Analysis of Demand Dynamics and Intermodal Connectivity in London Bicycle Sharing System

Nilufer Sari Aslam

MRES IN URBAN SUSTAINABILITY AND RESILIENCE 2014-2015

SUPERVISORS: PROF. TAO CHENG, DR. JAMES CHESHIRE

DEPARTMENT OF CIVIL, ENVIRONMENTAL & GEOMATIC ENGINEERING
UNIVERSITY COLLEGE LONDON

Abstract

Bicycle Sharing Systems (BSS) have gained increasing popularity in recent years due to their potential to increase cycle usage and improve the transit connectivity in cities. These systems offer a sustainable alternative to more traditional modes of urban transportation that have enormous benefits to health and environment. The most important factor for the success of a bicycle-sharing system is its ability to meet the fluctuating demand for bicycles and spaces across the network. There are facets of BSS resource demands that are linked to other modes of transportation due to the intermodal nature of bicycle share systems. These aspects have not been evaluated in much detail in the available literature. The aim of this study is to analyse the Demand Dynamics and Intermodal Connectivity in London Bicycle Sharing System (LBSS). This will provide key insights into the functioning of such a system and serves to provide policy makers with a wealth of data to explore many aspects of LBSS demand and usage patterns.

A three part analysis was carried out to inspect various aspects of LBSS demand. The first part focused on demand dynamics of individual docking stations as well as a formulation of demand and imbalance using travel flow data. The second part of the analysis looked at the source of demand imbalance due to intermodal transportation. London Tube/train flow data and LBSS data was used to highlight the relationship between the two networks. The last part of the analysis was to observe the on-ground realities of the systems and compare it against the information deduced from the data analysis.

Results highlighted the source of demand asymmetry in LBSS due to its role as an intermodal transportation alternative. It also demonstrated the self-balancing characteristics of large numbers of docking stations by means of flow data analysis. The work carried out in this study lays the foundation for future efforts to understand and forecast demand in BSS.

Keywords: Bicycle Sharing System (BSS), Demand imbalance, Demand asymmetry, Intermodal transportation and last mile commuting

Acknowledgements

I would like to express my deepest appreciation to all those who provided me the opportunity to complete this project. A special gratitude to my supervisors, Prof Tao Cheng and Dr James Cheshire, whose advice and encouragement, helped me to complete this project in short time period. Furthermore I would also like to take this opportunity to thank Mr Oliver O'Brien for sharing bicycle data with me and my friends, who have supported me during my project. Special thanks goes to my family, without their encouragement and support, I would have never been able to complete my work with the same determination and commitment.

TABLE OF CONTENTS

List of Figures

List of Tables

1. Introduction

Recent years have seen growing interest in bicycle-sharing systems (BSS) as an alternative to more traditional modes of urban transportation. These systems offer a sustainable option that has enormous benefits such as flexibility, physical activity, and support for intermodal transportation connectivity (Shaheen *et al*., 2010). Although BSS have been gaining steady popularity around the world, there are challenges in operating such a system. Unlike traditional modes of transportation, the operators of BSS have little or no control over the main resources of the system i.e. bicycle. The bicycles can be taken from one location and dropped at any of the other location in the network. This flexibility often results in a situation where some docking stations end up without any bicycles and some docking stations without any spaces. This problem is referred to as flow asymmetry and is a result of asymmetric demand. Traditionally the option available to the operators of the system is to reallocate bicycles in the network to improve availability of resources in the system.

Billions of rows of continuous and non-invasive data with spatial and temporal dimensions are now available in the public domain (Lathia et al. 2012; Blythe & Bryan 2007; Beecham & Wood 2013; Kusakabe et al. 2010; Blackwell & Sen 2012). In London alone, 30 million journeys are completed on Transport for London’s (TfL’s) network every single day (TfL 2013). The vast majority of these journeys are by bus and London Tube but a growing number of travellers make use of London’s bicycle sharing system (O’Brien *et al*. 2014). All these journeys leave behind a digital footprint, in the form of an electronic record. The rapid pace of technological advances and the availability of huge amounts of data from transport networks provides us with an opportunity to analyse problems such as demand asymmetry. Techniques such as data mining and machine learning can be applied to these problems to come up with innovative solutions.

In this study available data from origin to destination journeys from a number of transportation sources has been analysed in order to understand the problem of BSS demand asymmetry. Aspects of demand asymmetry in the context of BSS as an intermodal transportation alternative for last mile commuting have been explored in detail.

1.1 Aim and Objectives

London Bicycle share system (LBSS) provides a very flexible option for last mile transportation, by allowing users to take the bicycle from one station and drop it at another docking station. The flexibility of the system comes at the cost of the imbalance in the system at various times of the day. The aim of this study was to get a better understanding of the problem of demand imbalance by looking at it from the following aspects:

- Demand dynamics using the origin - destination journeys data
- Source of demand imbalance due to intermodal transportation
- Physical observations of the system

The objective of these analyses was to not only help provide understanding of the working of the LBSS but also offer insights that may be useful in the improvement of the quality of service.

1.1.1 Problem definition

London bicycle sharing system like other such systems suffers from the problem of demand imbalance. It means that at various times of the day some docking stations have no bicycles and some docking stations have no places to dock the bicycles. A continuous effort is required by the network operators to rebalance the network throughout the day and at night.

1.1.2 Importance of Study

Government policies in the recent years have placed significant importance on the environment, health, air quality and carbon emission (Case 2009). All these considerations are directly linked to the choices of last mile transportation.

The importance of this study is that it does not only look at the problem of demand imbalances at the level of the station but also takes into consideration the whole subject of last mile commuting. The significance of intermodal public transportation and its role towards the asymmetry of demand in BSS. This study can facilitate the adaptation of bicycle hire as a realistic and reliable mode of public transport. It can be useful to the planners of BSS in providing new insights into the working of the system. It can also provide new viewpoints to aid the BSS decision support system such as:

- Rebalancing activities
- New locations for docking stations
- Improvements in reliability of the system

- Improvements in quality of service

1.2 Background

In transportation, the last mile is referred to as the difficulty in getting people from transit stations to their workplace in the inner city (Shaheen & Cohen 2007). It is very specific to weekday office workers who travel from the suburbs and outer boundaries of the city into the city centres, but it can equally be relevant to people visiting cities for leisure, tourism and entertainment. The last mile commuting is relevant even for a city like London with its compact centre and extensive transportation infrastructure. Historically the options available to cover the last mile of commuting were buses and walking. The advent of Bicycle sharing systems has now offered a new low-cost and healthy public transport alternative to travel from the transit hub stations to the final destinations (DeMaio 2009).

1.2.1 Bicycle Sharing Systems (BSS)

A bicycle sharing system is an innovative scheme in which bicycles are made available for shared use to individuals for short journeys. Most BSS allow users to take bicycles from any location and return to any of the available locations scattered around the inner city. Over the last ten to fifteen years, bicycle-sharing systems have turned from being novelties in urban transportation into real options for public mobility in cities. These schemes are now an integral part of the urban transportation system for even large and complex cities like Paris and London (Midgley 2011).

Shaheen *et al* (2010) mentioned that bicycle sharing has evolved considerably from the time it first started in Amsterdam in 1965 (Table 1). The first generation of those schemes were called the "White bicycle Plan" that provided local population with free bicycling facility. The second generation of bike sharing program was built around 1996 using "coin-deposit locks" (Shaheen *et al.,* 2010). The popularity of the BSS was low until the advent of the third generation schemes when advanced technologies were first introduced and integrated into the system for flexible drop-off, pick-up and information lookup through mobile devices. The real growth in BSS coincided with the start of the century when the number of schemes ballooned from around 5 in 2001 to around 600 cities by 2013 with 0.6 million bicycles (Institute for Transportation & Development Policy (ITDP) 2013). There is an overlap between third and fourth generation schemes, with the latter harnessing the powers of real-time data for redistribution and service improvements (see Table 1).

Table 1: Bicycle sharing growth has undergone four evolutions (table constructed using information from Shaheen *et al*., 2010).

Bike sharing Generations	**Name**	**Component**	**Characteristics**
First generation	White bike (or free bike) systems	Bicycles	Distinct bicycles (usually by colour) Bicycles located arbitrarily Unlocked bicycles No charge for use.
Second generation	Coin-deposit systems	Bicycles Docking stations.	Distinct bicycles (colour or special design) Bicycles located at specific docking stations Bicycles with locks
Third generation	IT-based systems	Bicycles Docking stations Kiosks or user interface	Bicycles are distinct (colour, special design, or advertisements). Bicycles are located at specific docking stations. Bicycles have locks. Smart technology is used for bicycle check-in and checkout (Mobile phones, mag-stripe cards, or smartcards). Theft deterrents Programs are paid for as a membership service.
Fourth generation	demand-responsive, multimodal systems	Bicycles Docking stations Kiosks–user interface Bicycle distribution system	Distinct bicycles. May include electric bicycles. Specific docking stations that are more efficient (mobile and solar powered) Improved locking mechanism to deter bicycle theft Touch screen kiosks–user interface Bicycle redistribution system Linked to public transit smartcard

1.2.2 London bicycle sharing system (LBSS)

London bicycle sharing system (LBSS) was launched by Transport for London on the 30th of July 2010. According to Goodman and Cheshire, 2014, the scheme is available 24 hours a day, throughout the year. The bicycles can be taken from any station and returned to any other docking station. Initially, the users were required to register with an annual membership charge but those restrictions were later relaxed and casual use of the bicycles were encouraged. To hire a bicycle, users can either register online for an access key or else pay at docking stations by using a UK or international credit/debit card ('casual use', available since 3rd December 2010). Users initially pay for access to the bicycles, after which trips of 30 min or less are free but longer trips incur additional usage charges at a progressively increasing rate.

The scheme was launched in 2010 with 5000 bicycles located across 315 docking stations, across 45 km^2 of central London. This original zone included the entertainment centre of the West End, the business district of the City of London, and the leisure areas of Hyde Park and Regent' park. It also included some more affluent residential areas to the west and some more residential areas to the east. The scheme was later extended to the east of the city to cover a larger area of 65 km^2 (8th March 2012). The expanded scheme encompassed the business district of Canary Wharf in Docklands, characterized by city commuters working in financial services (Goodman & Cheshire 2014).

According to TfL, there are currently over 10,000 bicycles operational across 730 stations as part of London public bike sharing system (TfL 2015a) and the number of the docking stations are continually on the rise.

2. Literature Review

2.1 Bicycle Sharing System – Demand Imbalance

The flexibility of BSS where a bicycle can be taken from any station and dropped at any other station often results in the asymmetry of demand. This problem is well documented and researchers have looked into it from various angles.

According to Corcoran *et al*. (2010), there are two types of data analyses that have been used in bicycle sharing research:

- Docking station data analysis is based on the data captured at the docking stations. This includes total capacity of bicycles, available bicycles and available spaces.
- Journeys data analysis is based upon the origin to destination journeys data. Using this data, inflow and outflow of bicycles at any given docking station at any time interval can be calculated.

2.1.1 Docking station data analysis

According to Kaltenbrunner *et al.* (2010), the data collected at the docking stations at fixed intervals provide insights into spatial-temporal fluctuations in demand across a system. It also highlights the load at the station at any point in time. The study looked at the system's usage patterns based on the data collected over a period of time and developed a simple model to predict future trends. The results demonstrated the variances between weekday and weekend usage, and top usages at different times of the day, which depended on the docking station's closeness to the locations such as shopping centres, universities and workplaces. Although the analysis done based on docking station data can provide a good indication of demand fluctuations, it does not provide any information with regards to the movement trends (Kaltenbrunner *et al.*, 2010).

Lathia *et al*. (2012) published results from the London bicycle sharing system's docking station data. The paper studied the impact of changing the user-access policy in the London Barclays Cycle Hire scheme from registered only users to also include any casual users. The change in policy enabled anyone with a debit or credit card to have access to the service. Ease of access resulted in more people using the system and also greater weekend usage. (Lathia *et al*., 2012) Keeping the resources (bicycles and station capacity) constant and increasing just the user base by means of policy change, resulted in greater demand and greater turnover.

Fricker & Gast (2012) used the term ‘problematic station’ to define the demand imbalance of the station. In their study a stocastic model was created to gauge the impact of random user demand with the finite capacity of the docking station. The study sets out to quantify the impact of station capacity and ideal size of bicycle fleet in order to reduce the number of problematic stations. The results from the study suggest that the number of problematic stations decreases as the capacity of the docking station increases. (Fricker & Gast 2012). It also highlights the trade-off between cost consideration and the demand imbalance of the docking stations.

O’Brien *et al*., (2014) presented the first global view of BSS by analysing the data from 38 systems around the world. It compared the usage patterns and spatial- temporal distribution of bicycles within various systems across the globe. To achieve this, an extensive database consisting of geographical location and bicycle occupancy of each docking station within a particular system was created to chart the load of docking stations in the chosen systems. Many of the BSS exhibited similarity of usage peaks due to the morning and evening commuter rush hours. Based on the usage patterns, the authors were able to classify users into office commuters, leisure and tourist users. It was also observed that BSS in Asian cities generally have larger docking stations that facilitate commuter flows and temporal asymmetry (O’Brien et al. 2014). Larger docking station would provide greater buffer against demand asymetery but there are cost benefit implications of creating ever larger docking stations. This outcome is consistent with the data analysis from the LBSS study.

2.1.2 Journeys data Analysis

Journey data of origin to destination have been made available by some BSSs more recently. The availability of journeys data is still limited as compared to docking station capacity data (O'Brien *et al.*, 2014). The advantage of these types of data is that they capture the underlying mobility behaviour using origin to destination trips. It provides an opportunity to analyse travel demand by studying travel patterns and help optimize transportation efficiency. Flow based analysis in general is complex and more difficult to visualize than station-based analysis. Lines or curves are usually used to visualize flows as an indicator of usage on a given route between two docking stations. However, when flow data become bigger, lines representing flows will overlap and the underneath pattern will be more and more difficult to observe (Zhu & Guo, 2014). Different methodologies can be used in order to visualize such patterns more easily including aggregating of flows and use of different colours.

According to Corcoran *et al*., (2014), except for bicycle sharing data visualization in cities, there have not been many papers analysing using massive flow datasets with the aim to understanding the overall patterns of bike sharing behaviours (Corcoran *et al.*, 2014). This can be attributed to the limited availability of flow datasets and to a certain extent problems in data manipulation of big datasets that are necessary to enable analysis.

Zhu & Guo (2014) proposed a flow clustering method to extract clusters of identical flows and revealed aggregated flow patterns of trips in Shenzhen, China (Zhu & Guo, 2014). It extends the traditional hierarchical clustering method to aggregate and map large numbers of trips. It considers both origins and destinations in determining the similarity of two flows, which ensures that a flow cluster represents flows from similar origins to similar destinations and thus minimizes information loss during aggregation. Using techniques such as the spatial indexing, the new method is scalable to big flow datasets.

The types of demand for a BBS are demand for bicycle and demand for docking spaces. Both these types of demands may not be equal and often result in asymmetry. The operators of BSS are required to redistribute vehicles between stations to correct this asymmetry. Since future demand is not known exactly and is highly variable, the challenges faced by operators are amplified. Vogel *et al*. (2011) has also analysed extensive operational data from bicycle sharing systems in order to derive bicycle activity patterns. Activity patterns reveal imbalances in the distribution of bicycles and lead to a better understanding of the system structure. The study suggests that a structured data mining process can support planning and operating decisions for the design and management of bicycle sharing systems.

Borgnat *et al*. (2010) also used data mining to analyse the dynamics of bicycle movements in Lyon's BSS. Temporal patterns in the system-wide bicycle usage were examined. Weekdays showed usage peaks in the morning, at noon and late afternoon, whereas usage is concentrated in the afternoon on weekends. Furthermore, spatial patterns are examined by clustering bicycle flows between stations. Spatial and temporal dependencies exist between stations of clusters interchanging many bicycles (Borgnat *et al*., 2011). It suggests suitable planning of the location as the key to addressing the demand imbalance problem.

Lin and Yang (2011) presented an operation research based mathematical decision model to determine an adequate size and location of bicycle stations. Also the network structure of bicycle paths between the stations and travel paths for users between their origin and destination were determined. Simulated bicycle demand data are used for testing the model. Recent work focuses on building decision models without incorporating real world BSS behaviour.

2.2 Last mile and mixed mode transportation

A shared-vehicle system can be interpreted as an individual mode (for short trips) or as a vital segment of an intermodal route (for longer trips), (Nair *et al*., 2013). This relationship weighs heavily on the demand consideration of a BSS. Due to this reason the location of the BSS docking stations and its proximity to the transit hub is of paramount importance.

The decision with regards to the capacity and location of a docking station with respect to the train station has been considered by a number of authors. Shu *et al*. (2010) proposed a stochastic network flow model to support these decisions. In that model, the proposed design of a BSS in Singapore is based on demand forecasts of the Singapore MRT (Mass Rapid Transit). Lin and Yang (2011) have also considered a similar study, but structured it as a deterministic mathematical model. The model is created with the bicycle path network and mode sharing with other means of public transportation (Lin & Yang Ta-Hui 2011). The study suggests that an ideal design of BSS requires an integrated view of the available modes of the different transports.

In other studies, Chen *et al*. (2012) carried out a large-scale survey to evaluate the determinants of bicycle demand at several metro stations in Nanjing, China. Two transfer choice models were estimated to identify and quantify the determinants for bicycle transfer demand: one focused on current metro-walk trips, and the other on current metro-bus trips. The explanatory determinants were discussed, relative weights were computed and multiple linear regression analyses were applied to quantify the relationship based on the surveyed data. The results revealed that more than half of the users of metro preferred using shared bicycles if the walking distance is more than 15 minutes.

It has been argued that the presence of a shared-vehicle system increases the transit accessibility of the stations. It serves as a vital ‘‘last-mile’’ connection, the lack of which dissuades potential riders. They are strongly aligned with integrated transit systems explored in the past that also aim to increase the catchment area of transit (Davis, 2008; Shaheen *et al*., 2005).

