that they are likely to choose their usual route as long as some service is still provided along it, regardless of any frequency reduction or the crowding that may result. This, in addition with Railplan’s lack of explicit consideration of capacity constraints, means that the imperfect knowledge approach is likely to over-estimate link flows on critical links during a disruption. The simplified model, on the other hand, assumes that all passengers are aware of the disruption and assign a higher penalty to using the disrupted line when choosing a route, but are not aware of the behavior of other passengers, so its estimate should lie between that of the two Railplan model results.
This expectation is met on most of the critical inbound links shown in Table 4-5. However, there are some exceptions that provide additional insight into the assumptions underlying the different models. On the outer branches, passenger volumes are highly dependent on the weights and travel times of the overflow links. On the Rayners Lane branch, where due to comparatively low passenger volumes, no overflow links are added, passenger volumes in the simplified model significantly exceed those from either of the Railplan approaches. On the other hand, the South Western Railway Hounslow loop, which is represented by overflow links, is very close to the outer end of the Heathrow branch in the Hounslow area. Therefore, it presents an attractive alternative to the Piccadilly line and many passengers divert to it, decreasing the passenger flow on links such as South Ealing – Acton Town. This does not occur to the same extent in either Railplan model, likely because those models take its headways and capacity constraints into account, while the simplified model does not consider capacity constraints on overflow links. This is a compromise made to restrict the simplified model’s scope to the Underground; more precise calibration of mode shift may be possible through more careful specification of overflow link characteristics.

The increase in traffic on the South Kensington – Sloane Square link in the simplified model, compared to the other two models, is likely caused by the lack of explicit consideration of bus links in the simplified model. In this model passengers traveling from Knightsbridge, in particular, to central or eastern London must either board an eastbound Piccadilly line train one stop to Hyde Park Corner, ascend to the surface, and walk east until they reach a convenient transfer station, or board a westbound Piccadilly line train and change at South Kensington to the Circle and District lines. The latter path has a lower generalized cost, so most passengers use it, causing additional crowding on those links. In the base and imperfect knowledge Railplan models, on the other hand, these passengers would simply use a bus from Knightsbridge directly to central London, which is faster than walking or using the Circle/District lines, before boarding the Tube for the remainder of their journeys. The small number of stations with a similar phenomenon means that this effect has limited impact on the overall model results, but for some links it can overestimate crowding. In future work it may be desirable to include bus overflow links in limited situations to address this issue.

Notably, the capacity over the entire AM peak period is significantly higher than the total volume for most of the links in Table 4-5, despite the fact that some of these links have significant numbers of denied boardings, as shown in Table 4-9 below. This illustrates the effect that peakiness – fluctuation in demand within the peak period – has on the network. If passengers were evenly distributed throughout the three-hour period, almost all of the network’s links would have sufficient capacity to accommodate all passengers selecting that route, even on the disrupted Piccadilly line. The link from Manor House – Finsbury Park, however, would be overcrowded even if passengers were evenly distributed throughout the peak. This reflects the inability of passengers at that station to board trains that consistently arrive full. In the simplified model, the impact of this overloading is taken into account by extending the model to run past 10:00 AM with no new demand, to ensure that all passengers are accommodated, as discussed in Section 4.3.5.
4.4.2 Analysis

A better understanding of the simplified assignment model’s results can be obtained from examining the experience of an average passenger, based on the data presented in Table 4-3 in the previous section. This can be estimated by dividing the total number of passenger hours by travel time category by the number of passengers on the network. Considering unweighted travel times for the undisrupted case, an average passenger spends 3.4 minutes of their journey waiting for the first train to arrive at their origin or at a transfer station, 16.0 minutes traveling onboard a train, 0.4 minutes waiting for another train after being unable to board the first one, and 5.2 minutes walking within the station. Thus, considering unweighted travel time, the average passenger spends 25.0 minutes traveling on the network. This passenger’s generalized journey cost, which includes travel time as well as crowding- and walking- based discomfort, is 41.2 minutes. Such journey times are approximately consistent with what would be expected from the demand matrix, which states that over 80% of journeys (1.24 million of 1.5 million) in the morning peak begin in Zones 1-3, and over 90% (1.38 million of 1.5 million) in Zones 1-4, especially when the shortest-path assumption in the simplified model is considered. These figures, which suggest that passengers spend over a third of their time in the system not on a train, are useful for Underground decision-makers to choose where to focus investment meant to improve ambience and passenger experience, for example. Low-cost improvements to the in-station experience, such as a 2016 pilot project which encouraged passengers to stand on both sides of the escalator at Holborn (rather than standing on one side and walking on the other) and cut congestion by 30% (BBC News, 2017), may be justifiable.

