1 2 9 0

UNIVERSIDADE Đ
COIMBRA

Inês Daniela Cardoso Ferreira Frade

Bike-sharing Systems
Demand, Location and Impacts

Tese no âmbito do Programa Doutoral em Sistemas de Transportes orientada pela Professora Doutora Anabela Salgueiro Narciso Ribeiro e apresentada à Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Março de 2022

Faculdade de Ciências e Tecnologia
da Universidade de Coimbra

BIKE-SHARING SYSTEMS
DEMAND, LOCATION AND IMPACTS

Inês Daniela Cardoso Ferreira Frade

Tese no âmbito do Programa Doutoral em Sistemas de Transportes orientada pela Professora Doutora Anabela Salgueiro Narciso Ribeiro e apresentada à Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Março de 2022

1 2 9 0

UNIVERSIDADE Ð
COIMBRA

ACKNOWLEDGEMENTS

Á minha orientadora, Professora Anabela Ribeiro, pela disponibilidade, dedicação, apoio e confiança, bem como pelo espírito crítico que sempre me incutiu e por todo o conhecimento que me transmitiu.

Ao Professor Pais Antunes, pelo desafio que me lançou, pelo conhecimento científico partilhado e pela disponibilidade para o esclarecimento de dúvidas.

Ao Professor Christopher Zegras que me acolheu no MIT, desafiou o trabalho até então desenvolvido e me deu sugestões que valorizaram o meu trabalho.

Ao Professor Álvaro Costa por me ter sido o meu mentor na engenharia de transportes e mobilidade aplicada à realidade, por me ter dado a oportunidade de trabalhar em desafios tão interessantes e pela confiança para estar na linha da frente das soluções.

À Inês pelo poder de síntese, à Marta e ao Daniel que me ajudaram a dominar ferramentas SIG e com quem tirei muitas dúvidas e a toda a equipa da TRENMO por ter ajudado e influenciado a que este trabalho ficasse melhor.

Com experiência que adquiri na TRENMO, tenho a certeza que o trabalho que desenvolvi constitui uma proposta de valor para autoridades locais na implementação de sistemas de bike-sharing.

Ao Hugo Alexandre e Pedro Costa que me desafiaram no equilíbrio entre a eficácia e eficiência, na produção de valor faseado para a construção de um produto final sólido. E aos meus colegas da Worten que na fase final da escrita da tese tiveram comigo uma enorme paciência e compreensão.

Aos meus colegas e amigos do DEC obrigado pela incansável boa disposição e força que me deram nos momentos mais difíceis.

Á Mané, minha amiga de sempre, que me ajudou, que me ouviu e me apoiou, e que em todos as conversas me levantou os braços.

À minha família que me ajudou no equilíbrio pessoal. Aos meus pais, pelo apoio, compreensão, carinho e generosidade em todos os momentos da minha vida. Obrigado por sempre me terem dado as ferramentas necessárias para seguir o meu caminho e por sempre me terem ajudado. Todos os meus sucessos têm o vosso cunho!

Ao André para quem as palavras de agradecimento nunca serão suficientes, pelo ajuda e apoio, por estar sempre presente e disponível. Por me incentivar todos os dias a ser melhor. Não teria aqui chegado sem a tua ajuda. E ao António a quem espero contar esta história de resiliência e que, na sua inocência, me ajudou a manter o foco.

A todos estarei para sempre agradecida!

Em memória do meu avô Manuel do Rosário Frade e da minha avó Maria do Bonsucesso Cardoso que perdi durante esta caminhada. São referências de trabalho e perseverança e por quem continuo a ter uma grande admiração. Tenho a certeza que ficariam orgulhosos da minha chegada aqui!

ABSTRACT

Mobility is a cornerstone to territorial and social development in the world. Growing technological progress, the improvement of economic and social conditions over the last years has contributed to the change in mobility patterns currently centred on the excessive use of individual motorized transport.

The world is increasingly facing environmental problems caused by transport traffic. For that reason, the promotion of sustainable transport alternatives has been seen in the past few decades as one of the assets to reduce the negative externalities related to the transportation sector in urban areas.

The adoption of policies contributing to a modal shift and, consequently, improve the environment and people life quality are urgent and needed.

Bike-sharing systems increased their popularity consistently as transport alternatives in urban areas, and the number of bike-sharing schemes has grown significantly worldwide in recent years.

These systems' success depends on their implementation design. They must be capable of answering peoples' needs, maximizing the investment benefits, as these are the first concerns of the decision-makers.

This research focuses on bike-sharing system design. It intends to develop and provide strategic and practical methods and tools for transportation planners, policymakers, and investors' decision-making.

The main achievements of this decision support methodology are to define the potential demand relating it with the local characteristics, to design the system in terms of location of the stations, number of bicycles, and the dimension of the relocation process, considering the maximization of potential demand and the possible investment, and to estimate the environmental impacts.

As an outcome, two different approaches address the demand estimation on bike-sharing systems. The first approach provides a quick assessment adapted to local characteristics. The second methodology uses regression analysis to understand the variables that influence the demand using an existing bike-sharing system (Boston bike-sharing system).

The design of the system uses an optimization model that defines the location of the bicycle stations, the fleet size, the capacity of the stations, and the number of bicycles in each station, considering an initial investment lower than the given budget. In addition, it balances the annual cost of the system and the revenue, assuming the possibility of a supplementary budget. This budget from the system provider can cover any loss resulting from the shortfall between its operating cost and the revenue from the subscription charges.

Environmental impacts are estimated considering the traffic reduction resultant from a bike-sharing implementation, focused on small Particulate Matter (PM2.5). The results indicate a non-homogeneous relation between traffic reduction and emission reduction across the urban space due to the characteristics of the roads (such as street characteristics and driving conditions), achieving 12.5% of daily PM2.5 emissions in some urban roads.

The work produced in this thesis provides a tool for the design implementation of bike-sharing systems and constitutes a solid starting point for planning and implementing this transport mode.

RESUMO

A mobilidade é um fator fundamental no desenvolvimento territorial e da sociedade. No entanto, o desenvolvimento tecnológico e a melhoria das condições económicas e sociais nos últimos anos tem contribuído para a alteração dos padrões de mobilidade, atualmente muito centrada no uso do transporte individual motorizado.

O sector dos transportes é um dos principais contribuintes para os problemas ambientais atuais e, neste enquadramento, a promoção de alternativas sustentáveis ao uso do transporte individual é visto como uma das principais soluções para a redução das externalidades ambientais deste sector

É assim, incontornável a necessidade de adotar políticas que contribuam para esta alteração modal e, consequentemente, melhorar o ambiente e qualidade de vida das pessoas.

Os sistemas de bicicletas partilhadas são cada vez mais populares como alternativa de sustentável de transporte e o número de sistemas implementados tem aumentado, nos últimos anos, de forma significativa pelo mundo.

O sucesso destes sistemas depende da forma como a sua implementação é definida, como o sistema se adapta às necessidades das populações e como se maximizam os benefícios do investimento. Estes são os principais pontos de preocupação dos decisores.

Neste sentido, o trabalho de investigação apresentado foca-se na implementação dos sistemas de bicicletas partilhados. Pretende constituir uma ferramenta útil para o planeamento de transportes, decisores políticos e investidores no processo de decisão.

Os principais objetivos desta ferramenta são a definição da procura potencial devidamente relacionada com as características locais, o design do sistema incluindo a localização das estações, o número de bicicletas necessário e dimensionamento do processo de relocalização, considerando a maximização da procura potencial e possível investimento, e ainda o dimensionamento dos impactos ambientais.

Neste trabalho são apresentadas duas abordagens distintas à estimativa da procura de sistemas de bicicletas partilhadas: a primeira constitui uma rápida avaliação da procura que é adaptável às características do território, a segunda abordagem usa uma analise de regressão de forma a perceber as variáveis que influenciam o uso do sistema, através do estudo de um sistema já implementado (Sistema de bicicletas partilhadas de Boston).

O design do sistema é definido por um modelo de otimização que define a localização das estações de bicicletas e a sua capacidade, o número de bicicletas necessário em cada estação e, consequentemente, o tamanho da frota considerando um valor inicial de investimento inferior a um orçamento definido. Adicionalmente, este modelo faz o balanço entre o custo anual do sistema e o retorno financeiro do mesmo assumindo ainda um possível investimento regular para cobrir despesas necessárias.

Os impactos ambientais são estimados em função da redução de tráfego resultante da implementação do sistema de bicicletas partilhadas, focado nas partículas finas em suspensão (PM2.5). Os resultados demonstram um impacto não proporcional entre a redução de tráfego e a redução de emissões, uma vez que este impacto depende das características das ruas e das condições de circulação.

Em suma, acredita-se que o trabalho desenvolvido é constitui uma ferramenta importante na implementação e design dos sistemas partilhados de bicicletas. Constitui-se como um ponto de partida no planeamento e decisão deste modo de transporte.

CONTENT

FIGURES

TABLES

1 INTRODUCTION

1.1 Background and Motivation

Mobility is a cornerstone to the territorial and individual development of our society, the technological progress, the improvement of both economic and social conditions over the last years have contributed to the increase of individual motorized transport, changing mobility patterns. But the excessive use of private car has adverse impacts on the environment and on the quality of life.

Most of the developed countries present a high and growing car ownership rate. In the European Union there was an increase from 334 passenger cars per 1000 inhabitants in 1991 to 569 in 2019 (European Commission, n.d.) and, in the United States, 837 vehicles per 1000 inhabitants in 2018 (Statista, 2020). Regarding Portugal case, the Portuguese insurance authority (ASF - Autoridade de Supervisão de Seguros e Fundos de Pensões, 2020) presents 639 passenger cars per 1000 inhabitants in 2019, an addition of 46% to 2009 levels (437 passenger cars per 1000 inhabitants).

The motorization rate all over the world in 2014 is present in Figure 1, considering automobiles, SUVs, trucks, vans, buses, commercial vehicles, and freight motor road vehicles (and excluding motorcycles and other two-wheelers).

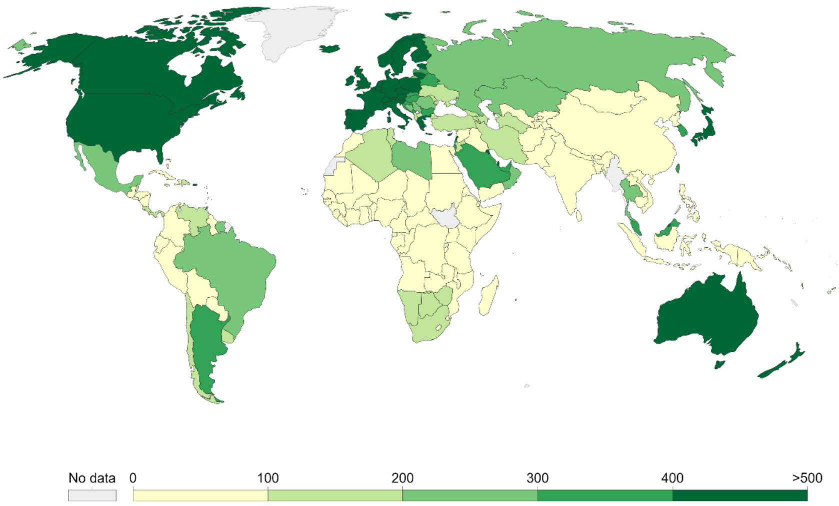


Figure 1 – Road motor vehicles per 1000 inhabitants in 2014

Source (*Our World in Data*, n.d.)

The high rate of car ownership, and its dependence on fossil fuels, turn the transportation sector into one the most pollutant sectors in the total of air pollutants emissions in Europe. It represents 27% of the emissions in Europe. According to European Environment Agency (EEA, 2019), urban motorized transportation is responsible for 40% of carbon dioxide emissions of road transport (and up to 70% of other greenhouse gas emissions).

The following figure shows the contribution of the transport sector to the total of pollutants emissions in Europe.

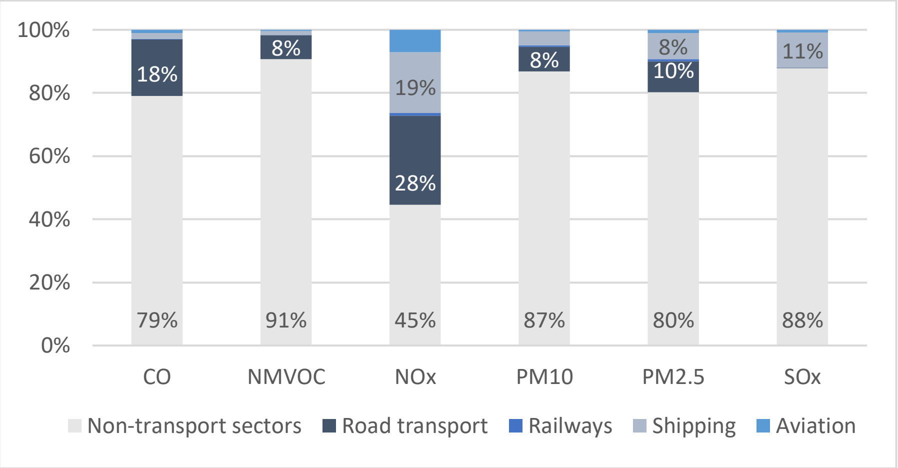


Figure 2 - Contribution of the transport sector to total emissions of the main air pollutants

Adapted from (EEA, 2019)

The exposure to pollutants is associated with the increased cases of cancer, respiratory and cardiovascular diseases (Chow, 2006; Künzli et al., 2000; Lim et al., 2012; Loomis et al., 2013; McGinnis, 1993; Pope & Dockery, 2006; Stephen S Lim‡, Theo Vos, Abraham D Flaxman, Goodarz Danaei et al., 2012).

According to the International Agency for Research on Cancer, human exposure to outdoor pollution is related *to increases in genetic damage that have been shown to be predictive of cancer in humans*, and it can promote cancer progression (IARC Monographs, 2015).

As referred by Künzli, it should be considered *air pollution and traffic-related air pollution as a widespread cause of impaired health*. Thus, the adoption of measures that has a consequence of the reduction of pollutants emissions will improve the life quality on the planet (Künzli et al., 2000).

According to Organisation for Economic Co-operation and Development (OECD/EU) (European Commission, 2010), about 168.000 to 346.000 premature deaths in 2018 in European countries can be related to small Particulate Matter ($PM_{2.5}$). The transportation sector is a relevant source of $PM_{2.5}$ emissions, particularly in urban areas (EEA, 2020a).

The concerns about environmental issues, their impacts on human health, and energy-saving stimulated sustainable policies worldwide.

The Kyoto Protocol (United Nations, 1998), signed on 11 December 1997, was the first step in the international commitment to greenhouse gas (GHG) emissions reduction. The aim of the European Union was a 20%-30% reduction of the emissions below levels of 1990 (base year) until 2020.

In December 2015, the Paris Agreement, a new international treaty, was signed, including green gas emissions reduction objectives. It works on a five-year objectives cycle to face the growing climate challenges until 2035 (United Nations, 2015).

More recently, on June 2021, the European Commission established the Climate Law that aims to turn Europe’s economy and society climate-neutral by 2050, investing in green technology and protecting the natural environment (European Commission, 2021a).

In this context, the main elements of the European strategy for low-emission mobility concerns are the increase of the transport system efficiency, the deployment of low-emission alternative energy for transport, and the transition towards zero-emission vehicles (European Commission, 2016).

The use of alternative power sources in transport, such as electric, led to the implementation of electric cars as substitutes for the internal combustion engine. Although controlling emissions, the new alternatives do not spare urban space and do not avoid congestion.

According to The World Bank Data (The World Bank, n.d.), in 2016, 54% of the world population lived in urban areas, and this proportion is higher in European Union – 76% or in North America – 82%. And the impacts of transportation tend to be severest in urban areas because of congestion and building density (Dias et al., 2018; Tchepel et al., 2012).

The European Commission emphasizes the importance of motorized single transport alternatives to promote sustainable mobility in economic, social, and environmental dimensions. And, at the same time, encouraging active travel (cycling and walking) contributing to healthy habits (Comission of the European Communities, 2007, 2011; European Commission, 2016, 2017).

Within the sustainable alternatives, promoting walking and cycling conditions turned to be of utmost importance, recently reinforced by the need to keep distance while in public spaces due to the Covid19 situation. Adding, and particularly during these times of Covid19, the maintenance of good individual health is fundamental. Active mobility behaviors play a relevant role in this health (Mueller et al., 2015).

On this framework, local authorities have a crucial role by ensuring accessibility and creating high-quality and efficient transport systems alternatives while reducing congestion and the excessive occupation of public space, pollution, and accidents (European Commission, 2017).

One of the best transport alternatives in urban areas that answer this environmental urgency is the bicycle. Bicycles are environmentally friendly, in terms of CO_2 emissions and noise. From the user point of view, they are cheaper, healthier, and for short distances in urban areas they can have some travel time gains and low space occupancy (European Commission, 1999; Qiu & He, 2018).

However, cycling and walking, face some problems related to longer travel distances or ascending slopes, carrying loads, and weather conditions. In any case, these are good alternatives to single motorized mobility in urban areas (Heinen et al., 2010).

The public bicycle service with electric bicycles and coordinated with transportation systems can address some problems associated with cycling. The main goal of bicycle-sharing systems is to provide public vehicles to individuals for traveling in a city. The service allows picking up (and dropping it down) a bicycle at different city points (stations) in coordination with other transport modes.

To plan a bike-sharing system as a mobility alternative for commuting travel is a challenging step to ensure the system's success and to contribute to a mobility revolution. Besides, it needs to address different users' needs.

The bike-sharing theme was not a popular subject on scientific research publications in 2011, the starting year of the present research work. However, worldwide interest in this topic became visible in recent scientific publications on bike-sharing, as reflected on Figure 3.

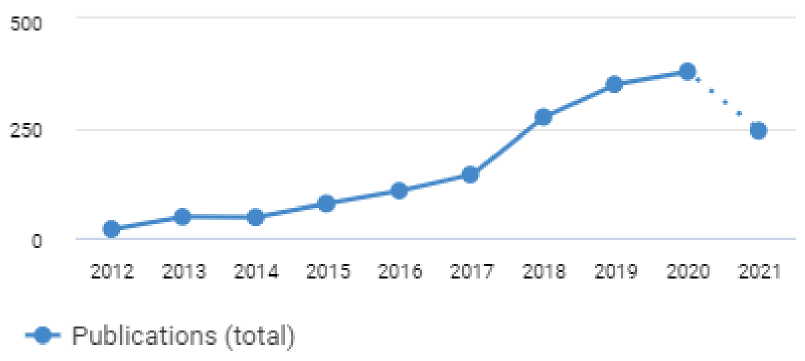


Figure 3 – Number of scientific publications on Bike-sharing

source: app.dimensions.ai (July, 2021)

The following sections describe the objective of this research work, the used methodology, and this thesis structure.

1.2 Objectives

This research focuses on bike-sharing planning and implementation. It intends to develop and provide strategic and practical methods to transportation planners, policymakers, and investors' decision-making.

The main goal of this decision support methodology is to determine the optimal location of bike-sharing stations according to a potential demand to the system, aiming additionally to measure some environmental impacts of the system.

Under this goal, some specific objectives orient this research in a segmented and sequential way:

- To study the demand of bike-sharing systems by understanding the characteristics that influence the systems usage, including user characteristics (as age, gender, or occupation), trip characteristics (as purpose or distance), and land use (work or commercial zones).
- To develop an optimization model that estimates the best location of the stations to maximize the potential demand to the system, considering economic investment constraints.
- To estimate the environmental impact of a bike-sharing system implementation through the consequent traffic reduction.

1.3 Research Questions and Global Methodological Approach

This thesis is related to the design of bike-sharing systems, promoting its implementation and assuming that these systems are one of the strategies to promote sustainable mobility.

A methodological strategy was developed based on assumptions to answer these objectives. As complex as it may seem, these systems are implemented nowadays without the expression of precise criteria, which implies, many times, the misplaced of the stations and the inadequate allocation of resources.

Moreover, in Southern European countries such as Portugal, where bike use is still at low levels, the implementation of bike-sharing systems requires studying its potential demand, dimensioning, and impacts. These studies' results can help to promote the best decisions for bike-sharing systems implementation.

Between the objectives and the methodological approach, some guiding research questions:

- Are bike-sharing systems a keystone within the present goals for sustainable mobility?
- What are the main lessons learned so far?
- What are the procedures to initiate its implementation?
- How can we identify the demand for these systems in countries where cycling mode is still at low levels?
- How can we dimension these systems assuming a certain level of demand and other restrictions?
- How can we convince politicians that these systems contribute to a better urban environment?

So the first methodological step was to study the history and the present characterization of bike-sharing systems across the world (chapter 2).

From this literature review, it was possible to be sure that the following methodological steps would be the most adequate to obtain, at least partially, some of the research questions answers.

The second step was to use some of the available knowledge to estimate the potential cycling demand for bike-sharing systems, showing the potential power of this approach for countries with low levels of cycling use (chapter3).

The city of Coimbra in Portugal is used as a case study. This choice also relates to the fact that, in Coimbra, the great majority of trips are short distances and the orographic characteristics are irregular (which increases the justification for the demand study). In this case, the demand estimation uses mobility behavior aggregate studies and other population and city characteristics.

And what could we learn from existing systems databases? Can those databases teach us something new about the usefulness of these systems' implementations? That is why the Boston system (Hubway/Bluebikes) database study is an example that can help future demand studies (3.4).

Although some information can be drawn from these databases, the usual multiple regression approaches have some limitations due to the databases' heterogeneous nature (associated with city characteristics). There are some limitations in finding the complete profile of users and other city characteristics that most likely relate to the bike-sharing system use.

However, these two demand studies' conclusions are helpful in the discussion of bike-sharing systems' use applicability and efficiency. Moreover, they can provide important clues on how to develop future demand studies for bike-sharing systems.

These demand studies are just a part of the problem. After demand identification, locating and dimensioning the system according to needs and resources is the following step (chapter 4). So, in this step, the development of a bike-sharing optimization model of covered demand maximization that locates stations and dimensions them according to financial resources and can be used independently from the demand estimation process partially solves the problem.

How to use bike-sharing systems to promote sustainable mobility? This promotion must be centered not only on adapting the bike-sharing system to the demand but most certainly to create substantial increases in the supply of bike-sharing.

Moreover, these implementations cannot be done independently of other sustainable mobility policies. In the process of showing to politicians that a bike-sharing system can act as a trigger towards a modal transfer from car to bicycle, it is also fundamental to demonstrate that some traffic restrictions are essential. Therefore, an exercise shows the advantages in emission reduction by transferring some trips from car to the bike-sharing system, using the previous models for demand estimation and stations locations and dimensioning, and adding models for traffic emissions estimation (chapter 5).

Not being enough to establish a precise methodology for bike-sharing systems implementation, this thesis aims to provide precious advice. Moreover, and as in the conclusions and future works chapter (chapter 6), bike-sharing implementations are efficient when the city infrastructures adapt to it (namely by limiting car traffic and its speed) and when bike-sharing connects with other transport modes.

But these are not objectives of this thesis, hoping that future works can fulfill the gap.

1.4 Thesis Structure and scientific production

This thesis is divided into 5 chapters, beyond the introductory chapter, each one included an introduction and synthesis and the chapter body are following described. The scheme presented in Figure 4 summarizes the thesis structure.

Chapter 2 provides a literature review on worldwide bike-sharing systems, the planning policies and guidelines to their implementation, the operation design features, and the factors that influence their success.

The specific literature is presented in each chapter according to the organization of the different themes helping to understand the scope of the correspondent theme.

Chapter 3 presents the study and analysis of demand on bike-sharing systems. The chapter has 3 main parts that correspond to a literature review on bike-sharing demand (section 3.2) and two different approaches on the demand evaluation: the first approach presents a methodology to estimate the demand considering external characteristics that affect bicycle usage (section 3.3) and the second analyzes the bike-sharing system of Boston based on 2014 database (section 3.4).

Section 3.4 includes a subsection of the literature review on a regression analysis that supports the analysis made, presented on 3.4.1.

Chapter 4 presents the location model of stations, the chapter is divided into a literature review on location models (section 4.2) the detailed description of the location model and its application in Coimbra study case.

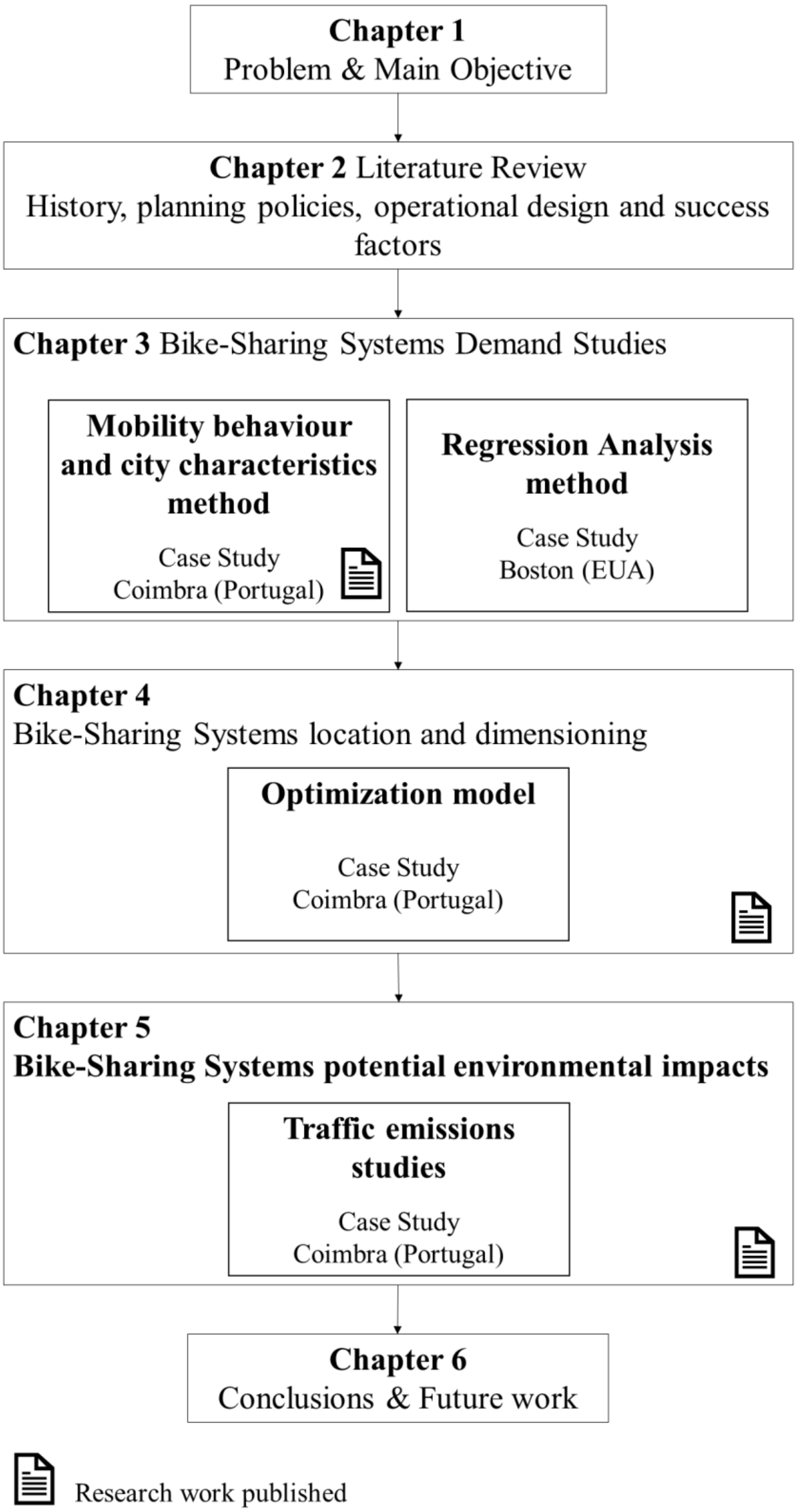


Figure 4 – Thesis work structure

The environmental impacts of bike-sharing usage are evaluated in chapter 5, and the chapter includes a literature review, the methodologic approach, and its application to the same case study on two specific scenarios.

The last chapter (6) is the conclusion of the thesis, including synthesis of thesis objectives, the methodologies, the models used, and the results of this research. It also includes an overview of possible future work in this area.

The work developed since the beginning of the research resulted in the following publications (ordered by latest publication date).

- Inês Frade, Anabela Ribeiro, Daniela Dias and Oxana Tchepel (2021): Bike-sharing systems implementation impact on emissions, for cyclist preferred routes in urban areas, *International Journal of Sustainable Transportation*, DOI: 10.1080/15568318.2021.1949076
- Inês Frade and Anabela Ribeiro (2015): Bike-sharing stations: A maximal covering location approach, *Transp. Res. Part A Policy Pract.*, vol. 82, 2015, doi: 10.1016/j.tra.2015.09.014.
- Inês Frade and Anabela Ribeiro (2014): Bicycle Sharing Systems Demand, *Procedia - Soc. Behav. Sci.*, vol. 111, pp. 518–527, 2014, doi: 10.1016/j.sbspro.2014.01.085.

2 BIKE-SHARING SYSTEMS REVIEW

2.1 Introduction

A bike-sharing system is a public shared used set of bicycles normally available in urban areas, focused on short-term basis trips for free or with a low fee. The service includes picking up and dropping off bicycles at different stations in an urban area.

The first bike-sharing system emerged in Amsterdam, the Netherlands, in 1965. Nowadays, several cities around the world have adopted public bicycle sharing systems as a transport option.

According to the Bike-sharing World Map[1], a total of 2003 bike-share programs are in operation worldwide, 300 being planned or under construction, and 809 were cancelled, suspended or closed. Globally there are almost 10 million bicycles (just in bike-sharing systems) available all over the world.

Figure 5 - The Meddin Bike-sharing World Map.

[1] "The Meddin Bike-sharing World Map." Russell Meddin, Paul DeMaio, Oliver O'Brien, Renata Rabello, Chumin Yu, Jess Seamon, Thiago Benicchio, Deng Han (ITDP) and Jacob Mason (ITDP). Accessed March 2021. http://bike-sharingworldmap.com/.

In this chapter, the general aspects of bicycle sharing are presented. This chapter summarizes the history and the evolution overview of the worldwide state of the art of bike-sharing systems and technological trends of these systems, their framework on the global, European, and Portuguese strategic guidelines, their operational design features, and the main success factors.

2.2 Bike-sharing systems – Historical background

For giving an accurate overview on the history of these systems it is important to distinguish between four generations of services (CHEN et al., 2018; DeMaio, 2009; Moon-miklaucic et al., 2019; S. a. Shaheen et al., 2010; Wang et al., 2010):

- 1st generation: free bike system,
- 2nd generation: coin-deposit systems,
- 3rd generation: information technology-based systems,
- 4th generation: the dock less and big data management possibilities.

The free bike-sharing system comprehends a set of bicycles (with unusual colors or shapes) available without costs to the user. Typically, their stations are located near public facilities. The system includes staff responsible for users' identification, reducing the needs of other human resources. The bicycles are free to be used by any user that needs them.

The first bike-sharing system emerged in Amsterdam, the Netherlands, in 1965. A set of fifty free bicycles, considered as the solution for traffic problems, were made available. However, the Witte Fietsen (white bikes) plan failed after its launch due to bicycle damages and thefts (Figure 6).

The same type of service happened in La Rochelle (France) in 1974. With support from the community, the bike-sharing service turned to be successful. Moreover, the Vélos Jaunes (yellow bikes) system still operates. This bike-sharing system is composed by 54 stations, 300 bicycles available all over the day, and 160 kilometers of cyclable spaces; the usage of bicycles is free in the first half-hour (Midgley, 2009).

In 1993, Cambridge (United Kingdom) implemented a free bike-sharing system called the Green Bike Scheme. However, as in the case of Amsterdam, the majority of the bicycles were stolen, and the program failed.

Figure 6 – Witte Fietsen (white bikes) in Amsterdam (Netherlands), 1965.
Source: (International Institute of Social History, n.d.)

In 1994 the first bike-sharing system in the United States was implemented in Portland (Oregon), called Yellow Bikes, and ended in 2001. In 1995 the Green Bike Program in Boulder (Colorado) operated with 130 bicycles but ended up closing as well. More recently, the city rethought the bike-sharing city system, now called the Boulder B-cycle, with 250 bicycles and 10 stations. This new system works now as a service of the second generation.

In Portugal, a free bike-sharing system in Aveiro, called Bugas, was launched in April 2000. It started with a stock of 350 bicycles spread over 33 parks all over the city. However, after the pilot period, some of the bicycles were vandalized or stolen. Currently, the system works as a less ambitious service with only one station and some bicycles. The difference between the previous type of service and the current coin-deposit systems is that there are some concerns about the location of the stations to ensure the efficiency of the operation, and the bicycles are not freely available.

The first bike-sharing using a coin deposit system launched in 1995 in Copenhagen (Denmark), called the Bycyklen (City Bike) – Figure 7. The users have to use a coin to unlock the bicycle in the docking stations. Currently, the Bycyklen system works from April to September, and there are 110 stations and about 2000 bicycles available.

Similar to the Copenhagen case, other bike-sharing programs emerged in Bycykler in Sandnes (Norway) in 1996, City Bikes in Helsinki (Finland) in 2000, and Bycykel in Arhus (Denmark) in

2005. And also in United States: the Yellow Bike Project in Minneapolis and St. Paul in 1996, and Madison (Wisconsin), Olympia (Washington), Austin (Texas), Princeton (New Jersey), Durham (New Hampshire), and Decatur (Georgia) between 1996 and 2002.

Figure 7 – Bycyklen (City Bike) in Copenhagen (Denmark), 1995.
Source: (Svenningsen, 2009)

Although some significant changes in the motorized transportation patterns in some cities, the coin-deposit system did not solve the theft problem. To overcome this problem, the third generation of bike-sharing emerged based on automatic services.

The automatic system uses 'smart' technology (mobile phones, mag-stripe cards, smartcards, or codes) to unlock the bicycles from the stations allowing the automatic identification of the users (for example, with a code). The casual users pay a security deposit to ensure the bicycle return, and its use is paid depending on the time interval of the usage. Typically this service is free in the first specified time interval, and the price gradually increases over time.

This system is simpler to manage in terms of human resources, but require a higher investment in technology. Some of the advantages of the technology introduction are the possibility of 24h service, the right displacement of stations in the city getting more evident, and the data collection about the usage of the service being easier (ConBici, 2007; S. a. Shaheen et al., 2010).

Most systems offer a choice of subscriptions: short-term subscription (1-day, 3-day or 7-day ticket) or long-term subscription (monthly or annual).

The Vélib' in Paris (France) – Figure 8, implemented in 2008, is one of the most popular (and largest) bike-sharing systems in Europe. It consists of a network with 1400 stations and about 19.000 bicycles (30% of them are electric) available. China is the country with the highest market for bike-sharing. In Hangzhou, in 2008, the local government launched a public bike-sharing system called Hangzhou Public Bicycle. In 2011, 60.600 bicycles were operating with 2.416 fixed stations, every 200 meters, in eight core districts (S. Shaheen & Guzman, 2011).

Figure 8 – Vélib' in Paris (France), 2008.
Source: Wikimedia Commons

The Bixi public bike-sharing system is a service of the third generation, developed by a company called Public Bike System Company (PBSC). This system is now in several cities from the United States of America, Australia, and England. At each station, the user needs to introduce his credit card to unlock the bicycles.

Most of these programs evolved into systems connected with an integrated traffic management system (intelligent transportation technology). It provides real-time information facilitating the adjustments between demand and supply, the need for electric bicycles, vehicle relocations, and the integration of several transport services in the same access subscription (public transportation or car-sharing).

In 2017, in Lisbon, the Gira Bike System – Figure 9 – started with a pilot system on Parque das Nações with 10 stations and 100 bicycles and now is working with a fleet of 600 bicycles and 81 stations. A system extension is planned in the forthcoming months for 1500 bicycles and 160 stations.

Figure 9 – Gira in Lisbon (Portugal), 2020.
Source: (Transportes & Negócios, 2020b)

The 4th generation of bike-sharing emerged with dock less bikes. The user can detect and unlock a bicycle using a mobile app, leaving the vehicle in a public place anywhere in a city or area where the system is working.

Ofo and Mobike – Figure 10, two companies born in Beijing (China) in 2015, are pioneers in dock less systems implementation. The dock less systems provide high flexibility from the user perspective since there is no need for stations with empty places at each moment. However, these characteristics also turn this possibility into a more challenging solution for public space use and vehicle relocation. It is also more demanding in terms of planning and regulations (CHEN et al., 2018; Chen et al., 2020; Moon-miklaucic et al., 2019).

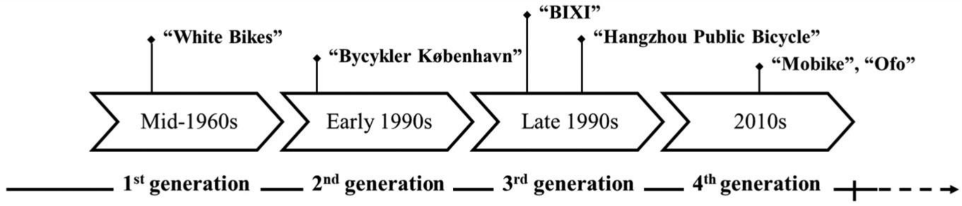
Figure 10 – Mobike in 2020.
Source: (Transportes & Negócios, 2020a)

However, the parking of the bicycles is also a determinant issue. Parking bicycles in unauthorized spaces can endanger other users of public spaces and expose these vehicles' vulnerability to thieves and damages. These are just some of the reasons why dockless systems ended up failing in many cities worldwide, despite of their initial attractiveness due to the apparent less investment (no formal stations).

Besides that the system has some relocation issues when bicycles are left away from the city center, thus the relocation occurs on a larger geographic scale.

In February 2018 two dockless systems were implemented in Lisbon, the oBike (with 350 bikes) and Jump from Uber (with 250 bikes). Both systems worked for just one month. The problems with irregular parking justified the withdrawal of all the bicycles, ordered by the city council (Pincha, 2018).

The Figure 11 summarizes the four generations of bike-sharing according to Chen et al. (Chen et al., 2020).



	Free bikes	Coin-deposit system	Docked IT-based system	Dockless IT-based system
Components	• Bicycles	• Bicycles • Docking stations	• Bicycles • Docking stations • Key/cards connected to payment • Mobile application	• Bicycles • Mobile application
Characteristics	• Free of charge • Users anonymous • Bicycles unlocked • Bicycles located at arbitrary places in a certain region	• Free of charge • Users anonymous • Bicycles locked in specific docking station • Coin as deposit used to unlock bicycles; deposit coin retrieved by returning bicycles to stations	• Usually paid as a membership service, or used temporarily • Users verified with real identity • Bicycles locked in specific docking station • Technology (smartcards, or user interface kiosks) used for picking up and dropping off bicycles in stations	• Pay a deposit for being a member (recently some operators have cancelled the deposit) • User verified with real identity • Charge based on usage time on mobile app or seasonal subscription • Bicycles locked • Bicycles located at arbitrary places • Smart locks, bicycle searching, locking and unlocking via specific mobile app
Cities	1-10	1-10	Over 1000	Over 200
Development paths	Originated in Amsterdam, the Netherlands	Originated in Copenhagen, Denmark, later expended to Europe	Originated in Portsmouth, UK, later expended across the world	First large scale implemented in Beijing, China, later expended to Asia, Europe and the Americas
Main financial resource	Public investment	Public investment	Public investment	Venture capital funding
Collaboration with public transit agency	NA	NA	• Through payment • Clustering the docks around public transport stations	Not yet

Figure 11 - Bike-sharing system generations (Chen et al., 2020)

2.3 Planning Policies – general issues

There is a growing interest in bike-sharing systems worldwide and a consequently growing concern about planning its implementation.

