Journal of Transport Geography, Volume 80, October. 2019, 102501

This is an author formatted version of this paper

Journal of Transport Geography

To download the published version visit
https://doi.org/10.1016/j.jtrangeo.2019.102501

Comparing bicycling and pedestrian mobility: Patterns of non-motorized human mobility in Greater Boston

Christian Bongiorno[a,b], Daniele Santucci[a,c], Fabio Kon[a,d,*], Paolo Santi[a,e], Carlo Ratti[a]

[a] *MIT Senseable City Lab, USA*
[b] *Politecnico di Torino, Italy*
[c] *Technical University of Munich, Germany*
[d] *University of São Paulo, Brazil*
[e] *IIT, CNR, Italy*

ARTICLE INFO

Keywords:
Urban mobility
Pedestrians
Walking
Cycling
Bike sharing
Non-motorized transportation

ABSTRACT

During the past 100 years, many large cities around the world prioritized individual transportation in cars over more sustainable and healthier modes of transportation. As a result, traffic jams, air pollution, and fatal accidents are a daily reality in most metropolis, in both developed and developing countries. On the other hand, walking and bicycling are effective means of transportation for short to medium distances that offer advantages to both the city environment and the health of its citizens. While there is a large body of research in modeling and analysis of urban mobility based on motorized vehicles, there is much less research focusing on non-motorized vehicles, and almost no research on comparing pedestrian and cyclist behavior. In this paper, we present a detailed quantitative analysis of two datasets, for the same period and location, covering pedestrian and bike sharing mobility. We contrast the mobility patterns in the two modes and discuss their implications. We show how pedestrian and bike mobility are affected by temperature, precipitation and time of day. We also analyze the spatial distribution of non-motorized trips in Greater Boston and characterize the associated network of mobility flows with respect to multiple metrics. This work contributes to a better understanding of the characteristics of non-motorized urban mobility with respect to distance, duration, time of day, spatial distribution, as well as sensitivity to the weather.

1. Introduction

Contemporary urban environments are highly affected by the consequences of car traffic: noise and air pollution, accidents, and fatalities are the most frequent phenomena that are detrimental to the urban experience, worldwide. Minimizing individual motorized mobility in cities is an objective of many municipalities and policy makers, to improve health conditions for citizens and increase environmental quality in urban spaces. One of the most compelling reasons to reduce automobile dependence is health: planning cities for health had been forgotten since the urban sanitarian movement in the mid-nineteenth century (Corburn, 2007). In the past two decades, a renewed interest in the connection between health and cities has arisen from concerns about obesity, physical inactivity, pollution, climate change, and road traffic injuries. Physical inactivity is one of the most important health challenges of the 21st century because of its influence on the most deadly chronic diseases. Therefore, transportation and planning policies promoting active modes of transportation, such as walking and cycling, as alternatives to private motor vehicles can contribute to improve health, with the potential of gaining further co-benefits such as congestion mitigation (De Nazelle et al., 2011). In this context, the concept of the walkable city has been widely investigated from different points of view. One of the most enlightening definitions has been given by Southworth (2005) who states that "Walkability is the extent to which the built environment supports and encourages walking by providing for pedestrian comfort and safety, connecting people with varied destinations within a reasonable amount of time and effort, and offering visual interest in journeys throughout the network." The advantages of a walkable city are numerous, are widely recognized, and affect multiple domains: walkable environments promote social and cultural inclusion, guarantee accessibility and safety, are linked to a higher attractiveness of urban spaces and better environmental quality. Numerous studies have analyzed correspondences between urban morphology and walking behavior (Handy et al., 2002; Forsyth et al.,

* Corresponding author at: MIT Senseable City Lab, 77 Massachusetts Avenue, Building 9-250, Cambridge, MA 02139, USA.
E-mail address: kon@ime.usp.br (Fabio Kon).

https://doi.org/10.1016/j.jtrangeo.2019.102501
Received 29 January 2019; Received in revised form 26 June 2019; Accepted 19 August 2019
0966-6923/

2008; Forsyth, 2015), highlighting the influence of safety, vegetation, compactness, and diversity. These studies have employed GIS datasets that describe urban characteristics and, more recently, datasets that collect walking behavior.

Nevertheless, even in appropriate environments, walking is severely limited by the distance it can cover. To close the gap between pedestrian and automobile mobility, bicycles represent a promising active mode of transportation. In recent years, the increasing presence of bike sharing systems in Europe, Asia, and the Americas has addressed the growing demand for non-motorized mobility. This increase in cycling not only fosters the foundations of a sustainable city, but it also brings crucial benefits for health and well being; it both replaces vehicle use (Rojas-Rueda et al., 2012) and promotes physical activity. Bike sharing systems also generate a large amount of data that enables quantitative studies of bicycle-based mobility. Bike sharing has grown rapidly in the past decade, since the first large-scale system was deployed in Paris in 2007. Currently, over 400 cities in the Americas, Europe, and Asia have at least one Bike Sharing System with the number of available bikes worldwide growing every month.[1]

Data from the USDOT 2017 National Household Travel Survey indicate that 35% of car trips in the USA were shorter than 2 miles, and almost half of them were less than 3 miles (USDOT, 2018). This is a distance that, in several situations, could be covered with either walking or cycling. The same survey shows that only 10.5% of the trips use walking as its mode while only 1% of the trips are based on bicycles. Thus, at least in the USA, there is a large room for improvement towards a larger share of non-motorized mobility modes, with its associated environmental and health benefits.

Several studies have already illustrated potentials of bike use and walking (Pucher et al., 2011; Griffin et al., 2014). The novelty of the present study is to perform a quantitative comparison of both mobility modes, in the same spatial and temporal domain. To this end, we employed datasets for the Greater Boston area covering the same period (May 2014 to May 2015). The analyzed area consists of approximately of 200 km^2 and total population of approximately 900,000 in 2016. Due to its compact structure, Boston is one of the most walkable cities in the United States (Vanky et al., 2017) and, also, a relatively bike-friendly city, having received a silver-medal award (LAB, 2018) from the League of American Bicyclists in 2017.

The anonymized human trace data was collected from an activity-tracking mobile phone application. The data, which includes about 260 thousand pedestrian trips of nearly 6 thousand users from May 2014 to May 2015, record GPS locations and walking behavior of individuals in the Boston metropolitan area.

Boston's pioneering bike-sharing system, Hubway, was launched in 2011 and it has been growing since then. In 2018, its name changed to BlueBikes and it now has over 1800 bicycles and 308 dock stations across Boston, Brookline, Cambridge, and Somerville. For the analysis described in this paper, we utilized open data from the Hubway system, which describes the origin, destination, and timestamps of each individual trip in the period under study, totaling nearly 800 thousand trips.

The main contribution of this paper is a novel data-driven analysis and comparison method between pedestrian and cycling mobility datasets. The method is illustrated with two datasets covering the same area and the same one-year period. We discuss the similarities and differences between the two modes of transportation, showing that they serve different, complementary purposes. We identify and quantify how non-motorized mobility is affected by weather, time, distance, and duration and characterize its spatial mobility flows in terms of network metrics. These findings can serve both as evidence for public policy and as a basis for further research in the field.

In the next section, we discuss related work on the analysis of pedestrian and cycling data. Then, we describe our datasets and methods. The following section presents a comparative analysis of pedestrian and cycling mobility patterns with respect to distance, duration, time of day, geospatial flows, and sensitivity to weather conditions. We finish by presenting our conclusions and discussing future work.

[1] See http://bikes.oobrien.com for a list of active bike sharing systems.