Holleczek (2013) has applied a new data mining technique to explore the spatial and temporal variations in transportation usage pattern and transportation options in Singapore. Data from cellular phone networks are integrated with the data from with public transport systems to analyse and visualize the urban mobility patterns (Holleczek et al. 2013). It categorized the transportation links as weak or underserved. The study demonstrates that the mode sharing in public transport increases throughout the day from 38% in the morning to 44% around mid-day and 52% in the evening. The study has demonstrated that multiple sources of data can be integrated to examine the mobility pattern and can also help address the problem of “last mile” of commuting.

2.3 Fleet management and redistribution in fourth generation BSS

The operators of the BSS are required to carry out on going redistribution activities in order to correct the spatial-temporal imbalance of the system. Before any redistribution activity is carried out, it is of paramount importance to understand the usage patterns; this can best be achieved by looking at the historical information from the docking stations. Nair *et al*., (2013) have applied an approach based on probabilistic demand characterization at each station based on historical data to understand the demand imbalance.

The understanding of BSS usage pattern and demand dynamics lead to better solutions to address the problem of asymmetry of the demands experienced at various times of the day at various docking stations. The majority of the work done in this direction is essentially to redistribute the resource i.e. bicycles, in such a way as to achieve better service levels, at the same time keeping the operational cost to a minimum. Traditionally, the solutions included redistributions using trucks but more recently, newer ideas have been gaining momentum. Crowd sourcing and dynamic pricing leverage the availability of real time information, social media and mobile technologies to provide new solutions to problems such as demand imbalance.

2.3.1 Relocation using Trucks

The problem of asymmetric flow of vehicles has been studied for vehicle sharing systems that could involve cars or bicycles (Kek *et al*., 2009; Cepolina E. M. and Farina A. 2012; Clemente *et al*., 2013). To generate redistribution plans to meet a target reliability level, others have applied mixed-integer programming (Kek *et al*., 2009), and multi-stage stochastic programming with recourse (Nair & Miller-Hooks, 2011).

2.3.2 Relocation in fourth generation BSS

Dynamic pricing and crowd-sourced solutions have been hailed as the answer in fourth generation BSS to the problem of demand imbalance. The following studies have investigated the crowd-based BSS balancing solutions (GoDCgo 2011; Pfrommer et al., 2014; Chemla et al., 2013; Fricker & Gast 2012).

V´elib' (2014) discussed how Paris provides an example of incentive based reallocation where users were offered extra minutes each time they returned a bicycle to an elevated station. Pfrommer *et al*. (2014) also introduced a dynamic pricing mechanism using model-based predictive control principles (Pfrommer *et al*., 2014). Singla *et al*. (2015) have presented a possible solution for demand imbalance using a crowd sourcing mechanism, which employs optimal pricing policies using the approach of regret minimization in online learning (Singla & Krause 2013). Crowd-sourced incentive based solutions can use the insights gained from the journeys data analysis to provide a better spatial-temporal view of demand imbalance.

3. Methodology

3.1 Introduction

This chapter provide the details of the analysis and a description of methodology used in the report. The study applies spatial and temporal data mining techniques on the available sources of travel data and visualizes the demand imbalance in the form of maps. This chapter is divided into three sections to cover the three type of analysis and experiments that have been carried out.

The first part of the analysis focuses on the demand considerations of LBSS. The demand in a BSS can be classified as the demand for bicycles or demand for spaces. Both of these demands have spatial and temporal dimensions. This chapter provides the formulations of demand measures constructed to describe the demand dynamics of individual docking stations using origin to destinations journey data.

The second part focuses on the source of demand imbalance due to intermodal transportation. The demand peaks in LBSS are driven by demand peaks in Oyster network and the flow of commuters between the two networks. This section of the report attempts to uncover the dependency of the usage between Oyster and LBSS networks and the demand volatility of LBSS due to this relationship.

The last part provides the details of the physical observations that were carried out for a number of important docking stations in LBSS. The on-ground observations will provide a link between information extracted from the data and on-ground reality of the system.

3.2 Data Description

The data considered for this study are obtained from multiple sources. This section of the report provides a brief description of the type and sources of data that have been used for the analysis. It also provides the detail about the data processing steps such as integration and aggregation of data in order to use them as an input into the analysis.

3.2.1 Oyster Data

Oyster is a smart card which is used to hold the credit and travel pass for the journeys that are carried out on the TfL network. With the help of Oyster cards, TfL is able to keep a record of individual journeys that are carried out using the card. The data for these journeys, excluding the personal details of the card users, were available for the analysis in this report. Available data were from 01 Jan 2012 to 31 Dec 2013 and did not include the journeys that are carried out on the buses.

Considering millions of journeys are carried out on the TfL network, the amount of data that are generated and stored is truly staggering. As a first step, the data were aggregated at 15 minutes and 1-hour intervals. Both the datasets were used for the further analysis. Table 2 provides a structure of the aggregated data.

Table 2: A structure of Oyster aggregated data

<table>
<tr><th>Variable</th><th colspan="4">Sample Values</th></tr>
<tr><td>Nls / Ofif</td><td colspan="4">National Location Code</td></tr>
<tr><td>Entry</td><td colspan="4">Entry data from Location Code</td></tr>
<tr><td>Exit</td><td colspan="4">Exit data from Location Code</td></tr>
<tr><td rowspan="2">Flag</td><td colspan="4">Flag donates data that is out of line with general norm</td></tr>
<tr><td>1= Entry Only</td><td>2= Exit Only</td><td>3= Entry and Exit</td><td>None</td></tr>
<tr><td>Day</td><td>1= Monday to Friday</td><td>2= Saturday</td><td>3=Sunday</td><td>4= Bank Holiday</td></tr>
<tr><td>Time</td><td colspan="4">Two datasets aggregated at 15 minutes and 1 hour time intervals</td></tr>
</table>

3.2.2 Bicycle Data

The access to the bicycle network is not recorded through Oyster card but through LBSS access key or credit card. Again the information stored for individual journeys is available without the specific details about the users. Electronic records of 17 million bicycle journeys carried out between 2012 and 2013 were accessible for the analysis.

Table 3: A structure of London Bicycle journeys data

Variable	Sample Values
Bicycle ID	1656
User ID	NULL
Station ID (Origin)	124
Station ID (Destination)	172
Date/Time (Origin)	Timestamp at origin
Date/Time (Destination)	Timestamp at destination

3.2.3 Bicycle Capacity Data

The third source of data for the project was the capacity updates available for the same period between 2012 and 2013. Capacity updates are sent out at an interval of 2 minutes from the bicycle docking stations. The updates include information such as available spaces and available bicycles and the timestamp. Table 4 provides the structure of bicycle updates dataset.

Table 4: A structure of London Bicycle capacity data

Variable	**Exemplified Values**
TfL ID	1
Available Bikes	3
Available Spaces	12
Total docks	15
Time Stamp	2012-01-02, 18:10

3.3 Data Manipulation

Along with the three data sources described in tables 2, 3 and 4, other sources included reference information such as station name and address, geographical location (latitude and longitude). All source, reference data and calculations were stored in a PostgreSQL database. The processing required for data manipulation such as aggregation was carried out within PostgreSQL and Valentino Studio. The SQL scripts are attached in Appendix 2.

Like any other data analysis study a certain degree of data processing was required in order to bring various datasets into the correct format.

- All the journeys terminating at a given station were aggregated under a single time period of 15 minutes and 1 hour.
- Journeys initiating from a given station were also aggregated under common time intervals. This provided two records per time period per station.
- Capacity updates were at 2 minute time intervals. It was linked with the rest of the aggregated data at the common interval points of 15/60 minutes.

3.4 Demand dynamics of LBSS

3.4.1 Introduction

In the planning of any type of transportation facility, it is necessary to have an estimate of how much it will be used. The knowledge about how the users will respond to price and service classification is critical to establish the best operating procedures. All these considerations come under the umbrella of travel demand. Demand for travel is different from the demand of other services, because it is a derived demand. People only undertake travel to facilitate other complex and varied sets of activities such as work, pleasure, shopping and home life. This observation links the study of travel demand with study of human behaviour and choices of technologies and urban development. It also brings into consideration the aspect of travel referred to as intermodal travel. Travel demand is a very complex subject and researchers tend to start with very simplified models to make any progress (Small K.A., 2007).

Facets of geo spatial demand on various BSS around the world have been analysed and visualized in a number of earlier studies (O'Brien *et al*., 2014). Large numbers of these studies have focused on the use of capacity update data that are available from the docking stations at frequent intervals.

- Capacity updates are useful to describe the load of the docking station at any given point in time (O'Brien *et al*., 2014), but it is not adequate to accurately define the demand.
- The capacity updates include changes that are due to bicycle balancing operations that hide the on-ground reality of bicycle usage, especially during the peak hours.
- Using the origin destination journeys data, it is possible to analyse bicycle usage more accurately across individual stations.
    - Origin-destination journeys and capacity data can be analysed to gauge the effectiveness of bicycle balancing operations.
    - It is possible to highlight how the BSS operators are trying to meet the service standards by keeping the system in a balanced state i.e. times of the operation activity for each station
    - Station service details such number of bikes added or removed during rebalancing operations

The analysis starts with the formulisation of the demand of bicycles and spaces in LBSS. The following definitions (Table 5) have been used in order to define demand in terms of journeys:

Table 5: Definition of the indicators

Symbol	Definition
C	Capacity of the docking station, the maximum number of bicycles that can be docked in the docking station. It is assumed that it will exclude any docking points that are not functional.
OUTF	Out Flow, the number of journeys starting from the docking station in a given time interval
INF	In Flow, the number of journeys ending at the docking station in a given time interval
Δt	Time interval used for the flow calculation. The value used for this analysis is 15 minutes and 60 minutes.
DF_B	Demand Factor of Bicycle
DF_S	Demand Factor of Space
NF	Net Flow is defined as the difference of incoming and outgoing bicycles. Its value is time interval dependent.
IMNF	Imbalance Net Flow is calculated using the maximum capacity of the docking station.

3.4.2 Demand Factor bicycle (DF_B)

Demand factor for bicycle is the proportion of outflow compared to the capacity of the docking station. The equation (equation 1) can be extended to give aggregate demand over longer periods such as days or weeks. This formulation of demand is specific for a given time interval.

$$DF_B = OUTF/C \qquad \text{(Equation 1)}$$

3.4.3 Demand Factor Space (DF_S)

As the name suggests, this defines the demand for bicycle docking spaces. In the presence of manual operations the demand for bicycles and spaces cannot merely be calculated as the opposite of demand for bicycles. It is defined as a ratio between the number of journeys ending at stations during a given time period and the capacity of the station (equation 2). The equation can be extended to give aggregate demand over longer periods such as days or weeks. This formulation of demand is specific for a given time interval.

$$\mathrm{DF_S} = \mathrm{INF/C} \qquad \text{(Equation 2)}$$

3.4.4 Imbalance Net Flow (IMNF)

Over a period of time, IMNF provides an indication of the disparity of the inflows and outflows in proportion to the size of the docking station (equation 3). It is simply defined as a ratio between net flow and capacity of the station during a given time period (equation 4).

$$\mathrm{NF} = \mathrm{INF} - \mathrm{OUTF} \qquad \text{(Equation 3)}$$

$$\mathrm{IMNF} = \mathrm{NF/C} \qquad \text{(Equation 4)}$$

3.4.5 Variations in Demand of Bicycles

The frequent capacity updates from the docking stations are impacted by the BSS rebalancing operations due to manual changes in the available bicycles and spaces. These changes are not representative of the actual user demand. Demand Factor bicycle (DF_B) is calculated and visualized as a measure of demand based on actual journeys that have taken place on the network. Demand factor is defined as demand between certain time intervals. Accurate calculation of demand over a time period is only possible by taking into account the actual flows generated by origin to destination journeys.

3.4.6 Variations in Demand of Spaces

Looking at capacity data, it appears that demand of bicycles and demand for spaces are reciprocal. When the number of bicycles at a station increases, the number of spaces decreases. In reality the demand of bicycle and spaces are completely independent of each other. Demand Factor Space (DF_S) is calculated based on journeys that are terminating at given docking stations during a defined time interval.

Maps for $\mathrm{DF_B}$and $\mathrm{DF_S}$ were created for morning and evening peak time intervals in order to demonstrate the variations in the demand of bicycle and spaces.

3.4.7 Demand and Self-Balancing Index

Demand discrepancy is a reflection of imbalance of the docking station. The objective of this analysis was to provide a network wide view of the demand discrepancies of bicycles and spaces at various times of the day.

An index of demand discrepancy called Self-Balancing Index was created using demand factor for bicycle and spaces. It provided a range of values and the docking stations were labelled accordingly. For the purpose of comparison it is necessary to scale the index value based on the duration of time interval considered.

Two maps were created at midday and end of the day to compare the demand discrepancy of bicycle and space.

3.5 Analysis of Source of Demand due to intermodal transportation

3.5.1 Introduction

The majority of journeys in any city are carried out over multiple modes of transportation. This is also true for London, where commuters have many options including shared bicycles to complete the last mile of their journeys. The vast majority of the commuter flow to LBSS originates from the transit hub stations such as Waterloo, King's Cross and Victoria (TfL 2015b).

Integrating and analysing the data from train and bicycle hire networks provides an opportunity to understand the source of this demand. The purpose of the analysis in this section was to quantify the strength of the relationship for commuter flow between TfL Oyster network and LBSS.

This part of the study started with the trend analysis of the time series of the two networks at different time intervals (daily, weekly and monthly), to understand obvious patterns in the data. To look further into this relationship, the Pearson correlation coefficient was calculated for time series data to quantify the strength of the relationship between the two networks. It was followed by linear regression to show the trend lines using docking station data as the dependent variable. The direction of the flow and the impact on the demand of bicycle and spaces can be defined as follows:

Oyster to LBSS flow: Because of the relative size of the networks, a large influx of Oyster users can impact significantly on the capacity utilization of a bicycle hire network. To show the strength of this relationship, the analysis considered exit (Tube/train stations) to exit (docking stations) data. This was considered as outflow from LBSS perspective

LBSS to Oyster flow: To gauge the strength of the relationship in the reverse direction, i.e. from bicycle on to train to establish if users coming into the station on bicycles are continuing with their commute *via* trains, the analysis considered entry (docking stations) to entry (train stations) data. This data flow was classified as inflow from the perspective of LBSS

3.5.2 Data Processing

The cycle hire data are for the individual journeys from one docking station (origin) to another (destination). In order to focus on the specific time windows it was decided that journeys should be aggregated into 15-minute time intervals. It resulted in two records per docking station per 15-minute period: one record for aggregate 'entry or inflow' terminating at the station and the second for aggregate 'exit or outflow' from a docking station.

Oyster data available for this analysis were aggregated at 15-minute intervals and were provided by TfL. In order to match the bicycle data, all the journeys terminating at a given station within a period were aggregated into one 'entry or inflow' record. All the journeys starting from a station within a period were aggregated in one 'exit or outflow' record. Table 6 provide a summary of the processed data:

Table 6: Processed data classification

Data	Direction
Aggregate Oyster	Exit or outflow
Aggregate Oyster	Entry or inflow
Aggregate bicycle docking station	Exit or outflow
Aggregate bicycle docking station	Entry or inflow

3.5.3 Study Area

Central London was the area studied for this analysis. It included all Oyster network stations and the docking stations on the LBSS.

3.5.4 Network Analysis

The commuter flow between the Oyster network and LBSS depended upon the proximity of the bicycle docking station to the Tube/train station. It was important to understand at what point the distance between the Oyster network and LBSS started to become relevant for intermodal transportation. According to Tran *et al*. (2015), 300m is considered a distance that people are willing to walk between two networks for intermodal transportation. BSS in Paris also used this guideline (Tran *et al*. 2015) to build a docking station every 300 meters for the first phase of its bicycle sharing system, as did London and New York. (Institute for Transportation & Development Policy [ITDP] 2013). Based on that information, this study also used 300m distances to identify the relevant docking stations near the train stations for intermodal commuter flow.

Network analysis closest facility (in ArcGIS) was applied to find out the closest/best route between two points i.e. docking station and the Tube/train station. The available data included 747 Docking stations and 163 stations and the objective was to identify the shortest walking distance between them.

3.5.5 Data Analysis

Data analysis included the temporal profile for the daily, weekly and monthly data for the selected stations and docking stations. The objective was to highlight the relationship between the usage patterns between the two networks.

<u>Daily Data (7 June, 2012)</u>

To understand the daily pattern of people flow between the two networks, data points were plotted for the three selected stations in Central London.

<u>Weekly Data (April, 2012)</u>

Results were plotted for the same three stations from the 23 April to the 29 April to show the separation between the weekend and weekday trends at hourly intervals.

<u>Monthly Data (April, 2012)</u>

The analysis for the month was carried out to highlight the relationship between docking stations and train stations over a longer duration.

3.5.6 Pearson Correlation

After investigating the daily, weekly and monthly trends, further insight was gained through calculating the Pearson correlation coefficient (defined in Equation 5). Correlation is a method for exploring the association between two continuous variables. Pearson's correlation coefficient (r) is a measure to determine the strength of this association.

- As a first step, a scatter graph was plotted to check the linearity of the relationship between two continuous variables. This method is only used when the relationship between the variables is linear.

- The closer the points are to the straight line, the higher the strength of relationship between the variables.

$$\text{Correlation}(r) = \frac{Cov(x,y)}{\text{std.dev}\,(\text{x}) \times \text{std.dev}\,(\text{y})} \qquad \text{(Equation 5)}$$

Where x is the number of the people using train stations at hour interval (train usage), and y is the number of the people using bicycle docking stations at an hour interval (bicycle usage). Covariance of x and x defined as Cov(x, y) is a measure of how much x and y change together. Std.dev(x) and Std.dev(y) define the dispersion of random variable x and random variable y.