Next, the changes in passenger time as a result of each disruption can be explored. Platform wait time (PWT) increases for the disrupted cases, as expected, by just over 10% network-wide. Part of this is caused by the Piccadilly line’s quadrupling of headways, on the section of the line that remains operational; total Piccadilly line PWT increases by approximately 55%, or from 8,649 hours to approximately 13,500 hours. The remainder is caused by passengers transferring to other lines; PWT on all other lines combined increases by roughly 5%, from 77,213 hours to roughly 81,500 hours. On train time (OTT) decreases, but as seen in Table 4-8 this decrease is more than compensated for by the increase in overflow link time. This reflects the mode shift that occurs during many large Underground disruptions: although passengers spend less time on the Tube, their journeys take longer overall because of increased use of other modes. (More detailed PWT and OTT data is presented in Tables 4-9 and 4-10 below). Crowding penalties increase by 12%, reflecting the increased volume-capacity ratio on the Piccadilly line and the spillover crowding that the suspension causes on the rest of the network. Walking times are also higher on the entire network, reflecting the degree to which passengers must make additional transfers to complete their trip, but also the use of auxiliary walk links in central London. In Scenarios 1, 2, and 3 respectively, the total passenger flow on auxiliary walk links is 21,792, 21,341, and 16,419. The lower value for Scenario 3 reflects the extension of residual Piccadilly line service from Kings Cross to Holborn which shifts passengers from overflow links – where after arriving at Kings Cross they are more likely to use
an auxiliary walk link than to descend into the Underground station and wait for a train – to the Piccadilly line. Scenario 2’s extension from Hyde Park Corner to Leicester Square, on the other hand, does not cause such a dramatic decline in auxiliary walk link usage because of the larger number of Tube alternatives on the western side of the network and the fact that the western-half overflow links connect to National Rail stations not on the Piccadilly line, thus helping disperse passengers around the network. As discussed below, it is clear that the choice of overflow and auxiliary walk links to model has a significant effect on model results.

Notably, for most categories the percentage change in passenger hours is nearly constant regardless of disruption. On the Underground, network-wide PWT, for example, increases by 10.8%, 10.7%, and 11.5% for Scenarios 1, 2, and 3 respectively; similarly, network-wide OTT decreases by 1.8%, 1.7%, and 1.5%. These relatively minor variations imply that it is not the length of the suspended section that is the primary determinant of passenger impact, but rather the existence of the suspension in the first place. The loss of ability to make through trips on the line as a result of the suspension, along with the overcrowding present due to the dramatic frequency reductions, play the dominant role in passenger impact compared with the distance over which the service is suspended.

Next, the impact of overflow links on model results is examined. Table 4-6 shows the total usage of inbound overflow links in each scenario, divided by which end of the Piccadilly line they serve, and compares these figures with the links at the inner end of the corresponding sections of the Piccadilly line itself.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Overflow links (east)</th>
<th>Manor House – Finsbury Park</th>
<th>Overflow links (west)</th>
<th>South Ealing – Acton Town</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undisr.</td>
<td>0</td>
<td>31,465</td>
<td>0</td>
<td>17,565</td>
</tr>
<tr>
<td>Scn. 1</td>
<td>13,966</td>
<td>17,457</td>
<td>6,361</td>
<td>11,097</td>
</tr>
<tr>
<td>Scn. 2</td>
<td>13,966</td>
<td>17,457</td>
<td>5,141</td>
<td>12,336</td>
</tr>
<tr>
<td>Scn. 3</td>
<td>10,993</td>
<td>20,436</td>
<td>6,362</td>
<td>11,097</td>
</tr>
</tbody>
</table>

Clearly, the presence of a disruption (and its corresponding frequency reduction) are primary factors in the usage of overflow links, with the service extension as a result of crossover installation playing a secondary role. On the eastern half of the line, when Piccadilly line service terminates at Kings Cross (in Scenarios 1 and 2) 44% of passengers divert to overflow links. Extending this service to Holborn (as in Scenario 3) reduces this to 35%. On the western half of the line, a similar phenomenon is observed – 36% of passengers divert to overflow links when the Piccadilly line operates to Hyde Park Corner (in Scenarios 1 and 3), but only 29% do so when the line is extended to Leicester Square (in Scenario 2).

Of course, these values have several caveats that must be considered. First, they are heavily dependent on proper calibration: a change in overflow link weighting or in passenger willingness to use a disrupted line could significantly alter these results. Though the values chosen for this
study are based on estimates of passenger behavior and preliminary validation and yield promising results, more careful calibration – with AFC data, passenger surveys, or other data sources – is necessary. Second, overflow links do not have headways or capacities, because they are approximations of many different services and modes that a passenger could use instead. Though such approximation is necessary to represent passenger behavior accurately while avoiding modeling the entire transit network, it remains a simplification that requires careful consideration during model setup and result interpretation. Third, the choice of which links to represent is also important, as the presence or absence of certain overflow links can alter model results. As discussed above, the simplified model’s lack of bus links near Knightsbridge station causes it to represent the Circle and District lines as more crowded than they are in reality. Likewise, it may be justified to include Crossrail and Thameslink links within the Circle line, as these lines’ integration with the Underground means that passengers are likely to divert to them during a disruption. This is a modeling decision that must be made carefully, and is dependent on the disruption being modeled and on observed route choice during past disruptions.