2. Related work

Combining bike and pedestrian data offers the opportunity to evaluate non-motorized mobility in a wider perspective, considering not only the recreational character of these modes of transportation. Unlike most vigorous physical activities engaged in for health or recreational purposes, walking and cycling can be undertaken for multiple purposes. Walking and cycling can be done for leisure, recreation, or exercise; for occupational purposes; and for basic transportation, including shopping or going to work (Saelens et al., 2003).

In fact, besides increasing physical activity, biking and walking are basic modes of transportation, particularly in dense urban environments. Rietveld (2001) points out that “transport statistics are usually formulated in terms of ‘main’ transport mode. This leads to a systematic underestimation of non-motorized transport modes. Even in the case of car trips, walking to and from the parking place is an inevitable element of the chain. The same holds true for walking and biking to the bus stop or the railway station. A consequence of this complementarity is that when the various trip elements are considered, the share of bike and walking is much higher.”

Still according to Rietveld (2001), in multimodal chains, pedestrians dominate the scene and bicycles are rather important in particular in train related public transport trips. Bike mobility is relevant for the activities at the end of train routes; non-motorized transport modes can be faster than bus or tram, especially when aspects like rescheduling costs and uncertainty costs are taken into account: because of their time-continuous character, these non-motorized modes do not give rise to the risk of missing a connection in a chain. Furthermore, individual characteristics such as age, income, and physical abilities play a role and income can be an essential feature to include in the evaluation.

Another frequently mentioned factor is infrastructure: bicycle paths may be essential to improve the convenience and safety of bicycle trips (Pucher et al., 1999) and could contribute to the community as a whole over the longer term. For example, a new bicycle path, if extensively used, may prompt or stimulate development and increased private investment (Krizek et al., 2009). Factors such as topography and points of interest also influence the “bikeability” of a certain region (Winters et al., 2013). Finally, social influences (family and neighbors) also have a significant impact on the willingness to walk and to cycle (Pikora et al., 2003).

Despite their increasingly recognized potential as a solution to several pressing problems, walking and cycling remain the most understudied – and least understood – modes of travel. Complicating the study of walking and cycling as modes of transportation is their frequent use for exercise and recreation rather than for travel (Krizek et al., 2009), mainly in the past decades. Despite individual preferences, replacing motorized mobility by cycling and walk brings up the necessity of considering these modes as part of the complex network of urban mobility. Because transport demands are diverse, increasing transport system diversity tends to increase efficiency and equity by allowing each mode to be used for what it does best (Litman, 2015). Conventional transportation analysis (and policy) often groups walking and cycling together as non-motorized modes, thereby implying that such travel serves similar purposes and markets. Both activities are human powered and entail greater direct exposure to environmental conditions than transit or cars (Krizek et al., 2009). Nevertheless, our studies help unveiling that, despite being non-motorized, cycling and walking present substantially different features and, as such, should be considered distinct transport modes.

Recent related works have used similar pedestrian datasets to understand human mobility patterns at a high resolution, both in a temporal and spatial scale. Vanky et al. (2017) examined the association between meteorological (weather) conditions in a given location and pedestrian trips frequency and duration with the use of geolocated digital data. These associations were determined for seasonality, urban microclimate, and commuting. While previous studies were often constrained to small spatial units of analysis, the increasing ubiquity of mobile devices offers opportunities to obtain new data to understand human activity. Leveraging these data offers an unprecedented opportunity to understand human mobility patterns at a substantial temporal and spatial scale, with a level of detail heretofore unavailable. Santucci et al. (2018) uses presence as an indicator for walkability by relating it to additional layers to provide an accurate model of the urban morphology. Their case study presents a methodology on how human walking activities can be sensed, quantified, and applied to determine the impact of the urban morphology and its effects on climate at a micro-scale. The study also reveals how people flows react to highly fluctuating microclimatic conditions and how pedestrians respond to the variability of the urban environment. Quercia et al. (2014) proposed an alternative algorithm to the shortest distance path to identify the most walkable routes based on qualitative measures of environmental perception extracted from Flickr metadata, validating their work with a controlled experiment with 84 participants.

There is a large body of research analyzing bike sharing data since providers started to release its data a few years ago. Most previous work on bike sharing data analysis focuses on usage patterns of individual bicycle dock stations (O'Brien et al., 2014; Faghih-Imani et al., 2014; Sarkar et al., 2015; Wang et al., 2016). Research that focused on mobility patterns many times focuses on the flows from one individual dock station to the other (Corcoran et al., 2014; Zhang et al., 2018). The problem is that stations are normally distributed unevenly across the city and each individual station is a too fine unit of analysis that does not provide an overall picture of city mobility dynamics. In a more recent study, Zhou (2015) used a clustering algorithm to group together trips connecting dock stations in Chicago and help identifying spatio-temporal patterns of biking behavior. Austwick et al. (2013) studied the complex network structure of the bike-sharing system on five different cities, and identified similarities in aggregated statistics such as the travel distances or the presence of clusters of dock stations.

To the best of our knowledge, no previous research has performed a comparative analysis of walking and cycling activity as presented in this paper. Similar (non-comparative) analysis in the past were based on small datasets or questionnaire data while our research is based on mathematical analysis of large datasets collected with GPS from real trips. In our research, we not only identify the influence of multiple metrics on walking and cycling, but we also quantify it based on data analysis. Finally, our comparisons were derived from the analyzed data using our flow-based method, which is also not prevalent in the literature.

A recent development in non-motorized mobility is the emergence of scooter-sharing systems, which became popular in 2018 and are now being deployed in a variety of cities. Preliminary work comparing a dockless scooter-share system with a bike-share system was performed by McKenzie (2019). The work is limited to analyzing the time and distance covered by the trips as well as the neighborhoods where they occur, but it already shows significant differences between the two modes, indicating that further research in this direction can be interesting.

3. Datasets and methods

The employed datasets refer to the Greater Boston urbanized area, Massachusetts. From a geographic point of view, the area is characterized by the presence of important natural features such as the coast to the Atlantic Ocean, the Charles River, numerous lakes and orographic variations. The urbanized area has different morphological patterns: while Boston's center has a dense mid to high-rise urban morphology derived from historic paths, the more peripheral areas have low density characteristics; in some cases, they can be associated with those of the urban sprawl. The urbanized area has a polycentric structure, where the centers of the different municipalities (i.e.: Harvard Square in Cambridge, Downtown and Backbay in Boston, Union square in Somerville) constitute the vital nodes.

To establish a connection between the urban geography and the employed data sets, we used an unbiased 100×100 grid that we distributed across the area, overlaying it to the urban structure.