3.5.7 Linear Regression

Regression is a statistical technique to establish the linear relationship between two or more variables. In its simplest form, regression shows the relationship between one independent variable (X) and one dependent variable (Y), as shown in equation 6. The direction and extent of that relationship is given by the slope parameter (β1), and in the absence of independent variables the status is given by the intercept (β0). An error term (u) captures the amount of deviation not predicted by the slope and intercept terms. The regression coefficient (R^2) determines how well the values fit the data (Campbell & Campbell 2008).

$$Y = \beta_0 + \beta_1 x + u \qquad \text{Equation 6}$$

In this study, linear regression was conducted by assuming docking stations time series data as the dependent variable and train station data as the independent variable.

3.6 Physical Observations

As part of the study, on-ground observations were carried out at a number of transit hub stations in order to better understand the dynamics of bicycle share usage. The flow of commuters from the Tube/train stations to the BSS was dependent upon the distance and location of the bicycle docking station. Three of the busiest commuter hubs in London i.e. Waterloo, King Cross and Victoria were selected for the physical observations of the system.

3.6.1 Average Timing/per rush hour

Efficiency of intermodal transportation depends to a large extent on the ease of transfer from one mode on to the other. Average time of 5 people to walk from the station exit to the docking station during rush hour was calculated as part of the analysis. This was based on the assumption that people choose the station exit nearest to the docking station if they intend to use the docking station.

3.6.2 Plan of observations

The observations were planned over three weekdays for each of the three areas during the commuter rush hours. The focus of the study were:

- To understand the layout of the train/Tube station and the location of the docking stations with respect to the station exits
- Distance and walking time from the Tube and train station exits to docking station.
- Frequency of usage and the direction of usage i.e. bicycle inflow and outflow.
- Capacity max-out point. It is defined as the time when the docking station runs out of bicycles or spaces.
- Observation of the bicycle rebalancing operation and its effectiveness. The bicycle rebalancing operations are generally triggered near the point of capacity max-out.

According to Gast (2011) docking stations were identified as a) Problematic stations (completely full or completely empty) b) Non-problematic stations (neither completely full nor completely empty) (Gast 2011). The classification of the station is time dependent.

Figure 1: Waterloo 2 without any bicycles

Docking Station: Waterloo 2

Date/Time: 29th June 2105 at 9.30 a.m.

Status: No bicycle

Classification: Problematic Station

Waterloo 2 docking station in a problematic state due to the lack of bicycles

Docking Station: Euston square

Date/Time: 07 July 2015 at 2.30 p.m.

Status: No space

Classification: Problematic Station

Euston Square docking station in a problematic state due to the lack of spaces

Figure 2**:** Euston square docking station

Figure 3 Non-problematic Station

Docking Station: Euston square

Date/Time: 08 June 2015 at 9.30 a.m.

Status: No space

Classification: Non-problematic Stations

Euston Square docking station in a non-problematic state. (enough bicycle and space available)

3.6.3 London Waterloo Area

Waterloo has the largest docking station with 126 docking points as well as being the busiest station (with 34,823 hires and docks made over this 6 week period, with an average of 1,161 hires and docks every weekday). There are a total of three docking stations in very close proximity to the train and Tube station as can be seen from the layout in figure 4.

The observations were carried out for Waterloo docking Stations 1, 2, 3 at starting at 8.00 a.m. on 29th June, Monday.

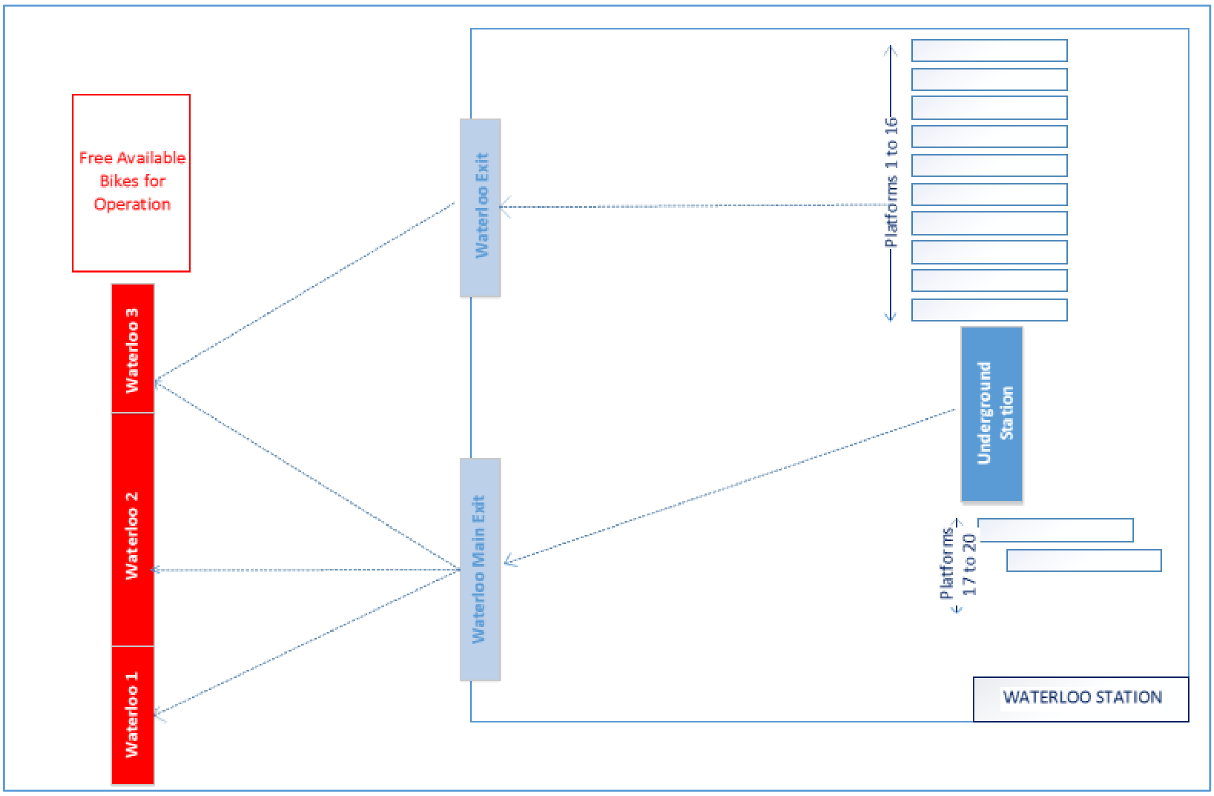


Figure 4: The layout of Waterloo station and respective docking stations (not drawn to scale)

3.6.4 London King's Cross Area

The docking stations (1-3) in the vicinity of London King's Cross were observed starting at 8 a.m. on Wed, 01 July 2015. Figure 5 presents the layout of the train station and respective docking stations.

1. Belgrove Street, King's Cross
2. St. Chad's Street, King's Cross
3. Northdown Street, King's Cross

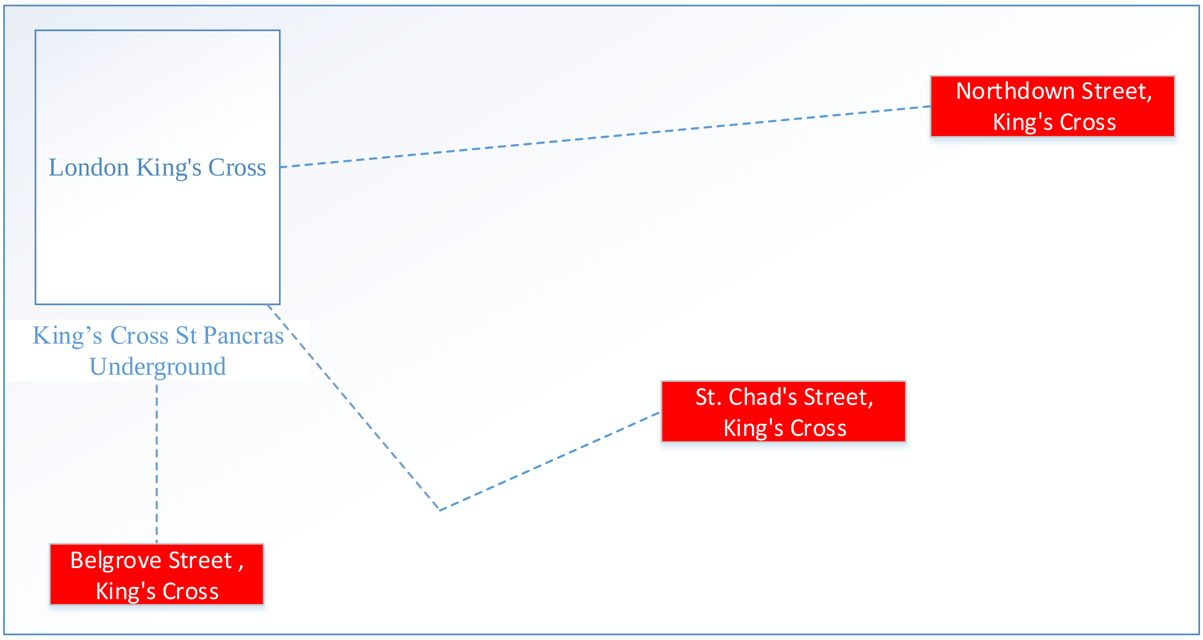


Figure 5: The layout of King's Cross stations & docking stations (not drawn to scale)

3.6.5 Victoria Station Area

Observations at Victoria Station were carried out on Monday, 29th June starting at 8:00 a.m. The docking stations observed were Cardinal Place and Ashley Place as shown in Figure 6

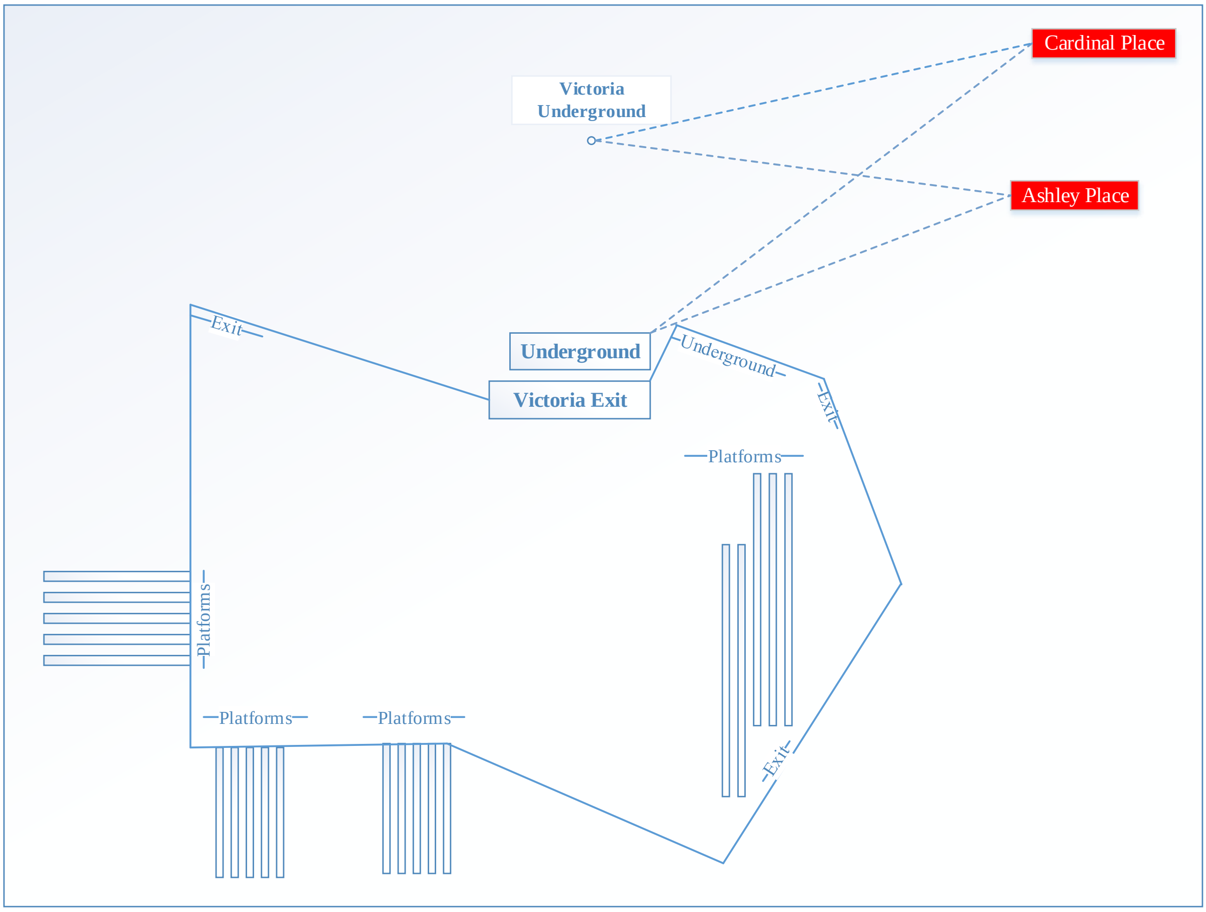


Figure 6: The layout of Victoria Station and respective docking stations (not drawn to scale).

4. Results

4.1 Introduction

This chapter is divided into three sub sections to cover the three part analyses that were carried out to understand the various aspects of demand imbalance in LBSS. The first section details the BSS usage analysis in London based on the journeys data in comparison to the information that can be extracted and presented using the docking stations capacity updates. It starts with demand factor based analysis of the individual docking stations followed by the network wide view of the demand imbalance for both bicycles and spaces.

The second part of the analysis section focuses on the sources of demand for LBSS. In an integrated urban transport system a mixture of modes of transport are generally used to complete a journey. In this section the relationship of LBSS and Oyster network flow is explored using a number of analysis techniques.

As with any other system, it is not possible to understand all aspects of the problem by merely looking at the data. It is often useful to observe the situation at ground level in order to better appreciate the complexities of the problem. The third section focuses on the results of physical observations at three of the busiest transit hubs. The transport hubs selected provide an ideal observation vantage point to understand the flow of users between networks and the operational aspects of bicycle docking stations.

Each section provides a summary of the results and the discussion around the usefulness of the new insight in the workings of the LBSS.

4.2 Demand dynamics of LBSS

4.2.1 Introduction

Capacity updates include information such as the number of available bicycles and the number of available docking spaces at a given time of the day. This information can be useful in establishing the load of a docking station at the time of the update. The load factor, which is the key measure used in some of the previous studies (O'Brien *et al.*, 2014), is the proportion of docking points in each docking station that currently has a bicycle available to hire. It is normally calculated from the number of bicycles and the number of free spaces in each docking station. The load in the ''load factor'' term therefore is a reference to a load of bicycles filling docking points – rather than a load of bicycles from the system being used on the streets (O'Brien *et al.*, 2014).

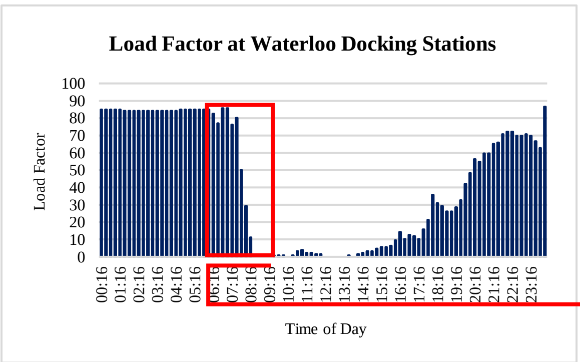


Figure 7: Load factor at Waterloo docking stations (1, 2 and 3) (7 June 2012).

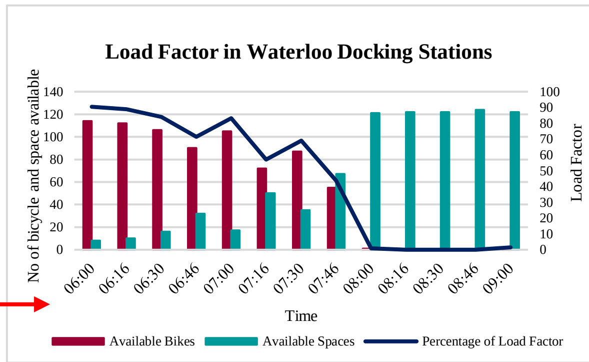


Figure 8: Load of Waterloo docking stations during morning peak (7 June 2012, 6:00 a.m.-9:00 a.m. data).

Figures 7 and 8 chart a typical working day using the capacity updates at Waterloo docking stations. As weekday commuters heavily use the station, the usage pattern is a reflection of the peak rush hour usage. By the middle of the morning rush hour (8:00 a.m. to 9:00 a.m.), the station is already running at maximum load; meaning that all available capacity has already been utilized. This picture changes only later in the day when the flow of commuter is in the reverse direction (after 2:00 p.m.).

Although figures 7 and 8 provide a very clear indication of the load, this information cannot be directly translated into demand due to the following reasons:

- A bicycle coming into the docking station and one bicycle going out of the docking station will not have any impact on the capacity update value, but it will impact demand.
- It is common knowledge that bicycle rebalancing activities are carried out by the operators to keep the system in a balanced state. If the analysis was only looking at the capacity updates, bicycles added to the docking station during the rebalancing operations may be wrongly categorized as demand for space.
- Similarly, bicycles removed from the docking station through rebalancing operations may be wrongly categorised as demand for bicycles.

4.2.2 Variations in the demand of Bicycles and Spaces

Demand of Bicycle (DF_B) and Demand of Space (DF_S) were calculated for all the docking stations in London (SQL code attached in Appendix 2), using the formulation explained in chapter 3. As an example, figure 9 shows DF_B and DF_S at every 15-minute intervals for Waterloo. Load of bicycles and spaces have an inverse relationship if considered from the perspective of capacity, but this is not the case when viewed from the context of demand. As can be seen from figure 9, there is not a direct or inverse relationship. This is because demands for bicycles are driven by journeys going out and demand for spaces is driven by journeys coming in and these two are completely independent variables.