Next, the line-by-line values for PWT and OTT are presented in Tables 4-7 and 4-8. For the purposes of this analysis, on sections where multiple lines share track - primarily along the Circle and Hammersmith & City lines - passengers are assigned to the Circle line on those links where it operates or to the District or Metropolitan where it does not. This is a rough approximation - obviously some passengers will use lines other than the Circle on common sections if those trains arrive first. For this model, which does not consider separate lines on common sections during route choice or overall performance metric calculation, such an approximation is used out of convenience because it allows a clear separation between passengers using the District or Metropolitan lines within Zone 1 (i.e. along the shared section with the Circle line) and passengers using them on the outlying sections of the network. Such a separation is useful to understand the difference between passenger behavior in Zone 1, where there are typically several possible Underground routes to a given destination, and that on the rest of the network, where there are usually fewer options available. Distributing passengers between shared lines based on those lines’ relative frequencies is possible, but less desirable in this context.

### Table 4-7. Line by line OTT for selected lines, in passenger hours (unweighted)

<table>
<thead>
<tr>
<th></th>
<th>Circle/H&amp;C</th>
<th>District</th>
<th>Metropolitan</th>
<th>Piccadilly</th>
<th>Victoria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undisr.</td>
<td>58,537</td>
<td>42,466</td>
<td>26,244</td>
<td>41,548</td>
<td>40,875</td>
</tr>
<tr>
<td>Scn. 1</td>
<td>64,244</td>
<td>47,042</td>
<td>26,491</td>
<td>18,679</td>
<td>44,025</td>
</tr>
<tr>
<td>Scn. 2</td>
<td>62,823</td>
<td>46,676</td>
<td>26,445</td>
<td>20,662</td>
<td>43,934</td>
</tr>
<tr>
<td>Scn. 3</td>
<td>64,203</td>
<td>47,037</td>
<td>26,490</td>
<td>20,934</td>
<td>43,094</td>
</tr>
</tbody>
</table>
Table 4-8. Line by line PWT for selected lines, in passenger hours (unweighted)

<table>
<thead>
<tr>
<th></th>
<th>Circle/H&amp;G</th>
<th>District</th>
<th>Metropolitan</th>
<th>Piccadilly</th>
<th>Victoria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undisr.</td>
<td>12,635</td>
<td>10,329</td>
<td>4,381</td>
<td>8,649</td>
<td>7,361</td>
</tr>
<tr>
<td>Scn. 1</td>
<td>13,470</td>
<td>11,525</td>
<td>4,921</td>
<td>13,512</td>
<td>7,957</td>
</tr>
<tr>
<td>Scn. 2</td>
<td>13,200</td>
<td>11,536</td>
<td>4,917</td>
<td>13,726</td>
<td>7,936</td>
</tr>
<tr>
<td>Scn. 3</td>
<td>13,459</td>
<td>11,525</td>
<td>4,921</td>
<td>14,466</td>
<td>7,769</td>
</tr>
</tbody>
</table>

As expected, large increases in PWT and OTT are observed on the District line, which shares several stations with the Piccadilly line in western London. Because the service pattern for each disruption considered here assumes 6 tph service from the Heathrow branch to central London, and 6 tph shuttling between Rayners Lane and Acton Town, many passengers arrive at Acton Town, are unable to board a Piccadilly line train to continue their journeys eastward, and divert to the District line, which is a slower route to central London. On the eastern end of the line, the Victoria line is the primary alternate route; connecting with the Piccadilly at Finsbury Park, as well as at Kings Cross and Green Park in central London. However, the Victoria line does not see as substantial an increase in PWT and OTT as the District line, because Finsbury Park is sufficiently far downstream that the majority of Piccadilly passengers have already been denied boarding farther upstream and have diverted to overflow links. On the Piccadilly itself, as expected, total OTT decreases, because the fourfold decrease in frequency means that fewer passengers are able to use the line, and PWT increases, because each passenger’s PWT increases in proportion to the frequency decrease. Results for other lines are a useful proxy for the degree to which those lines are used as alternatives to the Piccadilly. For example, the Waterloo & City line sees negligible increases in PWT and OTT, which suggests that few Piccadilly line passengers find it useful during a disruption. The Circle and Hammersmith & City lines see a large increase, on the other hand, likely because many passengers use them (or the District or Metropolitan lines on shared track sections) in Zone 1 to navigate around the Piccadilly line suspension.