The planning process, driven by the combination of political decision-making and technical analysis, must consider all the stakeholders in the process and their interests, as in any other planning process.

The transport plans develop strategies to shape the supply of the services, answering the expected demand, and, on the other hand, the spatial planning measures seek to influence travel behavior (Schwanen et al., 2004).

The planning criteria used directly depend on different and varying decision-making approaches in each country's government or decision unit.

Plans must support the economic viability of the systems and their efficiency, management and operation, increasing the safety and security of the users, improve the mobility and the accessibility of the people and goods, and improve the environmental conditions (Giuliano & Hanson, 2017).

The planning processes must encompass the following steps (J. de D. Ortúzar & Willumsen, 2001):

- formulation of the problem defining the objectives, standards, and constraints;
- collection of data needed for the characterization of the environment;
- construction of a model representing the data;
- generation of problem solutions;
- testing and identification of the possible solutions and
- implementation of the solution.

Another crucial step is to monitor the performance of the system and identify the strengths and weaknesses of the spatial planning policy. This strategy improves the planning process by identifying new solutions (Schwanen et al., 2004).

Policymakers should be the ones to promote bicycles as a transportation mode (Rietveld, 2004), since the planning policies and the models for the implementation of bike-sharing systems and services need to adapt to different local conditions (Bachand-marleau & Larsen, 2011).

As previously mentioned (section 1.1), the promotion of active transport modes, pedestrians and bicycles, needs to be strategically planned to reduce the environmental impact of transport sectors and improve the public space quality in urban areas.

During this research period, the European Commission published strategic documents showing clearly the importance of bicycles and active modes on sustainability:

- Green Paper (2007): Towards a new culture for urban mobility (Comission of the European Communities, 2007),
- White Paper (2011): Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system (Comission of the European Communities, 2011),
- The European Green Deal (2020) (European Commission, 2020);
- Climate Law (2021) (European Commission, 2021b).

The Green Paper (Comission of the European Communities, 2007) highlights the improvement of public transportation quality, the promotion of clean technologies and energy efficiency, encouragement of clean transportation alternatives such as cycling and pedestrian.

The White Paper (Comission of the European Communities, 2011) reinforces the gradual phasing out of motorized vehicles from the cities – halve the use of fuel motorized cars as urban transport by 2030 – to improve environmental conditions by promoting alternatives such as public transport, walking, and cycling. It also reinforces the fact that by facilitating walking and cycling they become an integral part of urban mobility and infrastructure design. One of its recommendations is about the importance of promoting information about available alternatives to individual transport. These alternatives include walk and cycle, car-sharing, and park & drive, among others.

Most recently, the European Green Deal (European Commission, 2020) focused on mobility transformation onto zero-emissions mobility scenario advice the technological improvement of vehicles, incentivizing the deployment of renewable and low-carbon fuels, and making sustainable alternatives widely available. These improvements will enable better modal choices.

The 3rd flag of this document set up the guidelines on making interurban and urban mobility more sustainable and healthy, putting cities at the forefront of this transition.

One of the main strategies is promoting the Mobility as a Service (MaaS) concept and globally the intermodality between transport modes. Moreover, acknowledging the importance of the shared services on this transformation, such as car sharing, bike-sharing, ride-hailing, and other forms of micro-mobility. Concerning cycling mobility, there is a growing awareness of the need to have safe bike lanes and that they should be double than they are now.

The global concern about soft modes promotion reflects in European Guidelines and Portugal through strategic planning documents.

The National Program of Spatial Planning Policy (*Programa Nacional da Política de Ordenamento do Território – PNPOT*) refers that it is crucial to improve the environmental performance of the transports sector and promote short distances transport alternatives and its coordination as public transport and shared services (as car and bike-sharing) mainly in urban areas (Direção-Geral do Território, 2019).

The Strategic Transport Plan (MOPTC, 2009) outlined guidelines for investments in the Portuguese transport sector between 2008 and 2030, reinforcing land-use policies level of coordination with principles of sustainable mobility minimizing the need for motorized travel, and promoting the use of active modes.

In 2010, Portuguese government created a working group to prepare the National Plan for the Promotion of Cycling and other Soft Modes (‘*Plano Nacional de Promoção da Bicicleta e outros Modos de Transporte Suaves*’) (Instituto da Mobilidade e dos Transportes, 2012).

Its main objective was to study strategies to promote the modal shift to soft modes. This plan includes the promotion of bicycle use, promoting the improvement of public spaces and infrastructures. This articulation requires the update of national regulations of traffic, spatial planning, and street design. This plan includes bike-sharing systems implementation in cities with more than 10.000 inhabitants (action III align f) as municipalities' responsibility. Being a local (municipal) responsibility, the bike-sharing systems are potentially included on local mobility plans: Action plans for sustainable urban mobility (*Planos de Ação de Mobilidade Urbana Sustentável – PAMUS*) – regional areas and Strategic Plans of urban development (*Planos Estratégicos de Desenvolvimento Urbano – PEDU*) – for municipal areas.

And finally, the Regulatory degree no. 131/2019 establishes the National Strategy for Cycle Mobility (Diário da República, 2019a). The plan is organized through six main elements of investment:

- Legal framework
- Research and development
- Infrastructures and intramodality – that includes the promotion of public bike-sharing systems
- Culture behaviors

- Monitoring and evaluation

So far, the design of bike-sharing systems benefits from the guidance of international good practices manuals or handbooks, such as:

- Bike-Sharing Guide (Gris Orange Consultant, 2009),
- OBIS Guide – Optimizing Bike-sharing in European Cities (Büttner et al., 2011),
- The Bike-Share Planning Guide (ITDP - Institute for Transporteation & Development Policy, 2013),

These guides reflect the main concerns on planning, based on shared successful systems features. So far, Portugal does not have an official guide plan to implement Bike-Sharing systems. As referred, it is a municipality's responsibility, although none published so far any criteria for its implementation.

2.4 Operational design

Although the implementation of these systems results from political decisions, the bike-sharing systems is a transport business so far. This means that the decisions on their implementation result from enterprises' profit interests and do not exactly correspond to the potential demand, in most cases.

The planning of bike-sharing lies on the definition of the following system's features:

Table 1 – Features on the Bike-sharing planning

Feature	Details
Business Model	Public/private/public-private partnership
System type	Classic bicycles/electric bicycles/ mix Station-based/dockless local/previous booking Technology & hardware
Pricing	Linear/progressive
Design	Location of stations(Station-based systems)/ Relocation/fleet size

Different assumptions impact the initial investment and maintenance but also the attractivity and potential demand of the system.

The implementation and operation of bike-sharing systems involve high costs. There are two types of costs associated with bicycle sharing services: startup (initial investment) and ongoing (maintenance) costs. The initial costs are associated with the system's implementation. They include planning and project, bicycle acquisition, stations, vehicles, and associated material for bicycle relocation. Including in these costs, there is also the staff to control and maintain the service, and communication and marketing. The ongoing costs are associated with the system maintenance. They include the relocation costs (vehicles and staff to operate it), control and management of the electronic system, equipment maintenance, and the replacement of the stolen and damaged bicycles (ConBici, 2007).

Different bicycle-sharing systems have variant annual costs.

In Table 2, it is possible to see that there are differences while comparing manual or automatic systems, considering the costs of staff, communication, and maintenance, according to Conbici (ConBici, 2007).

Table 2 – Estimated cost of each type of the systems per year (ConBici, 2007)

System type	Estimated Cost
Manual	1300-2400 €/year.bicycle
Automatic	1400-3900 €/year.bicycle

Some specialized services are needed to operate these systems under a business model. There are different types of bike-sharing service providers: public, public-private partnerships, or private companies (S. a. Shaheen et al., 2010), (Midgley, 2011).

Public transport agencies invest in bicycle-sharing services to improve the general public transportation systems of the city. In this case, public funding and revenue depend on fees and memberships. The public entities have autonomous decisions about the system, usually with non-profit motivations. Public transport agencies invest in bicycle-sharing services to improve the general public transportation systems of the city. In this case, public funding and revenue depend on fees and memberships. Many German, Chinese, and British bike-sharing programs are supporting by public agencies. The global perspective is that local governments and public authorities design bike-sharing programs for the well-being of the cities.

In Cambridge (United Kingdom), the bicycles of the 1st generation of bike-sharing (Green Bike Scheme) were abandoned on the streets. The government collected them and restored them, in order to put them again on the streets for public usage.

Even if governments are inexperienced to operate this type of system, there are some advantages with this local business model by the government: greater control over the system and special knowledge about the city transport system coordination (Wang et al., 2010).

Non-profit organizations provide bike-sharing services with the support of public agencies – public-private partnerships. Their revenues derive from membership, usage fees and sponsorships. The public bicycle systems are generally not profitable, as in the BIXI system that works on a non-profit model (Gris Orange Consultant, 2009). The BIXI system works with a combination of user fees, corporate sponsorship, and advertising licenses, and it does not need public funding.

The private sector offers bike-sharing services based on for-profit models with minimal government involvement. However, the private operators will need the locality's support to use public spaces (DeMaio, 2009). Nextbike in Germany works on this type of model.

The implementation of bike-sharing systems can result from an agreement between a private company and public authorities. As the *Vélib'* system in Paris was financed by the JCDecaux Group and as return, it was allowed by the municipality the grant of a substantial part of the out-of-home advertising hoardings. The company has also member system fees.

The business model of bike-sharing systems depends on business opportunities and policy framework. However, it impacts implementation decisions, such as the demand levels, station locations, or infrastructure adequacy. Public systems usually have smaller fleet sizes and a lower number of registered users when compared with other business model systems. However, this type of business models tend to have a better performance in terms of average daily trips per user and bicycle than private companies (S. a. Shaheen et al., 2010; Vassimon, 2016)

The systems tend to evolve to more high-level technological solutions that make the system friendliness to the user and the operation.

The bikes are robust to minimize wear and tear and facilitate maintenance, have a custom design to be easily identified and, the more recent versions, have a GPS to locate the stolen bicycles.

The fleet might use electric bicycles, including the correspondent stations with the necessary electric infrastructure to charge the vehicles.

The stations include docking points to park (and charge) the bicycles, as bicycles lock into the electronically controlled docking point.

The user unlocks the bicycle with a key, a card, a mobile app or a combination of both. In the first case, the station includes a rental terminal on each station. In more recent systems (with the mobile app), all the control and renting is online, and the station provides physical information and wireless to the user.

The system providers have to consider software to do the management of the system and the necessary interfaces.

The software does the system management, including station monitoring, redistribution planning, user registration, customer data management, billing/payment, and all the necessary information.

The pricing schemes include the registration and usage fees, some cases only charge usage fees. The registration fees allow for a short free period (abort 20 or 30 minutes) and growly increase in price after that period.

Most systems ask for a personal credit card number, which is also a guarantee for non-returned bikes.

The prices increase as shorter is the registration period and lower are the usage fees. The following table presents as an example the pricing scheme of two bike-sharing systems: Velib from Paris and GIRA from Lisbon.

Table 3 –Bike-sharing systems examples for different subscriptions

System	Types of subscription		Registration fee	Trip fee and base time	Additional fee by time period	
				0-30 min	30-60 min	+60 min
Velib (Paris)	Max	Classic bike	8,30 €/month	0 €	0 €	1 €/30 min
		Electric bike		0 €	1 €	1 €/30 min
	Plus	Classic bike	3,10 €/month	0 €	1 €	1 €/30 min
		Electric bike		1 €	2 €	2 €/30 min
	Libre	Classic bike	0 €/month	1 €	1 €	1 €/30 min
		Electric bike		2 €	2 €	2 €/30 min
	Séjour	Classic bike	15 €/ week	0 €	1 €	1 €/30 min
		Electric bike		1 €	2 €	2 €/30 min
	Découverte	Classic bike	5€ / 24 hours	0 €	1 €	1 €/30 min
		Electric bike		1 €	2 €	2 €/30 min
Gira (Lisbon)				0-45 min	45-90 min	+90 min
	Annual	Classic bike	25€ / year	0.10 €	1 €	2 €/ 45 min
		Electric bike		0.20 €	1 €	2 €/ 45 min
	Monthly	Classic bike	15€/month	0.10 €	1 €	2 €/ 45 min
		Electric bike		0.20 €	1 €	2 €/ 45 min
	Daily	Classic bike	2€/ 24 hours	0 €	2 €	2 €/ 45 min
		Electric bike		0 €	2 €	2 €/ 45 min

Globally, the pricing schemes are also coordinated with the other public transport modes of the city, promoting their mutual links.

Another common characteristic is that the potential demand of the system is estimated to define a sustainable planning project, namely its economic sustainability, through its design.

Section 3.2. includes a literature review on this theme. For the station-based systems, the location of the stations is a complex problem that influences system use. Different objectives or focus groups can address different approaches and solutions.

Section 4.2 is devoted to the deepening of this subject. Moreover, bike-sharing systems implementation should have a marketing plan. It will allow users attraction and will inform the population of this alternative.

2.5 Success factors

Some external conditions restrict the success of bike-sharing systems and their potential demand: factors related to the system itself and characteristics of the surrounding environment (urban and cultural) that discourage the generalization of non-motorized transport modes.

There are two main types of features. For the user, it is mandatory to feel a user-friendly environment, the permanent availability of vehicles (related to the scheme size, density, and working hours), the type of technology (it depends on the user preferences), and the price.

It is also mandatory to ensure the ideal transport operators, the best contracts, the ownership and maintenance responsibilities, and the financing sources. As a result, the best solutions can also guarantee some employment opportunities (Büttner et al., 2011; Gris Orange Consultant, 2009; Hunt & Abraham, 2007; ITDP - Institute for Transporteation & Development Policy, 2013; Zhang et al., 2015).

The spatial distribution of bicycles is crucial for the acceptance of bike-sharing systems. The user needs to feel that it is easy to find a vehicle on the trip origin and park space at the destination, to make bike-sharing a real transport alternative (Vogel et al., 2011).

One of the most determinant elements in the implementation of these systems is the location of the stations since it can compromise the success of the system.

Along the 24 hours, the demand tends to vary, typically decreasing during the night or in the winter season. Depending on the local the bike-sharing can close at night. These gaps in the night demand-supply might act as a deterrent some users.

Regarding pricing, it is possible to promote short trips with pricing structures that offer the first minutes of usage.

As characteristics of the implementation area that influences the use of bike-sharing (or globally, the bicycle), it is possible to point out the main ones: the lack of vehicles and road infrastructures and the consequent sense of insecurity, the geographic and weather conditions, population density, and other demographic and economic factors (An & Chen, 2006; Baltes, 1996; Dill & Voros, 2007; Duthie et al., 2010; S. L. Handy et al., 2010; Ozarks Transportation Organization, 2005; Schwartz et al., 1999).

The Figure 12 summarizes the main barriers to choose a bicycle as a transport mode.

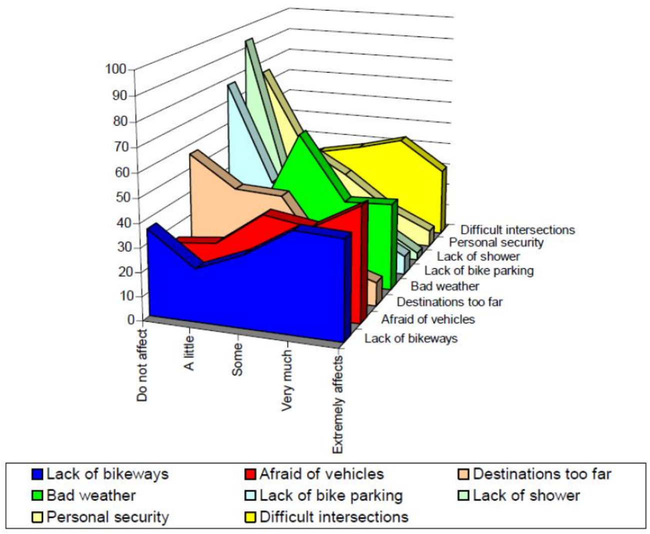


Figure 12 - Barriers to cycling (Ozarks Transportation Organization, 2005)

Bicycle users prefer to minimize the interaction with motorized traffic. Especially the inexperienced cyclists tend to prefer bicycle lanes instead of wide curb lanes. Another crucial characteristic of the bicycle infrastructure is the cycle path contiguity. There is a negative perception of a non-contiguous network that is discouraging cycling (*Bicycling and Transit - a Marriage Unrealized*, 2010; Stinson & Bhat, 2005; Taylor & Mahmassani, 1996).

Therefore, to promote the bicycle as a truly transport alternative, cities must be served by the needed cycling-friendly paths, forming a consolidated, connected, and supported network attractive to cyclists. The network should make public space friendly for riding a bicycle and motivate safe behavior in cyclists and other users.

The design of the best cycling infrastructure for each street will benefit from coordination between the road characteristics (average speed, traffic volume, and function), the network continuity, the appropriate pavement types, the signalization, and the crossings.

In urban areas with low traffic volumes and slow speeds (residential or commercial zones), the best solution for cycling is the creation of shared traffic areas.

For higher speeds and higher traffic volume urban zones, the recommended solutions include segregation from motor traffic for safety and comfort reasons. Between these two types of solutions, several others apply to other combinations of street/road characteristics.

The main characteristics, implementation conditions, good practices and other technical details are described at (Inês Frade et al., 2011; A. Ribeiro et al., 2012; Anabela Ribeiro et al., 2011).

Besides road infrastructure, it is relevant to provide safe parking spaces and facilities as a self-service repair station.

Weather conditions are a permanent constraint, but only extreme conditions (pouring rain or blistering heat) discourage cycling. The seasonal weather variations induce consequent variations on bicycle demand, low demand during the winter in cold cities and the summer in hot cities (Büttner et al., 2011).

However, Northern European countries such as Sweden - where bad weather is frequent - concentrate most urban cycling users.

33% of all journeys in Västerås, Sweden (a cold country), are made by bicycle. In Cambridge, United Kingdom, a wet country, cycling accounts for 27% of journeys (European Commission, 1999).

Steep slopes can make climbing difficult for cyclists, but the way down can lead to fast speeds that might be unsafe for cyclists or other users. The slope issues can be minimized if the right design recommendations are followed, establishing a maximum length according to the slope (S. Handy & Xing, 2011; Parkin et al., 2007; Stinson & Bhat, 2005).

As demographic factors the literature (Baltes, 1996; S. L. Handy et al., 2010) focus on gender and age. The female population and the older population are less willing to use the bicycle. On the other side, the student population in a city/area with universities or colleges tends to use the bike as the main transport mode.

Moreover, car ownership, qualified professional occupations, and higher incomes contribute negatively to bicycle use. And finally, people integrated into a social environment, mainly in the workplace, that support cycling mobility is more likely to use a bicycle to commute (S. Handy & Xing, 2011).

2.6 Synthesis and focus

Providing safe and convenient bike-sharing to citizens is a mandatory element in mobility and transport planning. Furthermore, the implementation of bike-sharing schemes promotes private bike use itself and enhances the image of cycling (DeMaio, 2009; Fishman et al., 2014; Woodcock et al., 2014).

The decision-making process about a new transport system implementation, such as bike-sharing, demands new knowledge about the new system: its operational characteristics, impacts, and success factors.

The new systems are of 3rd and 4th generations type: information in technology-based systems and dockless systems. Even though the dockless system seems to be more flexible, it introduces some challenges in regulatory standards, as referred on section 2.2. This research work focuses on bike-sharing schemes with physical stations suited to Portuguese public spaces law.

Bike-sharing systems are present in solutions for sustainable mobility and are a tool to address climate changes on the main Strategic Guidelines (Global, European and Portuguese strategies). Bike-sharing systems are tools to improve the quality of city life and the urban environment by making better use of urban spaces (European Commission, 1999).

Section 2.4 summarizes the key features of the bike-sharing project divided into four groups: the business model, the system type, the pricing, and the design. But some features might determine its success beyond planning the system, ensuring a potential demand answer, technological attractiveness, and globally good conditions for cycling.

3 DEMAND ANALYSIS ON BIKE-SHARING SYSTEMS

3.1 Introduction

The definition of the demand for bike-sharing systems is of crucial importance for bike-sharing systems design and consequent success. The demand must be estimated with the best accuracy possible in order to dimension the system.

The demand was estimated using two different approaches, both designed to obtain the potential demand to these systems according for the characteristics of the studied zone.

The first methodology (sub-chapter 3.2) – demand analysis for new systems implementation – relates the potential demand of bike-sharing systems with external characteristics that affect bicycle usage. The methodology was conceived during the years of 2012 and 2013 when bike-sharing become more significant as a transport alternative, as a research topic and due to the lack of literature (see Figure 3) and data on this subject driving the conception of a rough methodology that allows a quick assessment to the demand of these systems. This research work was published in the *16th EURO Working Group on Transportation Conference* proceedings (Inês Frade & Ribeiro, 2014).

The second approach (sub-chapter 3.4) – demand analysis on existing systems – presents a methodology based on spatial regression analysis to relate the generation/attraction of bike-sharing trips and the surrounding socio-economic characteristics. The method consists in the evaluation of a trips database of an implemented system and the local characteristics, considering socio-economic data, transportation data, employment information.

In order to provide an overview of the spatial regression analysis techniques and the measures for the goodness-of-fit, this section includes a literature review concerning this topic (not restricted to bike-sharing applications).

The methodology is applied to the Boston' bike-sharing system – the Hubway system.

The chapter is organized into four main sections, besides the introduction. Section 3.2 presents a literature review in bicycle sharing systems demand studies. Section 3.3 is presented the first bike-sharing demand methodology and its application to a study case in Coimbra (Portugal). Section 3.4 presents a literature review on spatial regression analysis, details the spatial regression methodology, and study of the Hubway bike-sharing system that evidenced some important clues on how a bicycle sharing system demand is influenced. Finally, the main conclusions of the bike-sharing demand evaluation studies are presented in section 3.5.

3.2 Literature Review

The demand of bike-sharing systems is a recent topic in the scientific community. When the presented research started (2011-2012) bike-sharing emerged as a trending topic. Bicycle as a transportation mode was, at that time, already seen as a transport alternative to promote in countries with reduced levels of bicycle use, and the scientific work published was scarce. The literature review presented in the following sections focused on the available literature on bike-sharing systems and bicycle demand.

To plan the responses to the traveler's needs and to estimate the transportation demand, and its variation over time is one of the biggest challenges of the urban transportation designers and general urban planners because the transportation system has a close interaction with the land use.

The general approach to urban travel demand is the four step model. The Four Major Stages or Four Step Model has been used to predict the number of trips made within an urban area from the population and urban activities information. It is divided into four phases: trip generation, trip distribution, modal split, and traffic assignment.

In short, the trip generation step estimates the total number of trips generated and attracted to each analysis zone; the trip distribution calculates the number of trips made between each pair origin-destination; the modal split determines the volume of trips that will use each available transportation mode, and the traffic assignment step assigns the trips onto a network (Hensher & Button, 2000; J. Ortúzar, 2000).

The demand for this type of service is not simple to characterize. The most common methods used in demand studies are, usually, made through stated or revealed preferences surveys (ConBici, 2007; dell'Olio et al., 2011; Monzón et al., 2010; J. Ortúzar, 2000; Taylor & Mahmassani, 1996) or it is estimated based on successful systems worldwide (Daddio, 2012; Krykewycz et al., 2010; New York City Department of City Planning (NYCDCP), 2009)

Dell'Olio et al. recommend that the determination of the potential demand of travellers using a stated preference household survey made by telephone, to find the personal profile of a potential user of public bicycles as well as the origin and destination of the journeys, the purpose of trips and the users' willingness to pay for the service (dell'Olio et al., 2011).

Several studies for planning bicycle-sharing services provide surveys in a representative and random population sample in order to characterize the users and potential users of the bike-sharing services (ConBici, 2007; Monzón et al., 2010; J. Ortúzar, 2000).

The Spanish Guide of Cyclist Mobility, *Guía de la Movilidad Ciclista* (Monzón et al., 2010), also presents two models to estimate the demand of potential users of the bicycle through discrete choice models. The first one is based on revealed choices surveys and the second one by stated preference choices surveys. Both models are developed from the surveys in Santander, Spain. According to the results of reveled choices surveys is defined the utility functions of each modal alternative. From the utility functions, it is possible to characterize the existent demand, which is very important to futures plans.

In the case of Santander, the Mozón et al. study concludes for bicycle mode that: the probability of cycling is higher in men than in women (regardless of age and income); when the age is lesser than 56 or income is lesser than 1200 € the possibility of bicycle choice is higher than the other cases; as in the other modes travel time affects the demand but the bicycle is the less influenced by this variable its choice probability decreases 0.58% when travel time increases 1% (Monzón et al., 2010).

From the results it is also calculated the demand elasticity for each one of the criteria: travel time, travel cost, access time, waiting time, and the time to destination.

The second method determines the utility function using the results from stated preference choices surveys. The results show that the most influential variables for the potential users to use bicycle sharing systems are the cost of the service, the weather conditions (rain or cold), and the lack of adequate infrastructure for cycling safely. Persons with less than 24 years, with income less than 1200 € or men are more willing to cycle than persons older than 24, with income higher than 1200 € or women.

It was calculated the elasticity demand for the criteria: travel time and travel cost. It was also related to the influence of travel time depending on the socio-economic characteristics (income, age and gender).

The traveler's decisions depend on the travel time, route, mode, origin, destination, and frequency. The level of demand is directly affected by the travel costs, land use, demographic characteristics of the population, household size, number of workers, income, transport mode ownership, and local culture (Bachand-marleau & Larsen, 2011), as referred in section 2.5.

The revealed preference survey is used to understand the user's choices from the available options. When it is introduced new conditions in the system, as a new transportation mode for instance, it is used the stated preference survey that evaluate people's responses to a not available option and that predicts their behavior.

As referred, the bike-sharing demand is also explained compering with other already implemented systems in other cities or in the same city in case of expansion, based on demographic, socioeconomic, and environment characteristics.

The demand of New York City bike-sharing system was designed using the user group patterns of successful bike-share programs: Velib'in Paris, Velo'v in Lyon and Bicing in Barcelona; from which three typical user groups were identified: commuters, recreational/errand riders and tourists.

The authors estimated the number of people in each potential user category in New York and applied to them different uptake rates (3%, 6% and 9%) to quantify the users of bike-share program. The uptake rates are defined based on London and Paris surveys (New York City Department of City Planning (NYCDCP), 2009).

Krykewycz et al. use a methodology to estimate the demand for a new bicycle-sharing program in Philadelphia – Pennsylvania (Krykewycz et al., 2010).

The authors' defined two market areas using raster based in geographic information system analysis and applied in the calculation of the trip rates for three Bike-sharing Systems, determined through surveys in Lyon, Paris (France) and Barcelona (Spain) in order to estimate the shift from other transport modes to bike-sharing, establishing different demand scenarios (low, middle and high). In the case of Seattle, the study was based in the Philadelphia study.

However, the market areas were defined considering a GIS dataset of weighted sum of indicators, that influenced the use of bike-sharing systems (population density, population density, job density, retail job density, commute trip reduction companies, tourist attractions, parks/recreation areas, topography, regional transit stations, bicycle friendly streets, streets with bicycle lanes and local transit stops). Rates observed in Lyon, Paris and Barcelona, to the defined market areas, were also applied (Gregerson et al., 2010).

Krizek & Stonebraker presented a methodology - developed for Puget Sound Regional Council in Washington in 2002 - that determines the total number of potential users of a bicycle station (in different scenarios) depending of the respective user groups, defined as: bicycle commuters who work within a quarter mile of the bicycle station; bicycle users who park their bicycles at transit stations and bicycle users who travel with their bicycles. The methodology relates the number of the users with the employment data, the number of transit trips, the bicycle use share within three miles of a proposed bicycle station, and the number of bicycle commuters to within a quarter mile of the bicycle station. The validation of this method was done considering the data of two existing bicycle stations and the methodology was considered reasonably accurate (*Bicycling and Transit - a Marriage Unrealized*, 2010).

In other studies, it is used revealed or stated preference surveys as methods for bike-sharing systems demand estimation (ConBici, 2007; dell'Olio et al., 2011; Monzón et al., 2010). In the cases of bike-sharing systems expansion, the revealed surveys can be very useful. However, in some cases the responses to the stated preference surveys can be strategic and may not reflect the real intentions of the interviewee. Surveys results must be used with care, mainly in the cases where similar services were not yet implemented.

In order to avoid the constraints caused by the surveys, it is very important to evaluate different bike-sharing systems, defining the profile of the users and potential users and the factors that can influence the demand (as the geographical conditions, the variation of demand during the day or over the seasons, and other characteristics such as age, sex, and/or job, etc.). Actually, monitoring the performance of a system, in order to identify the strengths and weaknesses of spatial planning policy, is other important step to improve the planning process identifying new solutions strategies (Schwanen et al., 2004).

The regression analysis is used to understand the relation between variables, as the characteristics of the environment and the use of bike-sharing. There are some studies of different bike-sharing systems that follow this approach, as Daddio, Rixey and Maurer & Maurer (Daddio, 2012; Maurer & Maurer, 2011; Rixey, 2012).

Daddio presented a regression approach to explain the station demand based on the demographic, socioeconomic, and built environment characteristics around each existing station measured within 400 meter walk distance from each station, using the data provided by Capital Bikeshare (bike-sharing system of Washington Metropolitan Area) (Daddio, 2012).

As result, the number of trips using bike-sharing system varies with the population between the ages of 20 and 39 (positive effect), the proportion of population that belongs to a race other than "white alone" (negative effect), the number of retail establishments selling alcohol (positive effect), the number of metro stations (positive effect) and the distance from weighted mean (ridership) from the center of full DC and CA Capital Bikeshare system (negative effect).

The trip generation includes information about the age and race of population, car ownership, the average income of the households, hotel rooms and workers who commuted by different modes; the trip attraction group considers the number of attractors (shopping centers, sports complexes, cultural/historic/civic sites, entertainment centers, museums, etc.), retails with alcohol licenses, colleges and parks in the surrounding areas, and transportation network characteristics include number of bus and subway stops, the length of bicycle infrastructure and the distance from system center.

A similar analysis was made by Rixey to three different cases: Capital Bikeshare from Washington (DC), Nice Ride MN from Minneapolis/St. Paul (Minnesota); and Denver B-Cycle in Denver (Colorado) by (Rixey, 2012).

The conclusions of this paper suggest that the station network effect is very important to explain the levels of usage of the system. The demand to the system (average monthly rentals by station) is also dependent on population density, retail job density, median income levels, share of alternative commuters and non-white population. However, the presence of bicycle in infrastructure seems to be less significant to the demand.

Maurer & Maurer developed a regression model that details the demand to the bike-sharing system of Minneapolis and apply it to Sacramento (California). The authors conclude that the demand is influenced by the income, racial composition, job density, high-earnings job density, commute patterns, and proximity to rail stations and parks (Maurer & Maurer, 2011).

As previously mentioned, when this research as initiated, the bike-sharing demand forecast was a topic under development and the approaches found in the literature also considered some research work that addressed strategies to estimate bicycle travel demand. These studies can be adapted into the demand of bicycle sharing systems or, at least, provide some guidelines in the determination of the number of users and their travel needs.

Krizek & Stonebraker presented a methodology, developed for Puget Sound Regional Council in Washington in 2002, that determines the total number of potential users of a bicycle station (in different scenarios) depending of the respective user groups, defined as: bicycle commuters who work within a quarter mile of the bicycle station and bicycle users who park their bicycles at transit stations and bicycle users who travel with their bicycles. The methodology relates the number of the users with employment data, number of transit trips, bicycle commuting mode share within 3 miles of a proposed bicycle station, and number of bicycle commuters to within a quarter mile of the bicycle station. The validation of this method was done considering the data of two existing bicycle stations. Ultimately, the methodology was considered reasonably accurate (*Bicycling and Transit - a Marriage Unrealized*, 2010).

The bicycle journeys to work as function of socio-economic, transport and physical variables is explored at (Parkin et al., 2007), considering an application in England and Wales to estimate the changes in levels of cycling use for 2012 London Olympic Games. It is based in a logistic regression model. As results, the authors had conclude that there are smaller proportions of bicycle in zones with more females and higher car ownership; the weather conditions, slopes and the physical conditions of the highway have impact of use of bicycle to work; the provision of infrastructure for cycle has a positive effect but only on zones with moderate slopes. This model is useful when already exists a bike-sharing service implemented; its application in other cities has to be done very carefully because, as previously mentioned, cycling greatly depends on the local culture (Bachand-marleau & Larsen, 2011).

A simple regression model to estimate cycle trip rates is presented by Barnes & Krizek, allowing to estimate the percentage of cyclists in an area where its value is not known. As results the authors present different bicycle users' rate for different geographic areas (metropolitan areas and over smaller areas such as specific parts of metropolitan areas) as consequence of formal policies and facilities. The model allows an analysis in different locations and geographic scales and provides a quick assessment of the solution to decision makers. Nevertheless, the model is too simplistic when it directly describes a relationship between the provision of bicycling facilities and the amount of bicycling that will take place, ignoring any other explanatory factors (Barnes & Krizek, 2005).

An overview of different approaches to determine the bicycle travel demand is presented by Turner et al. at the most interesting ones are following exposed. The Rhode Island Pre-ISTEA Study assumes that bicycle trips are generated in an area of influence of bicycle facility (based on typical walking distances to transit service - 0.8 km of distance). The generated trips are estimated using trip generation coefficients based on the characteristics of the population and that had been developed in Harrisburg, Pennsylvania; it considers 3 types of purposes: utilitarian (to work, to school and to personal business), recreational with recreational facilities destination and recreational without destination (to visit friends, riding in neighborhood and long distance) (Turner et al., 1997).

Other method has emerged for a Bikes-on-Bus Program from Metro-Dade Transit Agency, implemented in Dade County, Miami – USA, that assumes three different predictors of bicycle demand: location of transportation disadvantaged persons considering that a large number of bicycle trips would be made by low income groups that are neither elderly nor disabled; locations of bicycle commuters using census data, and demographic characteristics since bicycle usage depends of gender, race, and age. These methods have some weaknesses because they do not consider several important influencing factors as the presence of bicycle lanes, the climate and geography local conditions, etc.

The bicycle needs index identifies traffic survey zones with high bicycle use, and therefore, a need for bicycle facilities. It is based in a regression analysis considering as independent factors: the percentage of residents less than sixteen years of age, the number of hours worked per week, the percentage of land devoted to employment uses, population density, employment density, population density of residential land uses, and the ratio of workers to population. However, this regression has a low value of determination coefficient (0.42). This can be explained by the correlation between the explainable variables and there are other important factors that are not considered and that may influence the bicycle mode share (climate, topography, trip length, bicycle facilities, income, etc.).

The Latent Demand Score Method is a gravity model that provides a coefficient of potential demand for bicycle trips throughout a transportation network (in each arc of the network) based on the influence of generator/attractors points in the city on the number of bicycle trips in the surrounding roads segments. The coefficient of potential demand (called LDS) is calculated in function of the purpose of the trip, the number of generators or attractors in the city by trip purpose and their average trip (obtained by Trip Generation handbook), the effect of travel distance in on the number of trips (elasticity). One of the advantages of this model is that it resulted in geographic information system. This method was e developed for city of Decatur (Georgia) and it was already used in Baltimore (MD), Birmingham (AL), Philadelphia (PA), Tallahassee (FL), Tampa (FL), Phoenix (AZ), and Vero Beach, St. Lucie and Alachua (FL) and Westchester, Rockland and Putnam Cos. (NY).

As referred, the following sub-chapters present two different approaches to estimate de potential demand for a bike-sharing system.

3.3 Demand Analysis for new systems implementation

3.3.1 Global Methodology

The presented methodology focuses on the relation between the target public of bike-sharing, trip characteristics, and the physical characteristics of the city paths. As previously mentioned, bicycle usage is mainly mostly affected by the distance of the trip, the slope inclination, the purpose of the trip, and the lack of bicycle paths.

However, in an urban environment, all streets are adaptable for bicycle use, from minor to major improvements. Thus, the following methodology admits that the city where it will be implemented the bike-sharing is planning the streets adaptation to include bicycles with safety and comfort.

Therefore, the main advantage of this methodology is not only the demand quantification (which usually is made by applying a bicycle sharing users proportion to all the city trips – to all Origin-Destination pairs – only considering different purposes) but also modeling it according to the studied area.

The demand definition is studied considering two parts:

- quantifying demand based on other case studies – obtaining the proportion of bike-sharing users per trip purpose and
- defining, sequentially, the effect on demand caused by the trip characteristics (travel time between traffic zones) and physical city characteristics (slopes).

As the final result, it will be obtained an Origin-Destination matrix with bike-sharing proportions to the studied area. The main aspects of this methodology are presented in the next subsections and it will be applied to the case study of Coimbra. The methodology includes purpose, distance, and slopes as factors for bicycle use.

Purpose

The trip purpose influences the probability of using the bicycle (Marleau et al., 2011). For instance, the probability of using a bicycle for leisure trips is greater than for shopping purposes, because it can be difficult carrying shopping bags on a bicycle (Mcneil, 2011; PROBICI team, 2010).

The bike-sharing demand is also affected by the trip purpose, there are three typical user groups: commuters, recreational/errand riders and tourists. Thus, to each one it must be considered different initial rates of bicycle trips (R_n) per purpose (n) based on other study cases.

Distance

For short distances (between 2 and 8 minutes trips) in urban areas, the bicycle can be the most efficient transportation mode. However, while travel distance increases the competitiveness will be negatively affected, and consequently, the potential demand for this transport mode will decrease. In other words, the potential demand is affected by an elasticity which causes a fall in the percentage of bicycle trips when the distance travelled increases.

The elasticity is the ratio between the variation in the proportion of bicycle trips and the rate variation in travel time, between a reference situation and the desired point.

The elasticity varies with trip purpose too. The travel time has a different effect according to the travel purpose. For example, two extra minutes on a work journey travel time can significantly reduce the proportion of bicycle users, whereas in recreational travel it may be irrelevant (Heinen et al., 2011).

Thus, the percentage of bicycle trips for purpose as a function of travel time, R_{tn}, is calculated by equation (1): very short trips (less than 2 minutes trips) are made on foot thus there is no demand for public bicycles, while there is a range in terms of travel time where the demand of bicycle is maximum however there is an instant time from which it decreases being affected by the elasticity.

$$R_{tn}(t_i) = \begin{cases} 0 & t_i < t_{0n} \\ R_n & t_{0n} < t_i \leq t_{1n} \\ R_{tn}(t_{i-1}) + R_{tn}(t_{i-1}) \times E_n \times \dfrac{t_i - t_{1n}}{t_{1n}} & t_i > t_{1n} \end{cases} \tag{1}$$

Where t_{0n} and t_{1n} are the reference instants from each the proportion of bicycle trips starts increasing or decreasing, respectively, and they can vary by trip purpose, t_i is the time travel from each origin to each destination point, R_n is the initial rate of bike-sharing and E_n is the elasticity.