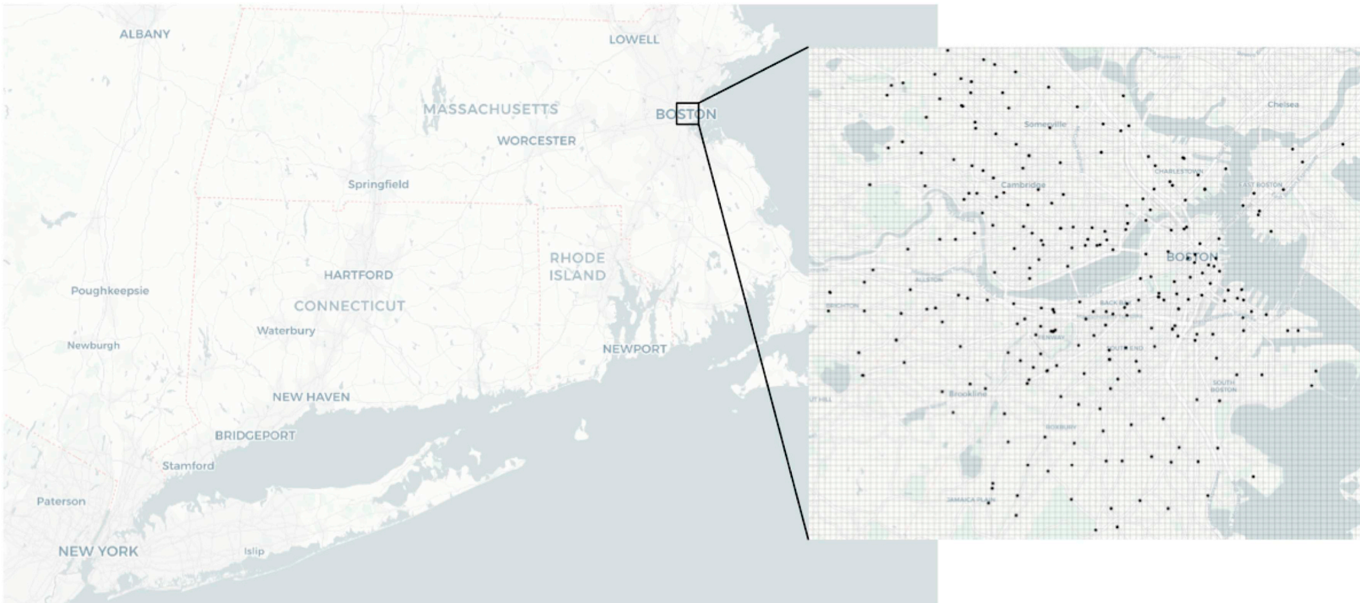
The walking dataset was collected from an activity-oriented mobile phone application. The data, which includes about 260,000 trips of over 6000 anonymous users from May 2014 to May 2015, record GPS locations and walking behavior of anonymous individuals in the Greater Boston area. The data indicates the trip start and ending coordinates with specific timestamps, associated to individual information through a unique user Id. For an additional anonymization, a random distance of up to 100 m was erased from the start and ending of each trip to avoid re-identification through reoccurring visited locations. The walking trajectories consisted originally of a set of geographic points with corresponding timestamps; due to the noise of the original GPS locations, the data were normalized using a map-matching algorithm based on the OpenStreetMap street network to rectify the mistaken trajectories. The Hidden Markov Map Matching algorithm (HMM) was used to match the measured longitudes/latitudes in human trace records to roads (Newson and Krumm, 2009). The HMM algorithm accounts for the GPS noise and the layout of the road network, and matches the GPS locations to corresponding streets with high accuracy. The uniqueness of this dataset consists of its size and resolution, although the tracked walking record cannot be attributed to specific population segments. The data anonymization procedure we applied due to privacy reasons also brings a contribution to literature and highlights the need of specific legislation to regulate this issue.

Bike sharing data was collected from the Bluebikes website, the largest Boston bike-sharing provider.[2] For each trip, we use the location and time of origin and destination. For the analysis involving distances and speed, we estimate the road distance between two bicycle dock stations by using the GraphHopper API (`http://graphhopper.com`) over OpenStreetMap map; in particular, we use the bike mode route planner, which provides bike-friendly routes. The bike routes suggested by the API is around 30% longer than the Euclidean distance, on average.

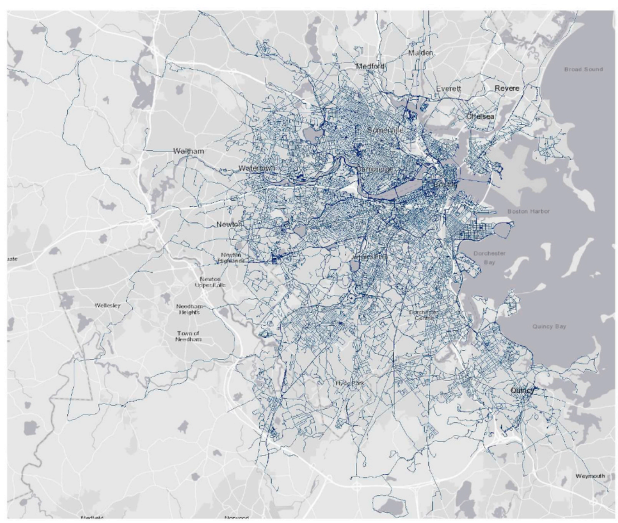
The cycling dataset does not contain a user ID for privacy concerns. The pedestrian dataset does contain a user ID and could support some analysis on the behavior of individual users. However, as the goal of this paper is to compare both mobility modes, we focused our analysis on trips and on the most relevant flows. While their large size brings the benefits of high resolution and coverage of the population, datasets collected through smartphone apps are usually lacking information about users due to privacy regulations. This is also the case with the pedestrian and bike sharing datasets used in this paper. Thus, we cannot attribute the observed mobility patterns to any specific population segment. On the other hand, it is known in the literature that datasets collected through smartphone apps are typically biased towards relatively younger, better educated, and more affluent population.

Both datasets are jointly processed in Jupyter Notebooks leveraging Python libraries for data analysis and geovisualization. To identify the relevant flows, we divide the Greater Boston area using a homogeneous 2-dimensional grid of size 100×100 and compute, for each of the 10,000 possible (origin, destination) pairs, how many trips were performed. Depending on the specific analysis under consideration, we filter the data with respect to different parameters such as trip duration,

[2] `https://www.bluebikes.com/system-data`



(a)



(b)

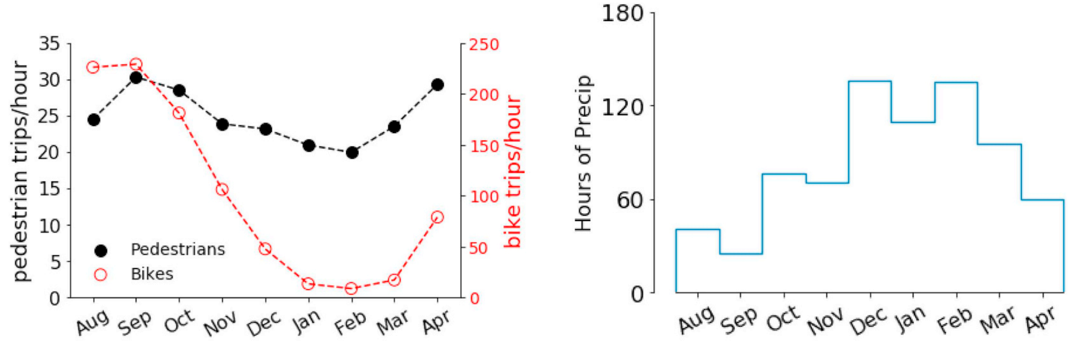
Fig. 1. The area under study: (a) the inset shows the 100 × 100 grid over the Great Boston and the location of the bike sharing stations; the inset covers an area of approximately 200Km^2. (b) pedestrian trajectories cover a wider area of approximately 400 Km^2 but relatively few of them are outside the area depicted in the inset of Figure a, which contains the central business district, commercial and educational areas.

distance, average speed, date, and time.

Fig. 1a shows the wide Greater Boston area and its location in the Northeast of the USA. Although both bike and pedestrian trips are spread throughout a large area of over 200 Km^2 in the region shown in the inset of Fig. 1(a), the large majority of the trips are concentrated in a more central area of approximately 50 Km^2 depicted in Fig. 5.