One further area of interest for Underground staff during any disruption is the volume of passengers, and particularly the number of people unable to board trains, across the network. The small size and high usage of many Underground stations means that they are vulnerable to overcrowding during disruptions; controllers and station staff always consider station crowding - which is in large part caused by denied boardings - during a disruption. The relevant links (i.e. those on the Piccadilly line and on primary diversion routes) with the most denied boardings (at the previous station) for the three disruption scenarios are presented in Table 4-9, along with the number of times a passenger is denied boarding on each. (If a passenger is denied boarding in two consecutive timebands, that passenger will be counted twice). This data is also presented for the same links on the undisrupted network for comparison purposes.

Notably, Scenario 3 sees higher denied-boarding volumes on the eastern section than the other two disruption scenarios. This is caused by the service pattern in Scenario 3, which has trains
Table 4-9. Relevant links with most denied boardings

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Undisr.</th>
<th>Scn. 1</th>
<th>Scn. 2</th>
<th>Scn. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnpike Lane Piccadilly</td>
<td>Manor House Piccadilly</td>
<td>0</td>
<td>22,516</td>
<td>22,516</td>
<td>32,012</td>
</tr>
<tr>
<td>(WB)</td>
<td>(WB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highbury &amp; Islington</td>
<td>King's Cross St Pancras</td>
<td>7,900</td>
<td>15,694</td>
<td>15,694</td>
<td>8,928</td>
</tr>
<tr>
<td>Victoria (SB)</td>
<td>Victoria (SB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euston Victoria (SB)</td>
<td>Warren St Victoria (SB)</td>
<td>0</td>
<td>11,774</td>
<td>12,644</td>
<td>8,944</td>
</tr>
<tr>
<td>Victoria District (EB)</td>
<td>St James's Park District (EB)</td>
<td>1,420</td>
<td>8,013</td>
<td>5,340</td>
<td>8,114</td>
</tr>
<tr>
<td>Wood Green Piccadilly</td>
<td>Turnpike Lane Piccadilly</td>
<td>0</td>
<td>7,781</td>
<td>7,781</td>
<td>16,286</td>
</tr>
<tr>
<td>(WB)</td>
<td>(WB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finsbury Park Victoria</td>
<td>Highbury &amp; Islington</td>
<td>1,452</td>
<td>4,862</td>
<td>4,862</td>
<td>1,027</td>
</tr>
<tr>
<td>(SB)</td>
<td>Victoria (SB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manor House Piccadilly</td>
<td>Finsbury Park Piccadilly</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>6,437</td>
</tr>
<tr>
<td>(WB)</td>
<td>(WB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

from Cockfosters operating as far as Holborn, as opposed to operating only to Kings Cross, which means that the line is more attractive to passengers because it serves more destinations directly. For the same reason, Scenario 3 sees lower denied boarding volumes on the Victoria line and 12% lower total volumes on overflow links - the Piccadilly line is on the shortest path for a greater number of passengers, so fewer passengers transfer at Finsbury Park or Kings Cross to the Victoria line or use overflow links to avoid the Piccadilly line entirely. This suggests a paradox with which Underground line controllers are very familiar – enhancing the service provided during a disruption might make it so attractive to passengers as to be dangerously crowded, requiring it to be shut down and further worsening the impact of the disruption.

Scenario 2, meanwhile, sees lower denied boarding volumes than the other disruption scenarios on the District line in the southern part of Zone 1, near Victoria station. When the Piccadilly line is truncated to Hyde Park Corner in Scenarios 1 and 3, passengers heading to Green Park, Piccadilly Circus, or Leicester Square must take the District line and change to northbound service at Victoria or Embankment. Scenario 2’s extension of the Piccadilly service to Leicester Square allows passengers to reach those stations directly and thus relieves the load on the District and Victoria lines near Victoria station. This relief of parallel lines is one of the main reasons for service extensions to be desirable, if these extensions do not cause excessive overcrowding on the affected line itself. As expected, Scenario 2 also sees a 6% decrease in overflow link usage as more passengers choose the Underground instead of shifting to other modes; this is lower than the decrease observed in Scenario 3 because, even when the Piccadilly line is truncated to Hyde Park Corner, there are still transfers available to other lines at six stations, allowing passengers to choose among several Underground options; in Scenario 3, on the other hand, aside from changing to the
already-crowded Victoria line at Finsbury Park, Underground passengers have no alternative but to remain on the Piccadilly line all the way to Kings Cross, which makes the overflow links more attractive by comparison.