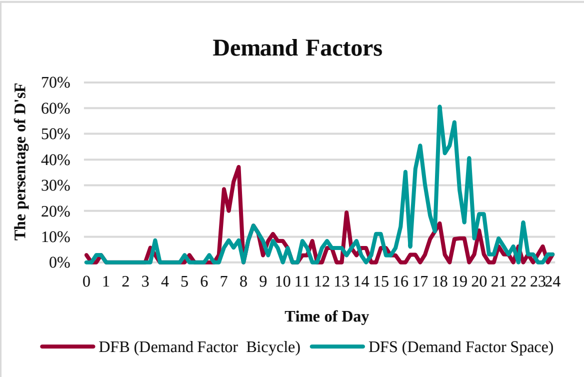


Figure 9: Demands at 15 min interval at Waterloo 1 docking station (20 June 2012).

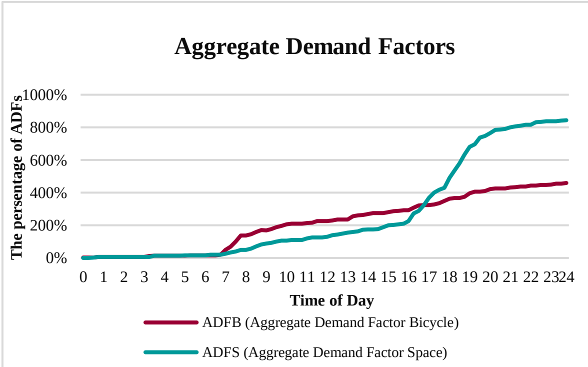


Figure 10: Aggregate demand Factors at Waterloo 1 docking station (20 June 2012)

Figure 10 provides aggregate demand, which is calculated incrementally over the whole day, by adding demand for every 15-minute interval. The two lines of aggregate demand follow very different trajectories and by the end of the day there is a substantial difference between the demand values of the bicycles and spaces. This is expected, as the inflows are 270 and outflows are only 147. A difference of this magnitude is an indication that the docking stations have an imbalance problem that is not automatically corrected during the day. It is also an indication that rebalancing operations are carried out for this station. This result has also been validated through physical observations of the docking station that are explained in detail later in this chapter. The aggregate demand can also be useful to:

- Identify the stations with the maximum demands discrepancy (bicycle and space) at a given time of the day.

- A station with less difference in the aggregate demands values of bicycle and spaces can be classified as self-balancing. This means no rebalancing operations are required for such a station.
- Another deduction from these charts is the overall turnover of bicycles at the docking station. High turnover is not reflected in the capacity updates but it is a reflection of high activity at the stations. This can help to determine possible new locations for the expansion of the service in the vicinity of the docking station.

Figure 11 presents the net flow at Waterloo 1 docking station. The net flow highlights the difference of incoming and outgoing flow from the station during a specified time interval.

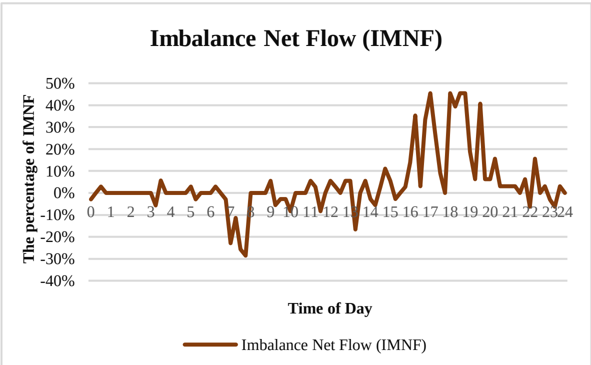


Figure 11: Imbalance at 15-minute interval at Waterloo 1, (20 June 2012).

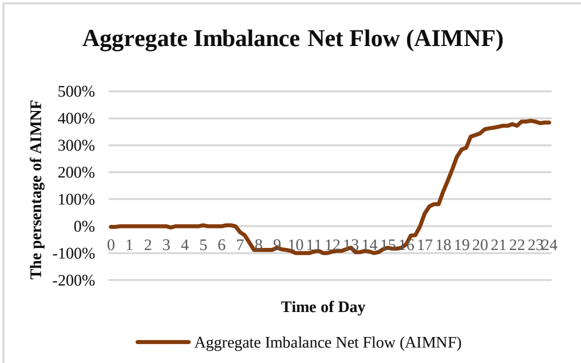


Figure 12: Aggregate imbalance net flow at Waterloo 1, (20 June 2012) is calculated incrementally during the day using the 15-minute imbalance.

As can be seen from figure 12, the positive inflow over the day resulted in significant positive imbalance (more bicycles coming in than going out) over the course of the day. For a station that is self-balancing at the end of the day will have the aggregate imbalance close to zero, which is not the case for Waterloo 1. With the limitation of fixed capacity at the docking station, the discrepancy in the number of in-coming and out-going bicycles appears confusing at first. The inflow is significantly higher than outflow from the afternoon starting at 4 p.m. As explained earlier, it can only be possible if we assume that significant manual operations are carried out at the stations where bicycles are removed from the docking station to make space for more bicycles. This was also confirmed through the physical observations of the system explained in detail in the next section.

The same analysis was also applied to the docking stations of Waterloo 2 and 3, King's Cross's closest docking stations ("Belgrove Street", "St. Chad's Street" and "Northdown Street") and Victoria's closest docking stations ("Ashley Place" and "Cardinal Place")

Belgrove, St Chad Street, King's Cross docking stations and Ashley place, Victoria station demonstrated an aggregate imbalance value closer to 0. This is an indication of a station that is self-balancing over the course of the day. The demand peaks in morning and evening are more than the available capacity of the station and it was observed through the aggregate demand of bicycles and spaces in the morning and evening rush hours. This suggests that bicycle rebalancing operations are carried out at the station to support rush hour operations. Again these results were verified through physical observations of the station.

The figures for the other stations (King's Cross and Victoria) are attached in Appendix 1.

4.2.3 Demand and Self-Balancing Index

The objective of the charts and maps in this section is to provide a network wide view of the demand dynamics. The analysis in this section is presented in the form of the maps of London with various measures drawn against the actual locations of the docking stations.

Table 7 shows some of the docking stations around the biggest transit hubs. DF_B (Demand factor bicycle) for each 15 minute interval is calculated for all docking stations in London. For example DBF_0815 represents DF_B for the interval 0800 to 0815. The first column i.e. Casa ID is the unique docking station identifier. Similarly, DF_S is calculated for the morning and evening rush hour.

Table 7: Calculation of Aggregate DFB

Casa ID	Dock Stations	DFB_0800	DFB_0815	DFB_0830	DFB_0845	(Aggregate DFB(per hour))
10004	St. Chad's Street, King's Cr.	0.09	0.17	0.57	0.13	0.96
10014	Belgrove Street , King's Cr.	2.1	2.57	0.19	0.19	5.05
10177	Ashley Place, Vict.	0.2	0.04	0.32	0.28	0.84
10316	Cardinal Place, Vict.	0.04	0.29	0.08	0.29	0.7
10593	Northdown Street, King's Cr.	0.1	0.05	0.14	0.38	0.67
40374	Waterloo Station 1	0.36	0	0.08	0.14	0.58
40374	Waterloo Station 2	0.09	0.04	0.04	0	0.17
40374	Waterloo Station 3	1.69	1.09	0.06	0.03	2.87
…	….	….	….	….	….	….

Figure 13 and 14 demonstrates demand of bicycles and spaces across London during the morning peak hours. It can be noticed that the demand of spaces is more evenly distributed in and around the financial districts, where there are more offices compared to the other parts of London studied. The demand of bicycles is more scattered and the stations around King’s cross and Waterloo reveal much higher demand variation. This visualization is consistent with other observations that majority of the demand for bicycles in the morning peak hours is clustered around the big transit hub stations around the outer boundaries of the city.

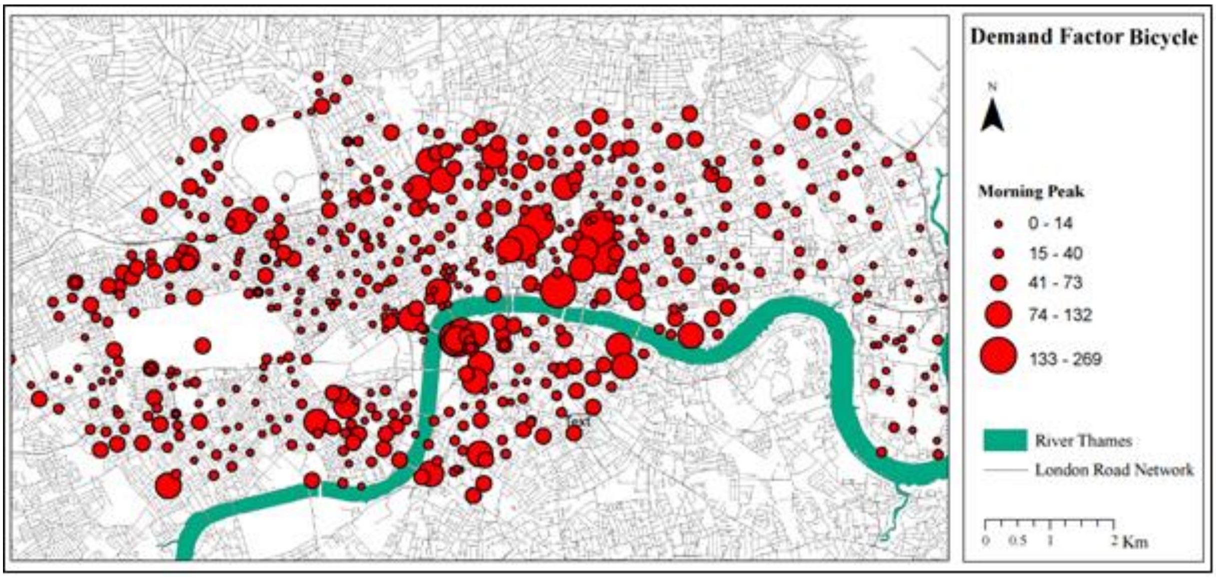


Figure 13: Demand of Bicycle ($\mathbf{DF_B}$) calculated for all the docking stations in London.

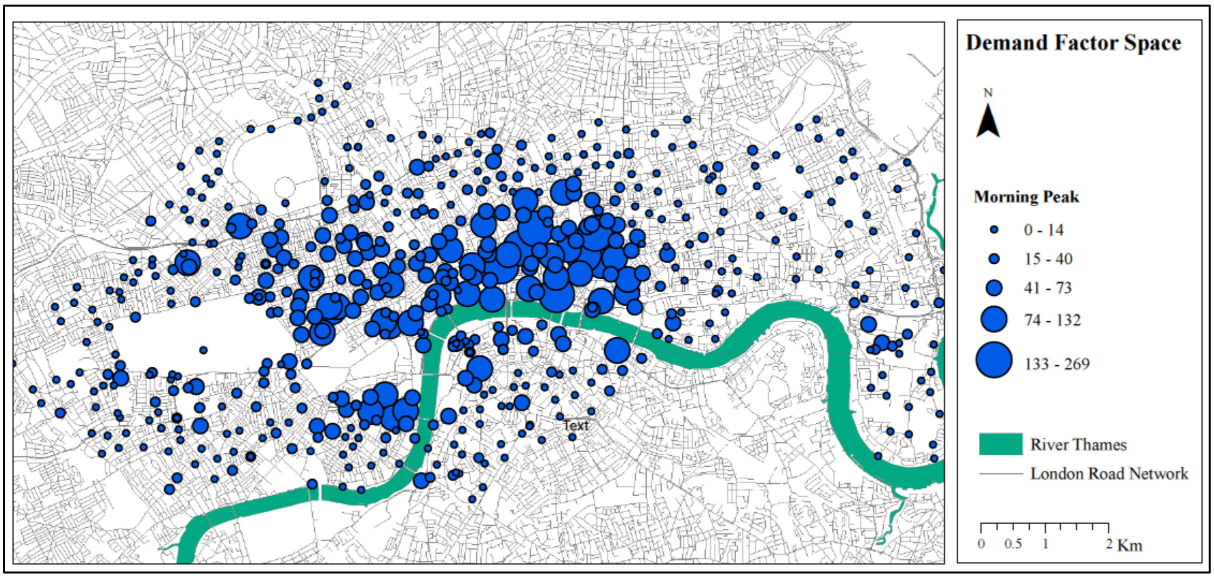


Figure 14: Demand of Space ($\mathbf{DF_S}$) calculated for all the docking stations in London

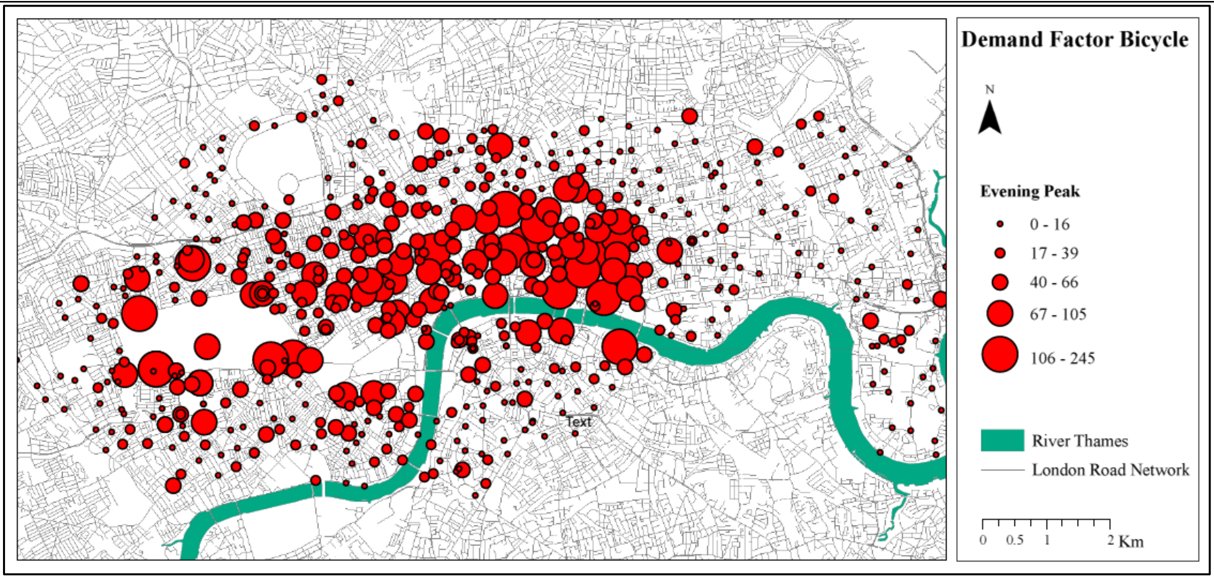


Figure 15: Demand of Bicycle (DF_B) calculated for all the docking stations in London

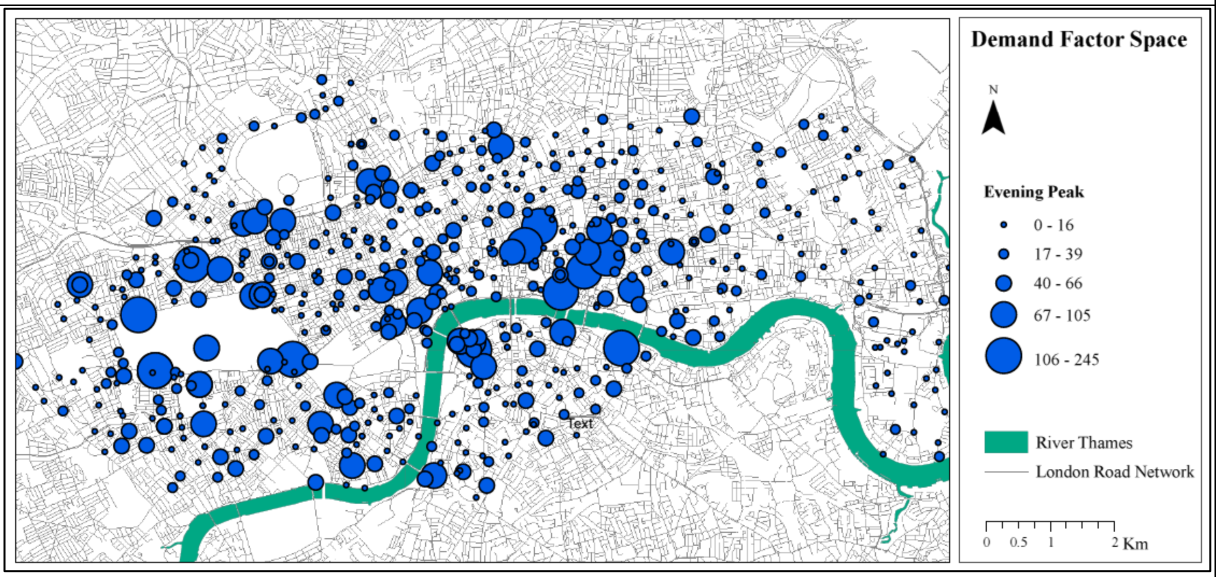


Figure 16: Demand of Space (DF_S) calculated for all the docking stations in London

Figure 15 and 16 presents a trend that is the reverse of what was observed during the morning peak hours. The DF_s here is more clustered around the outer boundaries of the city and near the transit hubs. The DFB on the other hand is evenly spread across the inner city.

Table 8 shows the calculation of self-balancing index for some of the docking stations around the busiest transit hubs. In the first column DFB calculated for each 15-minute interval and aggregate value is considered for each docking station at 24 hours. Similarly, DFs is calculated and aggregate values for the day are presented in the second column. As the last step, the difference between the two columns is used to calculate self-balancing index, explained in the methodology. For the comparison, midday index is calculated at 12 hours.