The initial values of E by purpose must be appropriate to each case study. They are strongly dependent on local conditions and personal behaviors to the use of bicycles. These behaviors should be estimated with field surveys or by benchmarking with other examples from around the world. One example is the elasticity values presented in Santander case (Monzón et al., 2010).

Slopes

The bicycle-sharing systems are specially adapted in the case of cities with steep slopes because the cyclists can use the bicycle in one direction and use other transport modes (such as buses) for the opposite direction. In these cases, the main problem of the sponsors of the bikes-sharing system is the relocation of bicycles that must be carefully designed.

The slopes contribute to the ability of a travel route for cycling, according to the American Association of State Highway and Transportation Officials (AASHTO Executive Committee, 1999), since grades greater than 5% are uncomfortable for many cyclists (because the ascents are not easy to climb and the descents induce excessive speeds). However, it may be used in short sections, and as a rule of thumb, the authors suggested the reference values of maximum road length for grades greater than 5% presented in Table 4.

The percentage of bicycle trips per purpose is affected by the differences of slopes between origin and destination mainly in cases of ascents.

In order to incorporate the effect of the slope in this methodology, each traffic zone is characterized by its roads grades, which means that if the zone has a lot of ascendant streets that do not obey the characteristics in Table 4, the demand for bicycle trips tends to decrease in trips with an undesirable destination.

Table 4 – Maximum extension to different slopes, adapted from (AASHTO Executive Committee, 1999)

Grade	Maximum extension
5-6%	240 m
7%	120 m
8%	90 m
9%	60 m
10%	30 m
>11%	15 m

The percentage of bicycle trips for purpose, $R_{sn}(s_i)$, is a function slope characteristic and it is calculated by equation (2).

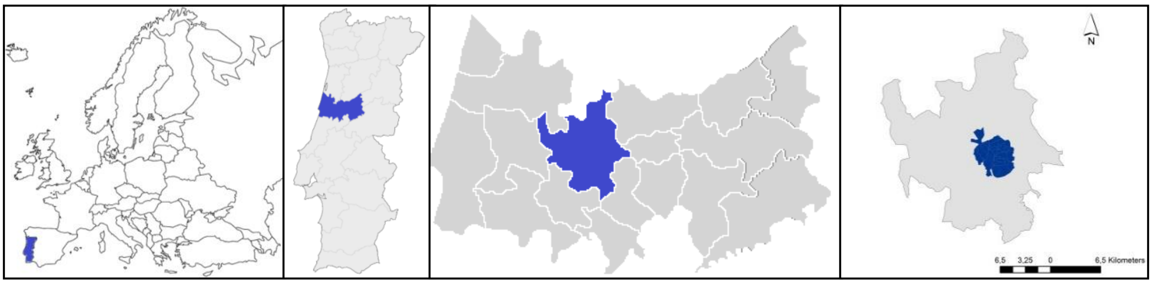
$$R_{sn}(s_i) = f_s \times R_{tn} \qquad \textbf{(2)}$$

Where f_s is a factor defined as a function of the undesirable routes proportion and R_{tn} is the percentage of bicycle trips for purpose as a function of travel time.

The study is applied to Coimbra city (Portugal). The city of Coimbra was chosen due to the availability of data, the Mobility Study of Coimbra had been published at 2009 (TIS.pt, 2009) and because its characteristics. It is a medium-size city with a large student population – the University of Coimbra has approximately 25 000 students. With a mixed geographic pattern – zones with high slopes like the university zone and flat zones like the Solum area and the river's side.

3.3.2 Case Study – Coimbra (Portugal)

Coimbra is located in the center of Portugal (Figure 13) and it had a population of more than 140 000 in 2011 Census.



(a (b (c (d

Figure 13 - Location of: (a) Portugal within Europe, (b) Coimbra district within Portugal, (c) Coimbra municipality within the district, and (d) Study area within Coimbra municipality

Coimbra does not meet a set of good infrastructural conditions to make bicycle as an optional transport mode. The lack of bicycle paths, bicycle supporting facilities and streets in steep slopes suggests that Coimbra is not suitable for cyclists. However, there are a lot of paths with soft slopes that can be easily adapted to meet the needs of a growing cycling population.

According the Coimbra's mobility study (TIS.pt, 2009), 42% of households had by then one car and about 45% had two or more cars, emphasizing the high motorization rate in the city – 522 cars per 1000 inhabitants compared with 473 cars per 1000 inhabitants in Europe in 2009 (Figure 1). Most daily trips are made by car (69%), and the bicycle is of very low importance. Anyway, this study present the bicycle as a forthcoming option, since 57% of trips have less than 4 km.
The municipal urban public transport services include buses, trolleybuses and one elevator. The mobility study say that 18% of daily trips in Coimbra are taken in public transport. The latest strategic plans for the city stress some points with implications for cycling (CIM Região de Coimbra, 2014, 2021; Deloitte, 2009; TRENMO, 2016, 2018).

3.3.3 Results

The mobility study divided the municipality of Coimbra into 61 traffic zones, 29 of which correspond to the city area of Coimbra.

The methodology is applied to this case study considering the origin-destination matrix for the traffic zones of the city of Coimbra, as well as the distances between each traffic zone and its physical characteristics, in a Geographical Information System built for this purpose.

As mentioned, the typical users of the bike-sharing systems are commuters, recreational/errand riders and tourists. In the Coimbra case study, the following rates of bicycle sharing trips per purpose were considered: (R_n) equals 3%, 9%, and 6% for commuters, recreational riders and tourists, respectively. These values were assumed considering a reference study for the New York City study case (New York City Department of City Planning (NYCDCP), 2009).

The influence of the distance between origin and destination is calculated by equation (1), and considering the values of Table 5. The values were based on the Santander (Spain) case – available information by the time.

Table 5 – Admitted values of t_{on}, t_{1n}, R_n and E_n

Purpose	t_{0n}	t_{1n}	R_n	E_n
commuters	1	8	3.0%	-0.08
recreational	1	10	9.0%	-0.01
tourism	1	10	6.0%	-0.01

To understand the slope effect of the Coimbra´s irregular orography, the routes in the road network were classified as suitable and unsuitable for cycling according to the relation between grade and extension of the roads presented in Table 6. The length proportion of suitable roads in each traffic area is presented in Figure 14.

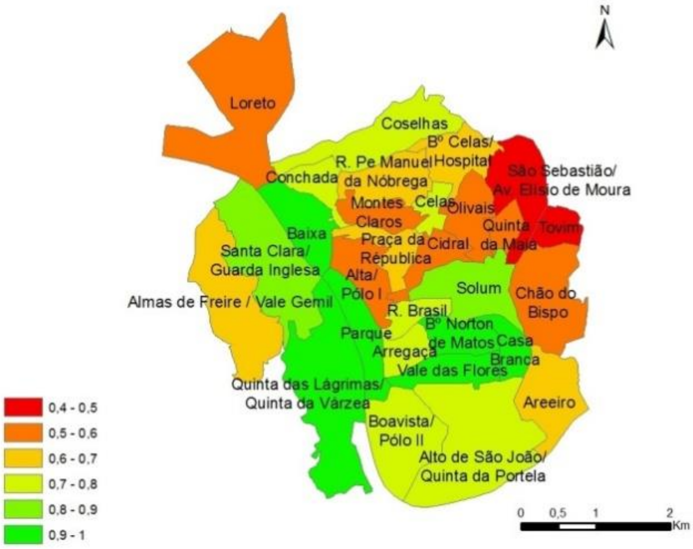


Figure 14 – Traffic zones of the study area classified l by the length proportion of suitable routes

As shown in the previous figure, the orography of the city is very irregular and, consequently, there are zones where is more comfortable to bike (green zones - were near to 100% of road extension respect the relation in) in terms of slopes, on the other hand, there are also zones where the major part of the routes are uncomfortable for cyclists (red zones - where only 40 to 50% meet Table 4).

The factor f_s of equation (2) was determined empirically, it should be adjusted through surveys. And it is presented in Table 6.

Table 6 – f_s values in function of rate of suitable routes.

Rate of suitable routes		fs
90%	100%	1
80%	90%	0.9
70%	80%	0.7
60%	70%	0.5
50%	60%	0.35
40%	50%	0.2

Figure 15 (a) presents the total number of trips attracted and generated in each, by all transport modes in the study area and Figure 15 (b) presents the results of this methodology application: all the trips simultaneously generated an attracted in each traffic zone, all day, by bike-sharing.

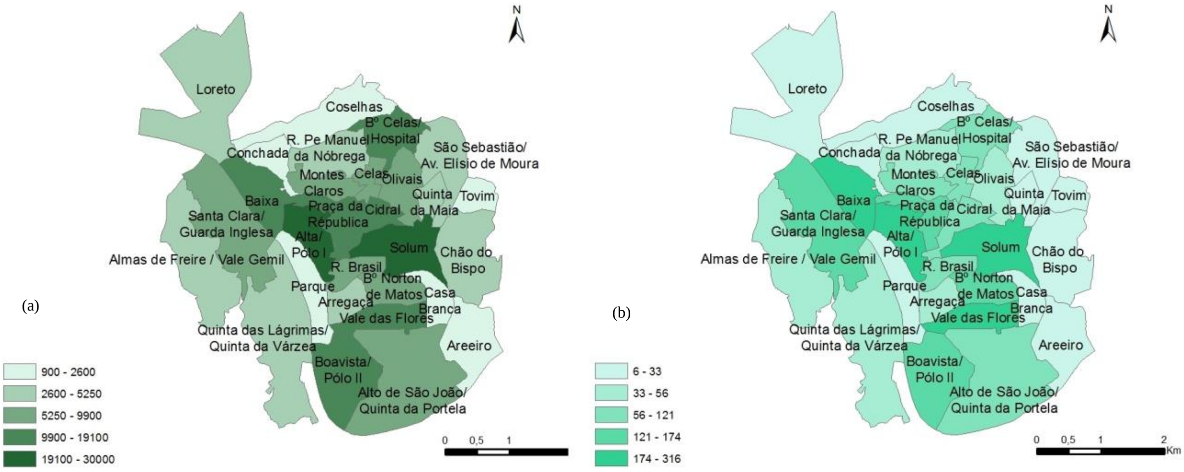


Figure 15 – (a) Total number of trips in traffic zones of study area by all modes; (b) The estimated number of trips in traffic zones of study area using bike-sharing.

From Figure 15 (a) it is possible to identify two traffic zones with a high number of generated and attracted trips (Alta e Solum), where the Alta area includes the Coimbra University and has about 30.000 daily trips.

After the application of the proposed methodology, it was possible to observe some changes in trip patterns: the east traffic zones (as Quinta da Maia or Chão do Bispo zones) lose some importance and the Baixa (downtown) and Vale das Flores area gained relevance, Figure 15 (b).

3.3.4 Synthesis

This section sets out a method for estimating the bike-sharing demand. It can geo-reference the demand, considering the characteristics of the city and of the trips. In terms of application of the methodology developed we can state that the potential bike-sharing demand pattern is different from the motorized trip pattern. This approach was illustrated by an application to the Portuguese city of Coimbra.

The main advantages of this approach are that it provides a quick assessment, and it can be adapted to other towns and cities according to its characteristics. The method can help in decision-making for transportation planners, policymakers, and investors. The method is useful in the full design of the system, including the location of bike-sharing stations and in the dimension of the fleet, as well as in the scheduling of the investments.

This approach uses estimations on potential bike-sharing use when bicycle use is still at low levels - and there is no bike-sharing system already implemented.

In fact, at these low levels, there is no significant data available to extract trends and specific proportions.

Instead, in cities with bike-sharing systems functioning, data availability can promote estimations for further extensions or implementations elsewhere. The following sub-chapter shows an approach considering this data availability.

3.4 Demand Analysis on existing systems

3.4.1 Literature Review on Regression Analysis

A regression model is a statistical technique used to analyze the relationship between variables and it could be defined as a simplified representation of the real world. It relates a single dependent variable (criterion) with one or a set of independent, or explanatory, variables (predictors). The case of more than one independent variable is called ‘multiple linear regression’.

The objective of a multiple linear regression is using the known values of the independent variables to predict the value of the dependent variable. It models the relationship between the variables by fitting an equation to the observed data (linear or nonlinear)

In a multiple linear regression model, the relationship between the dependent variable (Y_j) and the independent variables (X_{ij} $i=1,\ldots,k$) is assumed to be linear and is defined as:

$$Y_j = \beta_0 + \beta_1 X_{1,j} + \ldots + \beta_k X_{kj} + \varepsilon_j \quad j = 1,\ldots,n \qquad \textbf{(3)}$$

In the previous expression:

- the terms βj are called regression coefficients and it represents the extent to which the independent variable influences the dependent variable;

- ε_j is the disturbance (or error) term and represents the random error associated to the regression, it reads as identifying variables omitted from the model, measurement errors in the dependent variable, or random variation in the underlying data-generating process;
- β_0 is called the intercept and it represents the value of Y_j when $X_{ij} = 0, \forall i = 1,...,k$.

The parameters and coefficients of the model can be estimated trough the Ordinary Least Squares (OLS) method. The OLS method minimizes the sum of the residuals given by the difference between the observed values and estimated values of Y. But some assumptions are required:

1. There is a linear relationship between dependent variable and independent variables, as reflected in equation (3);
2. No relationship exists between two or more of the independent variables, meaning that all the independent variables in the model provide sufficiently independent information;
3. The error term:

 3.1. It has an expected mean value equal to zero

$$E(\varepsilon_n) = 0 \tag{4}$$

 3.2. The variance is constant across the observations– homoscedasticity

$$\mathrm{VAR}(\varepsilon_n) = \sigma^2 \tag{5}$$

 3.3. There is independency across observations in the error term elements

$$COV[\varepsilon_n, \varepsilon_j] = 0 \quad i \neq j \tag{6}$$

 3.4. It is uncorrelated with the independent variables

$$COV[\mathrm{X}_i, \varepsilon_j] = 0 \quad \forall_{i,j} \tag{7}$$

 3.5. It is normally distributed

$$\varepsilon_j \approx N(0, \sigma^2) \tag{8}$$

The goodness of fit of the resultant model is measured through a set of parameters: multiple coefficient of determination (R^2) and its adjusted value (*Adjusted* R^2), *F statistic*, and Akaike's information criterion (*AIC*).

The R^2 measures the proportion of the dependent variable explained by multiple regression model. And it is defined by the ratio between regression sum of squares and the total sum of squares, equation (9).

$$R^2 = \frac{\sum\left(\hat{Y}_i - \bar{Y}\right)^2}{\sum\left(Y_i - \bar{Y}\right)^2} = 1 - \frac{\sum \hat{\varepsilon}_i^{\,2}}{\sum\left(Y_i - \bar{Y}\right)^2} \tag{9}$$

The value of R^2 varies between 0 and 1 and the model is better as the closer the value of the R^2 is to 1. When $R^2=1$ the variation of independent variables explained all the variation of the dependent variable, and when $R^2=0$ the model does not explain any of the variation in dependent variable.

However, this coefficient is sensitive to the number of independent variables, increasing with the number of independent variables. To solve this influence, it is usually considered the adjusted R^2, which considers the number of degrees of freedom of the model.

The adjusted R^2 is calculated by the equation (10) where N is the number of records in the data sample and k is the number of independent variables.

$$\bar{R}^2 = 1 - \left(1 - R^2\right)\frac{N-1}{N-k} \tag{10}$$

Adjusted R^2 is a better goodness of fit measure than R^2 in models with numerous variables.

The F statistic tests the hypothesis that no one of the explanatory variables explains the variation of the dependent variable ($H_0\colon \beta_1 = \ldots = \beta_k = 0$). And it is used to test the significance of R^2. If the null hypothesis is accepted than it is expected that $R^2=0$.

In the context of multiple regression models, time and space may create correlated errors by revealing trends over time or over the territory that explain the DV variation. The presence of correlation, that is, error terms from different observations are correlated over time or space, will not affect the unbiased or consistency of the OLS regression estimators but it affects their efficiency.

The presence of multicollinearity (non-compliance of assumption 3.1), heteroscedasticity (non-compliance of assumption 3.2) and autocorrelation (non-compliance of assumption 3.3) suggest time and/or territory disturbances.

In these cases, it is possible to justify the definition of different types of regression:

- Time Series – regression between a set of variables across time in a spatial unit;
- Cross Section – regression between a set of variables in a moment in time, for a set of territorial units or individuals (spatial);
- Panel data – regression that combines spatial data with time data.

The following paragraphs summarize the methods to identify the non-compliance of the Classic Regression Model Assumptions concerning multicollinearity, autocorrelation and heteroscedasticity.

The resources of information for this review were (Anselin, 1988b, 1988a; Gujarati, 1996; LeSage & Pace, 2009; Maroco, 2010; Paradis, 2011; Pindyck & Rudbinfeld, 1981; Anabela Ribeiro, 2008; Wooldridge, 2002).

When the two or more independent variables (or combination of variables) are correlated with each other (non-compliance of assumption 3.1) there is a problem of **multicollinearity**.

If the regression model reveals multicollinearity cannot be estimated precisely: if the multicollinearity is perfect between independent variables (i.e., correlation equal to 1) the regression coefficients of the dependent variables are indeterminate and their standard errors are infinite and, in cases of presence (and not perfect) multicollinearity the regression coefficients possess large standard errors (in relation to the coefficients themselves).

The degree of collinearity can be detected through the pairwise correlation between variables, Variance Inflation Factor (VIF) and the Conditional multicollinearity Number, as explained below.

The pairwise correlation between variables of the database may identify correlation problems, even though there is not a establish boundary to preview collinearity, as a rule of thumb, correlations higher than |0.75| between variables suggests multicollinearity problems.

The Variance Inflation Factor (VIF) is defined as equation (11) where r_j is obtained for the regression of each independent variables on the other independent variables on a regression that does not involve the dependent variable.

$$VIF = \frac{1}{(1 - r_j^2)} \tag{11}$$

As reference, VIF values higher than 7.5 there could be redundancy among variables - presence of collinearity problems.

The Condition Index, defined as the square root of relation between the maximum and the minimum eigenvalue of the IV's correlation matrix, is also used to identify multicollinearity. As reference values, there is moderate to strong multicollinearity if Condition Index is between 10 and 30, and if it exceeds 30 there is severe multicollinearity.

Depending on the nature of the data and the severity of the problem, in the presence of multicollinearity, Gujarati (Gujarati, 1996) proposes different types of actions in order to alleviate it:

- Combining cross-sectional and time series data – as exposed forward;
- Remove one of the collinear variable(s), realizing that this decision may lead specification bias or specification error;
- Transforming variables;
- Adding new data, in some cases increasing the size of the sample (if possible) may attenuate the problem;
- Using extraneous or prior information that allows to relate the coefficients of correlated variables;
- Use non-linear equations;
- Finally, do nothing concluding that the regression model is statically insignificant and its not possible to explain the dependent variables with the set of considered variables.

The **heteroscedasticity** refers to the variance of disturbances not being constant (assumption 3.2) and affects the regression parameters precision, the t and F tests based on ordinary least squares can be highly misleading, resulting in erroneous conclusions about the model.

The detection of heteroscedasticity can be done using a graphical method doing scatterplots of model fitted values in function of residual squared; the identification of a pattern may indicate the presence of heteroscedasticity.

Figure 16 presents typical scatterplots of the estimated mean value of DV and residual squared. Figure 16.a suggests that no heteroscedasticity is present, in contrast with the other graphics, for instance, in the Figure 16.c it is possible to identify a linear relationship and in Figure 16.d and e suggest a quadratic relation.

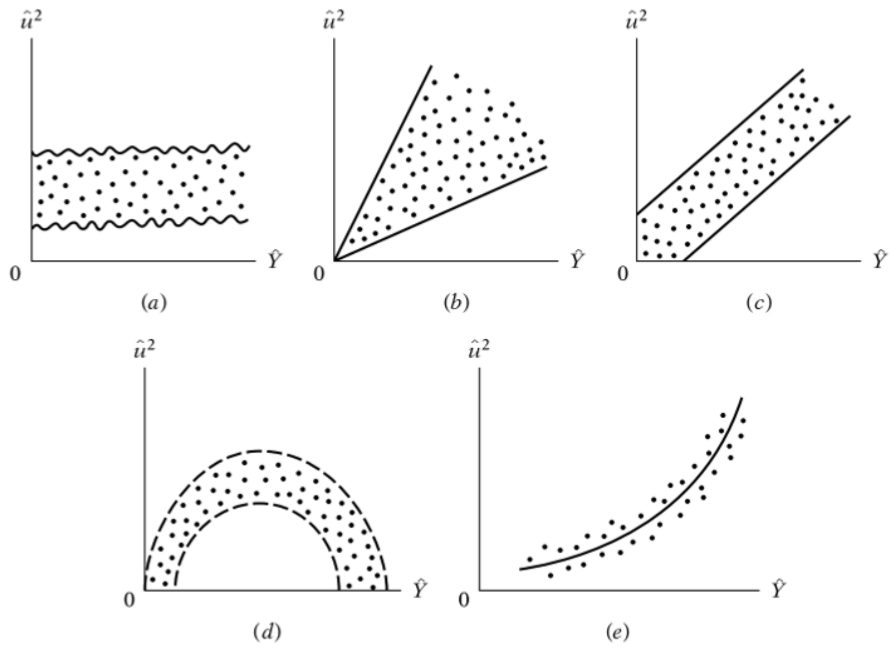


Figure 16 - Possible patterns of model fitted values vs. squared residuals (Gujarati, 1996)

If heteroscedasticity is detected, the plots of the disturbances versus the independent variables or partial variate plots allows to identifying where the problem occurs.

The Breuch–Pagan test is a Lagrange multiplier test is one of the tests used to identify for linear heteroscedasticity. This test considers as null hypothesis the variance of the residuals is constant (homoscedasticity), if the null hypothesis is rejected it is identified the heteroscedasticity.

The White test is a generalization of the Breusch–Pagan test allowing that the independent variable to have a nonlinear error variance for this reason is a more robust test. However, it has some issues when the model has many regressors. As the Breusch–Pagan test, the null hypothesis of White Test evaluates the (null) hypothesis of the homoscedasticity of the residuals against the hypothesis of unrestricted heteroscedasticity of the residuals.

The presence of heteroscedasticity might indicate the existence of spatial autocorrelation, which is confirmed (or not) in testing its presence.

When the variable is related by itself over time or space, there is autocorrelation and the assumption of uncorrelated regressors and disturbances (assumption 3.3) is violated.

The correlation over time (serial correlation) happens when a variable is time dimension, and its observations vary over different periods. In addition, a correlation over space occurs when the value in a space unit is influenced by the value of the variable in the neighborhood.

Autocorrelation measures the similarity, or correlation, between observations over time or over space. The plot of the residuals at time *t* against their value at time *t-1* or the plot of spatial data against its spatially lagged values – *Moran's Scatter plot* can provide useful information about autocorrelation and it identifies the type of correlation between the residuals.

Serial correlation is detected through a time sequence plot or Durbin-Watson test, combined with Breusch–Godfrey test in some situations, in case of spatial autocorrelation it is detected with Moran's I value combined with Lagrange Multiplier Lag and Error tests.

The Durbin-Watson statistic is defined by the ratio of the sum of squared differences in successive residuals to the residual sum of squares – equation (12).

$$d = \frac{\sum_{t=2}^{t=n} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{t=n} \hat{\varepsilon}_t^{\,2}} \qquad \textbf{(12)}$$

The interpretation of Durbin-Watson statistic results depends on two variables defined in function of the number of observations *n* and the number of independent variables – the lower bound d_L and the upper bound d_U (from the Durbin- Watson tables presented in Appendix D of (Gujarati, 1996)). The following figure shows the decisions related with the *d* value.

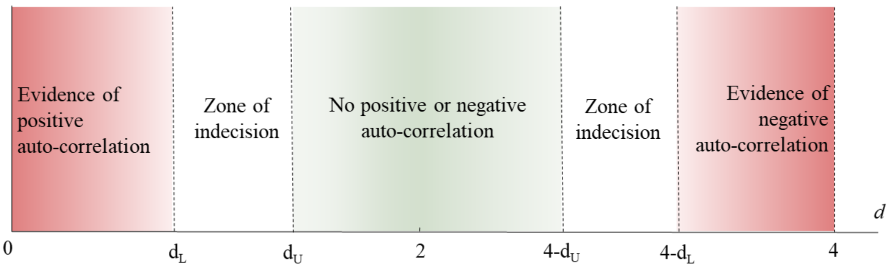


Figure 17 – Durbin-Watson statistic. Adapted from (Gujarati, 1996)

Considering as null hypothesis H_0*: No autocorrelation*, it is accepted if *d* is in the range]d_U; 4-d_U[and rejected if *d* is in]0; d_L[or]d_L; 4[. However, if *d* is in]d_L;d_U[or]4-d_U; 4-d_L[the test gives inclusive results. In fact, as mentioned by Pindyck & Rudbinfeld (Pindyck & Rudbinfeld, 1981), it is possible that the correlation of the errors is due to the autocorrelation of the independent variables and not to the serial correlation of the error terms.

In these cases, it is proposed the Breusch–Godfrey test, also known as Lagrange Multiplier test, in order to test serial correlation in these cases.

The null hypothesis of this test considers that the autocorrelation coefficients are equal to zero. The null hypothesis is H_0*: No autocorrelation*, against the possibility of the presence of autocorrelation (H_1).

Moran's scatter plot reveals the relation between the observations in the vector *y* and the average values of neighboring w_y.

It is necessary to define the neighborhood, it is logically expected that the closest observations are more likely to be similar than the distant ones. In order to quantify this relation, it is associated a weight to each territorial unit, in this context it is usual to define two types of matrices: rook and queen. Rook weight matrix uses only common borders to define neighbors and queen matrix includes all common points (boundaries and vertices) in the definition.

Figure 18 presents a random example of Moran's Scatter plot that articulates the values of a variable with spatial distribution (x ax) with the spatial lags of those values. To each value y in spatial unit i, there is a weighted average of the values of the variable y in the neighbors of i. The weights are the values of the matrix that reflect the neighborhood relations.

The different quadrants reflect different types of relations:

- the first sector (I) presents a positive relation, where high values of the variable are surrounded by high levels of the average of neighborhood;
- the second sector (II) presents negative relation, where low values of the variable are surrounded by high levels of the average of neighborhood;
- the third sector (III) presents a positive relation, where low values of the variable are surrounded by low levels of the average of neighborhood;
- the fourth sector (IV) presents negative relation, where high values of the variable are surrounded by high levels of the average of neighborhood.

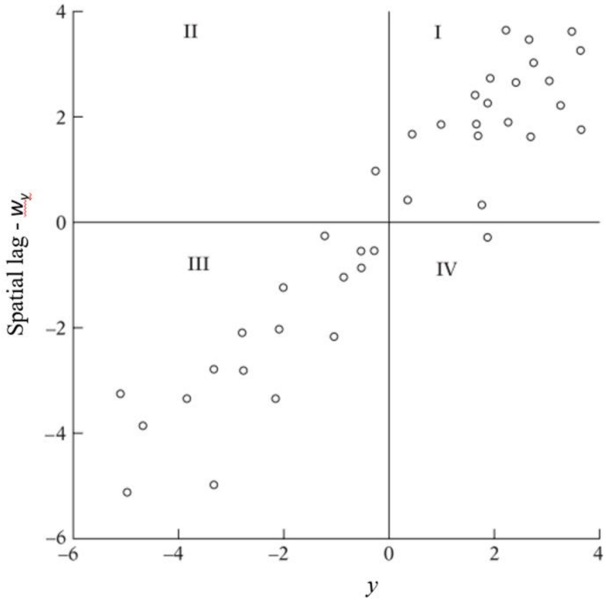


Figure 18 –Moran's Scatter plot example (Gujarati, 1996)

The slope of the trend line that adjusts the points of Moran's Scatter plot define the Moran's I coefficient, that is indicator of autocorrelation. The Moran's I value is determined through the expression (13), where n is the number of territorial units at the database, *wij* is the weight between observation y_i and y_j, and $\bar{y}$ is the mean of the variable.

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \textbf{(13)}$$

The value of *I* under the hypothesis of no autocorrelation is given by (14), Therefore, *I* greater than *I0* is positive autocorrelation, and lower than I_0 is negative autocorrelation (Paradis, 2011).

$$I_0 = \frac{-1}{n-1} \quad \textbf{(14)}$$

The presence of heteroscedasticity and autocorrelation weakens the OLS' efficiency in the estimation of regressions coefficients. In these cases, it is used the Generalized Least Squares – GLS to estimate the coefficients of the regression model.

GLS method transforms the regression variables in order to satisfy the standard least squares assumptions. The GLS estimator considers that each observation is weighted by a factor proportional to the error variance of the regression model.

As referred by Verbeek, the transformed model does not contain an intercept term because all the variables, including the intercept term are transformed. Thus, *the transformed regression is only employed to easily determine the GLS estimator and not necessarily has an interpretation of itself. That is, the parameter estimates are to be interpreted in the context of the original untransformed model* (Verbeek, 2008).

The dependent and independent variables of equation (3) are divided by the squared root of the error term variance – σ, transforming the equation (3) into the equation below where $\beta^{*}j$ are equivalent to the regression coefficients.

$$Y_j^* = \beta^*{}_1 X^*{}_{1,j} + \ldots + \beta^*{}_k X^*{}_{k,j} + \varepsilon_j^* \quad j = 1,\ldots,n \tag{15}$$

In GLS it is minimized the weighted sum of residual squares with the weight equal to $1/\sigma^2$.

The OLS estimator assumes that the explanatory variables are uncorrelated with error term, in other words, the error term has a random effect in the explanatory variables. GLS estimator, on the other hand, does not assume that the individual effects are uncorrelated with the regressors.

The Hausman Test is used to evaluate if the fixed effects and random effects estimator are significantly different. The null hypothesis of this test considers that the explanatory variables are uncorrelated with error term (Assumption 3.4 of OLS estimator). Thus, the rejection of the null hypothesis suggests the possible inconsistency of the random effects model and the possible preference for a fixed-effects specification.

This test evaluates the presence of omitted variables correlate with variable included in the model (LeSage & Pace, 2009).

The accomplishment of 3.5 assumption (the error term is normally distributed) is tested trough the Jarque-Bera test (as alternative tests is mentioned Anderson-Darling Test, Shapiro-Wilk Test, Ryan-Joiner Test or Kolmogorov-Smirnov Test).

Under the null hypothesis of Jarque-Bera test the residuals are normally distributed (*H_0: the residuals are normally distributed*), thus the assumption 3.5 is respected when it is not rejected the null hypothesis (p value of the JB test is high).

The following table summarizes of previous information.

Table 7 – OLS assumptions tests

OLS assumptions	Equation	Test	Rejection of the assumption indicates the presence of
There is a linear relationship between DV and IVs	$Y_j = \alpha + \beta X_j + \varepsilon_j \quad j = 1,\ldots,n$	-	-
The IV's are nonstochastic variables	-	Pairwise correlation Variance Inflation Factor (VIF) Condition Index Multicollinearity Condition Number	Multicollinearity
The mean value of the error term is zero	$E[\varepsilon_n] = 0$	-	-
The variance of error term is constant across the observations (homoscedasticity assumption)	$VAR[\varepsilon_n] = \sigma^2$	Graphical method Breusch–Pagan test White Test	Heteroscedasticity
The error term is independent across observations	$COV[\varepsilon_i, \varepsilon_j] = 0 \quad i \neq j$	Durbin-Watson test Breusch-Godfrey test - Lagrange multiplier test Moran's I Lagrange multiplier lag Lagrange multiplier error	Serial correlation (autocorrelation)
The error term is uncorrelated with the IV	$COV[X_i, \varepsilon_n] = 0 \quad \forall_{i,j}$	Hausman Test	Heterogeneity
The error term is normally distributed	$\varepsilon_j \approx N(0, \sigma^2)$	Jarque-Bera test	-

The analysis of Moran's I indicates the presence of autocorrelation. The Lagrange Multipliers statistics identify the type of autocorrelation, divided into two different tests:

- Lagrange Multiplier Lag (LML) to the model of spatial lag
- Lagrange Multiplier Error (LME) to the model error territorial

The null hypothesis evaluates the lack of autocorrelation at the dependent variable or at the error term, respectively.

There are robust versions of the previous statistics that tests the lag dependency in presence of missing error – Lagrange Multiplier Lag Robust (LMLR), and the tests for error dependence in presence of missing lag – Lagrange Multiplier Error Robust (LMER).

3.4.2 Methodology

Data availability from working bike-sharing systems allow the identification of the variables that influence the demand and how they influence it. The following methodology presented sets out a guideline to achieve this knowledge.

The main objective of this analysis is to identify the variable that can explain the number of generated and attracted bike-sharing trips in the studied system. As referred on the 3.2 section, physical, social, and economic framework influences the willingness to choose bike-sharing or other transportation modes.

Consequently, the identification of these variables, as well as their influence on the number of bike-sharing trips, can potentiate the forecast of the bike-sharing demand.

A spatial regression analysis is performed because the social and economic behavior has a territorial variability caused by the relation complexity between the variables all over the territory.

The Boston bike-sharing system, with an extensive database available, was chosen for this approach (the case study is presented on section 3.4.3).

This method can be used in different datasets, therefore enabling the utilization of this methodology for other study cases.

The methodology analysis compromises the steps presented on Figure 19. This process used the programs ArcGis® and GeoDa®.

It starts with the database construction, which must include all the essential information to the proper characterization of trips, such as:

- bike-sharing system – trips by origin and destination and date-time, location of the stations, users characteristics;
- socio-economic information – population by age, gender, education, job type, income, number of vehicles available, location of jobs and number of works;
- Mobility patterns – mode choice, time to leave home to go to work/school, and

- Transport system – public transport available, demand to each transport, location of facilities (bus stations, subway stations, car parking).

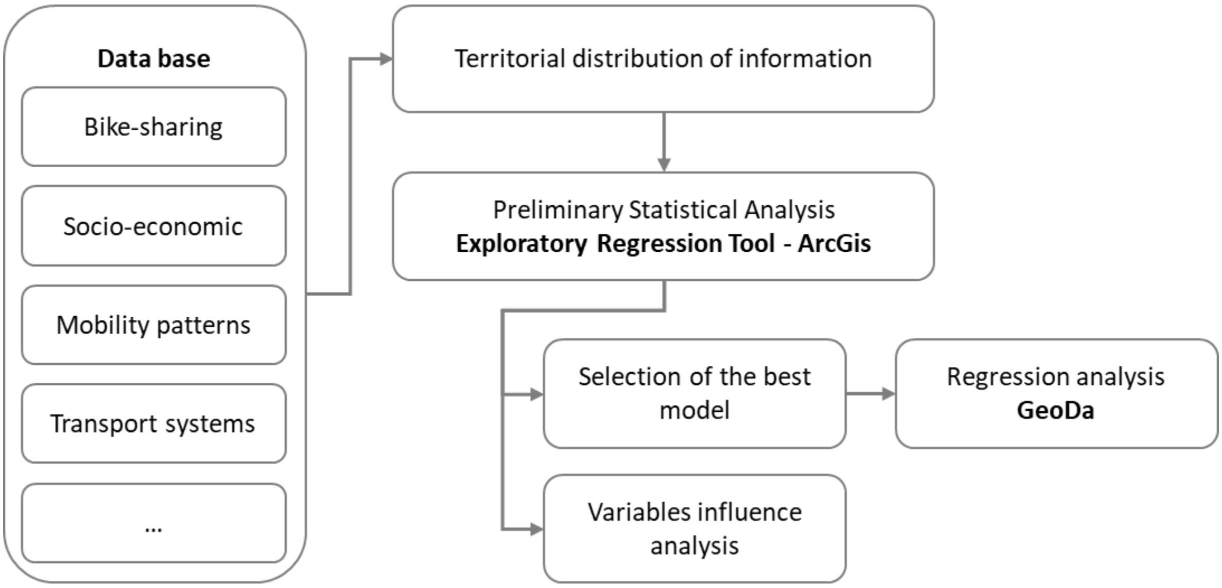


Figure 19 - Flowchart of regression analysis for demand estimation

In order to avoid bias in the distribution of the residuals, all of the variables are linearized before the construction of models.

The territorial distribution set is necessary to territorial uniformize the information. It can be tested different spatial approaches to find the most suitable to this problem analysis.

On the case study and within the GIS environment, different territorial distributions were tested: stations (points), Census areas, and grid division.

In order to find the best OLS model, given the available set of variables, it was performed the *Explanatory Regression* (ESRI, 2018a) tool from ArcGis ®.

The objective of the tool is to provide an idea of how the exploratory variables may explain the dependent variable – data mining of available information. It evaluates all combinations of possible explanatory variables that may explain the dependent variable.

The output of the exploratory regression tool is a report divided into three main sections: the global summary, the highest adjusted R-squared results, and the summary of variable significance (ESRI, 2020).

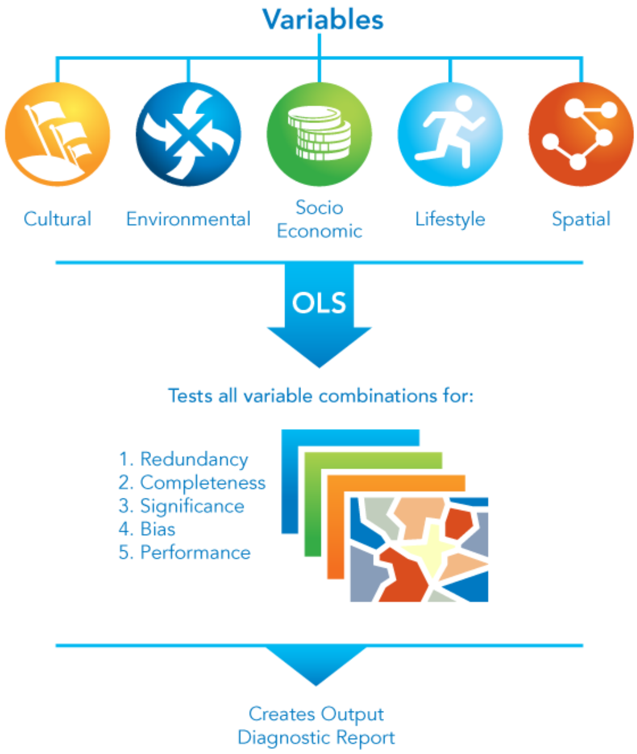


Figure 20 – Explanatory Regression tool of ArcGis® 10.3

Source: (ESRI, 2018a)

The Global Summary table lists the five diagnostic tests and the percentage of models that passed those tests assuming a cutoff (as a reference value) for each one.

Figure 21 presents an example of a summary table where it is possible to identify the diagnostic tests (adjusted R^2, explanatory variable coefficient p-value, VIF, Jarques Bera p-value, and the Spatial Autocorrelation p-value[2]) and the reference values of the cutoffs, the number of models tested (Trials) the number of models that passed the cutoff (#passed) and their corresponding proportion (% passed).

The models tested for the Spatial autocorrelation has to pass all the previous criteria, thus the number of trials is (normally) lower than the number of trials in the others tests.