Notably, the Piccadilly line link from Manor House to Finsbury Park has much lower denied boarding volumes than do the Piccadilly line links farther away from London. Although the volume from Manor House to Finsbury Park is obviously higher than that from Turnpike Lane to Manor House, for instance, the lack of passengers denied boarding at Manor House shows that most passengers originating there choose to use an overflow link. Indeed, the availability of a quick bus connection from Manor House to Kings Cross (which is absent at Turnpike Lane) diverts almost all passengers in Scenarios 1 and 2; in Scenario 3, however, the Piccadilly line extension to Holborn attracts many passengers back to the Tube and causes denied boardings at Manor House as well. This further underscores the influence, discussed above, that the specification of overflow links has on model results.

4.4.3 Computation Time

In terms of computation time, the simplified model significantly outperforms Railplan, though the magnitude of the difference is difficult to determine because the two models are run on different computing platforms. The simplified model is run on the author’s laptop computer, which has a 4-core Intel i7 CPU with 12 GB memory and runs Windows 10. On this computer, the simplified model takes less than twenty minutes to run two scenarios: one disrupted case and one undisrupted case. The Railplan models, on the other hand, are run on a dedicated TfL Windows server which has an Intel XEON E5-2697 CPU with 56 processors (Transport for London, 2019a), though to allow multiple analyses to be run in parallel, most model runs only use four processors. On this server, the base Railplan methodology requires one to two hours to run, and the imperfect knowledge methodology (which requires a base case to have been run as a prerequisite) takes approximately twenty minutes.

4.5 Conclusions

This chapter proposes a simplified passenger assignment model. The overarching goal of this model is to provide a tool for first-order sketch-planning analysis of the impact of unplanned disruptions on the London Underground. Most other tools in the literature are either too oversimplified to generate useful insight, or are so computationally intensive that they are infeasible to use for sketch planning-type analysis. This model aims to fill this gap.

To do this, the model makes several assumptions regarding passenger behavior and other modeling simplifications. The scope of the model is limited to the London Underground, with representation of other modes limited to overflow links for parts of the Underground with insufficient capacity during a disruption. The choice of which links to specify, and how much to weight them to account for the transfer, walking, and crowding penalties they incur, are assumptions that have important implications on model results. Though reasonable values are
chosen in this study, this warrants further study if the model is to be adopted on a larger scale. The network is considered in a quasi-dynamic manner, with demand and supply separated into fifteen-minute timebands, each of which is considered separately except for passengers who are denied boarding due to overcrowding. Passengers are assumed to be aware of the disruption and consider changing their route if their originally preferred one includes an affected link. All passengers are assumed to choose the shortest path between their origin and destination, and only deviate from it if they encounter overcrowding which makes it impossible to continue on their desired route in a reasonable time. As discussed in Section 4.3, these assumptions provide a reasonable approximation of passenger behavior during unplanned disruptions while maintaining the fast computation time that initially motivated the creation of the model.

This model shows promise in efficiently analyzing the impact of unplanned disruptions on the Underground. Its computation time provides a significant improvement over Transport for London’s existing modeling tools. It generates results related to link flow, passenger mode choice, and components of travel time that allow Underground stakeholders to assess the impact that a disruption will have on the network. Its estimates of disruption impact are consistently about 40% higher than those of Railplan, which is the current standard approach to modeling unplanned disruptions, because of the Underground-only scope, lack of passenger knowledge of other passengers’ behavior, and shortest-path assumptions that underlie the simplified assignment model. However, given the flaws in using the base Railplan methodology for modeling unplanned disruptions as discussed here and in Chapter 3, the simplified model’s results are likely to more accurately represent the actual passenger impact of unplanned disruptions.


5 Summary and Conclusions

This chapter summarizes the research, outlines the contributions of this thesis and suggests directions for future research in the areas of track layout analysis and modeling of major disruptions. First, the thesis is summarized, along with the most important findings. Next, several limitations in the modeling conducted here are discussed; although this work contains useful insight for Underground stakeholders, as well as other system operators and researchers, these caveats should be kept in mind. Then, recommendations to Transport for London are presented. Finally, several directions are proposed for future research.

5.1 Summary

The primary goal of this thesis is to develop a simplified assignment model to analyze the impact of major disruptions on the Underground, and determine the degree to which changes to track layout can help mitigate this impact.

First, the relationship between track layout and network performance during major incidents is discussed. Track layout features such as crossovers and sidings allow line controllers to maintain residual service on the unaffected portion of a line during a disruption; without these features, less residual service could be operated, resulting in a greater loss of passenger time. However, the mere presence of these track layout features on a line is insufficient; they must be well-located, have enough capacity to accommodate a reasonable fraction of the Underground’s high service frequency, and be convenient for controllers to use. The factors that affect a crossover or siding’s ability to satisfy these criteria are not well documented in the literature; instead, a wide range of site-specific constraints, controller practices based on experience, and ad-hoc judgements are used to determine where to place, or how to use, a crossover or siding.