To deal with the same time period, we restricted our analysis to the portion of the time covered by both datasets and we disregarded the period before August 2014, when the number of pedestrian trips was too small. Therefore, we considered all bike and pedestrian trips that occurred between 2014-08-01 and 2015-05-01. Furthermore, we removed all bikes trips that left and arrived at the same station. This resulted in 144,128 pedestrian trips and 795,531 bike trips. Workdays present similar patterns among themselves but they differ greatly from weekends, so we normally treat these classes separately. Within a single day, we investigate three different time periods: morning peak (from 7:00 to 10:00), lunch time (from 11:00 to 14:00) and afternoon peak (from 17:00 to 20:00) as their patterns differ significantly.

A strong variation in usage is observed in both mobility modes throughout the study period. The average number of trips per hour in our datasets reduces significantly during the winter months for both modes. In particular, bike trips are the most affected since the average number of trips per hour in the entire dataset drops from 230 in September to only 13 in February, as shown in Fig. 2a. A more moderate but still significant reduction of trip frequency is observed also in pedestrian trips, that range from 33 trips per hour in August to 20 trips per hour in February. Note that these numbers of trips per hour depend completely on the datasets we have: while the cycling dataset contains all of the bike-sharing trips of the period, the pedestrian dataset contains only a small portion of all pedestrian trips of the city, i.e., only for those pedestrians using the mobility app used to track their routes. Thus, we cannot compare the absolute number of trips in one dataset with the absolute numbers in the other; we can do only relative analysis.



(a) (b)

Fig. 2. (a) Average number of trips per hour for pedestrian and bike sharing on different months; (b) distribution of the number of hours of precipitation on different months.

To analyze the impact of weather conditions on non-motorized mobility, we obtained from `weather.com` the hourly precipitation report for the period under study. We considered as precipitation every event classified as snow, drizzle, Rain, Storm, Thunder, Sleet, Wintry Mix, and Precipitation. Fig. 2b shows the number of hours of precipitation per month occurred between 7 am and 7 pm, which is the time of the day in which the vast majority of the non-motorized trips happen. As expected, such events were more frequent during the colder part of the year, which is typical for the Boston area.

4. Comparative analysis

We now proceed with the analysis of both datasets, emphasizing the similarities and differences we found between the two modes of non-motorized mobility.

4.1. Distance and duration

In terms of coverage, pedestrian trips in our dataset are spread throughout a larger area of about 400 Km^2 while the bike trips are concentrated in the 200 Km^2 area covered by the bike dock stations. However, the pedestrian trips in the periphery are very sparse; the large majority of trips are concentrated in the more densely populated, commercial, and university areas around city centers as depicted in Fig. 5.

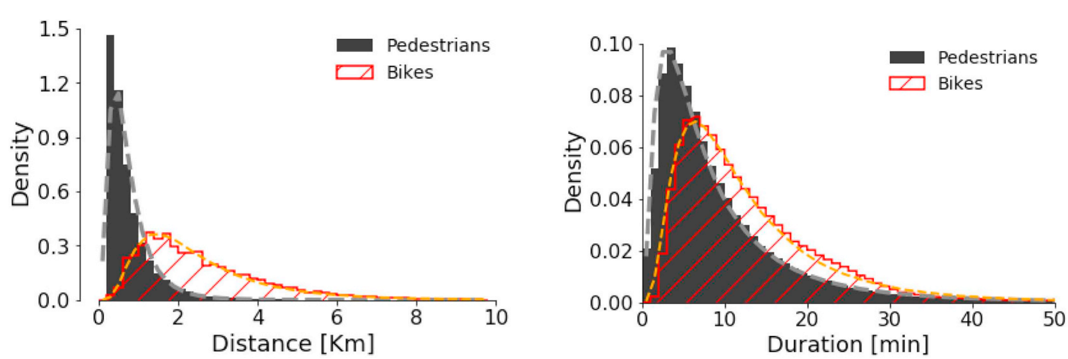
Fig. 3a shows the distribution of the distances travelled by pedestrians and bikes aggregated over the entire time period. We can see that pedestrian and bike trips typically cover a different range of distances. In fact, while the 2nd and 3rd quartiles (resulting in 50%) of pedestrian trips fall between 368 m and 948 m, with a median of 510 m (mean of 789 m), bike trips have 50% of the trips between 1.41 km and 3.42Km, with a median of 2.2Km (mean of 2.6 km).

Differently, the distributions of travel time, depicted in Fig. 3b, show a greater similarity between cycling and walking; a large portion of trips in the two categories falls in the range of 4 to 17 min and both of them present a log-normal distribution. More specifically, the 2nd and 3rd quartiles of pedestrian trips falls between 3.9 and 12.4 min with a median of 7 min, while the 2nd and 3rd quartiles of bike trips falls between 6.1 and 17 min, with a mean of 10.1 min. These findings corroborate the concept introduced in the 1980s by Yacov Zahavi of *travel time budget*, according to which travelers make available a constant amount of time for moving from one place to the other in their daily routine (Marchetti, 1994).

Undertaking a bike sharing trip has some overhead: it takes a few minutes to walk to a dock station, select a bicycle and use the App to request it, remove the bicycle from the rack and start the trip. Thus, it is normally not worth to use bike sharing for very short trips that take less than 5 min. Of course, the same does not happen for walking, explaining while, although the distributions present a similar log-normal shape, bike trips tend to be a little longer.

4.2. Time of the day

To characterize the typology of trips between pedestrians and bikes it is important to look at the temporal distribution of the trips over the day. Fig. 4 shows the distribution of trip departure times for both pedestrians and bikes. It is worth noticing the high level of similarity between both distributions over the entire day, except for the lunch



(a) (b)

Fig. 3. (a) Distribution of the distance covered by the trips (it can be approximated by the grey dotted line, a lognormal with parameters $\mu = -0.48$ and $\sigma = 0.66$, and by the orange dotted line, a lognormal with parameters $\mu = 0.76$ and $\sigma = 0.62$). (b) Distribution of the duration of the trips (it can be approximated by the grey dotted line, a lognormal with parameters $\mu = 1.9$ and $\sigma = 0.91$ and the orange dotted line, a lognormal with parameters $\mu = 2.34$ and $\sigma = 0.70$). Note that μ and σ are the mean and the standard deviation of the logarithm of the values of the distribution.

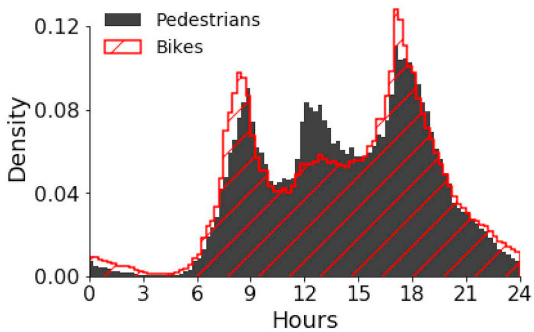


Fig. 4. Distributions of the departure times.

time, when walking shows a significantly higher frequency of trips. This observation supports the intuition that bike mobility is relatively more important for commuting than walking and that people in Boston tend to walk to lunch.

4.3. Spatial flows

To compare the spatial flows of the two modes of mobility, we aggregated the bike sharing and pedestrian trips into flows connecting different regions of the city. As mentioned before, we divided the greater Boston area by using a 100 × 100 grid, each Grid cell being a square of side close to 300 m. For each pair of grid cells, we counted the number of trips connecting those two regions. We then plot the resulting flows on the map, with an arrow whose width and opacity is proportional to the number of trips represented by that flow, so that flows with more trips are more prominent and flows with very few trips tend to disappear in the visualization.[3] Finally, to better visualize different mobility patterns, we divided the trips in workday or weekend and according to the time of the day. Fig. 5 depicts pedestrian mobility patterns on the left hand side and bike sharing patterns on the right hand side. The three rows present the patterns in the morning rush hour (7 am to 10 am), lunch-time (11 am to 2 pm), and afternoon rush hour (5 pm to 8 pm).