[2] The information about these diagnostic tests are presented on Spatial Regression analysis in section 3.4.1

```
************** Exploratory Regression Global Summary (O_TTD_LN) **************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff  Trials # Passed % Passed
             Min Adjusted R-Squared > 0,50 4160123   525631    12,63
            Max Coefficient p-value < 0,05 4160123   195033     4,69
                      Max VIF Value < 7,50 4160123  1822730    43,81
            Min Jarque-Bera p-value > 0,10 4160123   216261     5,20
Min Spatial Autocorrelation p-value > 0,10      18        0     0,00

------------------------------------------------------------------------------
```

Figure 21 – Explanatory Regression summary table example

The highest adjusted R-squared results section presents three best models (with the best performance on adjusted R^2 and VIF) for each possible combination number of exploratory variables, and, for each one, the results of statistic tests performed - adjusted R^2, AIC, Jarque-Bera, Breusch-Pagan, VIF and the Global Moran's I[3].

The summary of variable significance presents the behavior of each independent variable of the models tested - *Variables influence analysis.* In other words, it presents the significance of the variables with the correspondent proportion of models where it was considered significant and the proportion of models where they have a positive or negative influence on the dependent variable.

It provides the proportion of times that each variable is statistically significant and how it relates with de dependent variable (proportion of models with positive or negative influence). The variables that are consistently significant in the tested models are strong predictors of the dependent variable.

The report also presents a summary of variable significance that provides information about the consistency of the relationship between each independent variable with dependent variable - strong predictors will be consistently significant.

Additionally, the report includes the summary of multicollinearity with information about the relations between independent variables, the summary of residual normality, and the summary of residual spatial autocorrelation.

[3] The default spatial weights matrix file used to run the Global Moran's I is based on a weighted nearest neighbor conceptualization of spatial relationships.

The model identified by the Exploratory Regression tool at the highest coefficient of determination (R2) is detailed analyzed with GeoDa® within the OLS framework and to identify possible spatial trends using Maximum Likelihood Approaches

As referred, the analysis is performed considering three different spatial approaches: point data, census blocks, and a squared grid. It considers as dependent variables the total number of generated and attracted trips of each station or zone. Therefore, it is performed six analyses to the Boston' case study using this methodology :

- Generated trips through a points (stations) distribution;
- Attracted trips through a points (stations) distribution;
- Generated trips through a census areas distribution;
- Attracted trips through a census areas distribution;
- Generated trips through a grid distribution;
- Attracted trips through a grid distribution.

The territorial distribution selected entail necessary considerations in terms of territorial distribution of data that are following described.

- Stations distribution

  The characteristics of the surrounding zones are picked from a buffer of 400 linear meters of radius from each station.

  It is considered that the data distributed into areas (census blocks) is uniformly distributed and the contribution to the buffer is proportional to the area that intersect the buffer zone.

- Census areas distribution

  It aggregates the information attributed to points into areas, for instance the total number of trips an area is equal to the sum of trips from the stations within the area

- Grid areas distribution

  The grid territorial distribution was defined through the *Optimized Hot Spot Analysis* (ESRI, 2018b) tool from ArcGis®. That creates a fishnet map of statistically significant spatial clusters of high values (hot and cold spots) using the Getis-Ord Gi*[4] statistic. It evaluates the characteristics of the input feature class to produce optimal results.

The information is aggregated into the grid's zones adding the contributions of cells that results from the crossing of zones and grid shapefiles, and the information allocated into points is merged into areas.

3.4.3 Case Study – HUBWAY/Bluebikes System (and area)

Bluebikes is the name of the bike-sharing system that operates in four cities of the Great Boston metropolitan area: Boston, Brookline, Cambridge, and Somerville, in the State of Massachusetts – United States, see Figure 22. The system was called Hubway from 2011 until 2018.

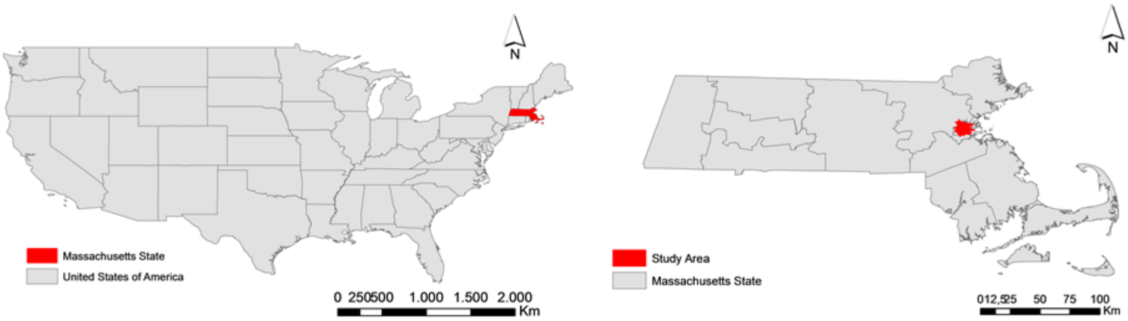


Figure 22 - Location of Massachusetts State in United States of America and location of the study area in Massachusetts State.

According to the United States Census Bureau these cities had 903.594 habitants in 2014. The study area is almost flat, it includes important universities and a big population of students, according to (Baltes, 1996; S. L. Handy et al., 2010) are great conditions to promote bicycle as a transportation mode. These characteristics can explain part of Hubway success.

The bike-sharing system was launched in July 2011 in the city of Boston with 610 bicycles, 60 stations and 3203 annual members (142289 trips). Currently it is operating with more than 3500

[4] Getis–Ord (Gi*) is a spatial statistic to identify hot spots. More information about this method at (Songchitruksa & Zeng, 2010)

bikes at over 325 stations through the Boston, Brookline, Cambridge, and Somerville (2020 data). According to Bluebikes' website page, on September, 2020 Bluebikes passed 12 million total rides.

The case study focusses on a database available online available under the *Hubway Data Visualization Challenge*, held in 2014. The database includes the records of every trips taken form July 28th 2011 until 1st December 2013, when the system was called Hubway. For that reason the bike-sharing systems is hereinafter referred as Hubway.

The database contains the origin and destination stations, date and hour as well as some information about the user. During this period, it was recorded 1.579.025 trips, 30 daily trips in average per station considering their operation days:

- The number of trips increases all over the years, as well as the number of stations, the following graphics presents the total number of trips per month (Figure 23) and the average of trips per operated station (Figure 24);

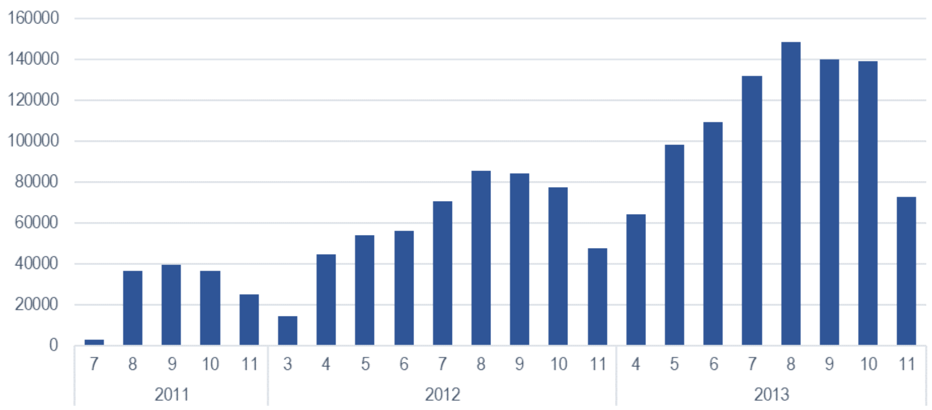


Figure 23 – Trips per month (from July 28th 2011 until 1st December 2013)

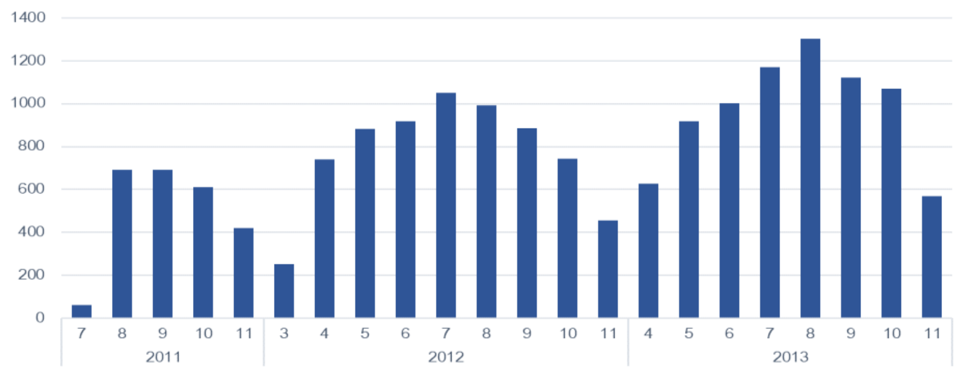


Figure 24 – Average BS trips per open station

- 70% of the trips are made by registered users and 30% by casual users;
- From the trips of registered users 75% are male and 25% female users;
- The trip takes in average 11 minutes on the system and the average daily distribution is presented in Figure 25, the yellow zones identify the peak hours (from 7:30am to 9:30 am and from 4:30 pm to 6:30 pm).

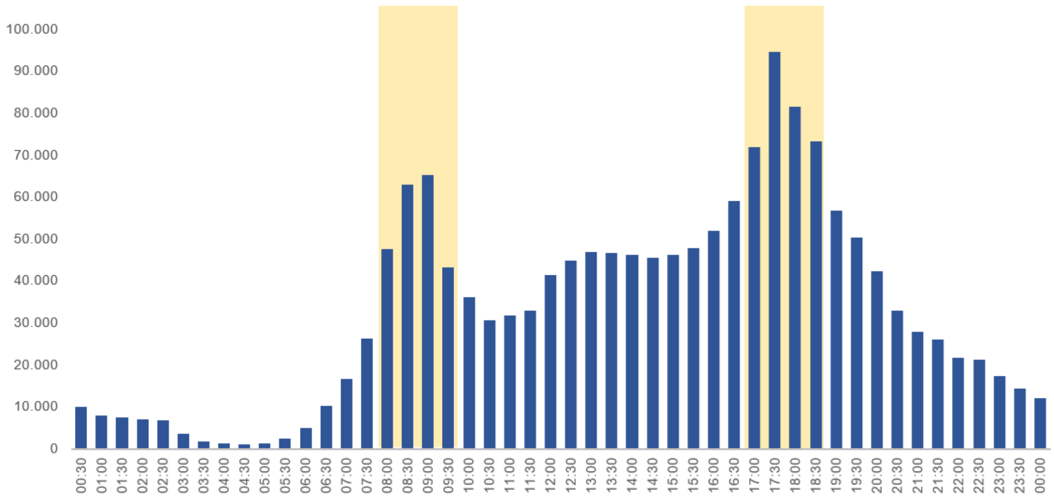


Figure 25 – Hour distribution of the trips recorded – beginning of the trips

- Until September 30th 2012 it was registered the age of the users, from the recorded data the age profile is presented in the Figure 26

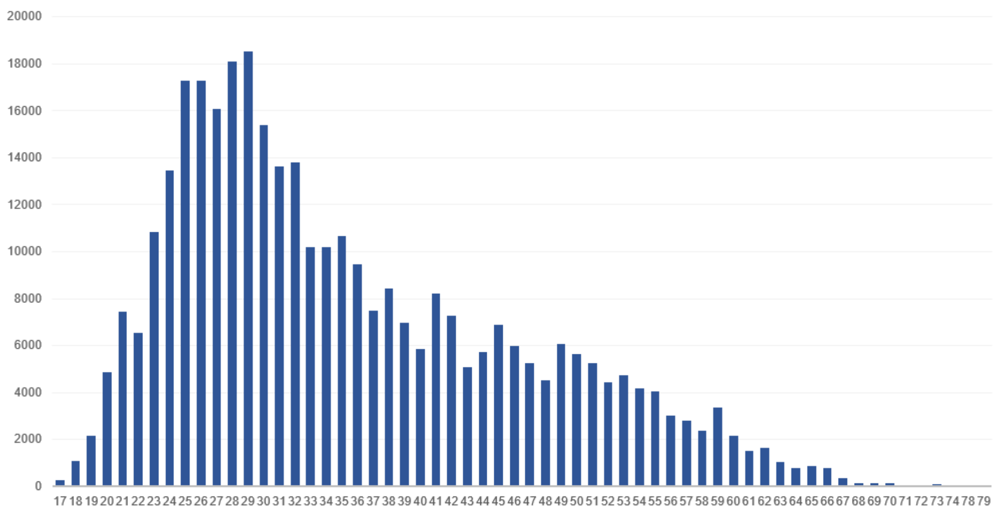


Figure 26 – Age profile of the trips recorded until September 30th 2012

The present analysis pretends to explain the trips generated and attracted to the bike-sharing system thought some variables. The available database includes:

- socio-economic information (population, medium household income, enterprises and employees, the travel time to work and the proportion of people that commutes by public transport or by bicycle),
- public transport data (bus stops and subways station and number of entrances on the subway),
- bike-sharing system operational information (such as stations and docks, the number of trips started and ended and the time and travel time associated at each trip).

The analysis is made considering the year of 2013, when a total of 904.675 trips were registered between April 2nd and November, 30th – when the system was operating. And the census data from U.S. Census Bureau (2009-2013).

The available database has the georeferenced information distributed in two ways: areal data and point patterns. The data is adjusted in order to define the regression model, as explained below.

To avoid scale problems in the interpretation of the results, the parameters are linearized using logarithmic (natural logarithm) transformations of the variables.

The Table 8 present and describe each variable of the database as well as the territorial distribution of the variable.

Table 8 – Database variables

Variable	Definition	Territorial distribution
O_TTRIP201	Total number of trips in 2013 – Origin at each station	Point
D_TTRIP201	Total number of trips in 2013 - Destination at each station	
OPDays2013	Number of operated days in 2013 at each station	
O_TTDay	**Average daily number of trips per operated day in 2013 – Origin at each station**	
D_TTDay	**Average daily number of trips per operated day in 2013 - Destination at each station**	
NEmployers	Employers	
Nworkers	Number of workers at each employer	

Variable	Definition	Territorial distribution
BusStops	Bus stops	
MBTAs	Massachusetts Bay Transportation Authority (MBTA) Subway stations	
MBTApaxE	Average daily number of entrances at each the MBTA station	
Pop2013	Total population	Census tract
Male2013	Total of male population	
P15_24Y201	Total population aged between 15 to 24 years	
P25_59Y201	Total population aged between 25 to 59 years	
P60_74Y201	Total population aged between 60 to 74 years	
TTHSH2013	Total of households	
TFam2013	Total of families	
FH	Total of family households	
NFH	Total of nonfamily households	
Il49k2013	Number of households with an annual income lower than $49,999	
I50_149k20	Number of households with an annual income between $50,000 and $149,999	
IM150k2013	Number of households with an annual income higher than $150,000	
TotEmploye	Civilian employed population with 16 years and over	
MBusScArt	Civilian employed population with management, business, science, and arts occupations	
Serv	Civilian employed population with service occupations	
SaleOf	Civilian employed population with sales and office occupations	
NConstMAin	Civilian employed population with natural resources, construction, and maintenance occupations	
ProdTran	Civilian employed population with production, transportation, and material moving occupations	
W16YM	Workers[5] with 16 years and over in households	
WatH	Workers who work at home	

[5] All the variables that mentions workers refers to “Workers with 16 years and over in households” and the information is based on the most often habits during a reference week of the survey.

People who used different means of transportation on different day of the week were asked to specify the one they used most often and people that use more than one means of transportation to get to work each day were asked to report the one used for the longest distance during the work trip.

Variable	Definition	Territorial distribution
WnatH	Workers who do not work at home	
WIT	Workers that use car, truck, or van as means of transportation to work (including company car but excluding taxicabs)	
WTP	Workers that use public transportation (excluding taxicab) as means of transportation to work	
Wwalk	Workers that walked as means of transportation to work	Census tract
Wbike	Workers that use bicycle as means of transportation to work	
Woth	Workers that use taxicab, motorcycle or others as means of transportation to work	
W0_459	Workers who leaves home to go to work between 12:00 a.m. and 4:59 a.m.	
W5_529	Workers who leaves home to go to work between 5:00 a.m. and 5:29 a.m.	
W530_559	Workers who leaves home to go to work between 5:30 a.m. and 5:59 a.m.	
W6_629	Workers who leaves home to go to work between 6:00 a.m. and 6:29 a.m.	
W630_659	Workers who leaves home to go to work between 6:30 a.m. and 6:59 a.m.	
W7_729	Workers who leaves home to go to work between 7:00 a.m. and 7:29 a.m.	
W730_759	Workers who leaves home to go to work between 7:30 a.m. and 7:59 a.m.	
W8_829	Workers who leaves home to go to work between 8:00 a.m. and 8:29 a.m.	
W830_859	Workers who leaves home to go to work between 8:30 a.m. and 8:59 a.m.	
W9_1159	Workers who leaves home to go to work between 9:00 a.m. and 11:59 p.m.	
TWl10	Workers that take less than 10 minutes to go to work	
TW10_14	Workers that take between 10 and 14 minutes to go to work	
TW15_19	Workers that take between 15 and 19 minutes to go to work	
TW20_24	Workers that take between 20 and 24 minutes to go to work	
TW25_29	Workers that take between 25 and 29 minutes to go to work	
TW30_34	Workers that take between 30 and 34 minutes to go to work	
TW35_44	Workers that take between 35 and 44 minutes to go to work	
TW45_59	Workers that take between 45 and 59 minutes to go to work	
TWM60	Workers that take more than 60 minutes to go to work	
VWnv	Workers 16 years and over in households with no vehicle available	
VW1v	Workers 16 years and over in households with 1 vehicle available	
VW2v	Workers 16 years and over in households with 2 vehicles available	
VWM3v	Workers 16 years and over in households with 3 or more vehicles available	
P18YHSh	Population 18 years and over with high school graduate or higher	
P18YBDh	Population 18 years and over with bachelor's degree or higher	

The spatial analysis considers 158 census areas, selected considering that the distance to the bike-sharing stations implemented is lower than 1 kilometer (0,62 miles). The study area is presented in the following map - Figure 27.

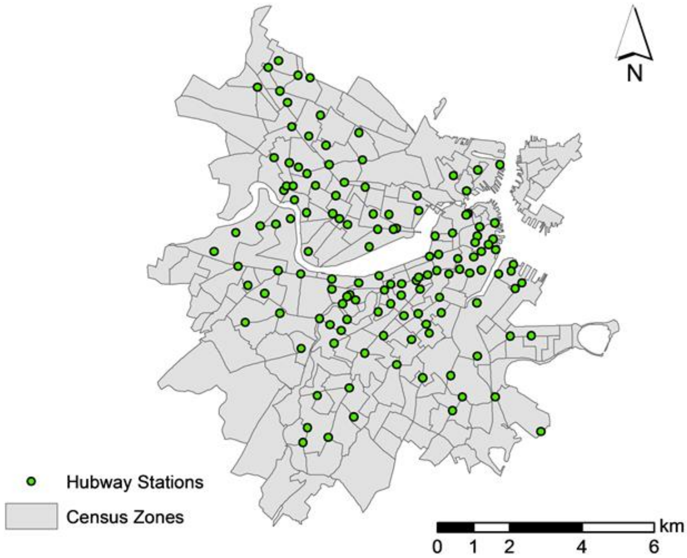


Figure 27 – Study case area, census zones and the location of hubway stations

The following sections presents the results obtained and the best models are analyzed.

3.4.4 Analysis and results

The Explanatory Regression tool from ArcGis® considered all the combinations from 1 to 5 independent variables of the set of variables presented at Table 8 – 56 variables. Thus 4.216.422[6] models were tested for each data distribution evaluated. The models with more than five variables presented high values of multicollinearity (condition index >30).

The same methodology is tested for the trips started and trips with destination in the stations of the Hubway stations for different spatial approaches. The results are present in the following subsections.

[6] Except the models that failed due to perfect multicollinearity.

Stations distribution

As mentioned the first territorial distribution tests the number of trips generated/attracted related to the characteristics of the surrounding zones considering buffers with 400 meters of radius from each station.

Figure 28 presents the distribution of the average number of trips per operated day started (Figure 28a) and finished (Figure 28b) in each station. There is a concentration of generated and attracted trips on downtown area where there is a higher concentration of facilities.

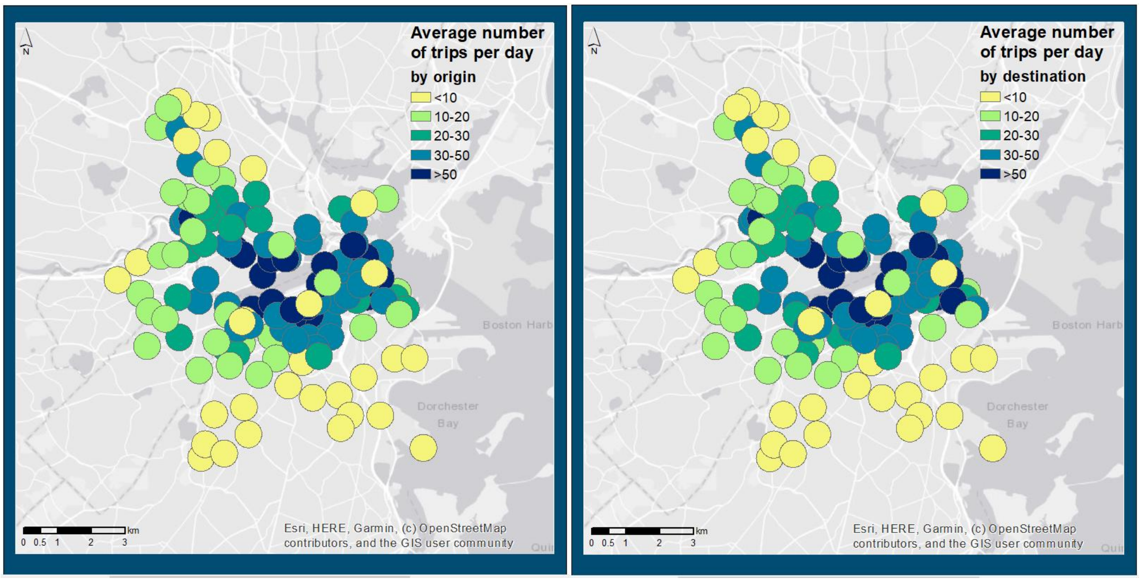


Figure 28 – Average number of daily trips by origin (a) and destination (b) - Stations models territorial distribution

The data from census zones are equally distributed throughout each zone to obtain the proportional (in area) characterization of each buffer, as explained on section 3.4.2.

Generated trips

The summary of the exploratory regression for the trips with <u>origin</u> in the stations (Table 9) presents five diagnostic tests, the cutoff considered in each diagnostic, and the number of models tried and models that passed each of those tests. Four (4) million models were analyzed, corresponding to all the combinations from 1 to 5 independent variables of the variables presented at Table 8 – 56 variables. The output of the exploratory analysis is presented at attachment I.1.1.

Table 9 – Exploratory regression global summary – stations model [origin]

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	4160123	525631	12.63
Max Coefficient p-value	< 0.05	4160123	195033	4.69
Max VIF Value	< 7.50	4160123	1822730	43.81
Min Jarque-Bera p-value	> 0.10	4160123	216261	5.20
Min Spatial Autocorrelation p-value	> 0.10	18	0	0

As the result of the ArcGis' tool, 13% of the models tested has an adjusted R^2 value higher than 50%, but only 5% of the models contain explanatory variables whose coefficients are statistically at the 95% confidence level (p-values smaller than 0.05).

44% of the models pass the multicollinearity criteria (VIF<7.5) and only 5% of them passes the Jarque-Bera criteria – an indication a normal distribution of the residuals.

Only 18 models passes all the criteria and are tested to spatial autocorrelation, however because there is no a neighborhood structure none of the models passes the criteria.

The results of exploratory regression identified the following variables as consistently significant in the models evaluated (Top five variables more significant).

Table 10 – Summary of Variable Significance – stations model [origin]

Variable	% Significant	% Negative	% Positive
WWALK	99.95	0.00	100.00
NCONSTMAIN	98.74	100.00	0.00
W830_859	97.73	0.18	99.82
NWORKERS	93.00	0.00	100.00
MDTAPAXE	88.11	0.17	99.83

Among the 4 million models analyzed, it is possible to conclude the isolated influence of these variables in the number of trips starting at each station. These variables were essentially the following:

- variable that indicates the number of people that used to walk to work (WWALK) is positively significant, areas where people walk have low distance trips and potential use of bike-sharing (central areas);
- the population with natural resources, construction, and maintenance occupations (NCONSTMAIN) is negatively significance, for this group travel distances tend to be longer to their working areas, and many times they have the transportation assured by the employer - they are not potential bike-sharing users;
- the variable that refers to workers who leaves home to go to work between 8:30 a.m. and 8:59 a.m. (W830_859) has positive sign, similarly as the people who walked to work, this group tends to work near home (since the working hour is at 9:00, usually), this means that in areas where people have low distance daily trips there is more tendency to potentially use the bike-sharing (central areas);
- The number of works (NWORKERS) in each buffer has also a positive influence, meaning that in areas with a higher concentration of jobs there is a higher number of tall buildings and, again, we are talking about central areas and central business districts (CBD);
- And the number of entrances on subway stations (MDTAPAXE) – reflects the global demand for a subway station and the existence of subway stations. Again, this is a characteristic of central areas. Bike-sharing station demand is higher at this type of POI (Point of Interest), showing its potential role as a solution to the first or last mile of a subway trip.

From all the possible combinations for these models the one with a higher adjusted R2 (0.65) includes three of the above consistently significant variables (WWALK, W830_859, and MDTAPAXE) with a positive relation and an additional variable POP2013 with a negative relation. These are the main factors affecting the trips with origin at the stations. Although the spatial regression is not analyzed in this case, the Geode regression tool was used just for basic regression validations upon the models, as shown in Table 11.

The model has a R^2 equal to 0.65 that reflects a model that explain 65% of the of the Hubway system through the stations’ approach.

Table 11 – Summary of results for the best model – stations model [Origin]

		Classical Model	
		z-value	**prob**
Stations model	Log likelihood	-115	-
	Akaike info	240	-
	Schwarz criterion	255	-
	R^2	0.65	-
	CONSTANT	5.955	-
	MBTApaxE	0.016	***
	Pop2013	-1.449	**
	Wwalk	0.848	***
	W830_859	0.652	***
	Normality of Errors Jarques-Bera	17.861	0.00013
	Heterocedasticity Breuch-Pagan	3.046	0.55024
	White test	11.7567	0.62584

The Jarques-Bera test is significant, which means that the model is biased, and some explanatory variables might be missing (the null hypothesis is refused, which admits normal distribution of errors). The Breuch-Pagan test indicates the presence of Heteroscedasticity.

The factors influencing the trips are the number of resident people that walk to work (WWALK), the number of persons that leave the house to go to work between 8:30 and 9 am (W830_859), and subway stations entrances (MBTApaxE). These three variables are on the top five influencing variables in the previous explanatory analysis.

The number of residents (Pop2013) is a variable with a negative impact on the number of bike-sharing trips, it reflects that areas with fewer residents are the areas with more potential bike-sharing trips departing (again, central areas).

Attracted trips

The exploratory regression for the trips with destination in the stations is presented in the Table 12 (the detail is presented at attachment I.1.1).

Table 12 – Exploratory regression global summary – stations model [destination]

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	4160123	429181	10.32
Max Coefficient p-value	< 0.05	4160123	197412	4.75
Max VIF Value	< 7.50	4160123	1822730	43.81
Min Jarque-Bera p-value	> 0.10	4160123	205848	4.95
Min Spatial Autocorrelation p-value	> 0.10	18	1	5.56

The obtained results are similar to the previous model in terms of volume of models that passes the referenced cutoffs. Most of the models represents less than 50% of the trips attracted and just 5% of the models has a normal distribution of the residuals.

It has identified by the tool one model that passes the spatial autocorrelation criteria – the model correspondent model (identified on annex I.1.1 – *Trips ended – Exploratory regression output*) does not commit the VIF cutoff (VIF=13.56).

The results of exploratory regression identified the following variables as consistently significant in the models studied (Top five variables more significant).

Table 13 – Summary of Variable Significance – stations model [destination]

Variable	% Significant	% Negative	% Positive
WWALK	99.90	0.00	100.00
W830_859	97.07	0.23	99.77
NWORKERS	95.04	0.00	100.0
MBTAPAXE	91.80	0.10	99.90
MBTAS	89.37	6.45	93.55

Among the exploratory regression approach models analyzed is possible to conclude that:

- Similarly to the results obtained on the origin model, the WWALK, W830_859, and NWORKERS variables contributes to the increase of the number of bike-sharing systems, which are variables related to the short distance trips,
- The existence of a subway station (MBTAS) and the number of entrances in it (MBTAPAXE) contributes positively to the demand for bike-sharing. Therefore, the results indicate that bike-sharing systems' success is related to proximity to a subway station.

Table 14 summarizes the best model achieved considering the trips finished at stations points.

Table 14 – Summary of results for the best model – stations model [destination]

		Classical Model	
		z-value	**prob**
Stations model	Log likelihood	-115	-
	Akaike info	243	-
	Schwarz criterion	260	-
	R^2	0.66	-
	CONSTANT	7.464	***
	MBTApaxE	0.023	***
	Pop2013	-1.640	***
	Wwalk	0.921	***
	Wbike	0.152	***
	W830_859	0.468	***
	Normality of Errors Jarques-Bera	18.772	0.000
	Heterocedasticity Breuch-Pagan	3.433	0.633
	White test	16.279	0.699

The results considering departure trips as the dependent variable are similar to the ones representing arrival trips.

In the present case, the model represents about 66% of the trips. However, the error distribution is not normal and is attached to an indication of heteroscedasticity.

The number of residents that walk to work and that go to work between 8:30 and 9:00, along with the subway demand, are the main significant exploratory variables.

The number of people that use the bike as a transport mode (Wbike) is also a significant exploratory variable. It indicates the bike-sharing trips arrival areas have more bike users. Again, we might be talking about central areas with shorter commuter distances and more bike commuters and that at the same time concentrate bike-sharing destinations.

Census areas (irregular areas)

A second approach is to estimate the influence of the total neighborhood characteristics (in which the station is) on the bike-sharing demand. And in this case, it will be possible to estimate a spatial regression: the influence of the neighbors at each zone. The data were analyzed through a territorial distribution equal to census tracts division, considering 158 census tracts areas, selected with a distance to the bike-sharing stations implemented is lower than 1 kilometer (0,62 miles). The information allocated into points is merged into areas, as previously referred.

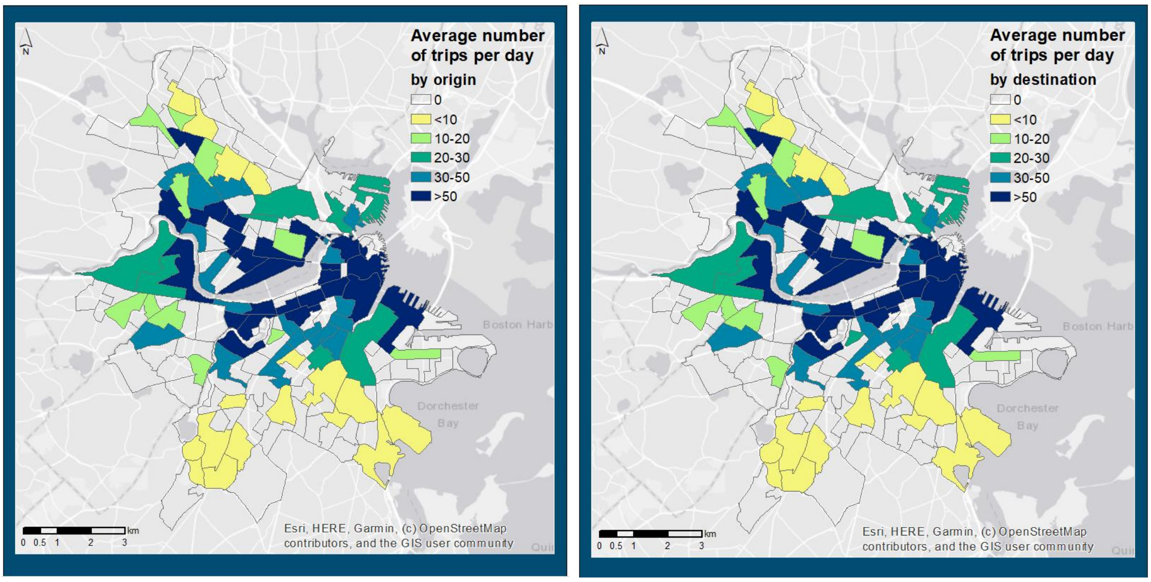


Figure 29 – Trips density by origin (a) and destination (b) - Census tracts models territorial distribution

As in the stations' case, the maps of generated and attracted trips are similar. It is possible to identify the Back Bay, the Downtown Crossing, the MIT zone and the Mid-Cambridge Area as the zones with more generated and attracted trips. And, on the other side, the North Cambridge and the south of Boston has a lower number of trips.

On bike-sharing it is possible to have zones/stations that generate more trips than attracting (or vice-versa) that systems tend to have a difficult relocation process – for instance, in high sloped zones when a bicycle (from the system) is a good option to go down but a not option to go up.

Generated trips

The exploratory regression results, using as dependent variable the trips that started at the stations, are summarized in Table 15.

Table 15 – Exploratory regression global summary – Census areas model [origin]

Search Criterion	**Cutoff**	**Trials**	**# Passed**	**% Passed**
Min Adjusted R-Squared	> 0.50	4163813	0	0.00
Max Coefficient p-value	< 0.05	4163813	7709	0.19
Max VIF Value	< 7.50	4163813	952563	22.88
Min Jarque-Bera p-value	> 0.10	4163813	2083	0.05
Min Spatial Autocorrelation p-value	> 0.10	18	9	50.00

Through the analysis to the exploratory analysis, it is possible to conclude that there is no models with R^2 higher than 0.5, among the 4 million trials tested. This means that with this data set does not explain the majority of the dependent variable variation.

Moreover, 95% of the models analyzed have the Jarque-Bera test statistically significant (the data does not form a normal distribution).

However, there is some heteroscedasticity probably related to the spatial distribution of data (spatial autocorrelation).

So, even if the R^2 is not significant at the required levels, there is an indication for some spatial autocorrelation.

Therefore, it is interesting to look at the variables consistently showing some significance, presented on next table.

Table 16 – Summary of Variable Significance – census areas model [origin]

Variable	% Significant	% Negative	% Positive
MBTAPAXE	85.09	0.01	99.99
MBTAS	80.73	7.06	92.94
MWORKERS	77.54	0.00	100.00
WWALK	67.93	0.95	99.05
VWNV	62.93	0.36	99.64

Between the models analyzed it is possible to conclude that:

- Census zones where there are subway stations (MBTAS) and those ones that registered more entrances (MBTAPAXE) have more probability of higher number of bike-sharing trips, similarly to the study considering the points approach, as expected;
- The number of workplaces (MWORKERS) seems to be also significant. There is a concentration of bike-sharing trips in areas where the workplaces are also more concentrated.
- Again, the use of walking to work (WWALK) seems to be related to the departing trips, probably related to central areas where the number of shorter trips is higher.
- A new variable that appears here is the non-availability of personal vehicles in the workers' household (VWNV). Car ownership tends to be a variable that negatively influences bike-sharing use.

The following table presents the model with highest adjustment identified by the exploratory ArcGis® tool to estimate the number of bike-sharing trips starting in the stations, estimated through the Geoda® software.

The statistical test of Jarque-Bera indicates a residuals normal distribution, Breuch-Pagan and White tests reveal the residuals are heteroskedastic. But, as previously mentioned, the value of R^2 has a weak adjustment to the dependent variable variation.

From the tests for spatial dependence diagnostics, it is possible to conclude that there is no spatial correlation between variables since all the tests are non-significant (p-value > 0.05).

Table 17 – Summary of results for the census areas model [origin]

		Classical Model	
		z-value	**prob**
Census areas model	Log likelihood	-401	-
	Akaike info	812	-
	Schwarz criterion	827	-
	R^2	0.21	-
	Constant	-7.503	***
	W630_659	-0.826	***
	W7_729	-1.023	***
	VW1v	1.244	***
	NEmployers	1.029	***
	Normality of Errors Jarques-Bera	4.216	0.12
	Heterocedasticity Breuch-Pagan	10.207	0.04
	White test	108.631	0.00
	Spatial Dependence Moran's I	-1.450	0.14
	LML	1.839	0.18
	LME	2.448	0.12
	LMLR	0.043	0.83
	LMER	0.652	0.42

However, they approach the level of significance of p=0.10, which shows some tendency for spatial dependence.

Besides VWNV, the significant independent variables are the following:

- People leaving home in the early morning (W630_659 and W7_729) tend to use other modes of transportation than bike-sharing, probably meaning that they have longer trips;
- The availability of just one vehicle in workers households (VW1v) increases the number of bike-sharing trips, meaning that some family members use other alternatives rather than the car;
- The number of employers (NEmployers) in each zone has a positive effect on the number of bike-sharing trips, certainly related also to the number of workplaces and jobs.

Attracted trips

It was not possible to obtain the full output of the exploratory regression for the model that considers the trips with a destination in each zone. The high amount of degrees of freedom tested requires large amounts of RAM and a computer with only 16Gb of RAM was used. Unfortunately, despite being tested multiple times, the solver maxed out the available RAM and the computer crashed. This lead to a situation where the output file was only partially populated.

The partial output is in annex I.1.2 and it presents the models with highest adjusted R-squared results. From this list it is possible to conclude that the value of R^2-adjusted is always lower than 0.3, which reveals a weak adjustment as in the previous model (with origin trips as dependent variable).

The following table summarizes the GeoDa® output for the best model identified on Exploratory Regression partial output.

As in the model using trips with origin in the stations, there is a normal distribution of the residuals at the same time they present a heteroscedastic behavior, according to the results of the Jarque-Bera and Breuch-Pagan tests. However, the adjustment of the model is not good explaining the number of trips starting in each zone.

Also, there is no spatial dependence between variables since all the tests are non-significant (p-value > 0.05).

The impact of independent variables is similar to the previous model presented, with the same variables being the significant ones, thus:

- People that early leave home to work (W630_659 and W7_729), contributes negatively to the number of bike-sharing trips;

- The availability of one vehicle in workers households (VW1v) and the number of employers in each zone is have a positive effect in the number of bike-sharing trips;
- In addition, the number of stations (MBTAs) also has a positive effect on the number of bike-sharing trips arriving, as in the point-station data distribution approach.

Table 18 – Summary of results for the best models – census areas model [destination]

		Classical Model	
		z-value	**prob**
Census areas model	Log likelihood	-457	-
	Akaike info	927	-
	Schwarz criterion	945	-
	R^2	0.28	-
	Constant	-6.627	***
	W630_659	-1.260	***
	W7_729	-1.720	***
	VW1v	1.994	***
	MBTAs	0.250	**
	NEmployers	1.546	***
	Normality of Errors Jarques-Bera	1.616	0.45
	Heterocedasticity Breuch-Pagan	11.23	0.05
	Spatial Dependence Moran’s I	-1.247	0.21
	LML	1.057	0.30
	LME	1.894	0.17
	LMLR	0.196	0.66
	LMER	1.034	0.31

Grid division

As referred, the grid is defined through *Optimized Hot Spot Analysis* tool from ArcGis®. It was considered the stations shapefile as input feature and the study area (Figure 27) as bound. The tool defined a fishnet with squares of 860 meters side, as presented at the Figure 30.

The information is aggregated into the grid's zones adding the contributions of cells that results from the crossing of zones and grid shapefiles, and the information allocated into points is merged into areas, as detailed on methodology (section 3.4.2).