Table 8: Calculation of self-balancing stations in London per day (created as an example)

Casa ID	Dock Station	($AggregatedDF_B$)	($AggregatedDF_S$)	Self-balancing Index (DFB-DFS)
10004	St. Chad's Street, King's Cross	4	3.67	0.33
10014	Belgrove Street , King's Cross	17.49	16.91	0.58
10177	Ashley Place, Victoria	3.76	3.64	0.12
10316	Cardinal Place, Victoria	5.2	4.6	0.6
10593	Northdown Street, King's Cross	1.99	1.61	0.38
40374	Waterloo Station 1	4.12	8.57	-4.45
40374	Waterloo Station	1.32	8.57	-7.25
40374	Waterloo Station 3	8.88	8.57	0.31
…	….	….	….	….

These observations lead to the hypothesis that during the course of the day the flow is back and forth between inner city and outer boundaries. Even though these variations cause imbalance at various times of the day, if left alone the majority of the docking stations have a tendency to revert to the equilibrium state. The following section demonstrates the demand discrepancies in the docking stations at various times of the day i.e. Midday and End of the day.

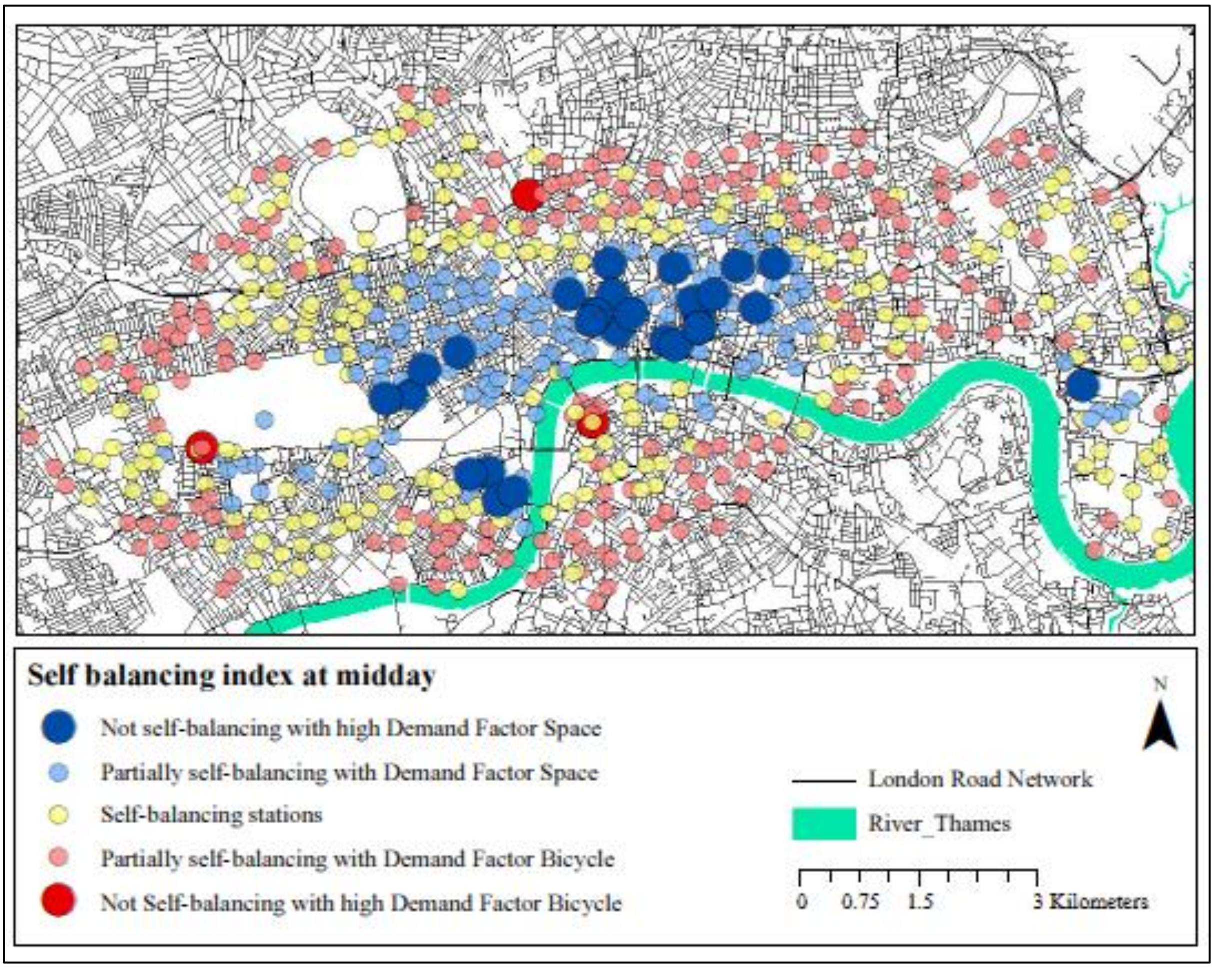


Figure 17: Self-balancing Stations at midday in London Network.

Figures 17 and 18 represent the demand imbalance calculations in such a way as to highlight the areas with varying degrees of demand imbalance. The formulation of demand imbalance is presented in chapter 3 with the docking stations being classified as a) Self balancing, b) partially self-balancing c) not self-balancing (or high degree of demand and supply discrepancy).

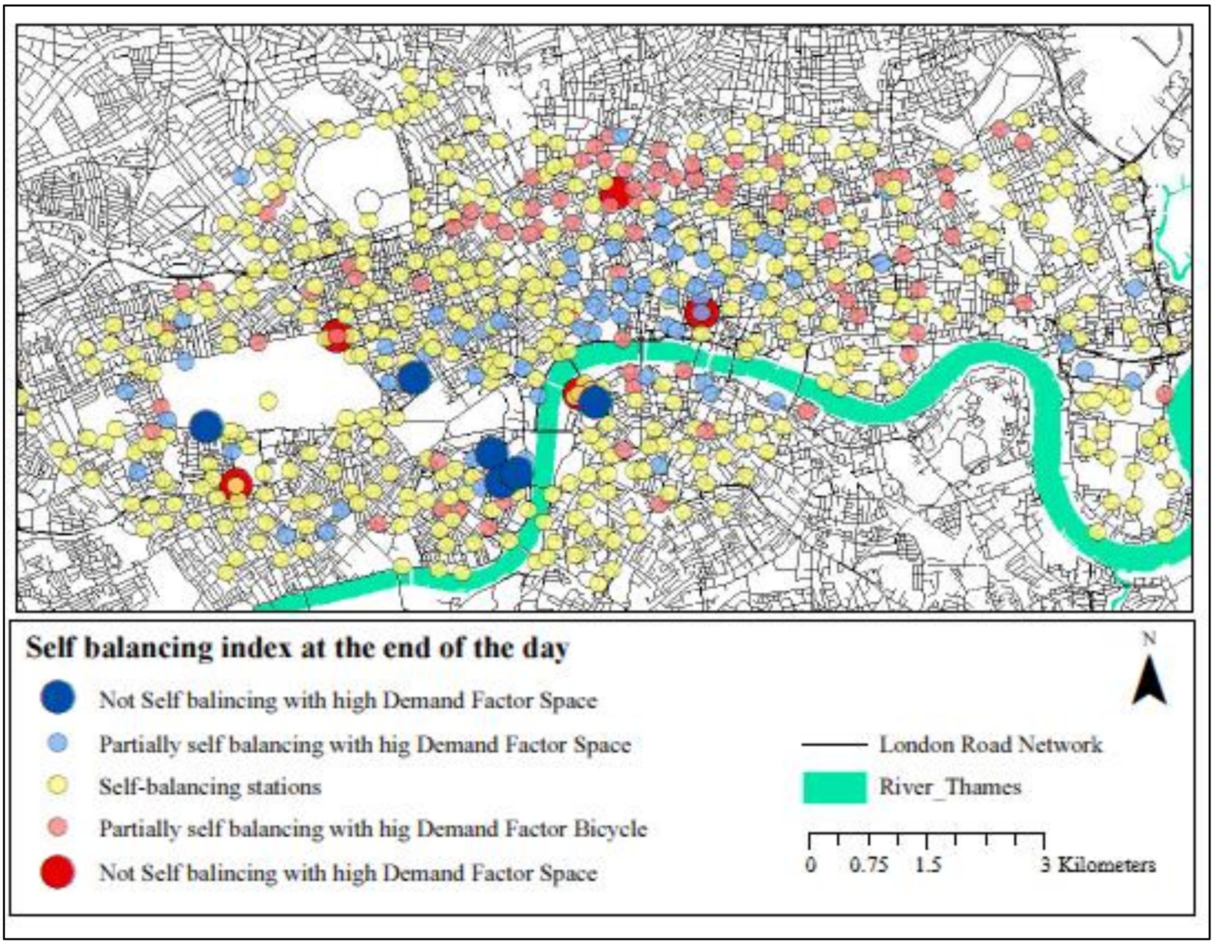


Figure 18: Self-balancing Stations at the end of the day (20 June 2012) in Central London.

Figure 17 shows the imbalance in the middle of the day (12 p.m., 20 June 1012) and it can be noticed that the vast majority of the stations demonstrates a high degree of asymmetry due to bicycles or spaces. The underlying assumption is that even though some docking stations appear to be imbalanced, if allowed to function without any interference they will eventually return to an equilibrium state. Figure 18 presents the same imbalance calculation at the end of the day and the majority of the docking stations have returned to the equilibrium state. The imbalance values are not calculated from the capacity update but represent actual journeys.

Table 9: Problematic docking stations example (not self-balancing docking stations with high DF_s and DF_B) during the midday and at the end of the day

Time	Not self-balancing with high DFB	Not self-balancing with high DFs
All Day	Cheapside, Bank	Curzon St, Mayfair
	Speaker's Corner Hyde Park	Waterloo 1,2, and 3, Waterloo
	Concert Hall Approach, South Bank	Smith Square Westminster
	Harwick Street, Clerkenwell	Horseferry Road, Westminster
	Natural History Museum, South Kensington	Place Gate, Kensington Gardens
Mid-Day	Waterloo 1,2 and 3, Waterloo	Wormwood St, Liverpool
		Norton Folgate, Liverpool Street
		Finsbury Square, Moorgate
		Fore St,Wood St, The Guildhall, Guildhall
		Algersgate, Barbican
		Queen Victoria St, St Paul's
	Belgrove Street, King's Cross	Stonecutter, Holborn
		Breams Buildings, Holborn
		Carey St, Holborn
		Red Lion St, Holborn
		Smith Square Westminster
		Horseferry Road, Westminster
		Grafton St, Mayfair
		Stanhope Gate, Mayfair

4.2.4 Summary of results

The demand of bicycle and spaces during weekdays present a consistent pattern. Docking stations near the transit hubs such as Waterloo, King's Cross and Liverpool Street see asymmetric demand of bicycles in the morning. The demand of spaces peak in the financial district and West End during the same time. This results in temporary imbalance of the system by midday. In the evening starting at 4 p.m. the demand trend reverses. The majority of the docking stations exhibit net flow which is close to 0. This observation suggests that the station is self-balancing over the course of the day.

Some docking stations such as those near to the Liverpool Street area exhibit high demand of bicycles and spaces at the same time. This results in high turnover and low imbalance throughout the day. This pattern is indicative of a location that is a transit hub as well as an area with offices.

All day analysis highlighted asymmetric DF_B at Bank, Hyde Park and South Kensington and asymmetric DF_s at Waterloo 1, 2 and 3, Westminster area, Mayfair and Kensington Gardens.

The results presented in this section demonstrate the usage patterns for an average weekday. The usage behaviour is dynamic and can be impacted by a variety of other factors such as weather, season and transport strike action.

4.3 Analysis of Source of Demand for LBSS

In order to understand the demand dynamics of LBSS it is important to understand the intermodal relationship of LBSS and the Oyster network. This section will present the results of the analysis that have been carried out to quantify the demand of bicycle and space in LBSS due to TfL Oyster network.

4.3.1 Study Area

Central London was the area studied for the report but the focus was on the three of busiest commuter regions due to their importance as intermodal transit hubs. Those areas were Waterloo, Victoria and King's Cross.

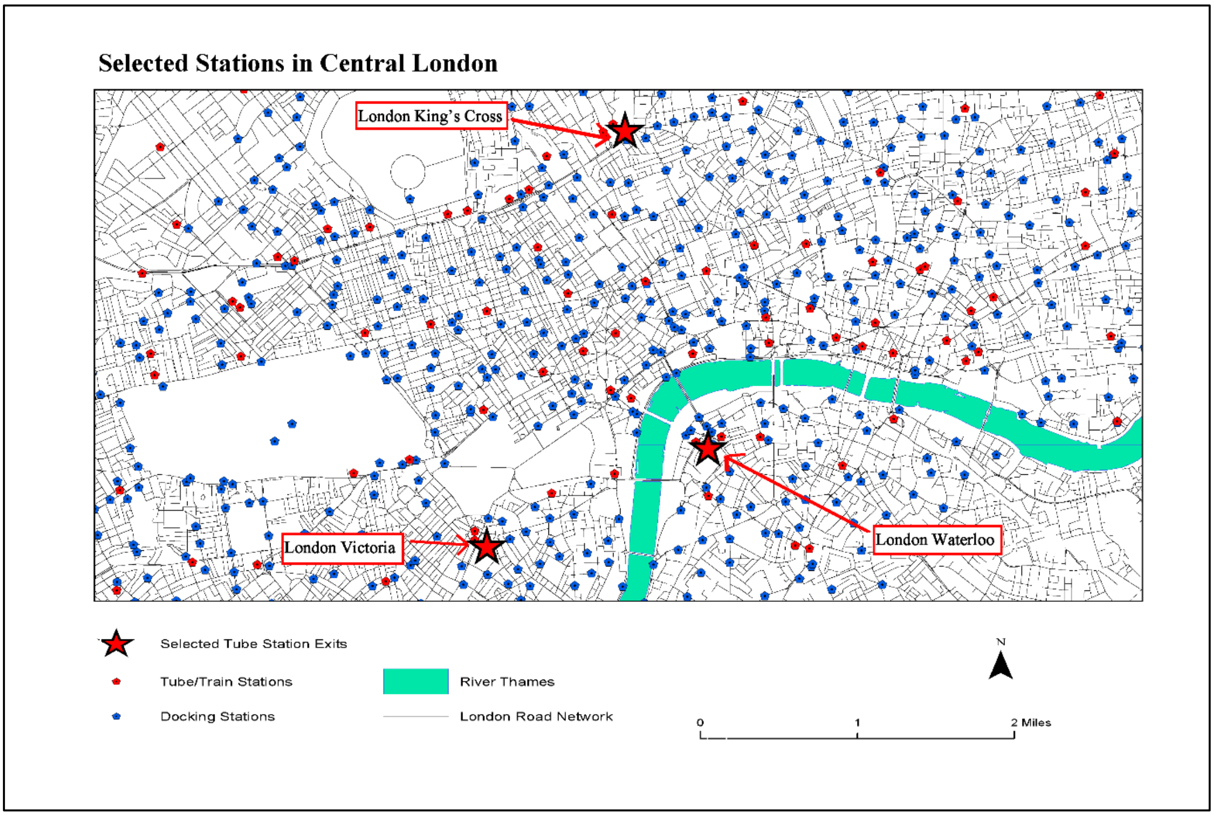


Figure 19: Selected regions in the study area i.e. London Waterloo, Victoria and King's cross.

The map shows the location of Oyster network and LBSS stations using the geographical coordinates

4.3.2 Network Analysis

In mixed mode transportation, the ease of transition between the modes is of paramount importance. As a first step to establish the relationship between Oyster network stations and LBSS docking stations, an analysis was carried out based on the proximity of the stations in the two networks.

'Closest facility' network analysis was conducted to find the docking stations in close adjacency to the Tube stations for the available data (747 Docking stations and 163 Tube Stations) and to identify the shortest path between them. A maximum of five docking stations were selected within 300m (walking distance) of the selected Tube stations.

In order to undertake a more detailed analysis, three Tube stations - London Waterloo, London King's Cross and London Victoria - were selected (shown as stars in Figure 1 and their details described below.

Table 10: The list of the and docking stations

Name	Station Exits	Bicycle Docking Stations
Waterloo	Shell Gates	
	Main Gates	Waterloo Station 1
	Auxiliary Gates	Waterloo Station 2
	W&C Validators	Waterloo Station 3
	Jubilee Gates	
Victoria	District Gates	Ashley Place, Victoria
	Main Gates	Cardinal Place, Victoria
King’s Cross	Main Entry Gates	
	Tube Gates	St. Chad's Street, King's Cross
	Thameslink Gates	Belgrove Street , King's Cross
	Northern Ticket Hall	Northdown Street, King's Cross

4.3.3 Data Analysis

4.3.3.1 Daily Data (7 June, 2012)

To understand the daily pattern of the relationship between the two networks, data points were plotted for the three selected stations in Central London.

The objective of integrating the data sets of two networks identified mixed mode users and their journeys. There are two possible scenarios:

1. A commuter leaves the oyster network (train/Tube) and takes a shared bicycle from the nearest LBSS docking station to continue the journey.
2. A commuter enters a docking station on a shared bicycle and leaves the bicycle and continue remaining journey on Oyster network.