These factors are explored and then used to identify three locations where a new crossover could be placed on the Piccadilly line, and the alternative service patterns that could be deployed using each, in case of a major unplanned disruption or planned weekend closure. A framework is proposed for the calculation of the benefits of a crossover in the case of disruptions and closures, which is applied to each proposed crossover. Railplan is used to model disruption scenarios for each crossover for different years, and with or without the addition of Crossrail 2; the results of the model are used to determine the degree to which each crossover could reduce the impact of a disruption on total passenger journey time. The benefits from each crossover are then aggregated based on the expected frequency of disruptions on each section of the line, and a preliminary business case is constructed. This business case analysis finds that there is no economic justification for the installation of any of the crossovers, when considering only the benefits in the event of major disruptions and closures. Additional uses for crossovers that could help justify the cost of a new crossover are also discussed.
The time-intensive modeling behind the track layout enhancement analysis motivates the exploration of other approaches to understanding the impacts of major disruptions. Alternative approaches to modeling transit networks in general, and major disruptions in particular, are explored. The current TfL practice in modeling disruptions is described, including metrics used to assess the service performance at any given time. The primary modeling-based approach, which uses the Lost Customer Hours metric and several models, including Railplan, provides a comprehensive description of the network performance, but is time- and computation-intensive and makes assumptions that weaken its effectiveness for understanding performance during major disruptions. These drawbacks are discussed, and the “imperfect knowledge” Railplan approach, which is under development by TfL to deal with some of these issues, is introduced. A case study comparing the two approaches for a closure on the Northern line is presented.

A simplified sketch-planning assignment model is developed for analyzing multiple scenarios quickly with a more realistic representation of passenger behavior during unplanned disruptions. Key features of the proposed model include:

- The model’s scope is restricted to the Underground, with the addition of walk links between some central London stations, and overflow links between some outlying stations and central London, to represent passenger mode shift
- All passengers are assumed to be aware of network crowding in normal conditions and of the disruption, but not of the resulting crowding
- The model does not seek to achieve equilibrium
- The model is quasi-dynamic, which means that supply and demand are considered static within the fifteen-minute timebands that comprise the study period (the three-hour AM peak period for the case study) except for passengers who are denied boarding.

To demonstrate the capabilities of this simplified model, it is applied to three partial line suspension scenarios on the Piccadilly line, and the results are compared with the results from both the base and imperfect knowledge Railplan models, both on an aggregate level and on a link level. The model operates faster than either of the Railplan models, and is easier and quicker to configure and use by TfL modelers, particularly once initial model overhead work (such as the development of rules for which overflow links to consider) is undertaken. The results of the simplified model are consistent with prior expectations on passenger rerouting during disruptions. For the scenarios considered - suspensions of the Piccadilly line in Central London - passengers divert to the District line in the west, and to the Victoria line in the north and central sections. Passengers also divert to other modes to travel to Central London, particularly if their branch of the Piccadilly line has reduced capacity; this is modeled through the inclusion of overflow links. The simplified model suggests that approximately 30-40% of passengers choose to change modes to get from outlying stations to central London during a major disruption.

The aggregate results for total passenger hours exceed those calculated by Railplan for each scenario by approximately 40%, for several reasons. First, unlike Railplan, the simplified model does not assume an equilibrium, and does not assume that passengers know about other passengers’ route choices. Railplan, by making these assumptions, assumes that passengers avoid
links that become heavily overcrowded during a disruption, which lowers estimated passenger impact. Second, the simplified model omits non-Underground modes, which artificially constrains passenger routing choices and increases link crowding, though the absence of some modes from the Tube map may discourage their use during a disruption. Third, all passengers are assumed to take the shortest path from origin to destination, which may be a valid assumption for commuters but still overloads certain links.

Compared to the imperfect knowledge approach, on the other hand, the simplified model provides a lower estimate of crowding on certain links. This is mostly caused by the imperfect knowledge approach’s assumption that approximately half of all passengers on the disrupted line are unaware of the disruption. Although passenger inattention, or poor announcement quality, mean that some passengers will always be unaware that a disruption has occurred, until confronted by it, the proliferation of smartphones and other real-time information media mean that this assumption is worth revisiting.

5.2 Limitations

The approaches developed here, both for track layout analysis and disruption modeling, have several limitations that must be recognized:

- **Consideration of only major disruptions and planned closures for track layout enhancement benefit.** The majority of crossover use is for one-off short-turns to get late trains back on schedule. Because of the large number of variables considered by controllers when deciding whether to make a train short-turn, and the large number of factors that influence the benefits of a short-turn on passengers, train operators, etc., the benefit of using crossovers for short-turns is not calculated here. This represents a strong underestimation of crossover benefit, particularly for those crossovers that are on outlying sections of the line and are thus most likely to be used for short turns.