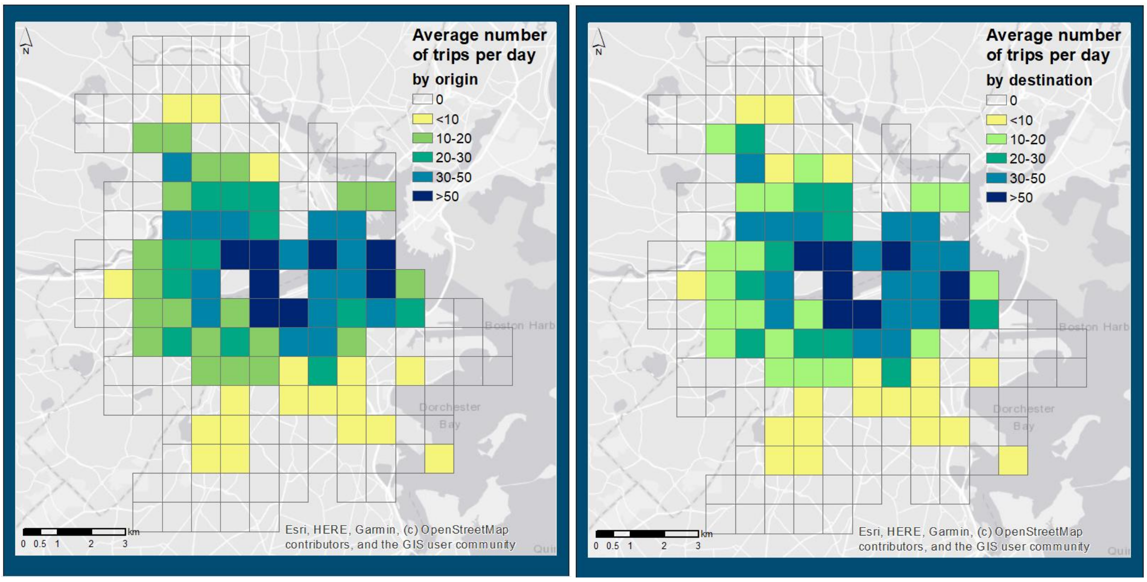


Figure 30 – Average number of trips per operated day by origin (a) and destination (b) – Grids' models territorial distribution

The pattern of generated trips and attracted trips is similar, as expected due to the previous analysis. There are small differences in the east zone of Boston Common and in the north of Cambridge. However, the differences are small in terms of trips and close to the scale threshold.

Generated trips

The Table 19 presents the summary of exploratory regression analysis considering as dependent variable the trips started in the grid cells.

The summary of the exploratory regression presents similar results to the census tract models analysis. There are no models with Adjusted R^2 higher than 0.50 – weak adjustment to the number of trips with origin in each station.

A low number of models reject the Jarque-Bera test – the data does not follow a normal distribution. However, there is evidence of spatial dependence.

Table 19 – Exploratory regression global summary – GRID model [origin]

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	4148046	0	0.00
Max Coefficient p-value	< 0.05	4148046	128037	3.09
Max VIF Value	< 7.50	4148046	155027	3.74
Min Jarque-Bera p-value	> 0.10	4148046	170708	4.12
Min Spatial Autocorrelation p-value	> 0.10	18	15	83.33

The following variables were identified as constantly significant in the models evaluated – the table presents the top five variables with more significance, contributing positively to the number of bike-sharing trips - Table 20.

Table 20 – Summary of Variable Significance – GRID model [origin]

Variable	% Significant	% Negative	% Positive
WWALK	99.91	0.00	100.00
MBTAPAXE	98.20	0.00	100.00
NWORKERS	95.25	0.00	100.00
MBTAS	94.40	6.86	93.14
VWNV	92.88	0.25	99.75

From the Table 20 is possible to summary some points that are similar with previous models:

- the number of resident people that walks to work (WWALK), meaning that bike-sharing trips are most likely to occur in areas where the trips have a short duration;

- the number of entrances on subway stations (MBTAPAXE and the number of subway stations (MBTAS)), meaning that bike-sharing trips are most likely to occur in areas where there is a subway and strong possibilities to intermodality;
- the number of workers at each zone (NWORKERS), meaning that areas with a high concentration of activities and workplaces are more likely to have bike-sharing trips;
- the number of workers resident in each zone with no vehicle available (VWNV) meaning that areas with a higher concentration of activities and workplaces are more likely to have less car ownership and bike-sharing trips are an alternative.

The Table 21 summarizes the comparison between one model achieved considering the analysis to the grid distribution data through a classic model, Spatial Lag model and Spatial Error model.

Estimating first the Classic Regression Model for the best model identified, the significant variables appearing are W830_8_860 (with a positive relation) and P18YBDh and Serv (with a negative one), besides WWALK and MBTApaxE.

In areas of higher literacy (P18YBDh - Population 18 years and over with bachelor's degree or higher), there are higher-income residents and higher car ownership (and fewer bike-sharing trips, as stated before). The same happens with the SERV variable analysis: people that work in service occupations tend to have higher incomes.

The error term follows a normal distribution, and there is strong heteroscedasticity due to spatial dependence, as shown in Moran's I significance.

The Classic Model presents an adjusted R^2 equal to 0.51, explaining 50% of the trips generated in each cell.

Considering Lagrange Multipliers (LML and LME) the significance appears only in the Lag model meaning that the spatial dependence is on the dependent variable.

The comparison between the three models through Log-likelihood, Akaike info, and Schwarz criterion values shows that the Spatial Lag is the best - the higher Log-likelihood and the lower information criteria.

It means that it is difficult to capture all the relations in only one model. The 50% explained is probably related to this dimension and heterogeneity. If from here, a reduced set of cells were considered, like the central ones, this proportion would increase, probably enormously. Even though, it is of outmost importance reflect on the variables influence on the number of trips.

Table 21 – Summary of results for the best models – GRID model [origin]

		Classical Model		Spatial Lag Model		Spatial Error Model	
		z-value	prob	z-value	prob	z-value	prob
GRID model	Log likelihood	-426	-	-423	-	-426	-
	Akaike info	865	-	861	-	864	-
	Schwarz criterion	883	-	882	-	883	-
	R^2	0.511	-	-	-	-	-
	ρ	-	-	0.301	***	-	-
	Λ	-	-	-	-	0.13668 4	-
	Constant	-1.261		-0.73535		-1.23718	
	Serv	-1.129	*	-0.89039	***	-1.16098	***
	Wwalk	2.274	***	1.6309	***	2.1389	***
	W830_8_860	1.746	***	1.36425	***	1.78195	***
	P18YBDh	-2.157	***	-1.58139	***	-2.06303	***
	MBTApaxE	0.132	***	0.11625 2	***	0.12975 9	***
	Normality of Errors Jarques-Bera	1.878	0.39				
	Heterocedasticity Breuch-Pagan	18.838	0.00	19.238	0.00	19.1414	0.00
	Spatial Dependence Moran's I	1.296	0.19	6.305	0.01	0.7126	0.40
	LML	5.806	0.02				
	LME	0.549	0.46				
	LMLR	11.004	0.00	-	-	-	-
	LMER	5.747	0.02				

Attracted trips

Table 22 shows the results of the exploratory regression analysis considering as dependent variable the trips with the destination in the grid cells defined. As before there are no models with R^2 higher than 0.5.

Table 22 – Exploratory regression global summary – GRID model [destination]

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	4148046	0	0
Max Coefficient p-value	< 0.05	4148046	128346	3.09
Max VIF Value	< 7.50	4148046	155027	3.74
Min Jarque-Bera p-value	> 0.10	4148046	180958	4.36
Min Spatial Autocorrelation p-value	> 0.10	18	15	83.33

The same variables are the most significant on the tested models for attracted (Table 20) and generated trips (Table 23).

Table 23 – Summary of Variable Significance – GRID model [destination]

Variable	% Significant	% Negative	% Positive
WWALK	99.91	0.00	100.00
MBTAPAXE	98.28	0.00	100.00
NWORKERS	95.77	0.00	100.00
MBTAS	94.44	6.86	93.14
VWNV	92.91	0.25	99.75

It is summarized the model achieved considering the analysis to the grid distribution data through a Classic model, Spatial Lag model and Spatial Error model in the Table 24.

The destination model performance and results are quite similar to the origin model results. Thus the model is adjusted to 51% of the cases (adjusted $R^2 = 0{,}51$), there is a normal distribution in the residuals and the residuals are heteroskedastic and it is detected the presence of spatial dependence.

Table 24 – Summary of results for the best models – GRID model [destination]

		Classical Model		Spatial Lag Model		Spatial Error Model	
		z-value	**prob**	**z-value**	**prob**	**z-value**	**prob**
GRID model	Log likelihood	-426	-	-423		-426	
	Akaike info	865	-	861		864	
	Schwarz criterion	883	-	882		883	
	R^2	0.512	-				
	ρ	-	-	0.299	***		
	λ	-	-			0.134	0.32
	Constant	-1.249	*	-0.727	0.30	-1.226	0.10
	Serv	-1.125	***	-0.887	***	-1.155	***
	Wwalk	2.278	***	1.636	***	2.144	***
	W830_860	1.741	***	1.361	***	1.775	***
	P18YHSh	-2.160	***	-1.587	***	-2.068	***
	MBTApaxE	0.133	***	0.117	***	0.131	**
	Normality of Errors Jarques-Bera	1.756	0.41	18.77	***	18.698	***
	Heterocedasticity Breuch-Pagan	18.399	0.00	6.241	**		
	Spatial Dependence Moran’s I	1.282	0.19	6.241	0.01248	0.688	0.41
	LML	5.751	0.01				
	LME	0.530	0.46				
	LMLR	10.982	0.00				
	LMER	5.762	0.01				

As in the case of origin models studied, the Spatial lag model is the one that presents the best adjustment, with lower values Log likelihood, Akaike info criterion and Schwarz criterion.

The influence of the variables are similar to the origin model in terms of variables and influence of them on the number of trips with destination in each zone.

3.4.5 Synthesis

None of the models analysed explains more than 60% of the bike-sharing trips, using the database available. This database reveals the proprieties of multicollinearity, heteroscedasticity, and autocorrelation. Therefore, these models should be used with caution as a way to predict the number of bicycle trips to expect, both departing and arriving, at points (stations) or areas.

Moreover, using the grid data distribution, it was possible to identify that part of this heteroscedasticity in the regression error term is due to clusters and outliers (Spatial Autocorrelation) in the variable’s distributions.

However, the combined Regression Analysis (Exploratory Tool and Spatial Regression) identified variables consistently significant through all the analyses as factors influencing bike-sharing trips. This consistency also seems to be relatively independent of the geographic distribution of the data or if it is an origin or destination bike-sharing trip, although there are some differences.

Other approaches were tested, considering peak time hours and only working days’ trips, but the conclusions remained very similar to the presented ones, not adding value to the outcomes.

Table 25 presents a synthesis of the results, where the main findings can be observed. The ‘+’ symbol indicates a positive correlation with the number of trips and the ‘-‘, a negative correlation. The database variables designation was substituted by their meaning, increasing the quick assessment of the findings.

Table 25 - Summary of Variable Significance – Results synthesis

Variable/Model	Point Departures	Point Arrivals	Census Tracks Departures	Census Tracks Arrivals	Grid Departures	Grid Arrivals
People that **walk to work**	+	+	+		+	+
People that work in industry or resources	+					
People that **leave home between 8:30 and 8:59**.	+	+			+	+
Number of Workers	+	+	+		+	+
Number of **entrances in the subway stations**	+	+	+		+	+
Population in 2013	-	-		-		
Number of **subway stations**		+	+	+	+	+
People that **bike to work**		+				
Households **with no car available**			+		+	+
Households with one car available				-		
People that leave home between 7:00 and 7:29.			-	-		
People that leave home between 6:30 and 6:59.			-	-		
Number of **employers** (companies or institutions)			+	+		
Population 18 years and over with bachelor's degree or higher					-	-
Civilian employed population with service occupations					-	-

Globally, city areas with high activity concentration, small commuting distances, and close to public transport facilities are the ones where one should expect higher bike-sharing trips demand. Focusing on the results it is possible to conclude that:

- The number of people that used to walking or cycling to work is positively significant, it suggests that this group is potentially the group that is willing to use bike-sharing – short distance trips;
- Workers who leave home to go to work between 8:30 a.m. and 8:59 a.m. have positive significance, similarly, as the people who used active modes to commute, this group tends to work nearly home and they have a short commute trip;
- Zones with more employers or a higher number of workers has also a positive impact on the number of bike-sharing trips, these trips could be related to last-mile trips for the people that use other transport and complete the trip with bike-sharing;
- The presence of subway stations and the demand for the subway have a positive effect on the bike-sharing trips validating that bike-sharing is a good option to last-mile for public transportation trips;
- The non-availability of a vehicle in the household has a positive influence on bike-sharing trips since people tend to use alternative transport modes.

The conclusions obtained are in line with the published research. The bike-sharing system’s users most frequently used the stations closest to either home (40%) or work (40%) (S. A. Shaheen et al., 2011).

As referred by Dell’Olio et al, some points that are potentially optimal locations for the stations: near to public facilities in the city, near to the most used stations and car parks to promote intermodality, on the flattest areas of the city, it should not overlap stations of different companies in the same location, and only where bicycle paths exist (dell’Olio et al., 2011).

Bicycles can work as a feeder and distributor service for public transport. Therefore, parks for bicycles and stations for bicycle sharing must be provided in public transportation terminals to complement the trips (Grava, 2003; Pucher & Buehler, 2008; Sörensen et al., 2012).

For further studies approaches, we must consider that central to these finds there are at least three variables absent from the database analysed:

- The distinction between a daily commuting trip (work or school) or a leisure trip;
- The trip length;

- The existence of free or cheap car parking.

This information would validate the results and must be included in future databases collections. In further studies and considering the variability of this database variables relations, other methods such as Instrumented Variables and Geographically Weighted Regressions approaches could be used. Even if this database was complete with all the variables needed, using a single regression for the entire Boston Bike-Sharing System cannot incorporate all the geographical variation in the relation between the variables.

It is important to mention that the chapter summarizes the research work done on the Hubway system analysis. Other approaches were done considering peak time hours, working days trips, etc. however the conclusions remained very similar to the presented ones not adding value to the outcomes presented.

There is always a little difference between the calculated R^2 adjusted in ArcGis® and GeoDa®, the difference in the analyzed models was not significant and never changed the conclusions obtained.

3.5 Conclusion

The prediction of a new transportation mode is always a challenge to a transport engineer. It is hard to foresee human behavior, which is even harder when the data available does not fit the needs of the problem.

This chapter presented two different approaches to address the demand estimation on bike-sharing systems, the first one, designed in a context of low published research on bike-sharing (Figure 3) and, specifically on bike-sharing demand, is a methodology focused on bike-sharing demand shaped by city characteristics that affect bicycle usage.

It provides a quick assessment and it can be adjusted to other territories despite being based on a global proportion of users that need to be adapted to the socio-economic and cultural characteristics.

The objective of the second present approach was the study of the Hubway system in order to understand the demand based on the socio-economic and cultural characteristics facing up the challenge of the first approach.

Although the results achieved are not conclusive in the prediction of bike-sharing demand, the information allowed to identify some of the city characteristics to consider in bike-sharing system demand and some of the potential bike-sharing user's profiles.

Systems monitoring is a step to improve the planning process for bike-sharing systems. With this approach, it can be possible to identify the strengths and weaknesses of spatial planning policy and the main characteristics that affect the use of the system.

4 LOCATION OF BIKE-SHARING STATIONS

4.1 Introduction

Bike-sharing is getting increasingly popular as a sustainable transport system as the number of bike-sharing systems grow significantly worldwide in recent years. One of the most relevant elements in the implementation of these systems is the location of the stations. As mentioned in chapter 2, the location of bike-sharing stations might compromise the success of the system.

Municipalities or public-private partnerships are usually responsible for implementing bike-sharing systems. The public investment in bicycle mobility (particularly bike-sharing) is complex because it is always subject to a budget. The main concern for public investment is to maximize the benefits through the design and implementation of bike-sharing systems. This work lays out a methodology to help with the decision-making of bike-sharing systems.

This work focus on bike-sharing systems with fixed stations. This chapter proposes an optimization model to design the bike-sharing system maximizing the demand covered with budgetary constraints.

It combines strategic decisions for locating and dimensioning bike-sharing stations (stations and number of bicycles) with operational decisions (bicycles relocation).

The model determines the optimal location of the bicycle stations, the fleet size, the capacity of the stations, and the number of bicycles in each station, considering an initial investment lower than the given budget. Moreover, it balances the annual cost of the system and the revenue assuming a possible supplementary budget from the system provider to cover any loss resulting from the shortfall between its operating cost and the revenue from the subscription charges.

The research work present in this chapter was published in Transportation Research Part A: Policy and Practice (I. Frade & Ribeiro, 2015).

The chapter is organized into four sections, besides the introduction. Section 4.2 presents a literature review on location models including general optimization models and location models applied on bike-sharing systems. Section 4.3 details the proposed optimization model for the design of a new bike-sharing network, as well as the assumptions upon which it is based. The model is then applied to the city of Coimbra (Portugal), and the data and results are presented in section 4.4. The last section 4.5 discusses the model formulation.

4.2 Literature Review – Location Models

Facility location is a strategic decision that depends on its initial goals. Locations are usually selected efficiently with the support of a particular type of optimization model, called facility location models, whose decision variables represent the location, the capacity, the coverage area of facilities considered, and, in this case, the relocation of bicycle stations (Daskin, 1995, 2008; ReVelle & Eiselt, 2005).

A facility location model can include different objectives, such as minimizing overall cost, minimizing transport cost, or maximizing demand coverage. These objectives correspond to solutions found through fixed-charge models, p-median models, and maximal covering models. Depending on whether capacity constraints apply to the facilities, the models classify them as capacitated or incapacitated. In bike-sharing stations, the literature reports different approaches to tackle station location with facility location models.

An optimization model is described by Lin and Yang which proposes an integer nonlinear program that determines the optimal location of docking stations, the bicycle lanes needed and what routes should be taken from each origin to each destination. It is based on cost minimization and assumes a penalty for uncovered demand. This model does not consider the relocation of bicycles; it assumes that bicycles and free spaces are always available in the stations, but this oversimplifies the problem (Lin & Yang, 2011).

The model presented by Lin et al. incorporates bicycle stock considerations as a hub location inventory model. Since the formulation presented is not computationally tractable, the authors proposed a greedy heuristic method to find efficient near-optimal solutions (Lin et al., 2011).

A mixed-integer linear program performed through a heuristic that optimizes the location of shared bike stations is presented by Martinez et al., assuming a fleet size and bicycle relocation calculation for a regular operating day. The focus of the method is to maximize revenue (Martinez et al., 2012).

Besides facility location models, the literature contains other methodologies to define the location of the stations.

Romero et al. consider a simulation-optimization method that relates public bicycles to private cars. The methodology is essentially a bi-level mathematical programming model that optimizes the location of public bicycle stations (Romero et al., 2012).

The García-Palomares et al. work proposes a GIS-based methodology to estimate the potential trip demand and its spatial distribution, the location of the stations (using location-allocation models), the station capacity and demand profiling for stations (García-Palomares et al., 2012). The balance of the bike-sharing systems problem, which considers the number of bicycles in each station and the optimal relocation routes, is discussed in (Lu, 2013; Raviv & Kolka, 2013; Sayarshad et al., 2012) stressing the importance of considering both the location problem and the relocation problem. The first sets out a robust fleet allocation model that generates the optimal daily allocation of bicycles to the stations and the redistribution flows of an implemented bike-sharing system, while minimizing the total cost (Lu, 2013). Raviv & Kolka present an inventory model to define the management of bike-sharing under the introduction of a user dissatisfaction function to assess the relocation service quality. The methodology focus is to find the initial inventory of the station that minimizes the dissatisfaction function (Raviv & Kolka, 2013). Finally Sayarshad et al., provides an optimization model to plan the relocation of bicycles in bike-sharing systems in small communities, assuming the maximization of the total benefit to the company (function of revenue and costs) (Sayarshad et al., 2012).

These works provide a good background for our study, but they miss some relevant points related to real-world implementations of these systems. As we know, public investment requires the maximization of the benefit, and in the case of bike-sharing, it also involves maximizing the number of users.

The maximization of demand coverage can be solved using maximal covering models, which are especially well suited to bike-sharing stations. These models were introduced by Church & ReVelle and their application makes it possible to determine the locations that maximize the covered demand, for a given number of facilities (Church & ReVelle, 1974).

The presented model for the location of bike-sharing combines strategic decisions with the system's size (stations and number of bicycles) establishment and with operational decisions (bicycles relocation). The model defines the optimal location of the bicycle stations, fleet size, station capacity, and bicycle quantities in each station. Moreover, using an initial investment achieves a balance between the annual cost of the system and the revenue. It also considers a possible supplementary budget given by the provider of the system to cover losses resulting from the shortfall between its operating costs and revenue from the subscription fees.

4.3 Modeling Approach

The optimization model presented below addresses some of the issues presented earlier in this thesis. The model is the backbone of the proposed approach as it defines the optimal design of a bike-sharing station network to maximize the demand covered while taking into account restrictions on the cost and level of service. It simultaneously determines the location of the stations, the number of bicycles that should be available in each station to maximize the demand (by defining the relocation operations across the zones), and the fleet size for each time period.

This optimization model articulates a demand study for the different zones in a city or urban area. The demand corresponds to the number of trips generated and attracted in each zone, being the zones small enough to guarantee walking distances inside each one.

The zones must be as small as possible for the solution to have the highest accuracy. The analysis to identify the most suitable distances is performed for each case in order to define it according to the locality and the available data.

Nevertheless, no zone should exceed 500 meters as the maximum distance between any two points regardless of its shape. In smaller zones, it is unlikely that someone will use the bike-sharing system to travel within a zone. The minimum length of the zone is defined through the possible dimension for the correct understanding of Origin-Destination (OD) trips matrix.

The models requires also a time-period OD matrix. The day is divided into periods which should match the same timeframe of the available data in each case study and articulated with the frequency of relocation activities. Specific periods may be very demanding both in terms of cost and human resources.

The objective of the model is to maximize the covered demand and the return on investment. On the revenue side, it considers a possible public investment contribution to the system and the revenue from the subscriptions. On the expenses side, it considers the relocation and maintenance costs (bicycles and stations).

The model is subject to capacity constraints to secure the coverage of demand, cost constraints based on net present value to satisfy the available budget, and domain constraints to ensure the viability of the variables.

The inputs of the model are: the demand for the system, maximum and minimum capacity of the stations, the price of the stations and bicycles and the relocations and maintenance costs, the total investment budget and the annual supplementary budget, as well as the discount and growth rate and the project's horizon years.

The model outputs include the number of stations in each zone and their capacity. The number of bicycles to locate and to relocate at each station and each time step, the total fleet size, the annual revenue, and the annual expenses are also outputs in this model.

The notation used to represent the sets, decision variables, and parameters used in the model is given below, in order of appearance. The available software programs, such as XPRESS®, can be used to solve the model.

Sets:

J: set of demand zones, indexed by i and j

T: set of time periods, indexed by t

Decision variables:

x_{ijt} : proportion of covered demand from zone i to zone j in time step t

y_i : is 1 if the bike station in zone i is opened and 0 otherwise

r_{ijt} : number of bicycles relocated from i to j at time step t

v_{it} : number of bicycles in zone i at the beginning of period t (needed to meet the demand in that zone)

z_i : number of docks in zone i

Parameters:

u_{ijt}: demand from i to j in time step t

ib: initial budget

sb: supplementary budget to cover loss resulting from the shortfall between operating costs and revenue from charges

i: investment needed to the implementation of the system

z_{min}: minimum station capacity

z_{max}: maximum station capacity

Tv: total fleet size of the system

cb: unit price of a bicycle

cs^f: fixed cost of a station

cs^v: variable cost of a station

cr^f: fixed unit relocation trip cost

cr^v: variable unit relocation trip cost

cms: maintenance cost of the bicycle station, including depreciation per year

cmb: maintenance cost of each bicycle, including depreciation per year

fa: annual user subscription

fm: monthly user subscription

fd: daily user subscription

dr: discount rate

gr: growth rate

n: project horizon (years)

f: income from the subscriptions

c: costs of the project in the project horizon (n)

b: benefits of the project in the project horizon (n)

The problem of determining the maximum coverage solution for locating bike-sharing stations is represented by the following model:

$$Max\ Z = \sum_{i \in J} \sum_{j \in J} \sum_{t \in T} (u_{ijt} \times x_{ijt}) \tag{16}$$

Subject to:

$$v_{it} = v_{i(t-1)} - \sum_{j \in J} u_{ij(t-1)} x_{ij(t-1)} + \sum_{j \in J} u_{ji(t-1)} x_{ji(t-1)} + \sum_{j \in J} r_{ji(t-1)} - \sum_{j \in J} r_{ij(t-1)} \quad \forall i \in J, j \in J, \in T \tag{17}$$

$$v_{i,1} = v_{i,T} \quad \forall i \in J \tag{18}$$

$$z_i \leq z_{\max} \times y_i \quad \forall i \in J \tag{19}$$

$$z_i \geq z_{\min} \times y_i \quad \forall i \in J \tag{20}$$

$$v_{it} \geq \sum_{j \in J} \left(u_{ijt} x_{ijt} \right) \quad \forall i \in J, i \in J, t \in T \tag{21}$$

$$v_{it} \leq 0.75 \times z_i \quad \forall i \in J, t \in T \tag{22}$$

$$v_{it} \geq 0.25 \times z_i \quad \forall i \in J, t \in T \tag{23}$$

$$\sum_{j \in J} r_{ijt} \leq v_{it} \quad \forall i \in J, t \in T \tag{24}$$

$$Tv = \sum_{i \in J} v_{it} \quad \forall t \in T \tag{25}$$

$$i = \sum_{i \in J} (cs^f \times y_i + cs^v \times z_i) + cb \times Tv \quad \forall i \in J \tag{26}$$

$$i \leq ib \tag{27}$$

$$c = 365 \times \left[cr^f \sum_{i \in J} \sum_{j \in J} \sum_{t \in T} r_{ijt} + cr^v \sum_{i \in J} \sum_{j \in J} \sum_{t \in T} d_{ij} r_{ijt} + cms \sum_{i \in J} y_i + cmb \times Tv \right] \times \left(\frac{1 - \frac{(1+gr)^{(n-1)}}{(1+dr)^{(n-1)}}}{dr - gr} \right) \quad \forall i \in J, j \in J, t \in T \tag{28}$$

$$f = (fa \times 0.5 + fm \times 12 \times 0.2 + fd \times 365 \times 0.3) \times \sum_{i \in J} \sum_{j \in J} \sum_{t \in T} (u_{ijt} \times x_{ijt}) \quad \forall i \in J, j \in J, t \in T \tag{29}$$

$$b = (sb + f) \times \left(\frac{1 - \frac{(1+gr)^{(n-1)}}{(1+dr)^{(n-1)}}}{dr - gr} \right) \tag{30}$$

$$-i - \left(\frac{c}{(1+dr)^1} \right) + \left(\frac{b}{(1+dr)^1} \right) \geq 0 \tag{31}$$

$$\sum_{j \in J} x_{ijt} \leq 1 \quad \forall i \in J,\ t \in T \tag{32}$$

$$x_{ijt} \leq y_i \quad \forall i \in J, j \in J, t \in T \tag{33}$$

$$x_{ijt} \leq y_j \quad \forall i \in J, j \in J, t \in T \tag{34}$$

$$r_{ijt} \geq 0 \quad \forall i = j \in J, t \in T \tag{35}$$

$$x_{ilt} \geq 0 \quad \forall i \in J, j \in J, t \in T \tag{36}$$

$$y_i \in \{0,1\} \quad \forall i \in J \tag{37}$$

$$v_{it}, z_i, r_{ijt} \in \mathbb{N} \quad \forall i \in J, j \in J, t \in T \tag{38}$$

The objective function (16) of this linear program maximizes the demand covered by the bike-sharing system.

Constraint (17) defines the number of bicycles available at a station in zone *i* in time step *t*. This is the balance of the bicycles available in the previous time step: difference between the bicycles leaving station *i* and arriving at station *i*, as well as the bicycles relocated from or to station *i*, assuming that the number of bicycles at the beginning and the end of the day is the same, constraint (18).

The capacity of any station is always the same as (or lower) than an established maximum capacity of the stations, constraint (19), and higher than a minimum, constraint (20). The number of bicycles available in station *i* in time step *t* has to be enough to meet the demand, constraint (21).

The stations should always have free parking places to ensure movement between stations and bicycles to cover the demand. Experience shows that free spaces must always be 25% of the station capacity. Thus at the beginning of *t* the number of bicycles in station *i* must be 75% of the capacity of that station, constraint (22), while there must be more than 25% of bicycles, constraint (23). But during each time period the model considers the possibility of fluctuations in the number of bicycles/parking spaces available.

The number of bicycles to be relocated from station *i* is lower than the number of bicycles in the station, constraint (24). Equation (25) determines the total fleet of the system.

The investment is the sum of the cost of the bike-sharing stations (defined as a function of the number of docks) and the cost of the bicycles (assuming that the implementation costs are included), equation (26), and the investment must be lower than the initial budget available *ib*, constraint (27).

The annual cost of the system includes the relocation cost, the maintenance cost and the vehicle depreciation cost, equation (28).

The income from subscriptions is determined from the annual and daily subscriptions, equation (29); it is estimated that daily 50% of the users have annual, 20% monthly and 30% daily subscriptions[7].

The benefits consider the supplementary budget (from public or other bodies) to cover any loss resulting from the shortfall between operating costs and the revenue from the daily charge, equation (30). The net present value for this problem must be greater than 0 to ensure a good investment, constraint (31).

The proportion of the demand from *i* to *j* that can be covered is no more than 1, constraint (32). Constraints (33) and (34) state that demand can only be served by installed bike stations.

Finally, equations (35) to (37) specify the domain of the decision variables.

Where it is possible to have smaller zones (census block size for instance) and thus meet the requirements presented above, the model locates none or one station per zone, thereby enabling the decision maker to choose easily where to locate the station in the zone.

However, special care should be taken where the zones defined by the demand study are not small enough to be considered in the terms specified; in other words they are big enough to have more than one station and the need to ride a bicycle within the zones is accepted, using the bike-sharing system. In these cases, the model must not consider $y_i \in \mathbb{N} \ \forall i \in J$ and equation (38) because it

[7] According to the manufacturers in this sector.

limits the number of bicycle stations to one per zone. Therefore, we did not consider this restriction in the case study presented in the next section.

4.4 Case study

The city of Coimbra, Portugal was chosen as a case study for this application. The case study and its demand forecast is presented on chapter 3.3.

It is concluded that 2291 of the daily trips (from a total of 122 253 trips) can be done using bicycles from the bike-sharing system, which is 2% of the trips.

In the application presented below, we assume a more optimistic scenario and demand is 2.5 times the demand defined in this study, thus 5728 of the daily trips can be done using bicycles from the bike-sharing system, which represents 5% of the total of trips.

The traffic zones considered are too large to satisfy the assumption that there are no bike-sharing trips inside the zone. Thus, as stated above, constraint (38) is not considered in this application. Five-time steps are considered in order to adjust to the available data.

The maximum bike capacity of the stations is assumed to be 20 and the minimum 10, because it is usually quite hard for bike-sharing systems to acquire land for stations with a capacity over 25 bicycles.

The cost of a station will depend on the number of docks. The fixed cost is taken to be €3000 with the price increasing by €500 per slot. The unit price of a bicycle considered in our study is €300[8] and other costs assumed were: the fixed relocation cost of €0.1, the variable relocation cost of €0.01 (per bicycle to be relocated), the annual maintenance costs of €100 per station and €50 per bicycle. A discount rate of 5% per year, a growth rate of 2% and a 15-year project horizon were assumed.

In scenario 1, the initial budget is €200 000 and there is no annual supplementary budget. The annual subscription is €40, the monthly subscription €10 and the daily subscription €3.

The optimal solution given by the model is presented in Figure 3. It covers 545 daily trips and locates 12 stations in the blue traffic zones. The zones have between 15 and 40 docks (227 in total) and the fleet has 168 bicycles in the system. The total investment is €199 900 (less than the available budget), the annual expenses are €183 656 and the annual revenue from the fares is €202 874 (more

[8] According to a Portuguese manufacturer (Órbitra), bicycles cost between €250 and €600.

than the annual expenses). Figure 31 shows the location of the stations, the number of docks in each zone and the number of bicycles in the first time step.

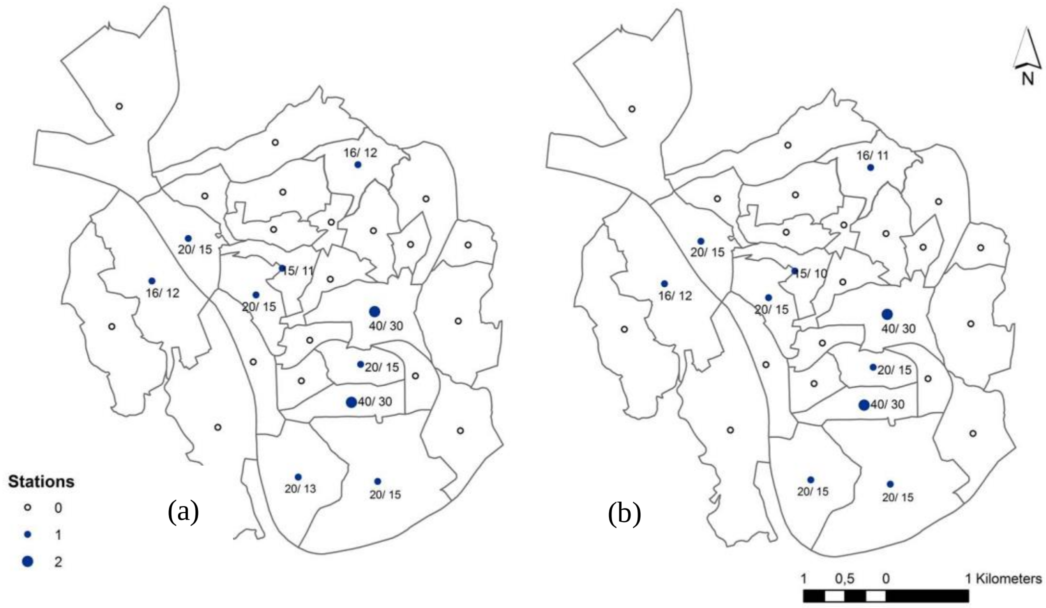


Figure 31 – Solution of scenario 1 (number of docks/number of bicycles in time step 1): (a) time steps 1,2,3 and 5 (b) time step 4

In scenario 2, the previous assumptions are the same but an annual supplementary budget of €50 000 is considered. The best solution, presented in Figure 4, covers 547 daily trips and locates 12 stations in the blue traffic zones, with only one station per zone. The zones have between 15 and 40 docks (227 in total) and there are 168 bicycles in the system, as before. The initial investment is €199 900, the annual expenses amount to €234 667 and the annual revenue from the charges is €203 686. In terms of station location, the solution is the same as for scenario 1.

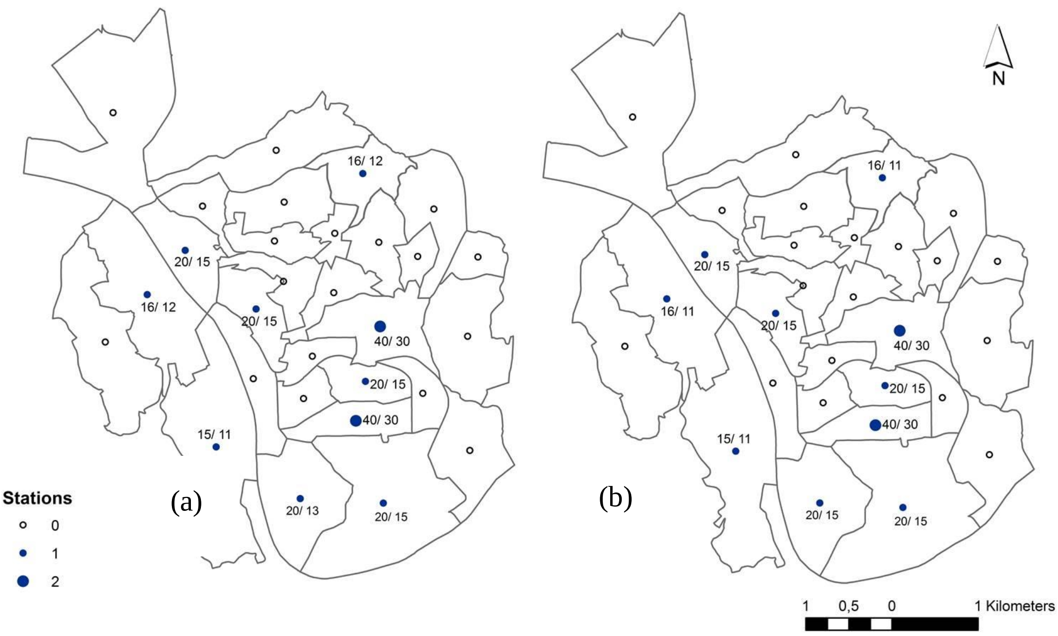


Figure 32 – Solution of scenario 2 (number of docks/number of bicycles in time step 1): (a) time steps 1,2,3 and 5; (b) time step 4

In scenario 3, the initial budget is increased to €300 000, there is no annual supplementary budget and all the other inputs are to the same as for scenario 1. The optimal solution covers 750 daily trips, locates 20 stations, the zones have between 14 and 52 docks (334 in total), and the fleet contains 243 bicycles. The initial investment is €299 900 and the annual expenses are €251 085 with annual revenue from the charges being €279 534. The location of the stations is presented in Figure 33.

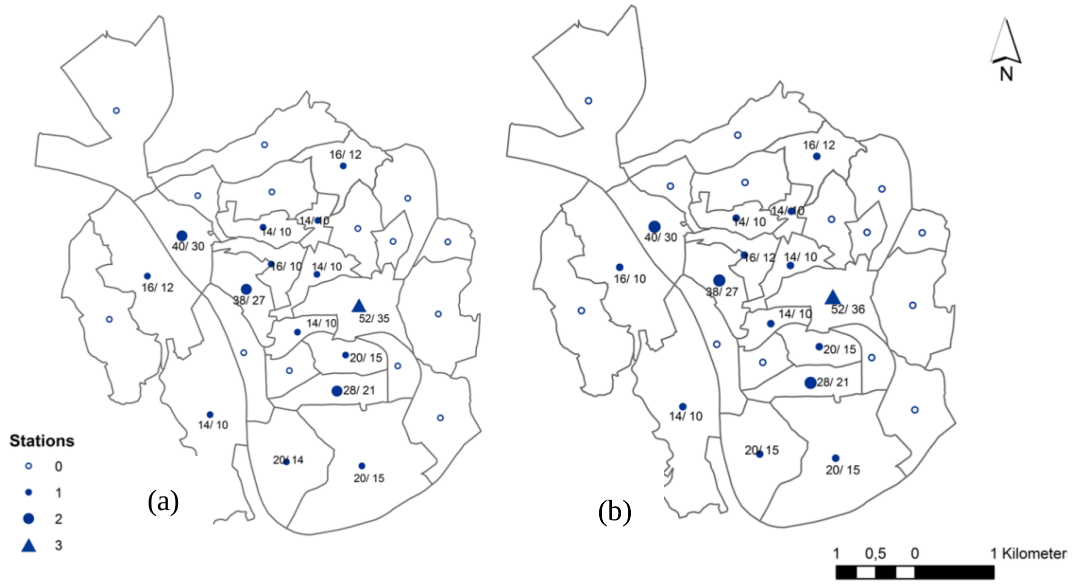


Figure 33 – Solution of scenario 3 (number of docks/number of bicycles in time step 1): (a) time steps 1,2,3 and 5; (b) time step 4

The last scenario (scenario 4) is similar to scenario 3 but the annual supplementary budget is €50 000. The optimal solution shown in Figure 6 covers 757 trips, locates 20 stations and the zones have between 14 and 52 docks with a fleet size of 336 bicycles. The initial investment is equal to the budget (€50 000), the annual expenses are €303 820 and the annual revenue from the charges is €282 160.