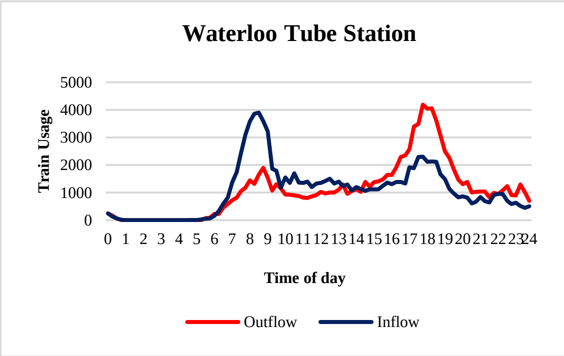


Figure 20: Waterloo Stations one-day data (5 April 2012)

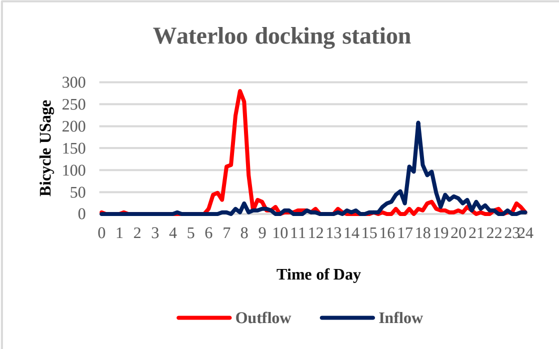


Figure 21: Waterloo Docking Stations, 24 hours data (5 April 2012)

Daily data for Waterloo stations are shown in figures 20 and 21. Morning outflow data (6 a.m.-10 a.m.) have significant demand peaks compared to afternoon (5 p.m.-9 p. m.) outflow data on both networks. This can be an indication that there may be a relationship in morning outflow data from Tube stations to docking stations. Afternoon inflow data (5 p.m.-9 p.m.) have also similar relationship from LBSS to Oyster network, but not as significant as that seen in the morning peak data. Plots for other stations i.e. London King's Cross and London Victoria are in the appendix section of this report. The strength of the relationship is explored further using Correlation coefficients and linear regression in later sections

4.3.3.2 Weekly Data

Results plotted for the same three stations from the 23 April to the 29 April show the separation between the weekend and weekday trends at hourly intervals. Figure 22 shows two peaks for Tube in both morning and afternoon and it can be seen the afternoon peak is higher than the morning peak. Figure 23 highlights the morning peak data for bicycle users. The afternoon peak is less obvious for the bicycle outflow as it may be due to national rail commuters using the Oyster network for the first leg of their journey.

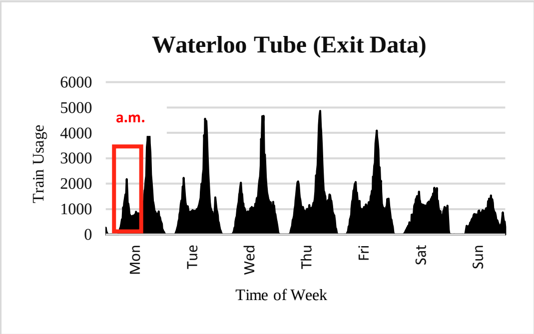


Figure 22 Weekday and Weekend data for Waterloo station. Highlighted peaks indicate possible relationship.

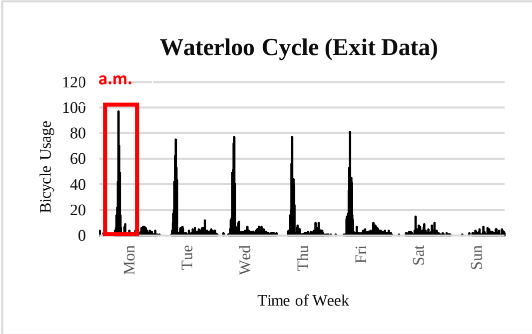


Figure 23 Weekday and Weekend data for the docking stations adjacent to Waterloo station. Highlighted peaks indicate possible relationship

4.3.3.3 Monthly Data (April, 2012)

The analysis for the month of April (2012) was carried out to highlight the relationship between docking station and train station over a longer duration. The usage pattern of office commuters is repetitive over the weekdays compared to the other users. Therefore, the results for the four-week period further emphasised the weekday and weekend trends visible in the weekly data.

4.3.4 Pearson Correlation

After investigating the daily, weekly and monthly trends, further insight could be gained through calculating the Pearson correlation coefficient (r). Hourly morning and afternoon peak monthly data are calculated in R Studio using Equation 5. The results are presented in Table 11.

Table 11: The results for Pearson Correlation coefficients in Waterloo Tube/Train Stations and Docking Stations.

Morning Peak (0600 a.m.-1000 a.m.)/Morning Activities				
Station to Docking	**0600-0700**	**0700-0800**	**0800-0900**	**0900-1000**
Waterloo (Outflow)	0.88534	0.90249	0.90056	0.39092
Evening Peak (0500 p.m. -0900 p.m.) /Evening Activities				
Station to Docking	**0500-0600**	**0600-0700**	**0700-0800**	**0800-0900**
Waterloo (Outflow)	0.23922	0.42573	0.27059	0.41861

Correlation Strength Classification

Description	Range
Very Weak	0.01 – 0.19
Weak	0.20 – 0.39
Modest	0.40 – 0.69
Strong	0.70 – 0.89
Very Strong	0.90 – 0.99

Correlation is an effect size and so it can describe the strength of the correlation using the guide suggested by Evans (1996). Correlation Strength classification in table 11 shows the range and descriptions for the absolute value of r.

According to Evans (1996), the range of the coefficients shows that there is a strong correlation from 6:00 a.m. to 9:00 a.m. in the morning peak. After 9.00 a.m., the correlation coefficient drops significantly (weak correlation). The reason might be because office commuters have either already reached their locations or they could not find bicycle/space in docking stations. On the other hand, there is weak and very weak correlation in the evening rush hours. The reason is that in the evening, the flow is in the reverse direction from LBSS to Oyster network. This can be demonstrated using inflow (from the docking station) to inflow (to the oyster network) correlation rather than outflow data from the Oyster network to the docking station.

4.3.5 Linear Regression

Correlation coefficient is non parametric and only indicates how two variables associate with one another, but it does not give an idea of the type of relationship between them. Regression models help investigate the relationship between a dependent variable y and one or more explanatory variables denoted x (Wooldridge n.d.) Linear regression was conducted by assuming docking stations time series data as the dependent variable and train station data as the independent variable.

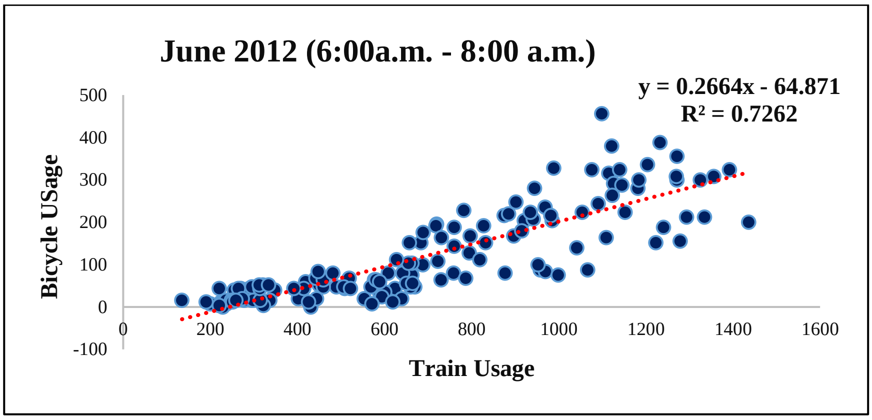


Figure 24: Linear regression for June 2012, a.m. peak (6:00 a.m.-8:00 a.m.)

Figure 24 shows two-hour intervals over a period of month excluding weekends. The trend line (red colour) demonstrate that the correlation between Oyster usage and bicycle usage. It is a good fit with a coefficient of determination + .72, explaining the variability of the data. It means that the increase in the number of commuters on the Oyster network is linked to the increase in the usage on LBSS.

4.3.6 Summary of results

This section of the analysis was set out to study the relationship between the bicycle hire network and the TfL Oyster network in order to understand the source of demand. The results were mostly as expected with the strength of the relationship closely linked to the rush hour commuting patterns. There was a dip in the correlation after the rush hour, which was very likely due to the drop in the number of commuters from the Oyster network

The outflow of the docking station is the variable related to DF_B. This variable is also very important in order to understand the relationship between two networks if two networks are in close proximity. It also applied to inflow, which controls DF_S in demand dynamics in Central London.

Daily/ weekly/ monthly data analysis simply shows that there is relationship. To understand the relationship between two networks, the data themselves were not enough. Because of this, Pearson correlation was applied. It was deduced that there was a strong relationship between two networks during commuter rush hours. Based on linear regression there was a positive relationship between two networks, meaning the increase in the Oyster flow results in an increase in LBSS flow.

4.4 Physical Observations

This section presents the results of the physical observations of the LBSS carried out for three of the busiest transit hubs in the central London.

4.4.1 London Waterloo Area

Figure 25 shows three of the closest docking stations in the vicinity of the transit hub Waterloo i.e. “Waterloo 1”, “Waterloo 2” and “Waterloo 3”. These stations were observed on 29th June 2012 during the morning and evening peak hours.

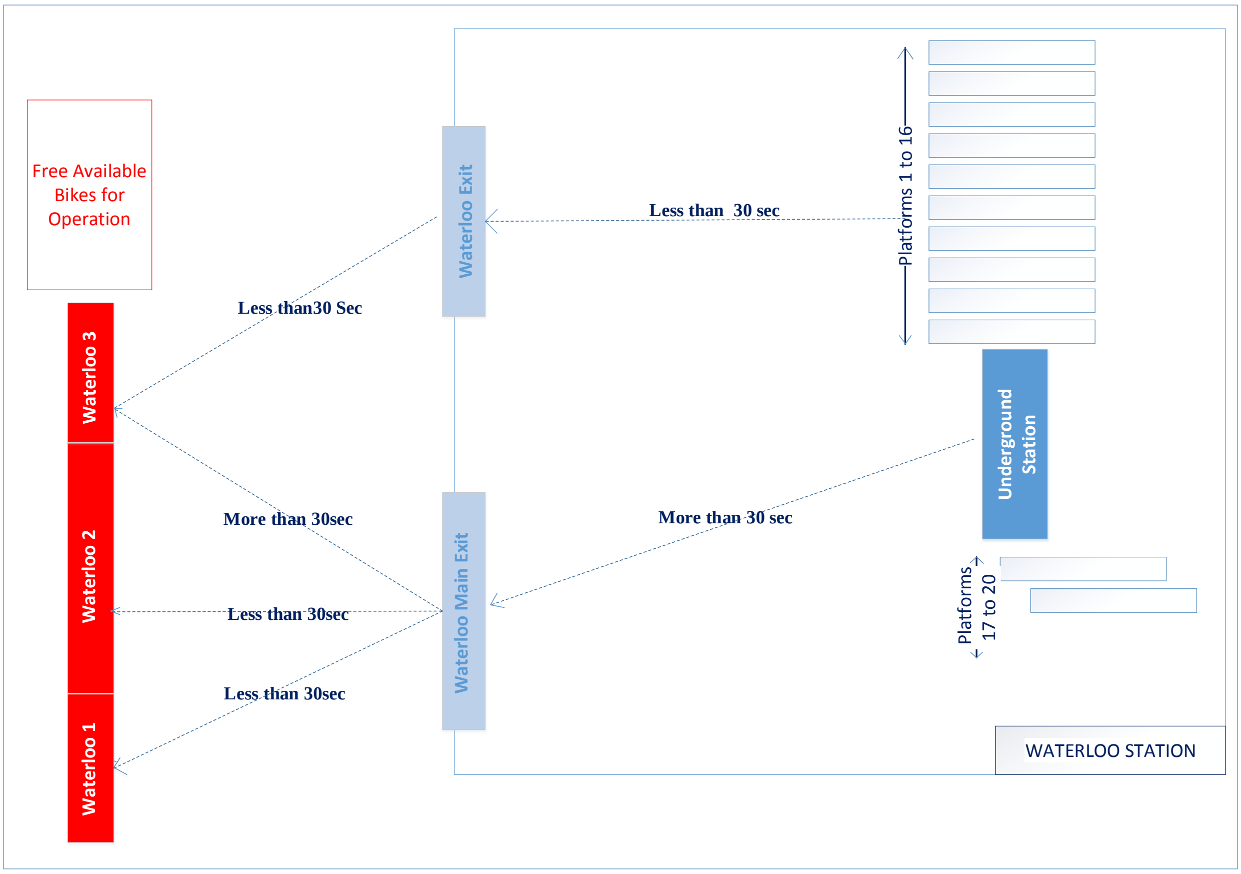


Figure 25: The layout of Waterloo station and respective docking stations (not drawn to scale)

In the layout, Waterloo transit hub Exits are divided into: Waterloo Exit and Waterloo Main Exit. Train commuters are more likely to use “Waterloo Exit” due to the short distance from platforms 1 to 16 to Waterloo docking stations. This path is also less likely to be crowded compared to Waterloo Main Exit. The walking time from all the platforms 1 to 16 to Waterloo 3 is approximately 1 min. “Waterloo Main exit” is the logical choice for Waterloo Tube users as well as platforms 17 to 20. They are more likely to prefer to exit from “Waterloo main exit” and the possible docking stations for users are Waterloo 1 and 2.

Surplus bicycles are stocked near Waterloo 3 docking station. Based on the additional requirements, bicycles are added to Waterloo 1, 2, 3 during the morning rush hour. Approximately 200 extra bicycles are added during the morning rush hour to the three waterloo docking stations by the operations team. Due to this reason the actual outflows from the stations are much larger than the available capacity. Because of the consistent surplus demand of bicycles at Waterloo during rush hour, operation teams provide additional bicycles and spaces for users. This observation is an indication that expansion of the facilities may be required in and around Waterloo area and this has been highlighted in a number of studies (Fricker & Gast 2012; O'Brien *et al*., 2014). The large capacity docking stations are less likely to be effected by temporal imbalance because the capacity provides a buffer against the variations in demand. This result is also confirmed by the data analysis carried out in this study.

4.4.2 London King's Cross Area

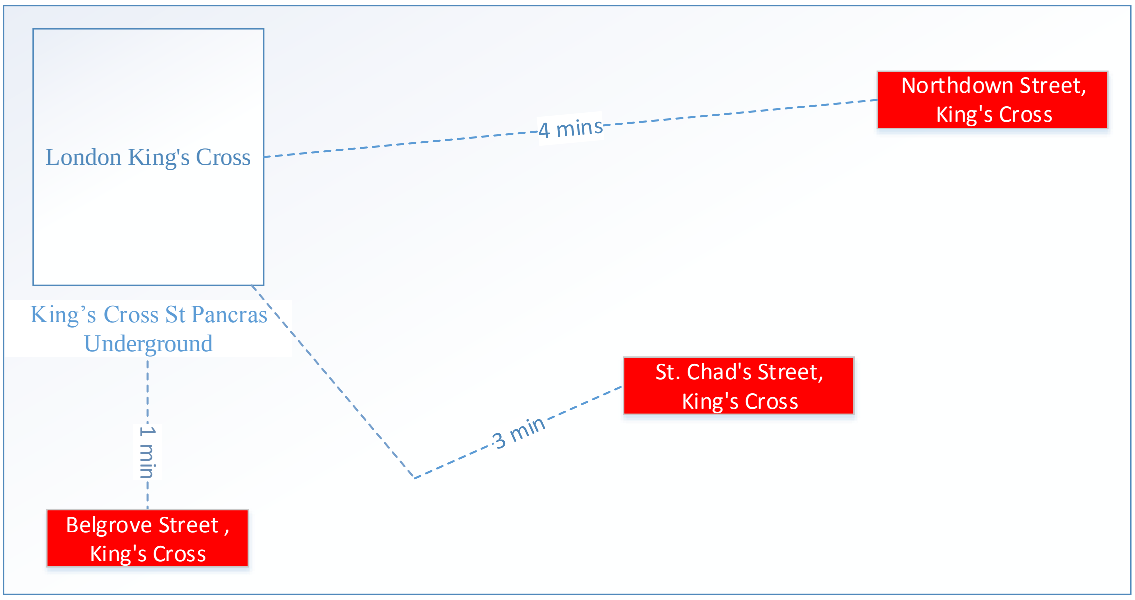


Figure 26: London King's Cross, London King's Cross St Pancras Station and their respective docking stations. (Not drawn to scale)

Three closest docking stations i.e. "Belgrove Street", "St. Chad's Street" and "Northdown Street" were observed during the day. The walking distance to the docking stations is 1min, 3 min and 3 min, respectively. Rebalancing operations teams are tasked with adding bicycles to the station in the morning peak hours and create space in the evening rush hours.