Observing Fig. 5, we can see that bike flows and pedestrian flows have different purposes. Both present a high concentration around the major subway and train stations. But bicycles are used mostly to connect the different cities in the Greater Boston and to connect neighborhoods that are 2 to 5 Km away from each other (see scale in the bottom figures). On the other hand, most pedestrian trips are shorter and are used to access locations in business and university districts. The bicycle trips are more concentrated on fewer flows while the pedestrian trips are more evenly distributed across all major business and university areas. On the one hand, we can see a few commuting-style pedestrian trips in the morning and afternoon covering longer distances but most trips are short, within the same area of the city. On the other hand, most bike sharing trips in the morning and afternoon peaks present a commuting style (from a residential area to business/education area in the morning and vice versa in the afternoon) or are connected to the major subway and train stations.

We can conclude that the overall shape of the flows across the city demonstrate that pedestrian trips and bike trips serve a different purpose with regard to urban mobility. Although they overlap in trips between 600 m and 1Km, they mostly complement each other providing alternative mobility modes for shorter (under 600 m, normally on foot) and longer (from 1Km to 4Km, normally on a bike) trips. Combined, they represent a real alternative for decreasing the number of car trips in contemporary cities since, for instance, as much as 1/3 of car trips in the USA are shorter than 3.2 Km (USDOT, 2018).

4.4. Network analysis

The network flows illustrated above can be modeled as a weighted directed network. The vertices of the network are the locations identified by the grid and the weight of a directed link w_{ij} is the number of trips between i and j in the considered time interval. In network science, the importance of a vertex can be measured by a centrality metric (Newman, 2018) such as degree, closeness, betweenness, or Page Rank. The betweenness centrality is more appropriate in assessing the centrality of a given location on a street network where the shortest path is, indeed, the most reasonable criterion to navigate a city. However, in our case, we are not taking into account the street network, and it is not obvious that a person would select the shortest path on the transition matrix of locations to reach a destination. In addition, it is worth noticing that the bike network is very dense, which implies that the shortest path connecting two nodes is in general the link itself. Here, we are dealing with a directed weighted network representing mobility flows within a city, thus Page Rank (Page et al., 1999) is a more appropriate metric. The Page Rank of a vertex i on a weighted directed network is the likelihood that a walker following a biased random walk on the network will arrive in i after an infinitely long walk. The larger is the value of the Page Rank, the more important will be the vertex with respect to the network flow.

We built both the bike and pedestrian networks by aggregating all the trips occurred between August and April. We then evaluated the Page Rank for all the vertices of the networks. However, to compare the results of both networks, we must take into account that just 142 of the vertices of the bike network have a degree larger than zero. This happens because not all the grid cells have a bike sharing station. Differently, 5,571 vertices of the pedestrian network have degree larger than zero. To visually compare the values, in Fig. 6a, we plotted only the Page Rank of the vertices with a degree larger than zero on both networks, i.e., the 142 vertices of the bike network that contains a station. One can observe a significant weak Pearson correlation of 0.34 between the Page Rank of the vertices of both networks, indicating that the most important locations in each of the two systems are quite diverse. In particular, by looking at the marginal distributions on Fig. 6a, the pedestrian locations (depicted as the histogram in the top of the figure) seem to have a high concentration in low values with few locations with a high value of Page Rank; differently, the respective locations on the bike network (depicted on the right-hand side of the figure) seem more uniformly distributed over a wider range of Page Rank values. That means that the pedestrian network presents a structure more concentrated on hubs than the bike network.

The hypothesis that the pedestrian network is more concentrated on hubs is confirmed by looking at the graphical representation of the Gini coefficient (Gini, 1936; Crucitti et al., 2006) depicted in Fig. 6b. The Gini coefficient is a measure of inequality commonly used in Economics. A specific point in the curve in Fig. 6b indicates that a certain fraction of the locations has a cumulative fraction of all incoming trips of the system. In the case of perfect homogeneity, such a curve would lie on the first bisectrix. The Gini coefficient is a number in $[0, 1]$ and is defined as two times the area between the first bisectrix and the curve. For the bike sharing network, the Gini coefficient is 0.37, whereas, for the pedestrian network, it is 0.52, confirming our hypothesis of a hub-based structure more pronounced for the pedestrian network.[4]

A hub-based structure, in contrast to a point-to-point structure, is a fundamental structure in the architecture of contemporary transportation systems. While the former is characterized by high connectivity among different locations with few direct connections, the latter requires a larger number of direct point-to-point links to produce a

[3] The width and opacity of pedestrian flows are relative to the total number of pedestrian trips only; the same for the bike-sharing trips.

[4] When considering the entire area of the pedestrian network, i.e., not only the area that overlaps with the bike network, the Gini coefficient is even larger, 0.75.

(a) Morning pedestrian flows

(b) Morning bike sharing flows

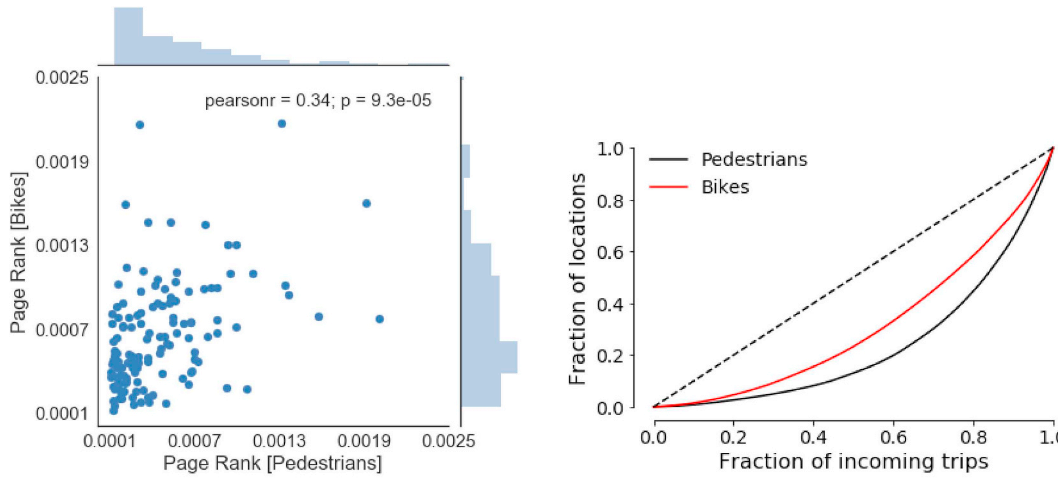
(c) Lunch-time pedestrian flows

(d) Lunch-time bike sharing flows

(e) Afternoon pedestrian flows

(f) Afternoon bike sharing flows

Fig. 5. Top mobility flows within Greater Boston.



(a) (b)

Fig. 6. (a) Scatter plot with marginal distribution of the Page Rank for the vertices with a degree larger than zero for both sets. A significant weak Person correlation is reported in the legend. (b) Graphical representation of the Gini coefficient for the considered locations, the dotted line is the hypothetical homogeneous distribution of the trips over the locations.

similar effect in connectivity. A specific kind of hub-based structure, the hub-and-spoke structure (O'Kelly and Miller, 1994) is the typical pattern followed by the large national airlines, with most of the flights connected to a major hub; differently, a point-to-point structure is typical of new low-cost companies (Alderighi et al., 2005). It is interesting to note that, in the case of the airline industry, the choice for the hub-and-spoke approach is a top-down planning decision, while for pedestrian mobility, the concentration around hubs is an emergent behavior from the individual decisions of walkers. It is indeed surprising that, for the bike-sharing mobility network, where the location of the dock stations might induce the emergence of hubs, data analysis shows that it is, in fact, less concentrated in hubs than the pedestrian network. A possible explanation for this phenomenon (yet to be verified) is that bike trips are used to reach a relatively larger variety of destinations (for example, for commuting and other activities), while pedestrian trips, during weekdays, are highly concentrated on business areas. Since a bike trip allows the rider to reach a wider area, the hubs tend to be less pronounced.