- **Sensitivity of disruption impact to modeling parameters and assumptions.** Many assumptions are made when developing service patterns for partial line suspension scenarios. For example, a terminal with a single crossover is assumed to be able to accommodate only 6 tph during unplanned disruptions. This fails to consider the possibility that changes to operating procedures could significantly impact service frequencies, and thus the passenger experience, during a disruption. For instance, if provision of additional station staff could increase the capacity of a terminal from 6 to 8 tph, the capacity of the line would increase and disruption impact would decrease, impacting the benefit of any track layout modifications.

In addition, there are several parameters in the simplified assignment model that have the potential to significantly impact its results. These include weighting of overflow links, the choice of which overflow and auxiliary walk links to add to the model, and parameters related to passenger route choice in conditions of severely limited capacity. Passenger route choice under severe capacity constraints is difficult to predict because it is generally based on incomplete information, personal preference for crowding, wait time,
and other factors, and knowledge of other modes. These are difficult to quantify and may change in coming years as passenger information provision rapidly evolves. This model assumes a general link-based penalty on disrupted lines for route choice because passengers are rarely provided with sufficient detail regarding service state; however, this is only an approximation of the node-based phenomenon of denied boarding. AFC data or dedicated surveys could be used to devise a better estimate of these three aspects of passenger behavior during disruptions.

- **Model scope limitation.** The simplified assignment model is limited to the London Underground, and uses overflow and auxiliary walk links to approximate the effect of passenger mode shifts. Although these links are added to the model to capture the greatest volumes of displaced passengers - those heading to Piccadilly line stations in central London - the breadth of the London transit network means that some passengers will choose alternative modes or routes that are not well represented by these links. Other passengers might defer their trips entirely, which is also not captured in this model.

- **Quasi-dynamic model nature.** The assumption that passengers who are not denied boarding remain within their initial timeband for the entirety of their journey is an approximation to avoid the high computation time inherent in a more dynamic model. Though the fifteen-minute timeband duration selected here is a reasonable compromise, because it is short enough to capture peakiness effects yet long enough relative to most passengers’ journeys for the above approximation to be acceptable, the assumptions underlying the propagation of passengers across timebands are worthy of further study.

### 5.3 Recommendations

This thesis aims to improve the TfL’s understanding of network performance during unplanned major disruptions on the Underground through the creation of a simplified assignment model, and to quantify the degree to which the track layout affects performance. There are several opportunities for TfL to enhance its procedures and models to incorporate the lessons learned from this research:

- **Consider revisiting operational practices to allow higher reversing capacity at temporary terminals during unplanned disruptions.** One of the primary reasons for the magnitude of passenger impact of the disruptions analyzed in this thesis is the reduction in frequency that results from the use of a temporary terminal. The frequencies assumed here are based on LU rules of thumb, which are derived from operational practices. Although a sensitivity analysis was not conducted here, it seems likely that increasing the residual frequency operated during a suspension will, by increasing line capacity and thus decreasing denied boarding volumes and passenger spillover onto other lines, significantly decrease disruption impact. Revising operational procedures to allow higher frequencies – for example, by stationing additional staff members at key locations – may have strong value for money, if this can be implemented in a safe manner.
Adopt, and continue development of, the imperfect knowledge methodology for strategic disruption analysis, and the simplified model for sketch planning. As recognized within TfL, the base Railplan approach is not well-suited for modeling the impact of unplanned disruptions, since it is designed primarily as an equilibrium-based strategic model and represents passenger behavior accordingly. While the existing methodology for calculating LCH is critical for consistency and backward compatibility with previous years’ incident records, to accurately represent the impact of unplanned disruptions it is important to use a methodology designed with this application in mind. In addition, although computing resources continue to become less expensive with time, the modeling and computing time needed to use Railplan to run the several thousand scenarios that comprise the incident library remains problematic.

Thus, it is to TfL’s advantage to adopt a simpler model that runs faster, requires less modeling set up, and represents the impact of unplanned disruptions accurately. The simplified model presented here meets these criteria, and using it for the creation of the incident library or first-order alternatives analysis for capital projects would benefit TfL and free up staff time for more strategic planning. This strategic analysis of unplanned disruptions, in turn, should be performed using a tool that is designed specifically for that purpose, such as the imperfect knowledge Railplan model. Further developing this model, for example to generate aggregate system-wide weighted passenger-hours figures, should enable its full scale use for disruption modeling, more accurately representing disruption impact and saving computation time.