4.4.3 Victoria Stations

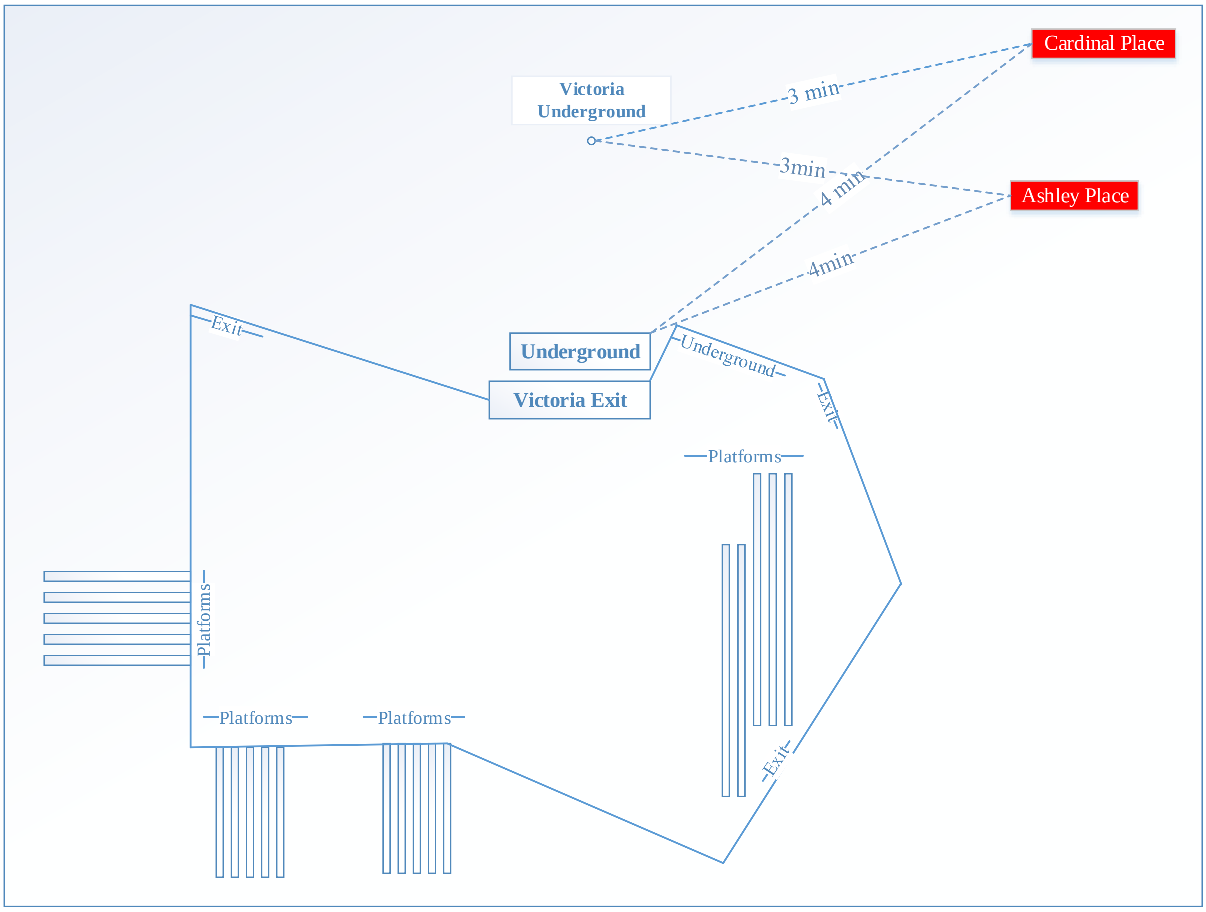


Figure 27: London Victoria with respective docking stations. (Not drawn to scale)

Two closest docking stations i.e. "Ashley Place" and "Cardinal Place" were observed during the day at Victoria Station on Monday, 29th June starting at 8:00 a.m. Two exit points from the Tube stations can be used to get to the docking stations and it takes approximately the same time (3 to 4 min) from each exit.

Reaching the nearest docking station from Tube/train exits is difficult due to ongoing engineering works and heavy traffic in the area. It is expected that the North and the South side of the station will be completed in 2016 and 2018 respectively (TfL 2015c). This will improve the intermodal connectivity of the transit hub significantly and result in increased demand in LBSS.

4.5 Summary of results

The objective of the physical observations of the docking stations was to get an insight to the workings that may not be very obvious from merely looking at the historical data from TfL. There are no data available for the operational activities carried out by the LBSS operators.

- Operational activity is largely limited to the large transit hub stations such as London Waterloo, Victoria and King's Cross. This observation agrees with hourly data analysis, which suggests that large numbers of the bicycle docking stations are self-balancing by the end of the day.
- Layout and ease of transition plays a very significant role in the usage of LBSS. For example due to the redevelopment of the area near Victoria station the accessibility to the bicycles is more difficult compared to other docking stations. The new layout of the station is expected to improve North and South side access to the station, and it is expected that it will improve the intermodal link with LBSS
- Operational difficulties include restrictions on the movement of bicycle storage trucks in the City of Westminster at night due to the restrictions on the movement of HGV imposed by council.
- Improvements in the methodology can include preparing a questionnaire and getting feedback directly from the users of LBSS. This will provide the user perspective of the system that can enable further improvement.

5. Discussion

It was shown that demand surge in the morning and evening rush hours was directly correlated to the demand surge on Oyster networks (trains and Tube). This is also noted in the user survey based study by Chen *et al*. (2012). The layout of the station along with its proximity is very significant in the way the commuters use BSS to complete the second leg of their journeys. Nair *et al*. (2013) also suggested that close coupling of transit and the bicycle sharing systems can lead to higher utilization. Intermodal trips with shared-vehicle segments can provide value-addition for users. Therefore, policies that seek to integrate the two systems can be profitable. Examples of such policies include seamless fare collection, preferential fares for transit users, and prime location of shared-bicycle stations.

The demand imbalance was created when some docking stations generated high demand of bicycles and low demand of spaces or *vice versa*.

It was observed transport hubs that were close to the business district created demand for bicycles and spaces during the same time interval. This included stations such as Liverpool Street. Large numbers of shared bicycle users are coming to the Liverpool street station area due to offices in the area, causing the demand of spaces there. At the same time due to the presence of Liverpool Street train and Tube stations, large numbers of commuters were using the docking stations in the area to complete the last mile of their commute. This led to a high demand for bicycles. The result of this high demand in bicycles and spaces during the same time interval means high turnover, which is highly desirable for the system. This result provides a very useful policy guideline for city planners. Transit hubs should have offices closer to them if cities want to facilitate sustainable modes of transportations such as shared bicycles. Unfortunately this also highlighted the problem at transit hubs such as Waterloo, Victoria and King Cross that are not as close to the many office areas. This was observed through the analysis of the data as well as the physical observation of the docking station. The mentioned transport hubs are serviced by operators in the morning and evening rush hours every weekday, to fill additional supply of bicycles and spaces.

In other significant results of the analysis using O-D journeys data, a large number of bicycle docking stations were shown to be self-balancing over the course of the day. It highlighted the need for less intervention in the running of the system and allow the natural flow of users to correct any imbalances. It was also observed that size of the docking station has significant impact on the imbalances as larger capacity accommodates temporal imbalances to some extent which supports the observations made by Fricker & Gast (2012) and O'Brien *et al*. (2014).

6. Conclusion

- It has been argued in this study that flow data (origin to destination journeys) provide better information to analyse demand in BSS.
- This is the first study of its kind to demonstrate with the help of flow data, the self-balancing characteristics of the docking stations.
- The report quantifies the strength of intermodal relationships between Oyster Network and LBSS. It was shown that demand surge in the morning and evening rush hours was directly correlated to the demand surge on Oyster network. This was also noted in physical observations.
- Data analysis noted that bicycle re-balancing operations are carried out at a number of transit hub stations during rush hours where the static capacity overwhelms the demand.

7. Future Work

Even though it has been demonstrated that a vast majority of the docking stations are self-balancing over the course of the day, it does not remove the need for reallocating bicycles for reasons such as adverse weather events and to increase the usage of the bicycles. The analysis carried out in this study can provide the basis for future work that may be required for any reallocation mechanism. It provides a very clear indication for areas with demand discrepancies at various times of the day.

The analysis provides a basis for forecast of LBSS journeys that can be used to drive future demand of the network. This is because O-D can very accurately highlight which areas have demand for bicycles and which areas have demand for spaces. That is the next logical step for this study. This type of information can help in decisions with regards to extensions of existing docking stations and planning for new ones.

This research can also be extended in several other directions. Alternative descriptions of the demand processes using a simulation model can be developed. One area for further work can be the analysis of bicycle re-balancing activities. Cost and benefit analysis of moving bicycles in the network to compare against other alternatives such as crowd sourcing.

The rate at which the resources (bicycle or spaces) are being consumed is the only measure of rush hour demand. A model can be built to extrapolate this trend for the whole of the morning and evening rush hour period to provide an indication of hypothetical maximum demand at the stations.

The model built for the analysis in this report can easily be extended to cover a longer period of up to 2 years. Using similar demand and self-balancing measures can enable the understanding of the seasonal variations in demand. It will also allow evaluation of longer term demand trends using inflows and outflows. Such a study can be of benefit to the operators and designers of bicycle sharing systems.

8. References

Beecham, R. & Wood, J., 2013. Exploring gendered cycling behaviours within a large-scale behavioural data-set. *Transportation Planning and Technology*, 00(00), pp.1–15. Available at: http://www.tandfonline.com/doi/abs/10.1080/03081060.2013.844903\nhttp://www.tandfonline.com.ezproxy.unal.edu.co/doi/full/10.1080/03081060.2013.844903#.Ul1Zm1BWw40. [Accessed: 21st August 2015].

Blackwell, M. & Sen, M., 2012. Large Datasets and You: A Field Guide. *Manuskript*, 14627, pp.1–8. Available at: http://www.mattblackwell.org/files/papers/bigdata.pdf\npapers3://publication/uuid/9B8CCBB4-E848-4D7E-9F6E-14AE5A96FC0C. [Accessed: 20th August 2015].

Blythe, P. & Bryan, H., 2007. Understanding behaviour through smartcard data analysis. *Proceedings of the ICE - Transport*, 160(4), pp.173–177.

Borgnat, P. et al., 2011. Shared Bicycles in a City: a Signal Processing and Data Analysis Perspective. *Advances in Complex Systems*, 14(03), pp.415–438.

Campbell, D. & Campbell, S., 2008. Statlab Workshop Introduction to Regression and Data Analysis with. *Analysis*.

Case, B., 2009. Cycle Hire Scheme Document 2 ST-PJ302C Cycle Hire Scheme – Business Case Submission. , pp.1–35.

Cepolina E. M. and Farina A., 2012. Urban car sharing: An overview of relocation strategies. *Urban Transport XVIII*, p.419. Available at: https://books.google.co.uk/books?hl=en&lr=&id=Alechd_pX2UC&oi=fnd&pg=PA419&dq=+Cepolina,+E.+M.,+and+Farina,+A.+2012.+Urban+car+sharing:+An+overview+of+relocation+strategies.+WIT+Transactions+on+the+Built&ots=43P4v-c224&sig=wNJXQKM4P3_BlgftAUFYEBBXIQA#v=on. [Accessed: 21st August 2015].

Chemla, D. et al., 2013. Self-service bike sharing systems : simulation , repositioning , pricing To cite this version :

Chen, L. et al., 2012. Determinants of Bicycle Transfer Demand at Metro Stations. *Transportation Research Record: Journal of the Transportation Research Board*, 2276(-1), pp.131–137. Available at: http://trb.metapress.com/openurl.asp?genre=article&id=doi:10.3141/2276-16. [Accessed: 21st August 2015].

Clemente, M. et al., 2013. The vehicle relocation problem in car sharing systems: Modeling and simulation in a Petri net framework. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7927 LNCS, pp.250–269.

Corcoran, J. et al., 2014. Spatio-temporal patterns of a Public Bicycle Sharing Program: the effect of weather and calendar events. *Journal of Transport Geography*, 41, pp.292–305. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0966692314001951. [Accessed: 1st August 2015].

Davis, U.C., 2008. Year 2008 EasyConnect : Low-Speed Modes Linked to Transit Planning Project. , (June).

DeMaio, P., 2009. Bike-sharing : History , Impacts , Models of Provision , and Future. *Journal of Public Transportation*, 12(4), pp.41–56. Available at: http://bike.cofc.edu/bike-share-program/history of bike sharing.pdf. [Accessed: 1st August 2015].

Fricker, C. & Gast, N., 2012. Incentives and regulations in bike-sharing systems with stations of finite capacity. *arXiv preprint arXiv:1201.1178*, pp.1–25. Available at: http://arxiv.org/abs/1201.1178. [Accessed: 18th August 2015].

Gast, N., 2011. Sizing , Incentives and Redistribution in Bike-sharing Systems.

GoDCgo, 2011. Capital Bikeshare's Reverse Rider Rewards. Available at: http://www.godcgo.com/reverse-rider-rewards.aspx. [Accessed: 21st August 2015].

Goodman, A. & Cheshire, J., 2014. Inequalities in the London bicycle sharing system revisited: impacts of extending the scheme to poorer areas but then doubling prices. *Journal of Transport Geography*, 41, pp.272–279. Available at: http://www.sciencedirect.com/science/article/pii/S0966692314000659. [Accessed: 21st August 2015].

Holleczek, T. et al., 2013. Digital breadcrumbs : Detecting urban mobility patterns and transport mode choices from cellphone networks. *arXiv preprint arXiv:1308.6705.*, pp.1–14.

Institute for Transportation & Development Policy (ITDP), 2013. The Bike-Sharing Planning Guide. , p.152. Available at: www.itpd.org. [Accessed: 10th August 2015].

Kaltenbrunner, A. et al., 2010. Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. *Pervasive and Mobile Computing*, 6(4), pp.455–466. Available at: http://dx.doi.org/10.1016/j.pmcj.2010.07.002. [Accessed: 21st August 2015].

Kek, A.G.H. et al., 2009. A decision support system for vehicle relocation operations in carsharing systems. *Transportation Research Part E: Logistics and Transportation Review*, 45(1), pp.149–158. Available at: http://dx.doi.org/10.1016/j.tre.2008.02.008. [Accessed: 1st August 2015].

Kusakabe, T., Iryo, T. & Asakura, Y., 2010. Estimation method for railway passengers' train choice behavior with smart card transaction data. *Transportation*, 37(5), pp.731–749.

Lathia, N., Ahmed, S. & Capra, L., 2012. Measuring the impact of opening the London shared bicycle scheme to casual users. *Transportation Research Part C: Emerging Technologies*, 22, pp.88–102. Available at: http://dx.doi.org/10.1016/j.trc.2011.12.004. [Accessed: 21st August 2015].

Lin, J.R. & Yang Ta-Hui, T.H., 2011. Strategic design of public bicycle sharing systems with service level constraints. *Transportation Research Part E: Logistics and Transportation Review*, 47(2), pp.284–294. Available at: http://dx.doi.org/10.1016/j.tre.2010.09.004. [Accessed: 1st August 2015].

Midgley, P., 2011. Bicycle-Sharing Schemes: Enhancing Sustainable Mobility in Urban Areas. *Commission on Sustainable Development*, 19th Sessi(8), pp.1–26.

Nair, R. et al., 2013. Large-Scale Vehicle Sharing Systems: Analysis of Vélib'. *International Journal of Sustainable Transportation*, 7(1), pp.85–106.

Nair, R. & Miller-Hooks, E., 2011. Fleet Management for Vehicle Sharing Operations. *Transportation Science*, 45(4), pp.524–540.

O'Brien, O., Cheshire, J. & Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography*, 34, pp.262–273. Available at: http://dx.doi.org/10.1016/j.jtrangeo.2013.06.007.

Pfrommer, J. et al., 2014. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems*, 15(4), pp.1567–1578.

Shaheen, S. a., Guzman, S. & Zhang, H., 2010. Bikesharing in Europe, the Americas, and Asia. *Transportation Research Record: Journal of the Transportation Research Board*, 2143(-1), pp.159–167.

Shaheen, S. & Cohen, A., 2007. Worldwide Carsharing Growth: An International Comparison. *Transportation Research Record: Journal of the Transportation Research*, 3483, pp.81–89.

Shaheen, S., Rodier, C. & Eaken, A., 2005. Improving California's Bay Area Rapid Transit District Connectivity and Access with Segway Human Transporter and Other Low-Speed Mobility Devices. *Transportation Research Record*, 1927(1), pp.189–194.

Singla, A. & Krause, A., 2013. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. *Proceedings of the 22nd international conference …*, pp.1167–1177. Available at: http://dl.acm.org/citation.cfm?id=2488490. [Accessed: 21st August 2015].

Small K.A., V.E.T., 2007. The Economics of Urban Transportation. Available at: https://books.google.co.uk/books?isbn=1134495714. [Accessed: 10th August 2015].

TfL. 2013. *Annual Report and Statement of Accounts Annual Report and Statement of Accounts.* [Online] [Accessed: 16th August, 2015].

TfL. 2015a. *Our Feeds.* London Transport London Available [Online] at: https://tfl.gov.uk/info-for/open-data-users/our-feeds#on-this-page-4. [Accessed: 16th August, 2015].

TfL. 2015b. *Santander Cycles*. , (February), London Transport London [Online], pp.2–6. Available at: https://tfl.gov.uk/cdn/static/cms/documents/sc-transparency-to-end-march-2015.pdf. [Accessed: 16th August, 2015].

TfL. 2015c. *Victoria Station Upgrade*. London Transport London [Online] Available at: https://tfl.gov.uk/cdn/static/cms/documents/vsu-project-update-july-2015.pdf [Accessed: 16th August, 2015].

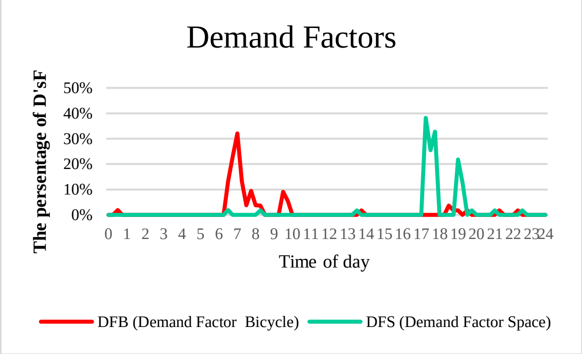
Tran, T.D., Ovtracht, N. & d'Arcier, B.F., 2015. Modeling Bike Sharing System using Built Environment Factors. *Procedia CIRP*, 30, pp.293–298. Available at: http://linkinghub.elsevier.com/retrieve/pii/S2212827115004692. [Accessed 16th August, 2015].

Wooldridge, J., The Simple Linear Regression Model. Available at: http://www2.warwick.ac.uk/fac/soc/economics/staff/vetroeger/teaching/po906_week567.pdf [Accessed: 13th August, 2014].

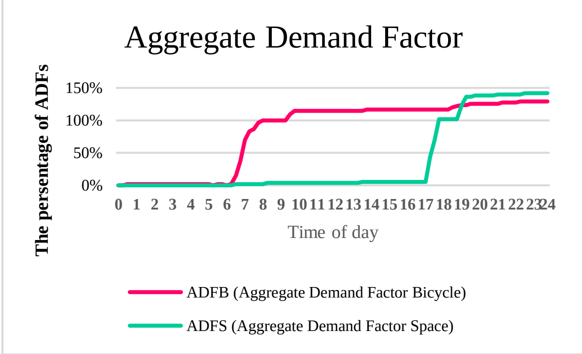
Zhu, X. & Guo, D., 2014. Mapping large spatial flow data with hierarchical clustering. *Transactions in GIS*, 18(3), pp.421–435.