To further characterize the network and verify how similar to a hub-and-spoke structure our networks are, we analyzed the correlation among the strengths of the nodes (where strength is the sum of the weights of the links of a node). Two possible kinds of correlations can be observed: assortative and disassortative mixing. In the case of an assortative mixing, hubs tend to be connected to other hubs and spokes to other spokes. In the case of disassortative mixing, hubs tend to be connected to spokes. An assortative mixing is typical of social systems such as coauthorship networks, whereas a disassortative mixing is typical of technological and biological systems such as the Internet or protein interaction networks (Newman, 2002).

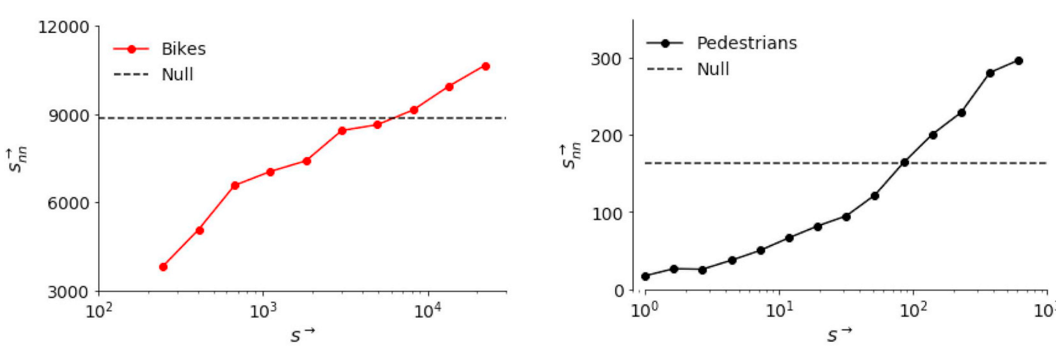
To derive a quantitative measure of this phenomenon, a common approach is to estimate the average degree of the nearest neighbors of the nodes of degree k, for each degree (Pastor-Satorras et al., 2001). If this quantity is positively correlated with k, then an assortative mixing is observed; on the contrary, if the correlation is negative, the system is disassortative. In our case, since we are dealing with a directed weighted network, such metric must be generalized by considering the strength. Given a network with N nodes and $s_i^{\rightarrow}$, the out-strength of node i, we can compute:

$$s_{i,nn}^{\rightarrow} = \sum_{j=1}^{N} \frac{w_{ij}}{s_i^{\rightarrow}} s_j^{\rightarrow} \tag{1}$$

that is the weighted average of the strength of the nearest neighbors of node i. An equivalent metric can be computed for the in-strength $s^{\leftarrow}$. In the case of the absence of the above mentioned correlation, the expected $s_{i,\,nn}^{\rightarrow}$ is a constant value:

$$E\left[s_{i,nn}^{\rightarrow}\right] = \sum_{i=1}^{N} \frac{s_i^{\leftarrow} s_i^{\rightarrow}}{L} \tag{2}$$

where L is sum of all link weights. This value refers to the expected value of $s_{i,\,nn}^{\rightarrow}$ in a random network that preserves the strengths $s_i^{\leftarrow}$ and $s_i^{\rightarrow}$ of the original network.



(a) (b)

Fig. 7. Both figures plot the average out-strength of the nearest neighbors of nodes of out-strength $s^{\rightarrow}$. The values of out-strength are grouped in logarithmic bins. The dotted line is the expected null value of no correlation. The figures show that both networks are an assortative mixing.

The analysis depicted in Fig. 7 shows that the structure of both networks display clearly an assortative mixing pattern. The same is observed by considering the in-strength. This behavior differs substantially from other transportation systems, such as airline networks, where the spokes are mostly connected to the hubs to improve the connectivity of the whole system with a small number of direct links (Bagler, 2008). In the two modes of non-motorized mobility analyzed here, hubs have a tendency to be connected to hubs, and spokes to be connected to spokes. A geographical analysis of node strengths indicates that hubs are concentrated in the business and commercial areas and spokes in the periphery and residential areas.

Another interesting aspect related to directed networks is reciprocity. The reciprocity on an unweighted network is defined as the fraction of bidirectional links over the total number of links. The reciprocity r is a number in $[0,1]$, where $r = 0$ indicates a network composed of only unidirectional links, whereas $r = 1$ indicates a network composed of only bidirectional links. The reciprocity was generalized for weighted networks by Squartini et al. (2013) as:

$$r = \frac{\sum_i \sum_{i \neq j} \min[w_{ij}, w_{ji}]}{\sum_i \sum_{i \neq j} w_{ij}} \tag{3}$$

Using the reciprocity metric we can investigate whether the mobility flow networks present a commuter behavior – e.g., flows in one direction in the morning and in the opposite direction in the afternoon – and to what extent.

To do that, we evaluated the reciprocity during the morning (7 am to 10 am), during the afternoon (5 pm to 8 pm), and also when aggregating morning and afternoon flows.[5] The reciprocity of the morning network for the pedestrians is 0.13, and for bikes is 0.18. Both values indicate networks composed mostly by unidirectional links. During the afternoon the networks show reciprocity of 0.14 and 0.28 for pedestrians and bikes respectively. That indicates a slight increase in the fraction of bidirectional links for the bike networks; this might refer to bikes being used for running errands such as going to the supermarket.

Interestingly, by aggregating the morning with the afternoon network, the reciprocity of the pedestrian network increases only to 0.14 whereas, for the bikes, the value of reciprocity reaches 0.39. This observation indicates that cycling is probably used much more for commuting than walking as almost 40% of the flows are bidirectional.

To study the local impact of commuting on the top flow links, we observed the inversion of the flow direction on the 200 links with the largest number of trips in the morning and afternoon for both pedestrian and bikes. Specifically, in Fig. 8 we show the scatter plot of the values $p^{\rightarrow}$, that is the fraction of trips going in a fixed direction on a specific link, in the morning and the afternoon. In the case of the bike network (Fig. 8a), a strong anticorrelation of -0.87 between morning and afternoon of $p^{\rightarrow}$ is observed; this confirms our global observation that the flows of the bike network tend to reverse the direction in the afternoon. Differently, the $p^{\rightarrow}$ for the pedestrian network has a significant positive correlation of 0.66, what indicates a strong stability of the directionality between morning and afternoon, confirming our global observation. This result confirms recent work on the asymmetry of pedestrian behavior performed with a different methodology (Malleson et al., 2018).

The low reciprocity of the pedestrian network does not necessarily imply that pedestrians are not commuters; in fact, they could select a different path or mode of transportation in the afternoon. However, the differences in flow reciprocity observed in our study constitutes a genuine and relevant difference between the two modes of transportation.

[5] We repeated this analysis with shorter morning and afternoon intervals around the rush hour, obtaining very similar results.