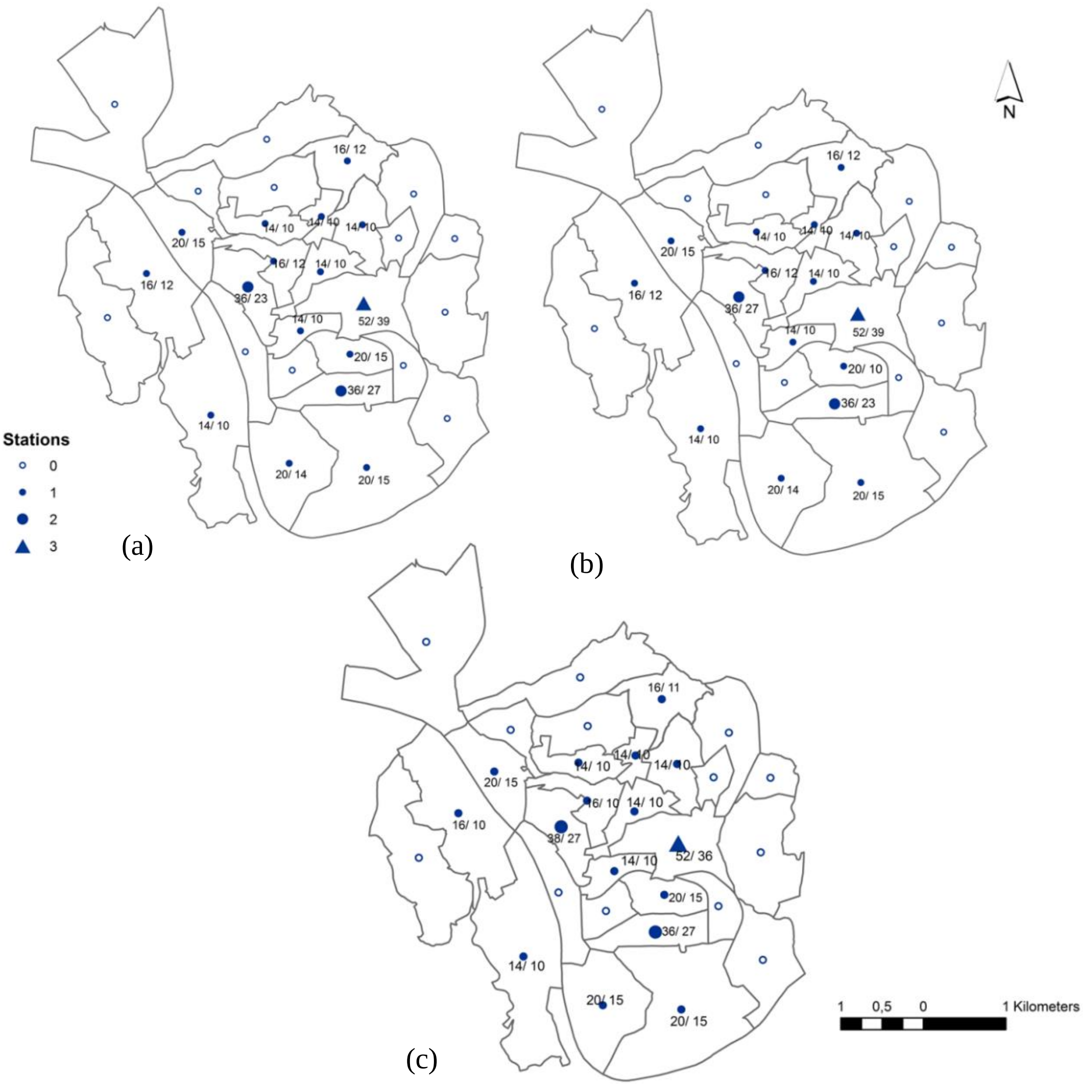


Figure 34 – Solution of scenario 4 (number of docks/number of bicycles in time step 1): (a) time steps 1,2 and 5; (b) time step 3; b) time step 4.

Table 26 summarizes the results obtained for the probed scenarios. We can conclude that by increasing the annual supplementary budget, the covered demand also increases as expected. Moreover, if the initial investment increases the solution covers more trips.

Table 26 - Comparison between scenarios

		Scenario 1	Scenario 2	Scenario 3	Scenario 4
Initial investment		€200 000	€200 000	€300 000	€300 000
Annual supplementary budget		€0	€50 000	€0	€50 000
Optimal solution	**Daily trips covered**	545	547	750	757
	Stations located	12	12	20	20
	Total of docks	227	227	334	336
	Total of bicycles	168	168	243	240
	Total investment	€199 900	€199 900	€299 900	€ 300 000
	Annual expenses	€183 956	€234 667	€251 085	€ 303 820
	Annual revenue	€202 874	€203 686	€279 534	€282 160

The relation between the initial budget and the covered demand is explored by solving several scenarios, with higher initial investment and not considering an annual supplementary budget. The variation results are in Figure 35. As expected, the amount of covered demand increases with an increase in the initial investment.

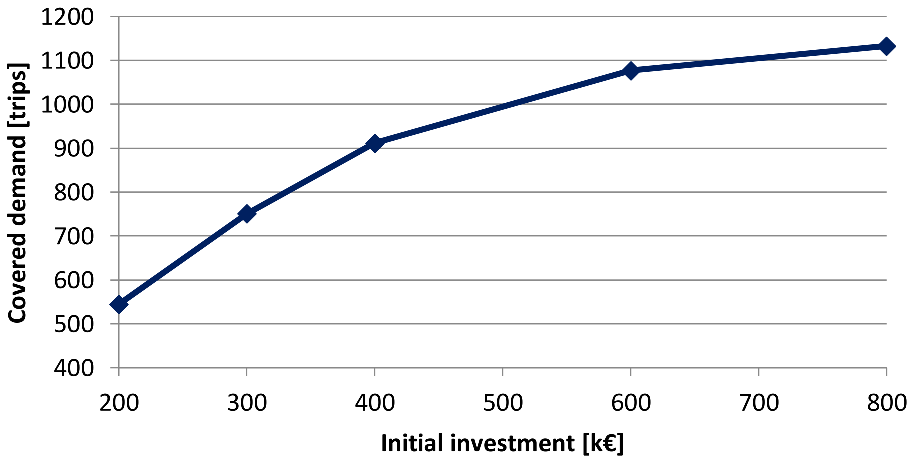


Figure 35 – Relation between the initial investment and the covered demand

Most bike-sharing systems are funded by the government or supported by advertising and sponsorship contracts. A more precise idea of the relations between investment and demand covered can help the decision makers.

4.5 Conclusion

The methodology described in this chapter allows determining the location of bike-sharing stations within an urban area. It takes into consideration the demand estimated using a maximum coverage model that also reflects constraints related to the budget and the level of service. The methodology can provide urban managers with good insights on how to design a bike-sharing system. The possibility to probe different scenarios can help decision-makers to choose the most suitable solution for their town or city. Additionally, no substantial initial investments in these systems should be made if there is no budget is available to sustain them in the long term. Therefore, the balance between initial investment and maintenance costs is included in a bike-sharing system optimization model.

The results of the Coimbra case study demonstrated that the model performs well. It indicates the best zones to locate the bike-sharing stations according to the selected criteria.

The model locates the stations per zone without providing the exact location of the station. For the accuracy in station location, the model admits coordination with a model minimizing the distance between people and stations.

Demand data used was defined based on the methodology presented in chapter 3.3, considering layers of city characteristics that affect the demand. Nevertheless, the location model presented is flexible for other demand studies.

In future research, the demand estimation can use the level of service. This level of service should be a dynamic component of the model estimated by the proximity of the stations, the number of available bicycles, and the cost to the user.

The demand should also have a scalable projection considering the variation of demand of the system, in light of the project horizon. It will also consider the need for inter-modality services, like the demand for bike-sharing services resulting from interaction with public transport services (considering public transport users who could potentially use public bicycles).

Other technological improvements such as electric bicycles and Internet of Things (IoT) frameworks can also represent changes in this base model.

The methodology outlined here can provide urban managers good insight into bike-sharing stations' location within demand and budget initial conditions. Therefore, it contributes significantly to the planning of future bike-sharing systems.

The implementation of these systems may also have other impacts such as the environmental impact on the city. The next chapter presents a methodology to measure the reduction of pollutants in cities considering the traffic reduction consequent of modal shift from motorized vehicles to bike-sharing.

5 BIKE-SHARING ENVIRONMENTAL IMPACTS

5.1 Introduction

Sustainable mobility has become a dominant theme nowadays, due, among other concerns, to the urgency in reducing emissions from transport. As referred, transportation is one of the greatest sources of air pollution in urban areas. Although CO2 emissions are one of the main transport-related environmental problems due to global warming, non-CO2 pollutants must be highlighted due to their harmful effects on human health.

This chapter analyses the implementation of a bike-sharing system (BSS) and its expected implications in emission reductions from road traffic. It focuses on Fine Particulate Matter ($PM_{2.5}$) due to its harmful effects on human health.

The methodology considers an integrated set of models: a) Potential demand estimation models for the use of bike-sharing; b) Mobility studies for the identification of the present travel behavior; c) Optimization models for the location of the bike-sharing stations and d) Traffic emission models, to estimate the reduction in emission due to an expected change in the modal share towards a car to bike-sharing modal shift, considering the present travel behavior.

The work produced intended to support and add value to the design of bike-sharing complementing the demand analysis and location of stations chapters. It was published in the International Journal of Sustainable Transportation (Inês Frade et al., 2021).

On exposure to pollutants subject, it was also made participation on Soft Modes Modeling in Urban Trips project funded by the Portuguese Foundation for Science and Technology (PTDC/ECM-URB/1407/2012).

That resulted in a collaboration on a published paper on Transportation Research Part D: Transport and Environment (Giménez-Gaydou et al., 2019), where is developed a tool that estimates the effort required and the exposure to pollutants on routes.

The chapter is organized into four sections, besides the introduction: section 5.2 presents a literature review on the transport emissions, Section 5.3 details the methodology applied on the evaluation of environmental impacts, the methodology is applied to the city of Coimbra (Portugal) and the results are presented on Section 5.4. and the main conclusions are presented in Section 5.6.

5.2 Literature Review

As referred on Chapter 1, air pollution problems are particularly evident in urban areas. More than 70% of European citizens living in cities in 2018 were exposed to $PM_{2.5}$ annual mean concentrations above the levels recommended by the World Health Organization (EEA, 2020b). Therefore, the definition and evaluation of emission reduction measures are of prime concern.

Transport (the single-use, motorized, internal combustion vehicle) is responsible not only for emissions but also for consuming urban space. This consumption leads to fewer conditions for a good urban life, seen to a broader extent. Therefore, there is an increasing importance given to the promotion of active modes of transportation (walking and cycling).

In European countries in the last two decades, several documents and directives point to the urgent need to promote sustainable transportation as part of the global policy to reduce emissions (as detailed on section 2.3). This urgency resulted in measures such as the implementation of traffic restrictions and the promotion of the modal transfer to more sustainable and active modes: extension of pedestrian areas, an increase of parking fees, cycling initiatives, and the creation of low emission zones, among others.

A cost-benefit analysis on active modes of transport, concerning health benefits and emissions reduction, located in two cities in New Zealand, stated that, following an investment in infrastructures related to pedestrian and cycling modes, the impacts fully justify the investment. Moreover, the benefit/cost ratio (over 10:1) is well in the range to justify the investment involved (Chapman et al., 2018). In a paper by Al-Rijleh et al., using 2012 emissions as a baseline, and by analyzing the traveling behavior of the existing traffic studies, as well as an analysis of the consumption per dwelling, it was possible to establish a set of scenarios of the hypothetical modal transfer. This analysis suggest that an 80% reduction in emissions is technically feasible through a combination of active transportation, cleaner fuels for public transit vehicles, and a significant market penetration of electric vehicles. Walking and cycling modes represent 20% in this 80% reduction (Al-Rijleh et al., 2018).

Moreover, transport policies that potentially lead to emission reductions should consider the combined use of transport and land-use policies. In these impact analyses, it is relevant to include the impact of changes in land-use or the implementation of new transport infrastructures. In this respect, some conclusions taken from a study carried out in New Zealand by Macmillan et al. (Macmillan et al., 2018). In this study, called Te Ara Mua - Future Streets study design, there was an evaluation of walking and cycling increase, searching for dynamic causal linkages between the built environment, local walking and cycling, and wellbeing.

These linkages highlight how policies might help in changing our common healthier future into environmental sustainability. It is fundamental to develop accurate modeling of these policies' benefits (health and environment). Works like the one by Hatfield et al. about non-recreational cycling affects the use of other transport modes or the one by Heine et al. about how travel behavior patterns capture policies' impacts are some of the examples. The health impacts of transportation are becoming a central issue on transportation policies(Hatfield & Boufous, 2016; Heinen et al., 2017)

Raza et al. analyze the health benefits and losses of a modal shift from cars to bicycles. The study quantifies the exposure to pollutants during the cycling trip, using ventilation rate, pollutant concentration, and trip duration data depending on the type of pollutant: particles, black carbon, or nitrogen oxides. In other words, there is a health risk associated with riding a bicycle, but it is essential to understand how exposure to different types of pollutants affects health. Sharply, it is necessary to recalibrate the positive health effects of modal shift with the impacts of inhalation of emissions. Thus, the authors consider an integrated approach in terms of health to properly assess the positive and negative aspects of the modal shift in question (Raza et al., 2018).

Bike-sharing systems have been implemented, fostering the objective of promoting cycling mobility and traffic and emissions reduction. However, there are just a few studies about the impacts of bike-sharing systems on emissions reduction.

The $PM_{2.5}$ impact on health appears in the study by Qiu and He (Qiu & He, 2018). After bike-sharing systems implementation in China, there was a decrease in $PM_{2.5}$ pollutants: 10,35 tons in the total emission amount and 2,5 μg/m3 in concentration (between 2015 and 2017). Furthermore, this paper demonstrates that $PM_{2.5}$ presents the highest impact on health. The reductions in mortalities and hospital admissions induced by $PM_{2.5}$ and PM_{10} are more significant than those from SO_2 and NO_2, confirming the great severity of particulate matter on public health. For example, in the case of mortality, SO_2, NO_2, $PM_{2.5}$, PM_{10} have, respectively, a reduction of health outcomes (number of cases) of less than 25, 59, 116, 87 (Qiu & He, 2018).

Zavala et al. suggest that fewer old cars imply the reduction of $PM_{2.5}$. However, there is no evidence that this impact is higher than the one resulting from the modal shift. In fact, and according to these authors, the effect resulting from the fleet renewal is always smaller than other policies because the congestion levels tend to keep the same (the number of vehicles stays the same). One way to go forward in $PM_{2.5}$ reduction is the modal shift from car to active transportation.

According to these findings, there is evidence that the $PM_{2.5}$ decrease between the existing situation and the situation after implementing a sustainable transport alternative (like a bike-sharing system) is a way to measure these systems' impacts on the environment and health.

It is acceptable to expect a shift between the car and the bicycle for short-distance trips if a bike-sharing system is available. However, it is fundamental to test for the effectiveness of bike-sharing implementations designed to stimulate a shift from car use to cycling. With a growing number of policy tools, projects and investments, its impacts must be verified (Zavala et al., 2013).

The main objective of this study is to estimate pollutants ($PM_{2.5}$) reduction after bike-sharing system implementation. That reduction estimation considers demand, station location, and urban road varying characteristics.

The estimation of a modal transfer impact on transport emissions is a complex process. The transfer from car to bicycle does not imply a straight change leaving aside other network effects like street infrastructure quality, intermodality possibilities, or the influence of possible future bicycle networks or infrastructures. However, a small change can affect air quality and health, and it is relevant to study those impacts alone. Considering that there is potential demand for this mode of transport for the next few years in countries with a small cycling mode share (like Portugal), the impacts on the environment and health also have a great potential to be considered a promotion tool. In other words, although the expected changes in modal shift and emission reduction are still small, they might increase exponentially if measures like bike-sharing systems complement other changes. In any case, a method developed to evaluate emissions reduction can address all magnitudes of changes.

Emission modeling is an alternative method to direct measurements. Different emission models for road traffic are currently available and have two crucial elements: emission factors and activity data, considering different levels of details and complexity (Smit et al., 2010). Those models depend on a large amount of data, and good results depend on their availability and quality (Grote et al., 2016).

Average-speed emission models are an alternative for urban scale studies. In these models, emissions are a function of average speed considering different vehicle technologies and allowing characterization of emissions for each road link within large networks. So, it counts not only with traffic flows but also with street specific characteristics. These models can relate traffic with emissions accounting for non-homogeneity.

Each street segment has its performance in terms of design and average speed, and non-homogeneity results in varying impacts across different streets, even if the traffic reduction is the same (Dias et al., 2018). The next section emphasis the methods used in the estimation of the $PM_{2.5}$ decrease that follows the implementation of a bike-sharing system, associated not with city areas but with streets.

5.3 Method

The emissions reduction is estimated by the traffic reduction caused by the modal shift that happens after a bike-sharing system implementation. Considering that the emissions depends on road characteristics, it is necessary to combine the demand model and a location model for bike-sharing stations according to traffic area characteristics and applying emission models to calculate before and after $PM_{2.5}$ emissions in each street.

The methodology considers an integration between models estimating the potential impact of a bike-sharing system on $PM_{2.5}$ emissions reduction. These models are:

a) Potential Demand Estimation Model for the use of bike-sharing, presented on sub-chapter 3.3;

b) Optimization Model for the Location of the bike-sharing stations, presented on chapter 4, and

c) Traffic Emission Model, to estimate the reduction in emissions due to an expected change in the modal share towards a car to bike-sharing' modal shift, considering the present travel behavior (Tchepel et al., 2012).

The Figure 36 summarizes the methodology applied in this study. Briefly, we follow a set of calculation steps:

- The potential for bike-sharing and the locations of the stations;
- The modal shift expected between car and bike-sharing system and following, the definition of different implementation scenarios;
- The allocation of the new traffic volumes to the road network;
- The emissions balance between the actual situation (with no system implemented) and with a bike-sharing system implemented accordingly with the different scenarios.

The bike-sharing system dimension and its stations' location are crucial aspects in the modal shift to be expected. City councils have reduced budgets and sometimes only a limited number of stations to be placed, and need to use location criteria.

The demand model, presented on section 3.3, estimates the bike-sharing demand in a city divided into small areas (as traffic zones) considering the trip pattern of the population and city characteristics. This pattern appears in an Origin-Destination (OD) matrix (it contains the number of trips between traffic zones).

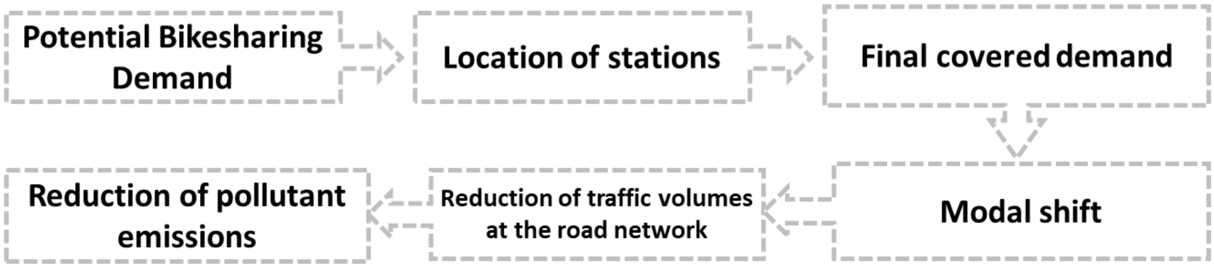


Figure 36 - Process of estimating emissions reduction resulting from the implementation of a Bike-Sharing System

The calculation of the traffic volumes and emission reduction in each road segment, before and after the bike-sharing system, comprehend the following steps:

1. definition of the OD matrix with the individual mode of motorized transport (IT) trips between the traffic zones and allocate the traffic volumes to the road network, relating the number of trips with occupation rate [as is characterization]
2. estimation of the OD matrix for the potential demand for the bike-sharing system,
3. definition of covered demand - through the location of the stations [different scenario building],
4. actualization of the initial OD matrix for individual transport, assuming that all the demand for the bike-sharing system comes from IT trips,
5. actualization of the allocation of IT traffic to the road network considering that the impact of the traffic reduction affects the shortest paths between the centroids of traffic zones [to be situation],
6. The emissions model considers the before and after traffic volumes, analyzing the difference between as-is and to-be scenarios.

Emissions reduction is related to modal shift, street characteristics, and driving conditions. It includes average vehicle speed, as in the Traffic Emission and Energy Consumption Model – Qtraffic (Tchepel et al., 2012). QTraffic is a mesoscopic model based on an average-speed approach following the updated European guidelines for emission factors (Dias et al., 2018). The QTraffic model considers a QGIS environment, a previously developed algorithm for line emission sources, and updated emission factors (Ole Kenneth, 2019). The model estimates emissions and energy consumption at the road segment level and includes detailed information about transport activity.

To estimate atmospheric emissions induced by road traffic, QTraffic requires information on three main sets of input data:

1. The road network of the study area (type, length, and gradient of each road);
2. The vehicle fleet composition (emission reduction technology, engine capacity, engine age, and fuel type – information extracted from national databases on the fleets - ACAP) and
3. Transport activity for each road (traffic volume and average vehicle speed).

This emission model structure is in Figure 37.

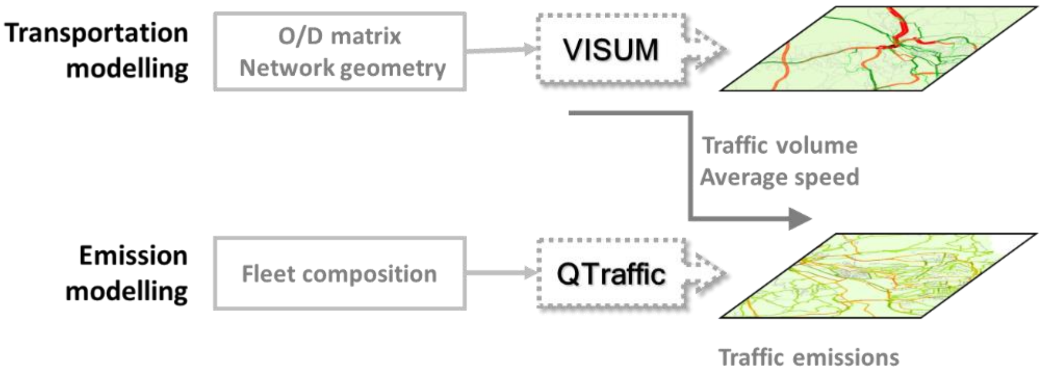


Figure 37 - Transport and emission modelling with QTraffic

The estimation of the traffic-related emissions for each road segment by the model follows the equation:

$$E_i = \sum\nolimits_k (EF_{ik} \times N_k).L \tag{39}$$

Where:
Ei: emissions of the pollutant i [g/km];

EF_{ik}: emission factor [g/km.veh] for pollutant i and vehicle technology k;

N_k : number of vehicles [veh] of technology k;

L: the road segment length [km]

The model provides quantitative information on traffic emissions for different pollutants, and the outputs for the study area directly compatible with GIS allowing spatial data analysis. In this analysis, only the $PM_{2.5}$ is calculated for the network and the correspondent before and after traffic volume.

5.4 Case study and results

As in Chapter 4, the methodology presented is applied to the city of Coimbra, in Portugal.

The initial the OD matrix with the individual mode of transport considered is the OD matrix that resulted from the mobility survey that occurred in Coimbra in 2008 (TIS.pt, 2009).

The OD matrix of potential demand to the bike-sharing system is calculated through the methodology presented on section 3.3. Likely the scenario used on location model (section 4.4) were it was assumed a optimistic scenario and where the initial demand estimated is 2 times the demand defined in the study presented on section 3.3.2, thus 5728 of the daily trips can be done using bicycles from the bike-sharing system, which represents 4% of the total of trips.

The result match related to the recent Portuguese government goals for 2025 (Diário da República, 2019b) that indicates 4% for bicycle modal share in urban areas.

Shortly, the 4% modal share of bicycle travel in cities seems feasible for Coimbra. It is also acceptable to consider that these trips will substitute car trips or IT (individual transportation trips in motorized vehicles) trips and imply emission reduction. Even when other effects are not included - like the net effect created by a network of new infrastructures or the possibilities for intermodality – it is possible to expect a certain number of trips transferred from the car into the bicycle. This modal shift affects the potential emission reduction at the street level, and it is possible to estimate this reduction by calculating the before and after level of emissions. Moreover, without emissions information for every street, we need models to estimate the current level of emissions and their reduction resulting from modal shift.

Assuming an average vehicle occupancy rate 0f 1.42 passengers/vehicle (data from the Coimbra Mobility Survey (TIS.pt, 2009)) it is possible to transform trips into traffic for emission calculations purposes, Figure 38.

Two different scenarios are established for the location of the bike-sharing stations consider the following criteria:

- Scenario 1: Locate stations to cover the zones with three (3) or more OD pairs, with 15 or more potential bike-sharing trips (observation of the number of trips in each OD pair).
- Scenario 2: Locate stations using the maximum covering approach developed in chapter 4

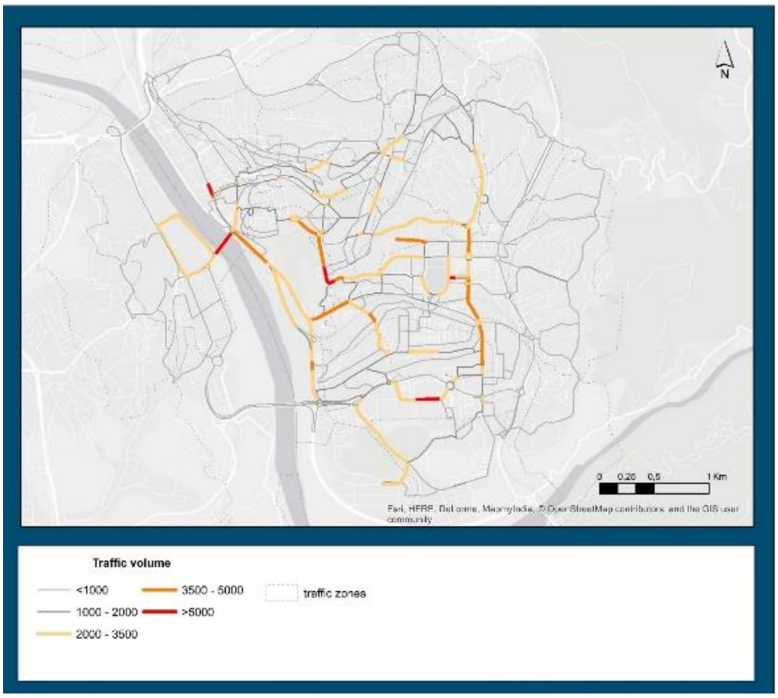
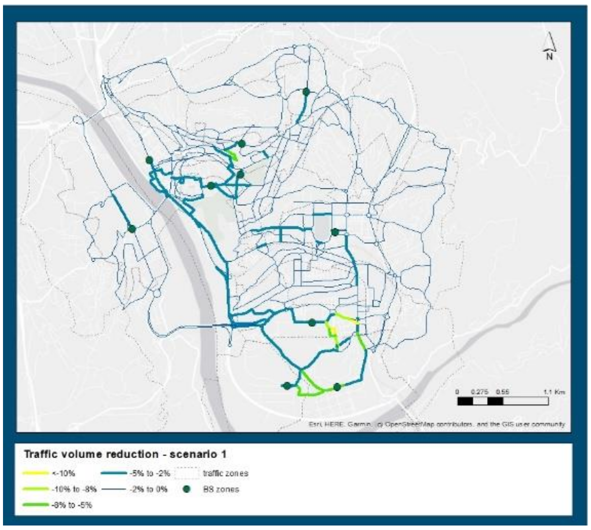


Figure 38 - Daily traffic volume in current situation

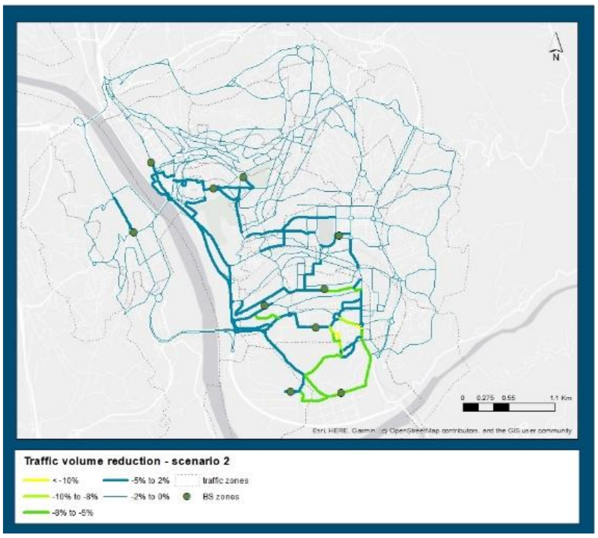
The resulting characteristics on the potential bike-sharing trips are as follows:

- Scenario 1:
    - Zones with bike-sharing stations (some zones can have more than one station depending on the level of demand): 10;
    - Potential bike-sharing trips covered: 816;
    - 1.6% of modal shift from car to bike-sharing resulting in minus 1.5% of IT traffic for the identified paths between served zones.
- Scenario 2:
    - Zones with bike-sharing stations (some zones can have more than one station depending on the level of demand): 10;
    - Potential bike-sharing trips covered: 907;
    - 1.8% of modal shift from car to bike-sharing resulting in minus 1.6% in IT traffic for the identified paths between served zones.

The IT traffic reductions obtained in each scenario and each road are presented in the following figure.



(Scenario 1)



(Scenario 2)

Figure 39 - Traffic reduction obtained for different scenarios of BS systems' location – scenarios 1 and 2.

By the comparison between Figure 39 and Figure 40, it is possible to observe small differences in the modal shift between the two scenarios. The bicycle modal share assumed is also weak (considering the expectations for Portugal 2025 - only 4%). Therefore the impacts are small. However, there are observable differences in the levels of IT reductions in each road, comparing the two scenarios: in scenario 2, there are more streets covered, and Scenario 1 has a higher impact in the north area of the city, whereas Scenario 2 has a higher impact in the south. It demonstrates that demand estimation is a determinant step in this modeling procedure.

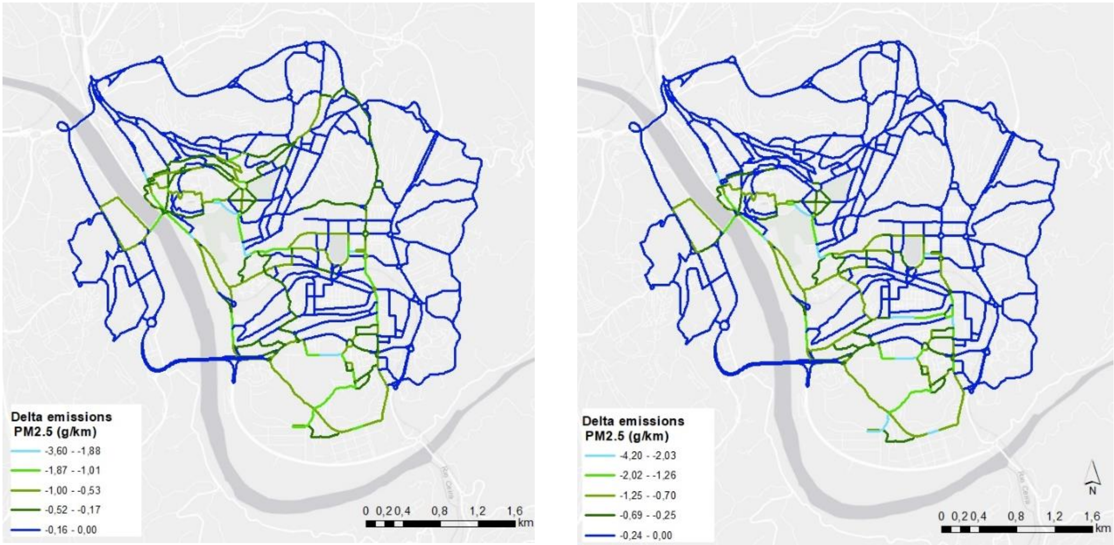
Both scenarios imply a reduction in car trips transferred to the bike-sharing system, consequently reducing emissions. However, these reductions are for the full extension of the paths or roads between the traffic areas. The emissions model considers that these differences impact differently along these paths because it is based on varying street characteristics and expected driving speed. Therefore, the traffic reduction on each road between traffic areas has a varying impact on emissions in a segmented way.

For the emission estimation model, the assumptions and criteria considered are the following.

- The car fleet composition for the study area available in the Portuguese Car Association (Associação Automóvel de Portugal, 2016) and

- The traffic volume and the vehicle speed for each road segment obtained by VISUM (upon information on the road network in the study area and data from a Mobility Study in Coimbra).

The next step is the emissions estimation. This estimation includes traffic levels, street characteristics, and driving conditions (Dias et al., 2018). The results for daily $PM_{2.5}$ emissions in the two scenarios compare with the reference situation (Figure 40).



(Scenario 1) (Scenario 2)

Figure 40 - Reduction of daily emissions for scenario 1 (a) and scenario 2 (b) in comparison with the reference situation

The emissions reductions obtained for Scenario 1 and Scenario 2 are, respectively, 3,6 g/km and 4,2 g/km. It corresponds to about 12,5% of daily $PM_{2.5}$ emissions at the road segment. Globally, the daily $PM_{2.5}$ emissions estimated for the Coimbra study area are about 1.7 kg and implementation of the bike-sharing system will contribute to a decrease of about 1,5% of these emissions.

5.5 Synthesis

A non-homogenous emission reduction obtained for the road network is evident in Figure 40, suggesting the importance of the characterization of transport activity data at the road segment level for emission quantification. This characterization includes street design and vehicle speed, rather than considering the modal shift between car and bike-sharing system and the consequent individual transport volume reduction homogeneous in each street or zone. Moreover, the bike-sharing implementation considered in Scenario 2 uses demand and location models specific for bike-sharing systems dimensioning. It implies a different demand and a minimum number of stations to cover that demand in each zone, responding to budget limitations.

This approach demonstrates that the impact of bike-sharing in emissions varies in each road segment considering both local variations on bike-sharing demand and location, and variations in emissions. Therefore, to increase and optimize this impact, the reinforcement of policies based on studies for the best locations for formal bike-sharing infrastructures and other road adaptations is an essential step.

The modest modal shifts considered and the assumptions (4% of all car trips will become bicycle trips in the bike-sharing systems) imply a modest impact on emissions, as expected. However, in a city where the current use of the bicycle share is 1% as in Coimbra, and almost no data on the use of the bicycle is available, these types of prospective studies are of utmost importance. The application shows the potentialities of the combination between:

- A potential demand model can help to adjust a global modal shift to the characteristics of the OD matrix and the characteristics of each street - and the resulting modal shift between car and bicycle appears in the shortest path that connects the traffic areas;
- An optimization model (and other criteria) that can help to locate the stations where the demand is higher, giving an accurate estimated number of trips covered considering the maximum coverage algorithm and reducing the number of IT trips in each street and
- An emission model that is not only based on traffic volume but also street characteristics and vehicle speed, giving a reduction in emissions with an indirect and more accurate distribution.

This approach and methods can help local authorities deciding which urban roads have a higher potential for the use of a bike-sharing system and therefore need to adapt in terms of infrastructure. It also identifies which urban roads benefit from higher $PM_{2.5}$ reduction, one of the hazard pollutants.

Finally, the obtained results through the combination of modeling tools, say much more than ‘fewer cars mean fewer emissions’ and thus highlighting the importance of modeling analysis.

These models count on trip information from some local mobility studies (from the city of Coimbra). However, this methodology can be applied and implemented anywhere, supporting the decision-makers on where to build the cycling infrastructures if they want to have substantial environmental impacts. This support is relevant in countries where the bicycle modal share is still too low (1%).

After the bike-sharing system implementation, one can expect that there will be trip transferability from car to bicycle. The method can account for this transferability despite the number of trips. This straight transferability from car to bicycle in a bike-sharing system considers that the urban roads can adapt with minor changes, including lower speeds (20 to 30 km per hour). However, the study does not include the impacts on the remaining cycling network or in the intermodality behaviors because it does not exist a formal cycling network in Coimbra.

For later stages of cycling implementation in Portugal, this method must consider the network impact at city-wide mobility and air quality levels.

5.6 Conclusion

This study's main objective is to estimate pollutants ($PM_{2.5}$) reduction after bike-sharing system implementation. That reduction estimation considers demand, station location, and urban road varying characteristics using a combination of modeling approaches from demand estimation to emissions reduction. The case study used was Coimbra in Portugal with modal transfers at still low levels according to Portugal's current stage in bicycle use.

The daily $PM_{2.5}$ emissions estimated for the Coimbra study area are about 1,7 kg. A bike-sharing system will contribute to a decrease of about 1.5% of these emissions, considering wide-central area impacts. However, the spatial distribution of the emissions reductions within the study area is not homogeneous. At the road segment level, it may achieve 12.5% (demonstrating the importance of the modal shift to reduce pollution). Moreover, it shows that new goals and policies, like bike-sharing in articulation with other policy measures, can contribute to these reduction levels increase. One can expect that with higher modal shift ratios, higher reductions will occur.

The link between these models, demand estimation and optimal location on one hand and traffic emission estimation on the other, can be an ally in the decision-making process for local authorities. This approach can support the implementation and articulation of policies, fostering together environmental quality and sustainable mobility. Bike-sharing systems implementation is an example of this articulation.

6 CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

Global warming impacts are forcing a change in the way countries are developing worldwide. The need for changes in mobility patterns, particularly on fossil fuels' dependence, is irrefutable.

The planning and implementation of sustainable alternatives to individual motorized mobility are central in reducing negative externalities related to the transportation sector, mainly in urban areas.

The international and national guidelines include these environmental concerns, focusing policies and investments on the needed modal shift to more sustainable transport modes (including active modes) to have a measurable impact on pollutant emissions.

Therefore, urban mobility plans and implementations must prioritize transport systems and modes that minimize global emissions. Moreover, ensuring the safety of users and the quality of city life.

The challenge of public decision-makers and transport engineers is the implementation of new efficient alternatives and also the change of habits and mobility culture by making this alternatives more attractive, safer and easier to use.

This work is focused, since its beginning, on the planning and design of bike-sharing systems as one of the main tools to operate those changes, namely in cities with low levels of active modes usage.