Incorporate Wifi-based route choice data into a comprehensive passenger demand model. Most ridership analyses on the Underground today rely on Oyster data, which records system entries and exits but do not provide direct evidence on passenger route choice. To maintain a ground-truth knowledge of route choice for models where this is important, some applications continue to use RODS data. The presence of two different sources of ridership data, neither of which comprehensively portrays passenger behavior on the network, causes confusion and reduces the accuracy of any model that depends on ridership data. Although the route choice algorithm used for Oyster-based modeling is calibrated to match results from RODS and other observed data, it is only the incorporation of Wifi data - which provides a measure of ground truth, as RODS does, while providing a large sample size - that will allow a comprehensive model of network loading to be developed. This model can also be validated against other data sources, such as railcar load-weigh data or CCTV-based passenger counting.

5.4 Future Research

This thesis introduces several topics that are worthy of more detailed examination than was possible here, and alludes to other questions that are beyond the scope of this research, but are still important to fully understand network performance during disruptions and the relationship
between track layout and network performance. Both areas studied here - track layout analysis and computationally efficient disruption modeling - are under-explored in the literature; additional study would help tailor a line’s track layout to its ridership patterns, incident characteristics, and controller preferences, and would also allow the consequences of unplanned disruptions to be predicted more quickly and accurately. Several attractive avenues for further research are:

- **Develop a methodology to quantify the benefits from other uses of crossovers and sidings, such as short-turns.** Short-turns are the most common reason for crossovers and sidings to be used, particularly on the outer ends of a line, because minor incidents that require short-turns occur much more frequently than the major disruptions that are modeled in this thesis. There are also other uses of crossovers - such as for engineering trains - that are more difficult to quantify, but are also important to the operation of a transit system. As such, any rigorous analysis of track layout must include the benefit - to passengers, train operators, line maintenance costs, or general line performance as appropriate - of these other uses. The frequency of these uses, particularly short turns, is easy to compute from NETMIS data, but converting this to a socioeconomic benefit is necessary.

- **Consider the nonlinear nature of delay perception in disruption modeling:** Passenger perception of delays is nonlinear with respect to delay time (Bates et al, 2001). In other words, a passenger perceives one delay of thirty minutes to be worse than six separate delays of five minutes each. This is not currently considered in TfL calculations of delay, such as LCH or EJT, or the models that underlie them, such as Railplan. Although extensive revealed preference research would be necessary to determine the exact relationship between delay duration and perceived disbenefit, implementing such nonlinearity in models would more accurately represent the impact of disruptions. Notably, a similar principle is already used in Railplan when considering crowding-based discomfort: the crowding penalty formula is quadratic, implying that passengers prefer to spend twice as long in a train that is half as crowded, to doing the opposite.

- **Explore passenger route choice under severe capacity constraints.** One of the primary reasons that the disruption impact for the Piccadilly line suspensions analyzed in Chapter 4 is so high is the frequency reduction required by the temporary termini. Under these conditions the capacity of the line is dramatically reduced, causing many passengers to be denied boarding, particularly on the outlying sections where there are few alternative routes to Central London onto which passengers can divert. The implications of this on passenger route choice, particularly in a non-equilibrium scenario with limited passenger knowledge of other passengers’ behavior, deserves careful consideration. Because passengers are not usually notified of by how much frequency is reduced, and do not know how many other passengers have chosen to change their route, it is difficult for them to anticipate the denied boarding time they will experience. The simplified model developed here approximates this as a link-based penalty for using the affected line, which is reasonable given that the limited information passengers typically receive (a “Severe Delays status”) is used for many types of incidents with many different types of impacts on passengers, but a closer
examination of the issue, perhaps with analysis of actual passenger behavior during disruptions, is warranted.

- **Explore mode shift during severe disruptions in more detail.** The simplified model makes several assumptions relating to mode shift. These are primarily affected by the choice of overflow and auxiliary walk links to specify, and the weights and travel times assigned to these links. As seen in Section 4.4, these links can significantly affect passenger route choice and aggregate model results; for example, the absence of bus links near Knightsbridge station causes some passengers to be diverted via the Circle and District lines, whereas in the real world these passengers would use buses or other surface transport. Even once rules for the inclusion of overflow links are established, it is important to quantify their travel time and weighting accurately to represent real world passenger behavior; otherwise, mode shift can be significantly over- or under-estimated, impacting model results. In addition, the lack of capacity constraints on these overflow links should be examined more carefully; though they are meant to represent a variety of rail, bus, and other modes that passengers may divert to, even the combination of all these modes cannot accommodate diverted flows from entire Underground lines without significant crowding. The influence of these factors on route choice and passenger experience can be measured through passenger surveys, revealed choice analysis of AFC data during disruptions, vehicle loadweigh data on non-Underground routes commonly used as alternates during disruptions, etc.
Bibliography