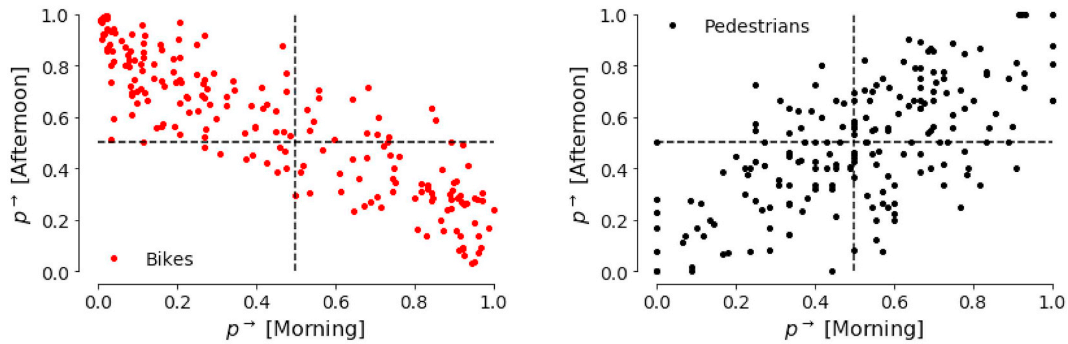
4.5. *Sensitivity to weather conditions*

To understand the impact of precipitation on non-motorized mobility, we must take into account that rainy days are not homogeneously distributed over the year. In Boston, they are more concentrated during the Winter period, as shown in Fig. 2b. Since we are interested in comparing the trip rates during the precipitation hours with the trip rates during fair (i.e., non-precipitation) hours, we must take into account the confounding effect of the temperature, which is normally lower during Winter. In fact, by focusing on all trips occurred between 7 am and 7 pm during weekdays, we observed a significant Pearson correlation between the number of trips per hour and the temperature for the bikes of 0.68 (p-value close to 0). Differently, the rate of pedestrian trips show a correlation of only 0.25 with the temperature (p-value of 10^{-52}), showing that walking is less affected by temperature than cycling.

As depicted in Fig. 2a, trip rate decreases significantly during winter, which is probably due to a mixed effect involving temperature and precipitation, the latter, mostly related to a perception of safety. If we want to compare the reduction of the trip rates due to precipitation events, first we must quantify the confounding effect coming from other covariates, such as temperature. If such effect is not negligible, then a statistical matching technique must be applied to rebalance the sets under study with respect to the other covariates (Stuart, 2010). One of the most common metric used to evaluate covariate bias is the standardized bias (SB) (Pan and Bai, 2015), defined as:

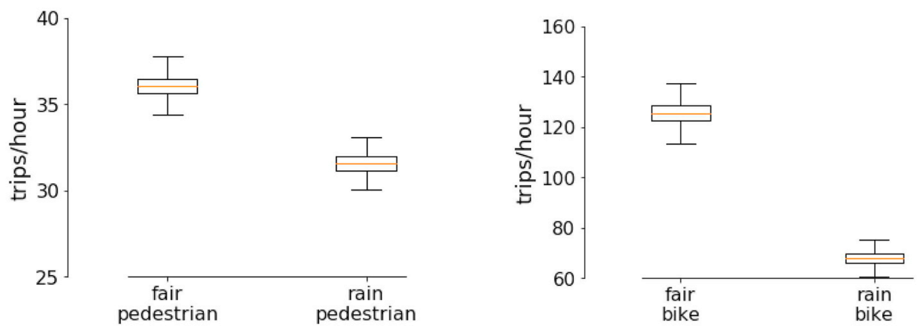
$$SB = \frac{|M_t - M_c|}{\sqrt{\frac{V_t}{2} + \frac{V_c}{2}}} \times 100\% \tag{4}$$

where the numerator is the mean difference in the covariate between treatment and control, and the denominator is the square root of the average of the variances of the related covariate. Typically, a value of SB larger than 25% is considered not acceptable. In our case, we considered the precipitation hours as the treatment and the fair hours as control. Thus, M_t and M_c are the average temperatures for the fair and precipitation hours respectively, and V_t and V_c are the related variances. The SB of the temperature covariate of pedestrians is 53%, whereas for the bikes is 42%. In both cases the SB is significantly high, and the comparison of the trip rates between treatment and control cannot be done directly without incurring in a bias. Thus, we used a bootstrap matching technique on temperature intervals (bins) to reduce the standardized bias (Dehejia and Wahba, 1999). In particular, we split the temperature in bins of 10 degrees Fahrenheit. For each bin of temperature t, we have four sets: ${n_f}^b(t)$ fair hour records for the bikes, ${n_f}^p(t)$ fair hour records for the pedestrians, ${n_p}^b(t)$ and ${n_p}^p(t)$ precipitation hour records for bikes and pedestrian, respectively. To rebalance the sets, we created bootstrap replicas of the records of the four sets with a sampling with replacement of $n(t) = \min\{{n_f}^b(t), {n_p}^b(t), {n_f}^p(t), {n_p}^p(t)\}$ elements from each set. We performed 10,000 bootstrap copies of the four sets by aggregating the bootstrap for every temperature bin. This procedure successfully reduced the SB of the temperature for both pedestrians and bikes. Specifically, the average value SB among the bootstrap copies was 1.5% for pedestrians and 1.4% for bikes, with a 95% confidence interval of (0.2%, 3.1%) and (0.1%, 3.0%) respectively. As a result, the average trip rates during the precipitation hours was significantly smaller than during fair hours, both for bikes and pedestrians, as shown in Fig. 9. In particular, precipitations imply a reduction of 13% of the pedestrian trip rate with a 95% confidence interval of (9%,16%). Differently, cycling was much more penalized by precipitation; for bikes, the reduction was 46% with a 95% confidence interval of (41%,51%), i.e., independent of the temperature, about half of the people decide not to take a bike trip when there is precipitation.



(a) (b)

Fig. 8. Both figures show the 200 links with the largest cumulative number of trips in the morning and afternoon. The x-axis represents the fraction of trips occurred in a certain direction in the morning; the y-axis is the fraction of trips occurred in the same direction during the afternoon. Note the clearly different patterns for bike sharing (Panel a) and pedestrians (Panel (b).



(a) (b)

Fig. 9. Both figures show the average trip rate during fair and precipitation hours for pedestrians (a) and for bikes (b). The boxplots were obtained with bootstrap matching with 10,000 replicas. The central orange line indicates the median, the box indicates the 25 and 75 percentile, the whiskers indicates the 1.5 interquartile range.

4.6. Sensitivity to variation of temperature

As we pointed out in the previous section, the external temperature has a strong influence on pedestrian and bike mobility. In this section, we aim to have a quantitative estimate of the impact on the number of trips per hour of an increase of temperature with respect to its average monthly value. It is important to consider variation of temperature with respect to the average temperature of each month and not the absolute value because the perception of what is a warm or a cool day varies a lot throughout the year. For instance, while $15^{o}C$ might feel cold for a Summer day, it feels warm for a Winter day. To do that, we estimate, from our dataset, the average temperature and the average number of trips for each weekday hour in each month (weekends show a very different pattern and were not considered in this analysis). Then, for a fixed month, we compute the Pearson correlation between the percentage increment of the number of trips per hour with respect to its average value and the increase in Celsius with respect to the average monthly value in the considered hour. Specifically, the increment of the number of trips per hour is defined as $\frac{n(t)-E[n(t)]}{E[n(t)]}$ where $E[h]$ is the average number of trips occurred in the hour of the instant t in the considered month, whereas the increase in temperature is simply $T(t) - E[T(t)]$, where $T(t)$ is the temperature in the instant t and $E[h]$ is the average temperature observed in the hour of the instant t in the considered month. To remove the effect of the weather, we considered only fair hours, i.e., without any kind of precipitation.