9. APPENDIXES

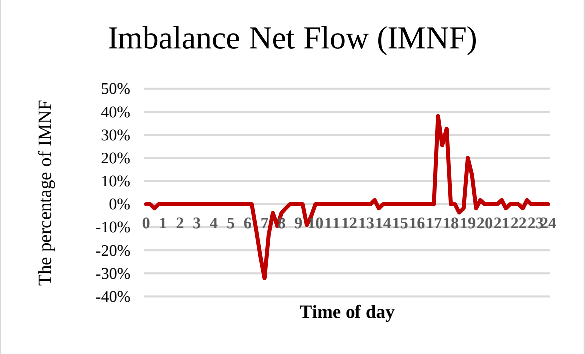
9.1 Appendixes



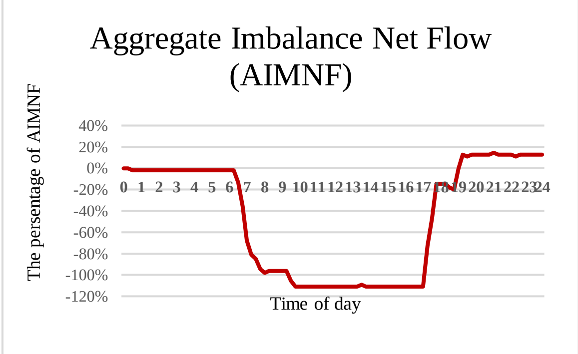
Demands at 15 min interval at Waterloo 2 docking station (20 June 2012).



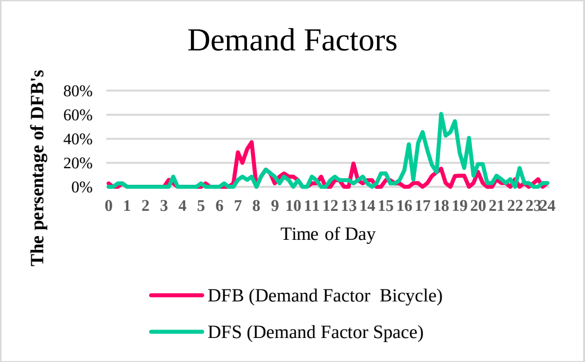
Aggregate demand Factors at Waterloo 2 docking station (20 June 2012)



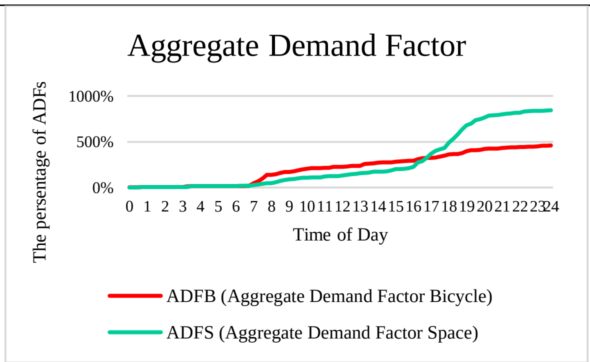
Imbalance at 15 minute interval at Waterloo 2 (20 June 2012).



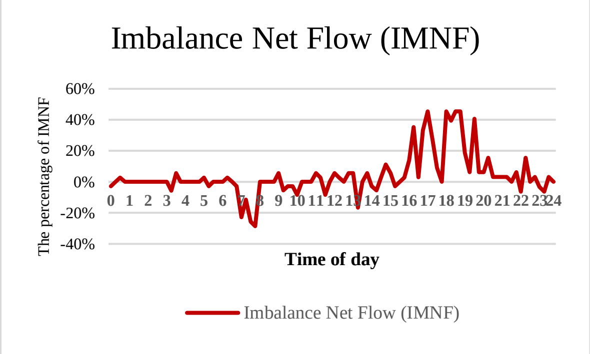
Aggregate imbalance net flow at Waterloo 2 (20 June 2012) is calculated incrementally during the day using the 15 minutes imbalance.



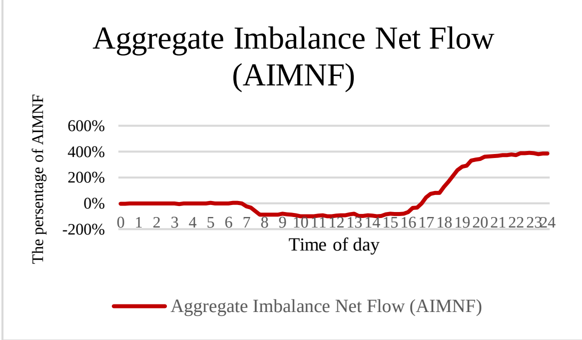
Demands at 15 min interval docking station (20 June 2012).



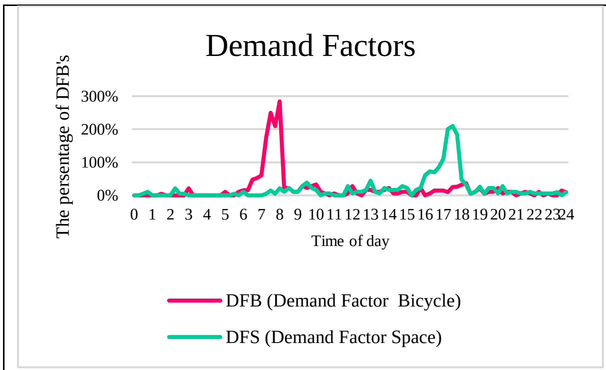
Aggregate demand Factors at Waterloo 3 docking station (20 June 2012)



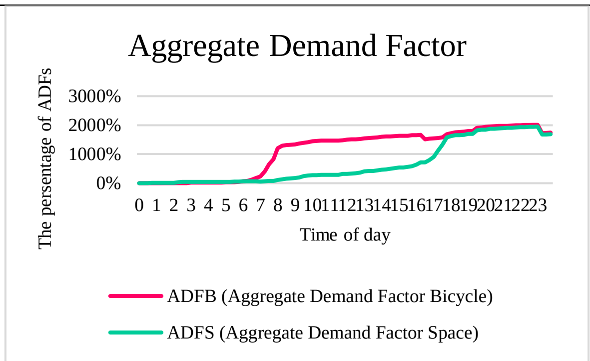
Imbalance at 15 minute interval at Waterloo 3 (20 June 2012).



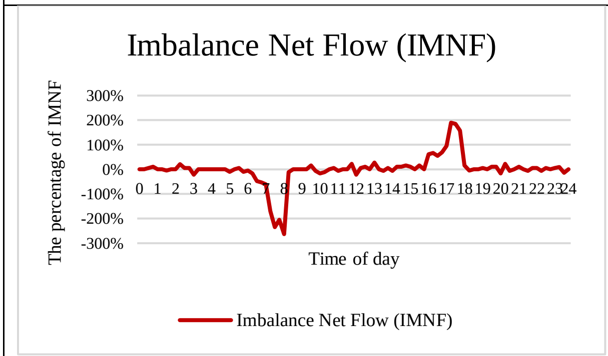
Aggregate imbalance net flow at Waterloo 3 (20 June 2012) is calculated incrementally during the day using the 15 minutes imbalance.



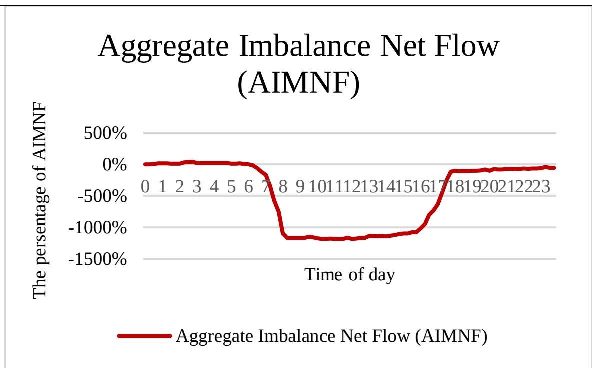
Demands at 15 min interval at Belgrove St, King's Cross docking station (20 June 2012).



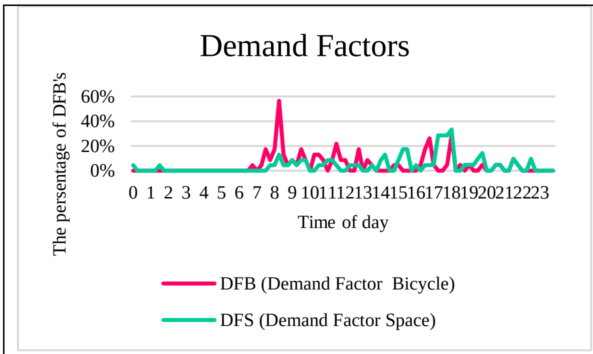
Aggregate demand Factors at Belgrove St, King's Cross docking station (20 June 2012)

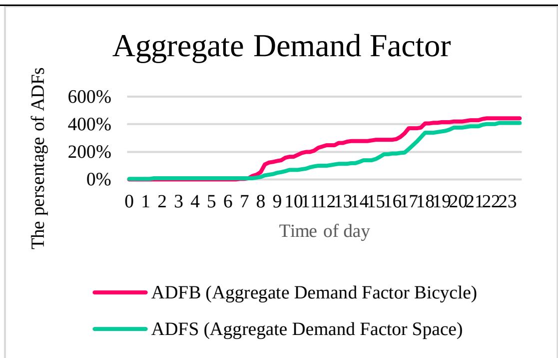


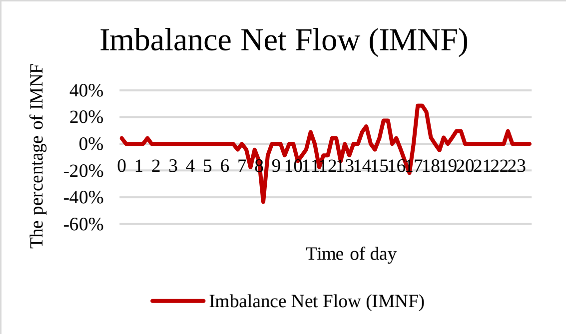
Imbalance at 15 minute interval at Belgrove St, King's Cross (20 June 2012).

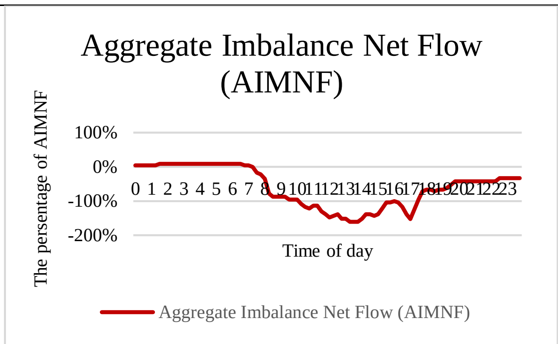


Aggregate imbalance net flow at Belgrove St, King's Cross (20 June 2012) is calculated incrementally during the day using the 15 minutes imbalance.

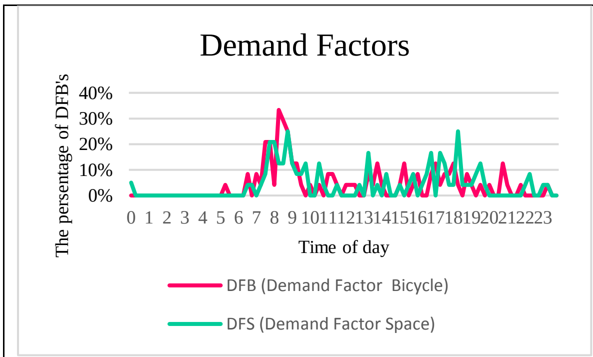
<table>
<tr>
<td>


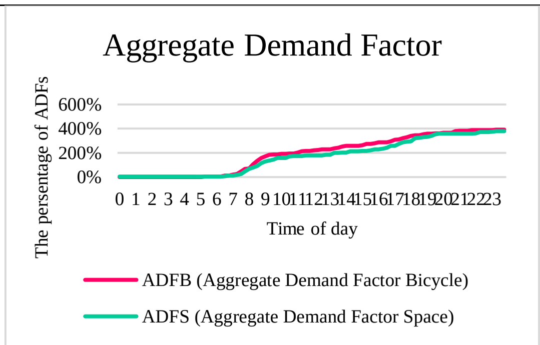
</td>
<td>


</td>
</tr>
<tr>
<td>Demands at 15 min interval at St Chads, King’s Cross docking station (20 June 2012).</td>
<td>Aggregate demand Factors at St Chads, King’s Cross docking station (20 June 2012)</td>
</tr>
<tr>
<td>


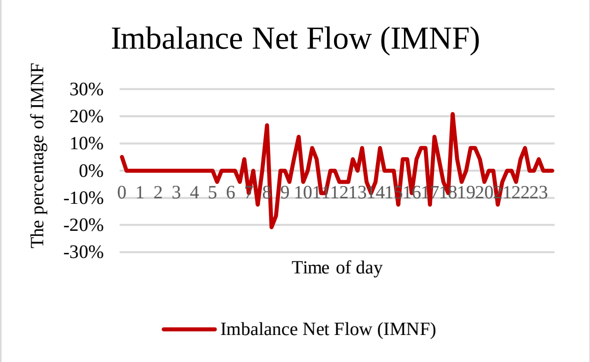
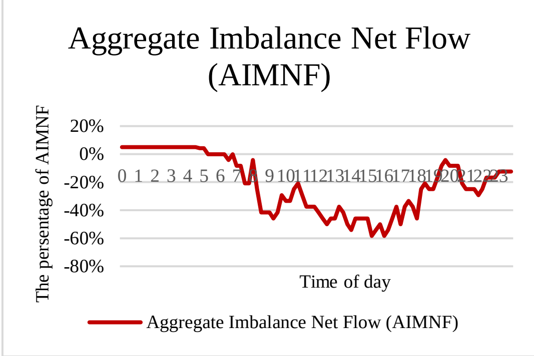
</td>
<td>


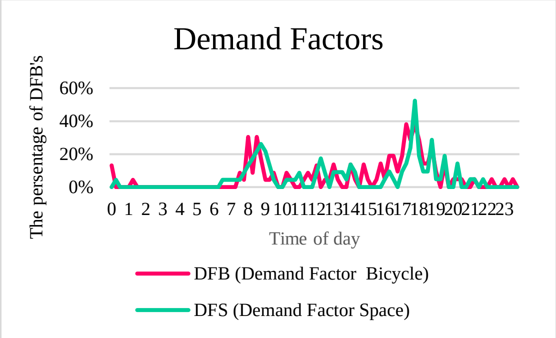
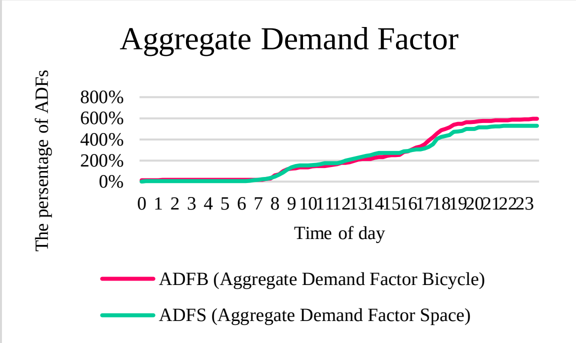
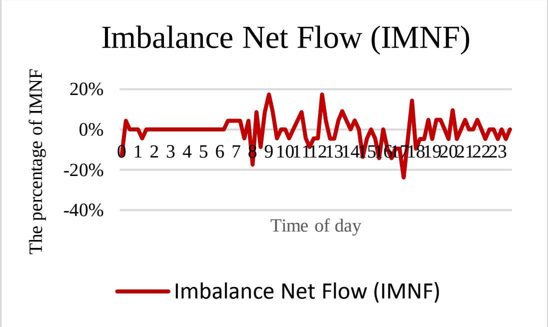
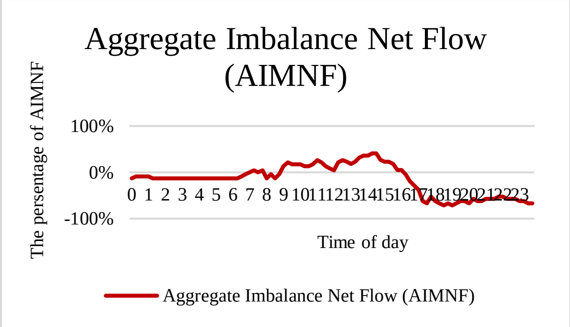
</td>
</tr>
<tr>
<td>Imbalance at 15 minute interval at St Chads, King’s Cross (20 June 2012).</td>
<td>Aggregate imbalance net flow at St Chads, King’s Cross (20 June 2012) is calculated incrementally during the day using the 15 minutes imbalance.</td>
</tr>
</table>

<table>
<tr>
<td>


</td>
<td>


</td>
</tr>
<tr>
<td>Demands at 15 min interval at Ashley Place, Victoria docking station (20 June 2012).</td>
<td>Aggregate demand Factors at Ashley Place, Victoria docking station (20 June 2012)</td>
</tr>
</table>

<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Imbalance at 15 minute interval at Ashley Place, Victoria (20 June 2012).</td>
<td>Aggregate imbalance net flow at Ashley Place, Victoria (20 June 2012) is calculated incrementally during the day using the 15 minutes imbalance.</td>
</tr>
</table>

<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Demands at 15 min interval at Cardinal Place, Victoria docking station (20 June 2012).</td>
<td>Aggregate demand Factors at Cardinal Place, Victoria docking station (20 June 2012)</td>
</tr>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Imbalance at 15 minute interval at Cardinal Place, Victoria (20 June 2012).</td>
<td>Aggregate imbalance net flow at Cardinal Place, Victoria (20 June 2012) is calculated incrementally during the day using the 15 minutes imbalance.</td>
</tr>
</table>