At a later stage, there was the opportunity to apply the methodologies developed into bike-sharing environmental impacts.

The objectives of this work were accomplished, namely through the publication of some of its results.

The potential bike-sharing demand research results on a methodology that provides a quick assessment of its forecast.

Additionally, the work developed on this subject using the case study of Boston (the HUB bike-sharing system) was an iterative process seeking the most significant local and systems use characteristics influencing the number of bike-sharing trips. Although no unique model justifies all the systems' demand variations, it was possible to find variables that consistently influenced demand. Nevertheless, the process of trying to find a unique model that could include all the significant predictors into one expression was not successful at high levels. Moreover, it shows that some variables are missing from this database and its weakness as a research tool.

The location model proposed uses an optimization method to design the bike-sharing system such that it maximizes the demand covered and takes the available budget as a constraint. It combines strategic decisions for locating bike-sharing stations and defines the dimension of the system (stations and number of bicycles) with operational decisions (bicycles relocation).

The model defines the location of stations bike-sharing stations, the system fleet size, capacity of the stations, and the number of bicycles in each station per period, with economic sustainability concerning since it considers an initial investment lower than the given budget. Additionally, it balances the annual cost of the system and the revenue. It assumes an additional budget from the system provider to cover any loss resulting from the shortfall between its operating cost (and the revenue from the subscription charges). The model is also adaptable to any systems expansion, since it may consider zones already served.

The last chapter complements the research work produced by laying out a methodology to measure the impacts of bike-sharing implementation on $PM_{2.5}$ emissions. The methodology articulates the research work developed in previous chapters and a traffic emissions model. It shows that the impacts are proportional to the modal share but not the same in all urban streets according to the local variations that these models' variables include.

Globally, the research work produced and presented in this thesis is a valuable tool on bike-sharing systems design, addressing the demand, location, and dimension estimation of stations and fleet, as well as measuring its environmental impacts.

6.2 Future work

Considering the conclusions outlined above and the evolution of the bike-sharing system during the last years, one can envisage the following topics as future work for global implementation or each research topic.

1- Integrated transport systems

Besides optimizing the planning and implementation of each transport mode, we need to focus on transport modes integration into optimized intermodal approaches. It is expected some alterations in the mobility patterns caused by the implementation of bike-sharing systems. These services capture users from others transport services such as bus transit, walking, autos, and taxis.

Furthermore, Shaheen et al. (S. A. Shaheen et al., 2011) suggest that bike-sharing acts both as a competitor and intermodal chain element, considering the available modal options.

For this purpose, the design of a bike-sharing system should consider its coordination with the available complementary public transportation services, providing attractive systems for other modes users, as last-mile solutions. Therefore, future demand and location estimation models must include public transport hubs as a predictor, even though it needs further validation through other bike-sharing systems operating.

2- Electric bicycles

Consider a fleet that integrates electric bicycles, as it is increasingly common on implemented systems. This consideration impacts the demand modeling since the negative impact of the slopes is lower, but the system implementation and maintenance costs will increase. The location modeling can also include electric bicycles, lowering some between-stations distances.

3- Environmental impact

The location models can include specific objectives or constraints related to the level of environmental impacts to achieve.

Thus, the location of stations can consider the maximization of demand covered and the environmental impact (through the reduction of emissions).

Moreover, the environmental impacts may constitute a constraint for the implementation and design of bike-sharing. A specific model constraint can help locate where mobility patterns changes are faster to obtain the desirable impacts.

4- Demand studies

Further demand studies on bike-sharing may include a global analysis of different bike-sharing systems databases. It may consider the characterization of different types of trips: commuted, leisure, or utility trips.

Moreover, these studies may include other socio-economic characteristics such as: population density, job density, and retail job density seem to relate more directly with bike-sharing trips. And, certain specific facilities such as tourist attractions, parks, recreational areas, regional transit stations, bicycle-friendly streets, streets with bicycle lanes, and public transit stations are more likely to consider the usefulness of a bike-sharing station nearby.

Finally, topography as a determinant factor must be included (Gregerson et al., 2010), as well as, the combination of trips purpose and its more likely duration.

The deep knowledge about the influent variables on bike-sharing systems demand can support an empirical methodology such as the method presented on section 3.3 that provides a quick and effective assessment of the demand.

5- Global impact of bike-sharing systems

Within a strategic city planning broad approach, bike-sharing systems implementation represents other benefits for the city, not measured by this study. Besides decreasing car trips and emissions, these beneficial impacts are, for example, health benefits to the population, urban space optimization, and, consequently, a general improvement in the quality of city life.

It would be interesting to evaluate the impacts of the systems on other areas that change from the implementation.

7 REFERENCES

AASHTO Executive Committee. (1999). *Guide for the Development of Bicycle Facilities*. American Association of State Highway and Transportation Officials.

Al-Rijleh, M. K., Alam, A., Foti, R., Gurian, P. L., Spatari, S., & Hatzopoulou, M. (2018). Strategies to achieve deep reductions in metropolitan transportation GHG emissions: the case of Philadelphia. *Transportation Planning and Technology*, *41*(8), 797–815. https://doi.org/10.1080/03081060.2018.1526879

An, M., & Chen, M. (2006). Estimating Non-motorized Travel Demand. *Transportation Research Record: Journal of the Transportation Research Board*, *07–2410,* 18–25.

Anselin, L. (1988a). Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. *Geographical Analysis*, *20*(1), 1–17. https://doi.org/10.1111/j.1538-4632.1988.tb00159.x

Anselin, L. (1988b). *Spatial Econometrics: Methods and Models* (Vol. 4). Springer Netherlands. https://doi.org/10.1007/978-94-015-7799-1

ASF - Autoridade de Supervisão de Seguros e Fundos de Pensões. (2020). *PARQUE AUTOMÓVEL SEGURO - Estatísticas*. www.asf.com.pt

Associação Automóvel de Portugal. (2016). *Automotive sales in Portugal (in Portuguese)*. www.acap.pt

Bachand-marleau, J., & Larsen, J. (2011). The much anticipated marriage of cycling and transit : But how will it work ? *TRB 2011 Annual Meeting*, *4869*(October 2010), 1–22.

Baltes, M. (1996). Factors Influencing Nondiscretionary Work Trips by Bicycle Determined from 1990 U.S. Census Metropolitan Statistical Area Data. *Transportation Research Record: Journal of the Transportation Research Board*, *1538*(1), 96–101. https://doi.org/10.3141/1538-13

Barnes, G., & Krizek, K. (2005). Estimating Bicycling Demand. *Transportation Research Record*, *1939*(1), 45–51. https://doi.org/10.3141/1939-06

Bicycling and Transit - a Marriage Unrealized. (2010). *303*, 1–14.

Büttner, J., Mlasowsky, H., & Birkholz, T. (2011). *Optimising Bike Sharing in European Cities*.

Chapman, R., Keall, M., Howden-Chapman, P., Grams, M., Witten, K., Randal, E., & Woodward, A. (2018). A Cost Benefit Analysis of an Active Travel Intervention with Health and Carbon Emission Reduction Benefits. *International Journal of Environmental Research and Public Health*, *15(5)*(962). https://doi.org/https://doi.org/10.3390/ijerph15050962

CHEN, F., TUROŃ, K., KŁOS, M., CZECH, P., PAMUŁA, W., & SIERPIŃSKI, G. (2018). FIFTH GENERATION OF BIKE-SHARING SYSTEMS – EXAMPLES OF POLAND AND CHINA. *Scientific Journal of Silesian University of Technology. Series Transport*, *99*(August), 5–13. https://doi.org/10.20858/sjsutst.2018.99.1

Chen, Z., van Lierop, D., & Ettema, D. (2020). Dockless bike-sharing systems: what are the implications? *Transport Reviews*, *40*(3), 333–353. https://doi.org/10.1080/01441647.2019.1710306

Chow, J. C. (2006). Introduction to the A&WMA 2006 Critical Review—Health Effects of Fine Particulate Air Pollution: Lines that Connect. *J. Air & Waste Manage. Assoc*, *6*(June), 707–708. http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=8&hid=11&sid=7dcd51dc-b458-4e0e-a3a1-02c4140ae95c@sessionmgr110

Church, R. L., & ReVelle, C. R. (1974). The maximal covering location problem. *Papers in Regional Science*, *32*(1), 101–118. https://doi.org/10.1007/BF01434264

CIM Região de Coimbra. (2014). *Estratégia Integrada de Desenvolvimento Territorial da Região de Coimbra 2014-2020*.

CIM Região de Coimbra. (2021). *Estratégia Integrada de Desenvolvimento Territorial da Região de Coimbra 2021-2027*.

Comission of the European Communities. (2007). *Green Paper - Towards a new culture for urban mobility*.

Comission of the European Communities. (2011). *White Paper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system*.

ConBici, C. en D. de la B. (2007). *Guía metodológica para la implantación de sistemas de bicicletas públicas en España*.

Daddio, D. W. (2012). *MAXIMIZING BICYCLE SHARING : AN EMPIRICAL ANALYSIS OF CAPITAL*.

Daskin, M. S. (1995). *Network and Discrete Location*. John Wiley and Sons.

Daskin, M. S. (2008). *What You Should Know About Location Modeling*. *January*. https://doi.org/10.1002/nav

dell'Olio, L., Ibeas, A., & Moura, J. L. (2011). Implementing bike-sharing systems. *Proceedings of the ICE - Municipal Engineer*, *164*(2), 89–101. https://doi.org/10.1680/muen.2011.164.2.89

Deloitte. (2009). *Plano Estratégico de Coimbra*.

DeMaio, P. (2009). Bike-sharing: History, Impacts, Models of Provision, and Future. *Journal of Public Transportation*, *12*(4), 41–56. https://doi.org/10.5038/2375-0901.12.4.3

Diário da República. (2019a). *Resolução do Conselho de Ministros 131/2019, 2019-08-02*. *2030*, 46–81. https://dre.pt/web/guest/home/-/dre/123666113/details/maximized?print_preview=print-preview

Diário da República. (2019b). *Resolução do Conselho de Ministros 131/2019, 2019-08-02*. *2030*, 46–81.

Dias, D., Humberto, J., Sá, E., Borrego, C., Fontes, T., Fernandes, P., Ramos, S., Bandeira, J., Coelho, M. C., & Tchepel, O. (2018). Assessing the importance of transportation activity data for urban emission inventories. *Transportation Research Part D*, *62*, 27–35. https://doi.org/10.1016/j.trd.2018.01.027

Dill, J., & Voros, K. (2007). Factors affecting bicycling demand: initial survey findings from the Portland, Oregon, region. *Transportation Research Record: Journal of the Transportation Research Board*, *2031*, 9–17. http://trb.metapress.com/index/rlq1710417402412.pdf

Direção-Geral do Território. (2019). *Programa Nacional da Política de Ordenamento do Território - Alteração Diagnóstico*. 225.

Duthie, J., Brady, J. F., Mills, A. F., & Machemehl, R. B. (2010). Effects of On-Street Bicycle Facility Configuration on Bicyclist and Motorist Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2190*(1), 37–44. https://doi.org/10.3141/2190-05

EEA. (2019). *Indicator Assessment: Emissions of air pollutants from transport*. available: https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-air-pollutants-8/transport-emissions-of-air-pollutants-8 [Accessed: 18-Jun-2021]

EEA. (2020a). Air quality in Europe - 2020. In *EEA Report* (Issue No 09/2020). https://www.eea.europa.eu//publications/air-quality-in-europe-2020-report

EEA. (2020b). Air quality in Europe - 2020. In *EEA Report* (Issue No 09/2020).

ESRI. (2018a). *Exploratory Regression*. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/exploratory-regression.htm

ESRI. (2018b). *Optimized Hot Spot Analysis*. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/optimized-hot-spot-analysis.htm

ESRI. (2020). *Interpreting Exploratory Regression results*. https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-statistics-toolbox/interpreting-exploratory-regression-results.htm#ESRI_SECTION1_2F7345C39ED6420784953D62979773F3

European Commission. (n.d.). *Eurostat - Transport statistics*. Retrieved March 27, 2021, from ec.europa.eu/eurostat/

European Commission. (1999). Cycling: the way ahead for towns and cities. In *Publications Office of the European Union* (Vol. 32, Issue 8). Publications Office of the European Union. https://doi.org/10.1038/5000250

European Commission. (2010). Health at a Glance 2009. In *Health at a Glance 2009*. OECD. https://doi.org/10.1787/9789264105133-ko

European Commission. (2016). *Towards Low-Emission Mobility. Driving the Modernisation of the EU Economy*. *17*, European Commission.

European Commission. (2017). European Urban Mobility; Policy Context. In *European Union*. https://doi.org/10.2832/827766

European Commission. (2020). *Sustainable and Smart Mobility Strategy – putting European transport on track for the future*.

European Commission. (2021a). *European Climate Law*. https://ec.europa.eu/clima/policies/eu-climate-action/law_en

European Commission. (2021b). *European Climate Law*.

Fishman, E., Washington, S., & Haworth, N. (2014). Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transportation Research Part D: Transport and Environment*, *31*, 13–20. https://doi.org/10.1016/j.trd.2014.05.013

Frade, I., & Ribeiro, A. (2015). Bike-sharing stations: A maximal covering location approach. *Transportation Research Part A: Policy and Practice*, *82*. https://doi.org/10.1016/j.tra.2015.09.014

Frade, Inês, & Ribeiro, A. (2014). Bicycle Sharing Systems Demand. *Procedia - Social and Behavioral Sciences*, *111*, 518–527. https://doi.org/10.1016/j.sbspro.2014.01.085

Frade, Inês, Ribeiro, A., & Correia, G. (2011). Methodology to determine a cycle network in a city. In CITTA 4th Annual Conference on Planning Research (Ed.), *Innovation in Governance and Decision Making in Planning* (pp. 1–21).

Frade, Inês, Ribeiro, A., Dias, D., & Tchepel, O. (2021). Transportation Research Record Bike Sharing Systems Implementation Impact on Emissions for Cyclist Preferred Routes in Urban Areas. *International Journal of Sustainable Transportation*, *0*(0), 1–9. https://doi.org/10.1080/15568318.2021.1949076

García-Palomares, J. C., Gutiérrez, J., & Latorre, M. (2012). Optimizing the location of stations in bike-sharing programs: A GIS approach. *Applied Geography*, *35*(1–2), 235–246. https://doi.org/10.1016/j.apgeog.2012.07.002

Giménez-Gaydou, D. A., Cupido dos Santos, A., Mendes, G., Frade, I., & Ribeiro, A. S. N. (2019). Energy consumption and pollutant exposure estimation for cyclist routes in urban areas. *Transportation Research Part D: Transport and Environment*, *72*, 1–16. https://doi.org/10.1016/j.trd.2019.04.005

Giuliano, G., & Hanson, S. (2017). *The Geography of Urban Transportation*. The Guilford Press.

Grava, S. (2003). *Urban Transportation Systems - Choices for Communities* (McGraw-Hill (Ed.)).

Gregerson, J., Hepp-buchanan, M., Rowe, D., Sluis, J. Vander, Wygonik, E., Xenakis, M., & McCormack, E. (2010). *Seattle Bicycle Share - Feasibility Study*.

Gris Orange Consultant. (2009). *BIKE-SHARING GUIDE*. Transport Canada. www.tc.gc.ca/urban

Grote, M., Williams, I., Preston, J., & Kemp, S. (2016). Including congestion effects in urban road traffic CO2 emissions modelling: Do Local Government Authorities have the right options? *Transportation Research Part D: Transport and Environment*, *43*, 95–106. https://doi.org/10.1016/j.trd.2015.12.010

Gujarati, D. N. (1996). *Basic Econometrics* (Mc Graw Hill (Ed.); 4th Editio).

Handy, S. L., Xing, Y., & Buehler, T. J. (2010). Factors associated with bicycle ownership and use: a study of six small U.S. cities. *Transportation*, *37*(6), 967–985.

https://doi.org/10.1007/s11116-010-9269-x

Handy, S., & Xing, Y. (2011). *Factors Correlated with Bicycle Commuting : A Study in Six Small US Cities Factors Correlated with Bicycle Commuting : A Study in Six Small U . S . Cities*. *January*. https://doi.org/10.1080/15568310903514789

Hatfield, J., & Boufous, S. (2016). The effect of non-recreational transport cycling on use of other transport modes: A cross-sectional on-line survey. *Transportation Research Part A: Policy and Practice*, *92*, 220–231. https://doi.org/10.1016/j.tra.2016.08.011

Heinen, E., Harshfield, A., Panter, J., Mackett, R., & Ogilvie, D. (2017). Does exposure to new transport infrastructure result in modal shifts? Patterns of change in commute mode choices in a four-year quasi-experimental cohort study. *Journal of Transport & Health*, *6*, 396–410. https://doi.org/10.1016/j.jth.2017.07.009

Heinen, E., Maat, K., & Van Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. *Transportation Research Part D: Transport and Environment*, *16*(2), 102–109. https://doi.org/10.1016/j.trd.2010.08.010

Heinen, E., van Wee, B., & Maat, K. (2010). Commuting by Bicycle: An Overview of the Literature. *Transport Reviews*, *30*(1), 59–96. https://doi.org/10.1080/01441640903187001

Hensher, D. A., & Button, K. J. (Eds.). (2000). *Handbook of Transport Modelling* (Volume 1). Pergamon.

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, *34*(4), 453–470. https://doi.org/10.1007/s11116-006-9109-1

IARC Monographs. (2015). *Outdoor Air Pollution - IARC Monographs on the evaluation of carcinogenic risks to humans* (International Agency for Research on Cancer - World Health Organization (Ed.); Vol. 109).

Instituto da Mobilidade e dos Transportes. (2012). *Ciclando. Plano de Promoção da Bicicleta e Outros Modos Suaves 2013-2020*.

International Institute of Social History. (n.d.). *Provo provoceert*. Retrieved July 25, 2021, from http://www.iisg.nl/collections/provo/b24-706-nl.php

ITDP - Institute for Transporteation & Development Policy. (2013). *THE BIKE- SHARE PLANNING GUIDE*.

Krykewycz, G. R., Puchalsky, C. M., Rocks, J., Bonnette, B., & Jaskiewicz, F. (2010). Defining a Primary Market and Estimating Demand for Major Bicycle-Sharing Program in Philadelphia , Pennsylvania. *Transportation Research Record*, 117–124. https://doi.org/10.3141/2143-15

Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, F., Puybonnieux-Texier, V., Quénel, P., Schneider, J., Seethaler, R., Vergnaud, J. C., & Sommer, H. (2000). Public-health impact of outdoor and traffic-related air pollution: A European assessment. *Lancet*, *356*(9232), 795–801. https://doi.org/10.1016/S0140-6736(00)02653-2

LeSage, J., & Pace, R. K. (2009). *Introduction to Spatial Econometrics* (CRC Press (Ed.)).

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., … Ezzati, M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8

Lin, J.-R., & Yang, T.-H. (2011). Strategic design of public bicycle sharing systems with service level constraints. *Transportation Research Part E: Logistics and Transportation Review*, *47*(2), 284–294. https://doi.org/10.1016/j.tre.2010.09.004

Lin, J.-R., Yang, T.-H., & Chang, Y.-C. (2011). A hub location inventory model for bicycle sharing system design: Formulation and solution. *Computers & Industrial Engineering*. https://doi.org/10.1016/j.cie.2011.12.006

Loomis, D., Grosse, Y., Lauby-Secretan, B., Ghissassi, F. El, Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Baan, R., Mattock, H., & Straif, K. (2013). The carcinogenicity of outdoor air pollution. *The Lancet Oncology*, *14*(13), 1262–1263. https://doi.org/10.1016/S1470-2045(13)70487-X

Lu, C.-C. (2013). Robust Multi-period Fleet Allocation Models for Bike-Sharing Systems. *Networks and Spatial Economics*. https://doi.org/10.1007/s11067-013-9203-9

Macmillan, A. K., Mackie, H., Hosking, J. E., Witten, K., Smith, M., Field, A., Woodward, A., Hoskins, R., Stewart, J., Van der Werf, B., & Baas, P. (2018). Controlled before-after intervention study of suburb-wide street changes to increase walking and cycling: Te Ara Mua-Future Streets study design. *BMC Public Health*, *18*(1), 1–13. https://doi.org/10.1186/s12889-018-5758-1

Marleau, J. B.-, Larsen, J., & Geneidy, A. E.-. (2011). CYCLE TRANSIT INTEGRATION AN OPPORTUNITY TO EXPAND SUSTAINABLE TRANSPORT OPTIONS. *Transportation Research*, 1.

Maroco, J. (2010). *Análise Estatística* (Edições Sílabo (Ed.); 3rd Editio).

Martinez, L. M., Caetano, L., Eiró, T., & Cruz, F. (2012). An Optimisation Algorithm to Establish the Location of Stations of a Mixed Fleet Biking System: An Application to the City of Lisbon. *Procedia - Social and Behavioral Sciences*, *54*(1965), 513–524. https://doi.org/10.1016/j.sbspro.2012.09.769

Maurer, L. K., & Maurer, L. K. (2011). *SUITABILITY STUDY FOR A BICYCLE SHARING PROGRAM IN SACRAMENTO , CALIFORNIA*. *859*.

McGinnis, J. M. (1993). Actual Causes of Death in the United States. *JAMA: The Journal of the American Medical Association*, *270*(18), 2207. https://doi.org/10.1001/jama.1993.03510180077038

Mcneil, N. (2011). Bikeability and the Twenty-Minute Neighborhood: How Infrastructure and Destinations Influence Bicycle Accessibility: How Infrastructure and Destinations Influence Bicycle Accessibility. *90th Annual Meeting of the Transportation Research Board*, *July 2010*.

Midgley, P. (2009). The Role of Smart Bike-sharing Systems in Urban Mobility. *Journeys*, *May*, 23–31.

Midgley, P. (2011). Bicycle-Sharing schemes: Enhancing Sustainable Mobility in Urban Areas. In *Global Transport Knowledge Partnership International Road Federation*.

Monzón, A., Rondinella, G., & Equipo Investigador PROBICI. (2010). *PROBICI. Guía de la Movilidad Ciclista. Métodos y técnicas para el fomento de la bicicleta en áreas urbanas*.

Moon-miklaucic, C., Bray-sharpin, A., Lanza, I. D. E. L. A., Khan, A., Re, L. L. O., & Maassen, A. (2019). *WORKING PAPER THE EVOLUTION OF BIKE SHARING : 10 QUESTIONS ON THE EMERGENCE OF NEW TECHNOLOGIES , OPPORTUNITIES , AND RISKS*. *January*.

MOPTC. (2009). *Plano Estratégico De Transportes 2008 - 2020*.

Mueller, N., Rojas-Rueda, D., Cole-Hunter, T., de Naszelle, A., Dons, E., Gerike, R., Gotschi, R., Int Panis, L., Kahlmeier, S., & Nieuwenhuijsen, M. (2015). Health impact assessment of active transportation: A systematic review. *Preventive Medicine*, *76*, 103–114.

New York City Department of City Planning (NYCDCP). (2009). *Bike-Share Opportunities in New*

York City.

Ole Kenneth, N. (2019). EMEP/EEA air pollutant emission inventory guidebook 2013: Technical guidance to prepare national emission inventories. *EEA Technical Report*, *12/2013*, 23.

Ortúzar, J. (2000). Estimating demand for a cycle-way network. *Transportation Research Part A: Policy and Practice*, *34*(5), 353–373. https://doi.org/10.1016/S0965-8564(99)00040-3

Ortúzar, J. de D., & Willumsen, L. G. (2001). *Modelling Transport*. John Wiley & Sons, Ltd.

Our World in Data. (n.d.). Retrieved March 27, 2021, from https://ourworldindata.org/grapher/motor-vehicle-ownership-per-1000-inhabitants

Ozarks Transportation Organization. (2005). Inventory of Bicycle Usage (Demand). In *Bicycle Survey Report*.

Paradis, E. (2011). Moran ' s Autocorrelation Coefficient in Comparative Methods. *ReCALL*, *2*(1), 1–9.

Parkin, J., Wardman, M., & Page, M. (2007). Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation*, *35*(1), 93–109. https://doi.org/10.1007/s11116-007-9137-5

Pincha, J. P. (2018). Nova rede de bicicletas partilhadas esteve na rua menos de um mês. *Público*. https://www.publico.pt/2018/03/06/local/noticia/nova-rede-de-bicicletas-partilhadas-esteve-na-rua-menos-de-um-mes-1805495

Pindyck, R. S., & Rudbinfeld, D. L. (1981). *Econometric Models and Economic Forecasts* (2nd editio). McGraw-Hill, inc.

Pope, C. A., & Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air and Waste Management Association*, *56*(6), 709–742. https://doi.org/10.1080/10473289.2006.10464485

PROBICI team. (2010). *Guía de la Movilidad Ciclista- Métodos y técnicas para el fomento de la bicicleta en áreas urbanas*.

Pucher, J., & Buehler, R. (2008). Making Cycling Irresistible: Lessons from The Netherlands, Denmark and Germany. *Transport Reviews*, *28*(4), 495–528. https://doi.org/10.1080/01441640701806612

Qiu, L.-Y., & He, L.-Y. (2018). Bike Sharing and the Economy, the Environment, and Health-Related Externalities. *Sustainability*, *10*(4), 1145. https://doi.org/10.3390/su10041145

Raviv, T., & Kolka, O. (2013). Optimal inventory management of a bike-sharing station. *IIE Transactions*, *45*(10), 1077–1093. https://doi.org/10.1080/0740817X.2013.770186

Raza, W., Forsberg, B., Johansson, C., & Sommar, J. N. (2018). Air pollution as a risk factor in health impact assessments of a travel mode shift towards cycling. *Global Health Action*, *11*(1), 1429081. https://doi.org/10.1080/16549716.2018.1429081

ReVelle, C. S., & Eiselt, H. a. (2005). Location analysis: A synthesis and survey. *European Journal of Operational Research*, *165*(1), 1–19. https://doi.org/10.1016/j.ejor.2003.11.032

Ribeiro, A., Frade, I., & Correia, G. (2012). Methodology to Measure the Potential od Bicycle Path. *5º Congresso Luso-Brasileiro Para o Planejamento Urbano, Regional, Integrado e Sustentável , PLURIS*, 97–112.

Ribeiro, Anabela. (2008). *As Infra-estruturas Rodoviárias e o Desenvolvimento Regional*. University of Coimbra.

Ribeiro, Anabela, Correia, G., & Frade, I. (2011). *Estudo para a inserção de uma rede de mobilidade ciclável na cidade de Tomar*.

Rietveld, P. (2004). Determinants of bicycle use: do municipal policies matter? *Transportation Research Part A: Policy and Practice*, *38*(7), 531–550. https://doi.org/10.1016/j.tra.2004.05.003

Rixey, R. A. (2012). Station-Level Forecasting of Bike Sharing Ridership: Station Network Effects in Three U.S. Systems. *TRB 2013 Annual Meeting*. https://doi.org/10.3141/2387-06

Romero, J., Ibeas, A., & Moura, J. (2012). A simulation-optimization approach to design efficient systems of bike-sharing. *Procedia-Social and Behavioral Sciences*, *2003*. http://www.sciencedirect.com/science/article/pii/S1877042812042449

Sayarshad, H., Tavassoli, S., & Zhao, F. (2012). A multi-periodic optimization formulation for bike planning and bike utilization. *Applied Mathematical Modelling*, *36*(10), 4944–4951. https://doi.org/10.1016/j.apm.2011.12.032

Schwanen, T., Dijst, M., & Dieleman, F. (2004). Policies for urban form and their impact on travel: the Netherlands experience. *Urban Studies*, *41*(3), 579–603. https://doi.org/10.1080/0042098042000178690

Schwartz, W., Porter, C., Payne, G., Suhrbier, J., Moe, P., & III, W. W. (1999). *Guidebook on Methods to Estimate Non- Motorized Travel : Supporting Documentation* (Issue July).

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), 159–167. https://doi.org/10.3141/2143-20

Shaheen, S. A., Zhang, H., Martin, E., & Guzman, S. (2011). *China ' s Hangzhou Public Bicycle Understanding Early Adoption and Behavioral Response to Bikesharing*. *March 2010*, 33–41. https://doi.org/10.3141/2247-05

Shaheen, S., & Guzman, S. (2011). Worldwide Bikesharing. *Romania*, *1*(300), 3. https://doi.org/10.1007/s13398-014-0173-7.2

Smit, R., Ntziachristos, L., & Boulter, P. (2010). Validation of road vehicle and traffic emission models – A review and meta-analysis. *Atmospheric Environment*, *44*(25), 2943–2953. https://doi.org/10.1016/j.atmosenv.2010.05.022

Songchitruksa, P., & Zeng, X. (2010). Getis-Ord Spatial Statistics to Identify Hot Spots by Using Incident Management Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2165*(December 2010), 42–51. https://doi.org/10.3141/2165-05

Sörensen, K., Vanovermeire, C., & Busschaert, S. (2012). Efficient metaheuristics to solve the intermodal terminal location problem. *Computers & Operations Research*, *39*(9), 2079–2090. https://doi.org/10.1016/j.cor.2011.10.005

Statista. (2020). *Statista - transport & Logistics Data*. www.statista.com

Stephen S Lim‡, Theo Vos, Abraham D Flaxman, Goodarz Danaei, K. S., Heather Adair-Rohani, Markus Amann*, H Ross Anderson*, K. G. A., Martin Aryee*, Charles Atkinson*, Loraine J Bacchus*, Adil N Bahalim*, K., Balakrishnan*, John Balmes*, S. B.-C., Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., … Ezzati, M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8.A

Stinson, M. A., & Bhat, C. R. (2005). *A Comparison of the Route Preferences of Experienced and Inexperienced Bicycle Commuters* (Issue 512).

Svenningsen, U. (2009). *Umebike's Blog*. Community Bike History. https://umebike.wordpress.com/

Taylor, D., & Mahmassani, H. (1996). Analysis of Stated Preferences for Intermodal Bicycle-Transit Interfaces. *Transportation Research Record: Journal of the Transportation Research Board*, *1556*, 86–95. https://doi.org/10.3141/1556-11

Tchepel, O., Dias, D., Ferreira, J., Tavares, R., & Isabel, A. (2012). *Emission modelling of hazardous air pollutants from road transport at urban scale*. *September*. https://doi.org/10.3846/16484142.2012.720277

The World Bank. (n.d.). *The World Bank Data*. Retrieved June 24, 2018, from www.worldbank.org

TIS.pt. (2009). *Inquérito à Mobilidade na Área de Influência do Sistema de Mobilidade do Mondego*.

Transportes & Negócios. (2020a, July). *Bicicletas "Gira" grátis para lisboetas*.

Transportes & Negócios. (2020b, July 10). *Bicicletas "Gira" grátis para lisboetas*. https://www.transportesenegocios.pt/bicicletas-gira-gratis-para-lisboetas/

TRENMO. (2016). *Plano Intermunicipal de Mobilidade e Transportes da Região de Coimbra*.

TRENMO. (2018). *Plano de Ação de Mobilidade Urbana Sustentável (PAMUS) na Comunidade Intermunicial da Região de Coimbra*.

Turner, S., Hottenstein, A., & Shunk, G. (1997). *Bicycle and pedestrian travel demand forecasting: Literature review*. *7*(2).

United Nations. (1998). Kyoto Protocol To the United Nations Framework Kyoto Protocol To the United Nations Framework. In *Review of European Community and International Environmental Law* (Vol. 7). https://doi.org/10.1111/1467-9388.00150

United Nations. (2015). *Paris agreement*.

Vassimon, P. P. De. (2016). *Performance Evaluation for Bike-Sharing Systems : a Benchmarking among 50 Cities*.

Verbeek, M. (2008). A Guide to Modern Econometrics. In *Text*. https://doi.org/10.1017/CBO9781107415324.004

Vogel, P., Greiser, T., & Christian, D. (2011). *Understanding Bike-Sharing Systems using Data Mining : Exploring Activity Patterns*. *20*, 514–523. https://doi.org/10.1016/j.sbspro.2011.08.058

Wang, S., Zhang, J., Liu, L., & Duan, Z. (2010). Bike-Sharing-A new public transportation mode:

State of the practice & prospects. *... and Management Sciences (...*, 222–225. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5563463

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., & Goodman, A. (2014). Health effects of the London bicycle sharing system: health impact modelling study. *BMJ (Clinical Research Ed.)*, *348*(February), g425. https://doi.org/10.1136/bmj.g425

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. *Booksgooglecom*, *58*(2), 752. https://doi.org/10.1515/humr.2003.021

Zavala, M., Barrera, H., Morante, J., & Molina, L. T. (2013). Analysis of model-based PM2.5 emission factors for on-road mobile sources in Mexico. *Atmosfera*, *26*(1), 109–124. https://doi.org/10.1016/S0187-6236(13)71065-8

Zhang, L., Zhang, J., Duan, Z. Y., & Bryde, D. (2015). Sustainable bike-sharing systems: Characteristics and commonalities across cases in urban China. *Journal of Cleaner Production*, *97*, 124–133. https://doi.org/10.1016/j.jclepro.2014.04.006

I. Attachments

```
--------------------------------------------------------------------------------

Table Abbreviations
AdjR2 Adjusted R-Squared
AICc  Akaike's Information Criterion
JB    Jarque-Bera p-value
K(BP) Koenker (BP) Statistic p-value
VIF   Max Variance Inflation Factor
SA    Global Moran's I p-value
Model Variable sign (+/-)
Model Variable significance (* = 0.10; ** = 0.05; *** = 0.01)

--------------------------------------------------------------------------------
```

I.1. Models Outputs

I.1.1. Stations

Trips started – Exploratory regression output

```
******************************************************************************
Choose 1 of 56 Summary
    Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0,41 305,05 0,16  0,02 1,00 0,00  +WWALK***
 0,23 340,27 0,00  0,08 1,00 0,00  +TWL10***
 0,22 342,47 0,00  0,00 1,00 0,00  +NWORKERS***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 2 of 56 Summary
         Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0,58 263,12 0,00  0,82 2,06 0,00  -POP2013***  +WWALK***
 0,57 264,46 0,00  0,74 1,20 0,00  -SERV***  +WWALK***
 0,56 266,81 0,00  0,64 2,02 0,00  -SALEOF***  +WWALK***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 3 of 56 Summary
                 Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0,63 246,63 0,00  0,92 3,60 0,00  -POP2013***  +WWALK***  +W830_859***
 0,62 250,72 0,04  0,79 6,18 0,00  -TOTEMPLOYE***  +WWALK***  +W830_859***
 0,62 250,99 0,00  0,60 4,08 0,00  -POP2013***  +MBUSSCART***  +WWALK***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 4 of 56 Summary
                        Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0,65 242,01 0,00  0,83 13,24 0,00  -I50_149K20***  +WWALK***  -P18YHSH***  +P18YBDH***
 0,64 243,46 0,00  0,93  3,62 0,00  +MBTAPAXE**  -POP2013***  +WWALK***  +W830_859***
 0,64 243,63 0,06  0,37  7,97 0,00  -POP2013***  -I50_149K20**  +WWALK***  +W830_859***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 5 of 56 Summary
                                 Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0,67 235,71 0,00  0,59 13,56 0,05  -I50_149K20***  +WWALK***  +WBIKE***  -W9_1159***  +P18YBDH***
 0,66 236,77 0,00  0,39 13,55 0,01  -I50_149K20***  +WWALK***  +WBIKE***  -P18YHSH***  +P18YBDH***
 0,66 238,57 0,00  0,35 21,03 0,04  -POP2013***  +P25_59Y201***  -I50_149K20***  +MBUSSCART***  +WWALK***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
```

```
************** Exploratory Regression Global Summary (O_TTD_LN) **************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff  Trials # Passed % Passed
             Min Adjusted R-Squared > 0,50 4160123   525631    12,63
            Max Coefficient p-value < 0,05 4160123   195033     4,69
                      Max VIF Value < 7,50 4160123  1822730    43,81
            Min Jarque-Bera p-value > 0,10 4160123   216261     5,20
Min Spatial Autocorrelation p-value > 0,10      18        0     0,00

------------------------------------------------------------------------------
```

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
WWALK	99,95	0,00	100,00
NCONSTMAIN	98,74	100,00	0,00
W830_859	97,73	0,18	99,82
NWORKERS	93,00	0,00	100,00
MBTAPAXE	88,11	0,17	99,83
P18YBDH	87,31	3,06	96,94
TWM60	86,97	99,15	0,85
MBTAS	85,15	6,34	93,66
IM150K2013	84,11	2,81	97,19
TW15_19	83,13	8,05	91,95
WIT	83,03	95,87	4,13
TW10_14	82,71	6,87	93,13
NEMPLOYERS	82,32	5,65	94,35
TWL10	80,93	8,37	91,63
MBUSSCART	78,93	12,32	87,68
TW30_34	78,07	94,13	5,87
VWNV	77,94	11,18	88,82
WATH	76,10	2,79	97,21
W6_629	72,56	97,53	2,47
WTP	71,67	91,04	8,96
SERV	70,21	90,34	9,66
TW35_44	63,14	86,78	13,22
W7_729	62,32	89,86	10,14
W16YM	58,53	41,24	58,76
TOTEMPLOYE	58,47	41,00	59,00
P18YHSH	57,23	31,27	68,73
WNATH	57,11	53,62	46,38
POP2013	57,01	70,86	29,14
VW2V	55,72	85,37	14,63
NFH	55,25	16,80	83,20
TW45_59	49,76	86,84	13,16
I50_149K20	48,94	62,29	37,71
TFAM2013	48,53	82,73	17,27
FH	48,53	82,73	17,27
MALE2013	47,27	58,93	41,07
W8_829	46,97	27,43	72,57
P15_24Y201	44,87	52,39	47,61
W9_1159	44,55	47,49	52,51
P25_59Y201	44,13	56,12	43,88
PRODTRAN	42,95	87,47	12,53
VW1V	42,57	37,87	62,13
TTHSH2013	41,74	48,00	52,00
SALEOF	41,38	62,59	37,41
IL49K2013	38,79	61,58	38,42
BUSSTOPS	37,56	98,52	1,48
W0_459	37,53	86,06	13,94
TW20_24	35,16	28,50	71,50
W5_529	35,03	98,06	1,94
W730_759	32,86	40,95	59,05
WBIKE	29,59	22,21	77,79
P60_74Y201	26,71	45,30	54,70
TW25_29	25,17	46,51	53,49
W630_659	20,12	64,29	35,71
VWM3V	14,90	56,13	43,87
W530_559	6,69	64,75	35,25
WOTH	4,71	42,37	57,63

Trips started – OLS regression GeoDa output

```
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  StationDataBase_LN
Dependent Variable  :    O_TTD_LN  Number of Observations:  131
Mean dependent var  :     3.00127  Number of Variables   :    5
S.D. dependent var  :    0.991276  Degrees of Freedom    :  126

R-squared           :    0.653066  F-statistic           :     59.2954
Adjusted R-squared  :    0.642052  Prob(F-statistic)     :4.57374e-028
Sum squared residual:     44.6589  Log likelihood        :    -115.393
Sigma-square        :    0.354435  Akaike info criterion :     240.787
S.E. of regression  :    0.595345  Schwarz criterion     :     255.163
Sigma-square ML     :    0.340907
S.E of regression ML:    0.583873

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT        5.9554       0.884232        6.73511     0.00000
          MBTApaxE     0.0167674     0.00730834        2.29428     0.02343
           Pop2013      -1.44982       0.168862       -8.58585     0.00000
             Wwalk      0.848112      0.0946916        8.95656     0.00000
          W830_859      0.652311       0.141148        4.62147     0.00001
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   65.123493
TEST ON NORMALITY OF ERRORS
TEST                  DF           VALUE             PROB
Jarque-Bera            2            17.8610          0.00013

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF           VALUE             PROB
Breusch-Pagan test     4             3.0455          0.55024
Koenker-Bassett test   4             1.9009          0.75397
SPECIFICATION ROBUST TEST
TEST                  DF           VALUE             PROB
White                 14            11.7567          0.62584
============================== END OF REPORT ================================
```

Trips ended – Exploratory regression output

```
****************************************************************************
Choose 1 of 56 Summary
    Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.39 313.96 0.17  0.00 1.00 0.00  +WWALK***
 0.23 343.87 0.00  0.00 1.00 0.00  +NWORKERS***
 0.21 346.45 0.00  0.05 1.00 0.00  +TWL10***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
****************************************************************************
Choose 2 of 56 Summary
         Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.56 271.10 0.02  0.77 2.06 0.00  -POP2013***  +WWALK***
 0.55 274.13 0.00  0.83 1.20 0.00  -SERV***  +WWALK***
 0.54 276.35 0.03  0.83 2.06 0.00  -MALE2013***  +WWALK***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
****************************************************************************
Choose 3 of 56 Summary
                Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.61 256.63 0.02  0.97 3.60 0.01  -POP2013***  +WWALK***  +W830_859***
 0.60 259.98 0.00  0.61 4.08 0.00  -POP2013***  +MBUSSCART***  +WWALK***
 0.60 260.01 0.01  0.51 3.08 0.00  -POP2013***  +WWALK***  +WBIKE***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
****************************************************************************
Choose 4 of 56 Summary
                         Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.63 250.90 0.00  0.81 13.24 0.01  -I50_149K20***  +WWALK***  -P18YHSH***  +P18YBDH***
 0.63 251.45 0.00  0.97  3.62 0.00  +MBTAPAXE***  -POP2013***  +WWALK***  +W830_859***
 0.63 252.08 0.00  0.52 13.31 0.02  -P15_24Y201***  -I50_149K20***  +WWALK***  +P18YBDH***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
****************************************************************************
Choose 5 of 56 Summary
                               Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.65 243.79 0.00  0.58 13.56 0.12  -I50_149K20***  +WWALK***  +WBIKE***  -W9_1159***  +P18YBDH***
 0.65 245.95 0.00  0.96  4.33 0.04  +MBTAPAXE***  -POP2013***  +WWALK***  +WBIKE***  +W830_859***
 0.65 246.00 0.00  0.36 13.33 0.00  +NWORKERS***  -POP2013***  -I50_149K20***  +WWALK***  +P18YBDH***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
****************************************************************************
```

```
************* Exploratory Regression Global Summary (D_TTDAY_LN) *************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff  Trials # Passed % Passed
             Min Adjusted R-Squared > 0.50 4160123   429181    10.32
            Max Coefficient p-value < 0.05 4160123   197412     4.75
                      Max VIF Value < 7.50 4160123  1822730    43.81
            Min Jarque-Bera p-value > 0.10 4160123   205848     4.95
Min Spatial Autocorrelation p-value > 0.10      18        1     5.56

------------------------------------------------------------------------------
```

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
WWALK	99.90	0.00	100.00
NCONSTMAIN	97.99	100.00	0.00
W830_859	97.07	0.23	99.77
NWORKERS	95.94	0.00	100.00
MBTAPAXE	91.80	0.10	99.90
MBTAS	89.37	6.45	93.55
TWM60	87.85	99.37	0.63
P18YBDH	87.40	2.88	97.12
NEMPLOYERS	85.46	6.03	93.97
IM150K2013	83.34	2.89	97.11
WIT	82.66	95.96	4.04
TW15_19	81.67	8.50	91.50
TW10_14	81.39	7.04	92.96
TWL10	79.55	8.55	91.45
MBUSSCART	78.32	12.64	87.36
TW30_34	77.47	94.21	5.79
VWNV	76.90	11.36	88.64
WATH	75.34	2.78	97.22
W6_629	74.08	98.04	1.96
WTP	70.15	90.71	9.29
SERV	69.81	90.58	9.42
TW35_44	62.07	86.67	13.33
W7_729	61.28	89.87	10.13
TOTEMPLOYE	57.61	41.59	58.41
W16YM	57.60	41.78	58.22
POP2013	57.51	72.36	27.64
VW2V	56.99	86.81	13.19
P18YHSH	56.71	31.44	68.56
WNATH	56.14	54.07	45.93
NFH	55.83	15.50	84.50
TFAM2013	51.59	85.84	14.16
FH	51.59	85.84	14.16
I50_149K20	49.24	63.28	36.72
TW45_59	48.80	86.44	13.56
MALE2013	47.49	60.78	39.22
W8_829	47.12	25.92	74.08
W9_1159	44.16	48.12	51.88
P25_59Y201	44.15	57.68	42.32
P15_24Y201	43.98	52.96	47.04
PRODTRAN	43.82	88.50	11.50
TTHSH2013	41.57	48.12	51.88
VW1V	41.43	40.53	59.47
SALEOF	40.42	62.01	37.99
IL49K2013	37.88	61.29	38.71
BUSSTOPS	36.26	97.69	2.31
TW20_24	35.85	26.40	73.60
W0_459	35.29	85.64	14.36
W5_529	32.64	97.45	2.55
W730_759	31.80	42.73	57.27
WBIKE	30.09	21.63	78.37
P60_74Y201	26.86	50.44	49.56
TW25_29	24.41	46.93	53.07
W630_659	20.85	65.66	34.34
VWM3V	14.60	55.30	44.70
W530_559	6.07	59.53	40.47
WOTH	5.09	46.56	53.44

Trips ended – OLS regression GeoDa output

```
>>05/23/21 22:30:36
REGRESSION
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  StationDataBase_LN
Dependent Variable  :  D_TTDay_LN  Number of Observations:  131
Mean dependent var  :     2.99562  Number of Variables   :    6
S.D. dependent var  :     1.00354  Degrees of Freedom    :  125

R-squared           :    0.660847  F-statistic           :      48.713
Adjusted R-squared  :    0.647281  Prob(F-statistic)     :9.29464e-028
Sum squared residual:     44.7446  Log likelihood        :    -115.519
Sigma-square        :    0.357957  Akaike info criterion :     243.038
S.E. of regression  :    0.598295  Schwarz criterion     :     260.289
Sigma-square ML     :    0.341562
S.E of regression ML:    0.584433

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT       7.46267       0.989481          7.542     0.00000
          MBTApaxE      0.023273     0.00742059        3.13627     0.00213
           Pop2013      -1.64093       0.179914       -9.12061     0.00000
             Wwalk      0.921901      0.0996639         9.2501     0.00000
             Wbike      0.152269      0.0552088        2.75806     0.00669
          W830_859      0.468953       0.155161        3.02236     0.00304
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   76.134122
TEST ON NORMALITY OF ERRORS
TEST                  DF           VALUE             PROB
Jarque-Bera            2            18.7728          0.00008

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF           VALUE             PROB
Breusch-Pagan test     5             3.4333          0.63350
Koenker-Bassett test   5             2.1463          0.82855
SPECIFICATION ROBUST TEST
TEST                  DF           VALUE             PROB
White                 20            16.2791          0.69916
============================== END OF REPORT ================================
```

I.1.2. Census Tracts

Trips started – Exploratory regression output

```
******************************************************************************
Choose 1 of 56 Summary
    Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0,10 829,92 0,03  0,00 1,00 0,46  +NWORKERS*
 0,08 831,72 0,03  0,00 1,00 0,52  +NEMPLOYERS
 0,05 837,33 0,00  0,10 1,00 0,88  +MBTAPAXE***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 2 of 56 Summary
          Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0,12 827,04 0,00  0,00 1,07 0,48  +MBTAPAXE*  +NWORKERS*
 0,11 828,02 0,02  0,00 1,02 0,20  -W530_559**  +NWORKERS**
 0,11 828,11 0,00  0,00 1,05 0,53  +MBTAPAXE*  +NEMPLOYERS
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 3 of 56 Summary
               Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0,14 824,24 0,01  0,00 4,88 0,12  -WIT***  +VW1V***  +NWORKERS***
 0,14 824,67 0,01  0,00 3,35 0,17  -W630_659***  +VW1V***  +NWORKERS**
 0,14 824,86 0,01  0,00 5,00 0,10  -WIT***  +VW1V***  +NEMPLOYERS**
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 4 of 56 Summary
                        Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0,19 815,03 0,12  0,00  8,65 0,07  -W630_659***  -W7_729***  +VW1V***  +NEMPLOYERS***
 0,18 817,19 0,02  0,00  8,58 0,11  -W630_659***  -W7_729***  +VW1V***  +NWORKERS***
 0,18 818,04 0,06  0,00 13,05 0,08  -W630_659***  -TW20_24***  +VW1V***  +NEMPLOYERS***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 5 of 56 Summary
                            Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0,21 813,54 0,04  0,00  8,76 0,09  -W630_659***  -W7_729***  +VW1V***  +MBTAPAXE  +NEMPLOYERS**
 0,20 814,07 0,04  0,00  8,66 0,10  -W630_659***  -W7_729***  +VW1V***  +MBTAS  +NEMPLOYERS**
 0,20 815,13 0,08  0,00 11,05 0,06  -W630_659**  -W7_729***  +VW1V***  -VW2V  +NEMPLOYERS***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
```

```
************* Exploratory Regression Global Summary (O_TTDAYLN) **************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff  Trials # Passed % Passed
             Min Adjusted R-Squared > 0,50 4163813        0     0,00
            Max Coefficient p-value < 0,05 4163813     7709     0,19
                      Max VIF Value < 7,50 4163813   952563    22,88
            Min Jarque-Bera p-value > 0,10 4163813     2083     0,05
Min Spatial Autocorrelation p-value > 0,10      18        9    50,00

------------------------------------------------------------------------------
```

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
MBTAPAXE	85,09	0,01	99,99
MBTAS	80,73	7,06	92,94
NWORKERS	77,54	0,00	100,00
WWALK	67,93	0,95	99,05
VWNV	62,93	0,36	99,64
W630_659	60,23	99,90	0,10
VW1V	59,91	2,75	97,25
W830_859	59,14	0,78	99,22
NEMPLOYERS	50,95	0,18	99,82
VW2V	43,89	95,13	4,87
NFH	41,76	9,39	90,61
MBUSSCART	39,85	14,15	85,85
TW45_59	36,98	97,53	2,47
W530_559	34,68	100,00	0,00
WIT	33,47	97,56	2,44
TW10_14	29,80	8,61	91,39
W9_1159	25,89	12,67	87,33
TTHSH2013	23,16	35,00	65,00
W8_829	22,97	29,44	70,56
TW35_44	22,30	94,10	5,90
W7_729	21,34	89,47	10,53
P25_59Y201	18,36	76,07	23,93
P18YHSH	16,75	43,98	56,02
P18YBDH	15,60	24,32	75,68
WNATH	14,97	80,02	19,98
IL49K2013	14,66	19,19	80,81
TFAM2013	14,18	70,46	29,54
FH	14,18	70,46	29,54
I50_149K20	14,12	47,98	52,02
VWM3V	13,88	99,97	0,03
W16YM	12,96	76,83	23,17
TW15_19	12,81	37,25	62,75
TW20_24	12,21	59,77	40,23
TOTEMPLOYE	11,50	42,11	57,89
TWM60	10,90	74,38	25,62
W730_759	9,99	37,94	62,06
TW30_34	7,29	70,80	29,20
POP2013	6,27	67,22	32,78
MALE2013	5,75	54,70	45,30
SALEOF	5,45	44,74	55,26
WTP	5,43	55,09	44,91
TW25_29	5,27	37,36	62,64
SERV	5,23	47,27	52,73
P15_24Y201	4,34	28,24	71,76
P60_74Y201	2,78	60,07	39,93
BUSSTOPS	1,88	0,00	100,00
WBIKE	1,06	0,00	100,00
TWL10	0,32	15,01	84,99
IM150K2013	0,13	30,78	69,22
W6_629	0,06	97,80	2,20
NCONSTMAIN	0,03	8,49	91,51
WATH	0,01	7,67	92,33
W0_459	0,00	98,68	1,32
PRODTRAN	0,00	84,19	15,81
WOTH	0,00	0,06	99,94
W5_529	0,00	99,08	0,92

Trips started – OLS regression GeoDa output

Classic model

```
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  BGRIDataBaseDAY_LN2
Dependent Variable  :   O_TTDayLN  Number of Observations:  158
Mean dependent var  :    -3.44864  Number of Variables   :    5
S.D. dependent var  :     3.45972  Degrees of Freedom    :  153

R-squared           :    0.214393  F-statistic           :     10.4385
Adjusted R-squared  :    0.193854  Prob(F-statistic)     :1.67414e-007
Sum squared residual:     1485.75  Log likelihood        :    -401.237
Sigma-square        :     9.71075  Akaike info criterion :     812.475
S.E. of regression  :     3.11621  Schwarz criterion     :     827.788
Sigma-square ML     :     9.40345
S.E of regression ML:      3.0665

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT      -7.50278       0.895757       -8.37591     0.00000
          W630_659     -0.826008       0.224132       -3.68537     0.00032
            W7_729       -1.0253       0.283284       -3.61934     0.00040
              VW1v       1.24392       0.276864        4.49288     0.00001
        NEmployers       1.02944       0.216502        4.75487     0.00000
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   20.482184
TEST ON NORMALITY OF ERRORS
TEST                  DF           VALUE             PROB
Jarque-Bera            2             4.2160          0.12148

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF           VALUE             PROB
Breusch-Pagan test     4            10.2070          0.03708
Koenker-Bassett test   4            16.3774          0.00255
SPECIFICATION ROBUST TEST
TEST                  DF           VALUE             PROB
White                 14           108.6314          0.00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : BGRI_QueenMatrix
   (row-standardized weights)
TEST                          MI/DF        VALUE          PROB
Moran's I (error)            -0.0787       -1.4501        0.14702
Lagrange Multiplier (lag)       1           1.8395        0.17501
Robust LM (lag)                 1           0.0431        0.83549
Lagrange Multiplier (error)     1           2.4480        0.11768
Robust LM (error)               1           0.6516        0.41956
Lagrange Multiplier (SARMA)     2           2.4911        0.28779
============================== END OF REPORT ================================
```

Trips ended – Exploratory regression output

```
******************************************************************************
Choose 1 of 56 Summary
    Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.12 952.92 0.12  0.00 1.00 0.93  +NWORKERS*
 0.10 955.95 0.12  0.00 1.00 0.75  +NEMPLOYERS
 0.08 960.10 0.00  0.55 1.00 0.31  +MBTAPAXE***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 2 of 56 Summary
          Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.16 946.99 0.08  0.00 1.07 0.88  +MBTAPAXE**  +NWORKERS*
 0.15 948.95 0.05  0.00 1.05 0.74  +MBTAPAXE**  +NEMPLOYERS
 0.14 949.32 0.17  0.00 1.02 0.44  -W530_559**  +NWORKERS**
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 3 of 56 Summary
                Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.18 943.22 0.25  0.00 4.88 0.28  -WIT***  +VW1V***  +NWORKERS***
 0.18 943.32 0.04  0.00 4.86 0.09  -WNATH***  +WWALK***  +NWORKERS**
 0.18 943.34 0.07  0.00 1.09 0.50  -W530_559**  +MBTAPAXE**  +NWORKERS*
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 4 of 56 Summary
                       Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.24 933.54 0.74  0.00  8.65 0.16  -W630_659***  -W7_729***  +VW1V***  +NEMPLOYERS***
 0.23 935.15 0.09  0.00 23.46 0.23  -W630_659**  -MBTAS***  +MBTAPAXE***  +NWORKERS***
 0.23 935.33 0.41  0.00  8.58 0.23  -W630_659***  -W7_729***  +VW1V***  +NWORKERS***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 5 of 56 Summary
                              Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.26 929.21 0.40  0.00  8.76 0.21  -W630_659***  -W7_729***  +VW1V***  +MBTAPAXE*  +NEMPLOYERS**
 0.26 929.93 0.13  0.00 24.93 0.02  +WWALK***  -W630_659***  -MBTAS***  +MBTAPAXE***  +NEMPLOYERS**
 0.26 930.07 0.45  0.00  8.66 0.23  -W630_659***  -W7_729***  +VW1V***  +MBTAS*  +NEMPLOYERS**
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
```

Trips ended – OLS regression GeoDa output

Classic model

```
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set           :  BGRIDataBaseDAY_LN2
Dependent Variable :   D_TTDayLN  Number of Observations:  158
Mean dependent var :    -1.81121  Number of Variables   :    6
S.D. dependent var :     5.17343  Degrees of Freedom    :  152
```

```
R-squared           :    0.282350  F-statistic           :     11.9605
Adjusted R-squared  :    0.258743  Prob(F-statistic)     :9.01745e-010
Sum squared residual:     3034.78  Log likelihood        :    -457.661
Sigma-square        :     19.9657  Akaike info criterion :     927.322
S.E. of regression  :     4.46829  Schwarz criterion     :     945.697
Sigma-square ML     :     19.2075
S.E of regression ML:     4.38263

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT      -6.62748         1.4676       -4.51586     0.00001
          W630_659      -1.25987       0.324021       -3.88824     0.00015
            W7_729      -1.71965       0.406557       -4.22979     0.00004
              VW1v       1.99436       0.397102        5.02228     0.00000
             MBTAs      0.249601       0.106016        2.35437     0.01983
        NEmployers       1.54645       0.321416        4.81136     0.00000
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   21.737891
TEST ON NORMALITY OF ERRORS
TEST                  DF           VALUE             PROB
Jarque-Bera            2             1.6157          0.44582

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF           VALUE             PROB
Breusch-Pagan test     5            11.2329          0.04695
Koenker-Bassett test   5            14.9284          0.01067
SPECIFICATION ROBUST TEST
TEST                  DF           VALUE             PROB
White                 20           113.0756          0.00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : BGRI_QueenMatrix
   (row-standardized weights)
TEST                          MI/DF        VALUE          PROB
Moran's I (error)            -0.0693       -1.2465        0.21259
Lagrange Multiplier (lag)       1           1.0567        0.30397
Robust LM (lag)                 1           0.1962        0.65783
Lagrange Multiplier (error)     1           1.8941        0.16874
Robust LM (error)               1           1.0335        0.30933
Lagrange Multiplier (SARMA)     2           2.0902        0.35165
============================== END OF REPORT ================================
```

I.1.3. GRID

Trips started – Exploratory regression output

```
******************************************************************************
Choose 1 of 56 Summary
     Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.29 918.79 0.00  0.01 1.00 0.00  +WWALK_1***
 0.25 926.98 0.00  0.00 1.00 0.00  +VWNV_1***
 0.24 928.80 0.00  0.00 1.00 0.00  +W830_8_860***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 2 of 56 Summary
            Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.44 880.77 0.11  0.01 21.49 0.84  +WWALK_1***  -P18YHSH_1***
 0.42 887.59 0.03  0.01 13.72 0.18  -P25_59Y2_1***  +WWALK_1***
 0.41 888.77 0.05  0.01 28.75 0.10  +WWALK_1***  -W9_11_1160***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 3 of 56 Summary
                   Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.48 871.78 0.19  0.00 22.64 0.93  +WWALK_1***  -P18YHSH_1***  +MBTAPAXE_1***
 0.47 875.41 0.14  0.00 22.69 0.92  +WWALK_1***  -P18YHSH_1***  +MBTAS_1**
 0.46 878.26 0.20  0.01 22.57 0.82  +WWALK_1***  -P18YHSH_1***  +NWORKERS_1**
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 4 of 56 Summary
                          Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.48 870.64 0.23  0.00 23.35 0.86  +WWALK_1***  +WBIKE_1**  -P18YHSH_1***  +MBTAPAXE_1***
 0.48 871.06 0.17  0.00 39.00 0.93  +WTP_1***  +WWALK_1***  -P18YHSH_1***  +MBTAPAXE_1***
 0.48 871.51 0.28  0.00 23.34 0.89  +WWALK_1***  -P18YHSH_1***  +NWORKERS_1  +MBTAPAXE_1***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 5 of 56 Summary
                                   Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)    VIF   SA   Model
 0.50 867.60 0.27  0.00 563.19 0.82  -TTHSH201_1***  +NFH_1***  +WWALK_1***  -P18YHSH_1***  +MBTAPAXE_1***
 0.50 868.23 0.39  0.00  27.64 0.53  -SERV_1***  +WWALK_1***  +W830_8_860***  -P18YBDH_1***  +MBTAPAXE_1***
 0.49 869.10 0.25  0.00 124.51 0.68  -MBUSSCAR_1**  +WWALK_1***  +W830_8_860**  -P18YHSH_1***  +MBTAPAXE_1***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
```

```
************* Exploratory Regression Global Summary (O_TTripD_1) *************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff  Trials # Passed % Passed
             Min Adjusted R-Squared > 0.50 4148046        0     0.00
            Max Coefficient p-value < 0.05 4148046   128037     3.09
                      Max VIF Value < 7.50 4148046   155027     3.74
            Min Jarque-Bera p-value > 0.10 4148046   170708     4.12
Min Spatial Autocorrelation p-value > 0.10      18       15    83.33
```

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
WWALK_1	99.91	0.00	100.00
MBTAPAXE_1	98.20	0.00	100.00
NWORKERS_1	95.25	0.00	100.00
MBTAS_1	94.40	6.86	93.14
VWNV_1	92.88	0.25	99.75
TWL10_1	92.52	7.48	92.52
W830_8_860	89.67	0.92	99.08
WBIKE_1	87.43	0.02	99.98
NEMPLOYE_1	78.64	7.15	92.85
WOTH_1	77.34	0.00	100.00
P15_24Y2_1	66.42	14.98	85.02
TW10_14_15	60.79	6.82	93.18
NFH_1	60.30	9.83	90.17
TWM60_1	58.78	88.45	11.55
WATH_1	57.64	3.34	96.66
WIT_1	57.49	94.65	5.35
W9_11_1160	52.48	21.73	78.27
TW15_19_20	49.53	17.48	82.52
SALEOF_1	49.22	13.40	86.60
P18YHSH_1	47.17	41.04	58.96
TW20_24_25	46.19	17.39	82.61
VW2V_1	45.60	88.43	11.57
P25_59Y2_1	43.97	74.67	25.33
TTHSH201_1	42.77	44.78	55.22
IL49K201_1	42.55	20.83	79.17
TFAM2013_1	42.39	79.59	20.41
FH_1	42.39	79.59	20.41
W7_729_730	40.84	74.39	25.61
MALE2013_1	39.48	30.49	69.51
TW45_59_60	37.18	93.52	6.48
W16YM_1	34.73	46.84	53.16
I50_149K_1	34.46	53.61	46.39
WNATH_1	34.34	47.41	52.59
MBUSSCAR_1	33.39	32.25	67.75
VW1V_1	33.30	39.81	60.19
P18YBDH_1	32.78	35.52	64.48
POP2013_1	31.57	50.14	49.86
W8_829_830	31.29	53.91	46.09
W730_7_760	30.95	30.13	69.87
TOTEMPLO_1	30.55	41.65	58.35
TW25_29_30	30.48	30.87	69.13
WTP_1	30.28	33.03	66.97
TW30_34_35	29.15	69.66	30.34
SERV_1	26.52	69.71	30.29
TW35_44_45	21.61	64.75	35.25
P60_74Y2_1	19.27	34.49	65.51
IM150K20_1	16.82	24.28	75.72
VWM3V_1	12.12	88.14	11.86
PRODTRAN_1	10.58	32.33	67.67
W630_6_660	9.24	62.13	37.87
W5_529_530	7.99	99.23	0.77
W0_459_460	5.35	4.24	95.76
NCONSTMA_1	1.61	84.88	15.12
W6_629_630	0.57	49.67	50.33
BUSSTOPS_1	0.08	5.57	94.43
W530_5_560	0.03	60.13	39.87

Trips started – OLS regression GeoDa output

Classic model

```
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  GRIDDataBaseLN
Dependent Variable  :  O_TTripD_1  Number of Observations:  161
Mean dependent var  :    -2.42506  Number of Variables   :    6
S.D. dependent var  :     4.90278  Degrees of Freedom    :  155

R-squared           :    0.511411  F-statistic           :      32.448
Adjusted R-squared  :    0.495650  Prob(F-statistic)     :1.53082e-022
Sum squared residual:     1890.84  Log likelihood        :    -426.751
Sigma-square        :      12.199  Akaike info criterion :     865.501
S.E. of regression  :      3.4927  Schwarz criterion     :      883.99
Sigma-square ML     :     11.7444
S.E of regression ML:       3.427

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT      -1.26144       0.721238         -1.749     0.08227
            Serv_1      -1.12915       0.303036       -3.72611     0.00027
           Wwalk_1       2.27498       0.323275        7.03727     0.00000
        W830_8_860       1.74613       0.350903        4.97609     0.00000
         P18YBDh_1       -2.1573       0.353786       -6.09776     0.00000
        MBTApaxE_1       0.13196      0.0415349        3.17707     0.00180
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   23.232463
TEST ON NORMALITY OF ERRORS
TEST                  DF           VALUE             PROB
Jarque-Bera            2            1.8784          0.39094

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF           VALUE             PROB
Breusch-Pagan test     5           18.8382          0.00206
Koenker-Bassett test   5           23.9654          0.00022

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : GRID_QueenMatrix
   (row-standardized weights)
TEST                          MI/DF        VALUE          PROB
Moran's I (error)             0.0324        1.2958        0.19503
Lagrange Multiplier (lag)       1           5.8063        0.01597
Robust LM (lag)                 1          11.0044        0.00091
Lagrange Multiplier (error)     1           0.5486        0.45889
Robust LM (error)               1           5.7467        0.01652
Lagrange Multiplier (SARMA)     2          11.5530        0.00310
============================== END OF REPORT ================================
```

Spatial Lag model

```
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : GRIDDataBaseLN
Spatial Weight      : GRID_QueenMatrix
Dependent Variable  :  O_TTripD_1  Number of Observations:  161
Mean dependent var  :    -2.42506  Number of Variables   :    7
S.D. dependent var  :     4.90278  Degrees of Freedom    :  154
Lag coeff.   (Rho)  :    0.301342

R-squared           :    0.537161  Log likelihood        :    -423.598
Sq. Correlation     : -            Akaike info criterion :     861.196
Sigma-square        :     11.1254  Schwarz criterion     :     882.766
```

```
S.E of regression   :     3.33548

-----------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
-----------------------------------------------------------------------------
      W_O_TTripD_1      0.301342         0.1052        2.86447     0.00418
          CONSTANT     -0.735355       0.709189       -1.03689     0.29978
            Serv_1     -0.890396       0.306222       -2.90768     0.00364
           Wwalk_1        1.6309       0.354178        4.60476     0.00000
        W830_8_860       1.36425       0.368187         3.7053     0.00021
         P18YBDh_1      -1.58139       0.376687       -4.19817     0.00003
        MBTApaxE_1      0.116252      0.0400269        2.90435     0.00368
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE        PROB
Breusch-Pagan test                       5        19.2377     0.00174

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : GRID_QueenMatrix
TEST                                     DF      VALUE        PROB
Likelihood Ratio Test                    1         6.3050     0.01204
============================== END OF REPORT ================================
```

Spatial Error model

```
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : GRIDDataBaseLN
Spatial Weight      : GRID_QueenMatrix
Dependent Variable  :  O_TTripD_1  Number of Observations:  161
Mean dependent var  :   -2.425055  Number of Variables   :    6
S.D. dependent var  :    4.902785  Degrees of Freedom    :  155
Lag coeff. (Lambda) :    0.136684

R-squared           :    0.514980  R-squared (BUSE)      : -
Sq. Correlation     : -            Log likelihood        : -426.394313
Sigma-square        :     11.6586  Akaike info criterion :     864.789
S.E of regression   :     3.41447  Schwarz criterion     :     883.277

-----------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
-----------------------------------------------------------------------------
          CONSTANT      -1.23718       0.752708       -1.64364     0.10025
            Serv_1      -1.16098       0.319625       -3.63232     0.00028
           Wwalk_1        2.1389       0.338523        6.31835     0.00000
        W830_8_860       1.78195       0.371085        4.80198     0.00000
         P18YBDh_1      -2.06303       0.368774       -5.59429     0.00000
        MBTApaxE_1      0.129759       0.041464        3.12944     0.00175
            LAMBDA      0.136684       0.133578        1.02325     0.30619
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE        PROB
Breusch-Pagan test                       5        19.1414     0.00181

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : GRID_QueenMatrix
TEST                                     DF      VALUE        PROB
Likelihood Ratio Test                    1         0.7126     0.39858
============================== END OF REPORT ===============================
```

Trips ended – Exploratory regression output

```
******************************************************************************
Choose 1 of 56 Summary
     Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.29 918.94 0.00  0.01 1.00 0.00  +WWALK_1***
 0.25 927.11 0.00  0.00 1.00 0.00  +VWNV_1***
 0.24 928.96 0.00  0.00 1.00 0.00  +W830_8_860***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 2 of 56 Summary
            Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.44 880.77 0.11  0.01 21.49 0.85  +WWALK_1***  -P18YHSH_1***
 0.42 887.60 0.03  0.01 13.72 0.19  -P25_59Y2_1***  +WWALK_1***
 0.41 888.72 0.05  0.02 28.75 0.11  +WWALK_1***  -W9_11_1160***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 3 of 56 Summary
                   Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.48 871.59 0.19  0.00 22.64 0.94  +WWALK_1***  -P18YHSH_1***  +MBTAPAXE_1***
 0.47 875.27 0.14  0.00 22.69 0.93  +WWALK_1***  -P18YHSH_1***  +MBTAS_1**
 0.46 878.15 0.21  0.01 22.57 0.83  +WWALK_1***  -P18YHSH_1***  +NWORKERS_1**
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 4 of 56 Summary
                          Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)   VIF   SA   Model
 0.48 870.53 0.24  0.00 23.35 0.87  +WWALK_1***  +WBIKE_1**  -P18YHSH_1***  +MBTAPAXE_1***
 0.48 870.88 0.17  0.00 39.00 0.92  +WTP_1***  +WWALK_1***  -P18YHSH_1***  +MBTAPAXE_1***
 0.48 871.25 0.28  0.00 23.34 0.90  +WWALK_1***  -P18YHSH_1***  +NWORKERS_1  +MBTAPAXE_1***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************
Choose 5 of 56 Summary
                                   Highest Adjusted R-Squared Results
AdjR2   AICc   JB K(BP)    VIF   SA   Model
 0.50 867.39 0.28  0.00 563.19 0.80  -TTHSH201_1***  +NFH_1***  +WWALK_1***  -P18YHSH_1***  +MBTAPAXE_1***
 0.50 868.13 0.42  0.00  27.64 0.55  -SERV_1***  +WWALK_1***  +W830_8_860***  -P18YBDH_1***  +MBTAPAXE_1***
 0.49 869.01 0.27  0.00 124.51 0.67  -MBUSSCAR_1**  +WWALK_1***  +W830_8_860**  -P18YHSH_1***  +MBTAPAXE_1***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model
******************************************************************************

************* Exploratory Regression Global Summary (D_TTripD_1) *************

               Percentage of Search Criteria Passed
                    Search Criterion Cutoff  Trials # Passed % Passed
             Min Adjusted R-Squared > 0.50 4148046        0     0.00
            Max Coefficient p-value < 0.05 4148046   128346     3.09
                      Max VIF Value < 7.50 4148046   155027     3.74
            Min Jarque-Bera p-value > 0.10 4148046   180958     4.36
Min Spatial Autocorrelation p-value > 0.10       18       15    83.33
```

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
WWALK_1	99.91	0.00	100.00
MBTAPAXE_1	98.28	0.00	100.00
NWORKERS_1	95.77	0.00	100.00
MBTAS_1	94.44	6.86	93.14
VWNV_1	92.91	0.25	99.75
TWL10_1	92.54	7.48	92.52
W830_8_860	89.60	0.95	99.05
WBIKE_1	86.97	0.03	99.97
NEMPLOYE_1	79.80	7.15	92.85
WOTH_1	77.19	0.00	100.00
P15_24Y2_1	66.38	15.00	85.00
TW10_14_15	61.36	6.73	93.27
NFH_1	60.41	9.78	90.22
TWM60_1	59.01	88.60	11.40
WIT_1	57.45	94.63	5.37
WATH_1	57.15	3.44	96.56
W9_11_1160	52.34	21.79	78.21
TW15_19_20	49.39	17.57	82.43
SALEOF_1	49.17	13.40	86.60
P18YHSH_1	47.11	41.08	58.92
TW20_24_25	46.14	17.37	82.63
VW2V_1	45.61	88.44	11.56
P25_59Y2_1	43.94	74.81	25.19
TTHSH201_1	42.82	44.62	55.38
IL49K201_1	42.56	20.81	79.19
TFAM2013_1	42.31	79.60	20.40
FH_1	42.31	79.60	20.40
W7_729_730	40.89	74.49	25.51
MALE2013_1	39.37	30.56	69.44
TW45_59_60	37.90	93.71	6.29
W16YM_1	34.68	46.89	53.11
I50_149K_1	34.40	53.76	46.24
WNATH_1	34.29	47.46	52.54
MBUSSCAR_1	33.30	32.26	67.74
VW1V_1	33.29	39.74	60.26
P18YBDH_1	32.66	35.65	64.35
POP2013_1	31.52	50.24	49.76
W8_829_830	31.29	53.80	46.20
W730_7_760	30.90	30.30	69.70
TOTEMPLO_1	30.53	41.61	58.39
TW25_29_30	30.45	30.85	69.15
WTP_1	30.23	33.09	66.91
TW30_34_35	29.21	69.90	30.10
SERV_1	26.46	69.62	30.38
TW35_44_45	21.61	64.77	35.23
P60_74Y2_1	19.20	34.54	65.46
IM150K20_1	16.58	24.38	75.62
VWM3V_1	11.90	87.73	12.27
PRODTRAN_1	10.38	32.66	67.34
W630_6_660	9.25	62.23	37.77
W5_529_530	7.87	99.21	0.79
W0_459_460	5.19	4.37	95.63
NCONSTMA_1	1.54	84.06	15.94
W6_629_630	0.60	49.21	50.79
BUSSTOPS_1	0.10	5.36	94.64
W530_5_560	0.03	59.74	40.26

Trips ended – OLS regression GeoDa output

Classic model

```
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  GRIDDataBaseLN
Dependent Variable  :  D_TTripD_1  Number of Observations:  161
Mean dependent var  :    -2.42724  Number of Variables   :    6
S.D. dependent var  :     4.90234  Degrees of Freedom    :  155

R-squared           :    0.511648  F-statistic           :     32.4788
Adjusted R-squared  :    0.495894  Prob(F-statistic)     : 1.4754e-022
Sum squared residual:     1889.58  Log likelihood        :    -426.697
Sigma-square        :     12.1909  Akaike info criterion :     865.394
S.E. of regression  :     3.49154  Schwarz criterion     :     883.883
Sigma-square ML     :     11.7365
S.E of regression ML:     3.42586

-----------------------------------------------------------------------------
       Variable      Coefficient      Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
          CONSTANT      -1.24998       0.720998       -1.73369     0.08496
            Serv_1      -1.1253        0.302936       -3.71464     0.00028
           Wwalk_1       2.27767       0.323168        7.04796     0.00000
        W830_8_860       1.74156       0.350786        4.96474     0.00000
         P18YBDh_1      -2.16082       0.353668       -6.10974     0.00000
        MBTApaxE_1       0.133281      0.0415211       3.20996     0.00161
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   23.232463
TEST ON NORMALITY OF ERRORS
TEST                 DF           VALUE             PROB
Jarque-Bera           2             1.7560          0.41562

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                 DF           VALUE             PROB
Breusch-Pagan test    5            18.3998          0.00248
Koenker-Bassett test  5            23.2445          0.00030

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : GRID_QueenMatrix
   (row-standardized weights)
TEST                         MI/DF        VALUE          PROB
Moran's I (error)            0.0318        1.2828        0.19957
Lagrange Multiplier (lag)      1           5.7508        0.01648
Robust LM (lag)                1          10.9819        0.00092
Lagrange Multiplier (error)    1           0.5305        0.46639
Robust LM (error)              1           5.7616        0.01638
Lagrange Multiplier (SARMA)    2          11.5124        0.00316
============================= END OF REPORT ================================
```

Spatial Lag model

```
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : GRIDDataBaseLN
Spatial Weight      : GRID_QueenMatrix
Dependent Variable  :  D_TTripD_1  Number of Observations:  161
Mean dependent var  :    -2.42724  Number of Variables   :    7
S.D. dependent var  :     4.90234  Degrees of Freedom    :  154
Lag coeff.   (Rho)  :    0.299788

R-squared           :    0.537126  Log likelihood        :    -423.576
Sq. Correlation     : -            Akaike info criterion :     861.153
Sigma-square        :     11.1242  Schwarz criterion     :     882.723
```

```
S.E of regression   :      3.3353

-----------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
-----------------------------------------------------------------------------
      W_D_TTripD_1      0.299788       0.105283        2.84745     0.00441
          CONSTANT     -0.727083       0.709148       -1.02529     0.30523
            Serv_1     -0.887424       0.306145        -2.8987     0.00375
           Wwalk_1       1.63674        0.35435        4.61897     0.00000
        W830_8_860       1.36134        0.36807        3.69859     0.00022
         P18YBDh_1      -1.58758       0.376812        -4.2132     0.00003
        MBTApaxE_1      0.117721      0.0400304        2.94078     0.00327
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE        PROB
Breusch-Pagan test                       5        18.7777     0.00211

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : GRID_QueenMatrix
TEST                                     DF      VALUE        PROB
Likelihood Ratio Test                    1         6.2413     0.01248
============================== END OF REPORT ================================
```

Spatial Error model

```
----------
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : GRIDDataBaseLN
Spatial Weight      : GRID_QueenMatrix
Dependent Variable  :  D_TTripD_1  Number of Observations:  161
Mean dependent var  :   -2.427238  Number of Variables   :    6
S.D. dependent var  :    4.902341  Degrees of Freedom    :  155
Lag coeff. (Lambda) :    0.134238

R-squared           :    0.515090  R-squared (BUSE)      : -
Sq. Correlation     : -            Log likelihood        : -426.352895
Sigma-square        :     11.6538  Akaike info criterion :     864.706
S.E of regression   :     3.41377  Schwarz criterion     :     883.194

-----------------------------------------------------------------------------
       Variable       Coefficient     Std.Error       z-value    Probability
-----------------------------------------------------------------------------
          CONSTANT      -1.22631        0.75164       -1.63151     0.10278
            Serv_1       -1.1557       0.319119       -3.62154     0.00029
           Wwalk_1       2.14461       0.338027         6.3445     0.00000
        W830_8_860       1.77577       0.370479        4.79318     0.00000
         P18YBDh_1      -2.06856       0.368268       -5.61699     0.00000
        MBTApaxE_1      0.131246      0.0414412        3.16703     0.00154
            LAMBDA      0.134238       0.133751        1.00364     0.31555
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE        PROB
Breusch-Pagan test                       5        18.6982     0.00219

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : GRID_QueenMatrix
TEST                                     DF      VALUE        PROB
Likelihood Ratio Test                    1         0.6883     0.40675
============================== END OF REPORT ================================
```