Fig. 10 depicts the result of this analysis on both systems. To assess the statistical significance of the metric, we estimate a 95% bootstrap confidence interval (DiCiccio and Efron, 1996) on the Pearson correlation indicated by the bands in Fig. 10. As result, during the Summer months, the increment of the number of trips of the pedestrian is negatively correlated with the increase of temperature, it means that a rise in temperature implies a decrease of trips; differently, the increment of the number of bike trips seems to be not significantly correlated with the temperature increment. During the Autumn and until the beginning of the Spring, the increment of the number of bike trips is significantly correlated with the increment of temperature; it means that a rise in temperature implies an increase of cycling trips. It is worth noticing that only during the winter period we observed this clear correlation for pedestrians, whereas during Autumn such association seams less robust.

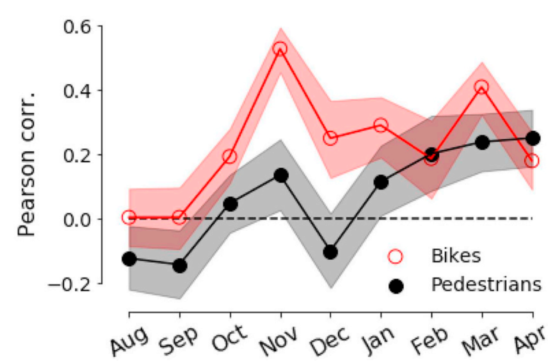


Fig. 10. The Pearson correlation coefficient between the increment of trips with respect to the month average and the increment of temperature with respect to the month average. Bands show the 95% bootstrap confidence interval.

Table 1 shows the percentage of increment in the number of trips described by the linear law

$$\frac{n(t) - E[n(t)]}{E[n(t)]} = \beta(T(t) - E[T(t)]) \tag{5}$$

The β parameter measures the average percentage increment of trips associated with an increment of $1^{o}C$. For example, in November, for the bike system, an increment of 1 °C is associated with an increment of 3.8% in the number of trips in relation to its monthly average. Thus, we could say that a mild November day – for example, with a temperature

Table 1
The table shows $100 \times \beta$, which is the average percentage variation of the number of trips associated with an increment of 1 °C. CI indicates a 95% bootstrap confidence interval on β. The number in bold indicates a value of the β significantly different from 0, according to the CIs reported in the table.

Month	% Bike-sharing trips variation	CI	% Pedestrian trips variation	CI
Aug	0.0	(−0.8, 0.9)	**−1.5**	**(−2.7, −0.3)**
Sep	0.0	(−0.5, 0.5)	**−1.0**	**(−1.7, −0.3)**
Oct	**1.3**	**(0.7, 1.8)**	0.4	(0.4, 1.2)
Nov	**3.8**	**(3.1, 4.4)**	**0.8**	**(0.2, 1.5)**
Dec	**2.7**	**(1.4, 3.9)**	−1.0	(−2.2, 0.2)
Jan	**3.1**	**(1.9, 4.5)**	**1.2**	**(0.1, 2.2)**
Feb	**2.1**	**(0.7, 3.4)**	**1.9**	**(0.8, 2.9)**
Mar	**3.1**	**(2.5, 3.8)**	**1.5**	**(0.9, 2.1)**
Apr	**3.2**	**(1.6, 4.7)**	**2.3**	**(1.5, 3.1)**

of 10^oC above average – would tend to present an increase close to 38% in the number of bike sharing trips. It is worth noticing that the monthly evolution of β seems to be consistent with the observation drawn for the Pearson correlation coefficient.

In summary, these results show that, on the one hand, during hot months, people tend to walk less when the temperature increases but this change does not affect cycling. On the other hand, in colder months (October to April) people tend to use significantly more bike sharing when the temperature increases. During these colder months, there is also an increase, but of smaller magnitude, in pedestrian trips when the temperature increases.

5. Discussion

Regarding the specific data in the analyzed datasets, we can summarize the most relevant findings as follows:

- The duration of trips in both modes of transportation is similar, typically from 4 to 14 min, following a log-normal distribution.
- Bike trips are typically longer than the pedestrian ones; bike trips are often used to connect different parts of the city such as different neighborhoods, while pedestrian trips are normally used to connect different blocks within the same or adjacent neighborhoods.
- Pedestrian trips have frequency peaks at morning commute time (around 9:00 am), afternoon commute time (around 6:00 PM), and at lunch time (around 12:30 PM). Bike sharing trips present peaks at similar morning and afternoon commute times but do not present a significant peak at lunch time.
- The network of pedestrian mobility flows presents and hub-based structure more pronounced than the bike sharing flow network. Both networks show assortative mixing.
- The cycling flows present a very significant bidirectional pattern with opposite directions in the morning and afternoon rush hours, indicating a heavy use of bike sharing for commuting. For the pedestrian flows, this pattern was not present.
- Bike trips are more affected by severe weather conditions (particularly precipitation) than walking. This may hinder the use of cycling as an alternative means of transportation during certain periods of harsh weather.
- In the Summer, citizens tend to walk less during hotter days (while cycling seems not to be affected). In the Winter, cyclists tend to use bikes significantly more during days in which the temperature is relatively mild (while pedestrians are less affected by this factor).
- Overall, the strong sensitivity to the weather condition poses a severe limitation to non-motorized mobility, especially during the winter season.

Thus, if cities wish to promote active modes of transportation as a means to provide a better environment for its citizens, the sensitivity to the weather, mainly in the case of cycling, must be somewhat addressed; otherwise, the benefits would be limited only to a certain part of the year (around 8 months in the case of Boston).

The fact that pedestrian infrastructure (i.e., sidewalks and crosswalks) are of much higher quality and pervasive in the area of study than their cycling counterpart (bike paths and streets shared with cars) is probably the most important reason why walking is less affected by rain and temperature. Cyclists probably feel relatively less safe than pedestrians when riding under rain and when there is snow or ice on the streets. In fact, there is evidence that improving the cycling infrastructure make people more willing to use bicycles, in general. In fact, Bill et al. (2015) experienced with the use of realistic visualizations to promote active travel modes in a new walking and cycling route in Glasgow. By working with focus groups, a frequent remark from participants was that the quality of the cycling infrastructure, in particular, the appearance of protection is considered important in encouraging a wider uptake of cycling. With a different methodology, based on regression modeling with urban morphology and survey variables, Rybarczyk and Wu (2014) showed that the space syntax of the city built environment has a influence on bicycle mode choice. A hypothesis to be tested in future studies is that, even in periods of harsh Winter, when the cycling infrastructure is very good and unobstructed, a relevant level of cycling trips are still preserved.

Our finding that the pedestrian flow network presents a more pronounced hub-based structure than the bike sharing network can be counter-intuitive as the city network of sidewalks is completely point-to-point and relatively homogeneous throughout the city. Thus, even though pedestrians can potentially walk anywhere, they do not; they tend to base their trajectories on anchor points, the "hubs". Past studies have investigated the correlation between street network configurations and pedestrian behavior using various methods and showing different results (Kang, 2017). In particular, studies on street configuration have found that diverse street attributes generate higher walking volume. While early investigations emphasized the physical features of street networks in relation to walking behavior, later research confirmed the significance of street density and link structures for walking activities (Cervero and Kockelman, 1997; Crane and Crepeau, 1998; Lee and Moudon, 2006). Further research is needed to better understand what characteristics make up these hubs; one can hypothesize that major points of interest in the city comprise most of these hubs.

6. Limitations

A limitation of our work is the fact that we are using bike sharing trips as a proxy to the overall use of bicycles in the urban space. In many cities, however, bike sharing represent only a fraction of the total cycling trips. It might be the case that cyclists who have their own bikes follow different mobility patterns. This could be mitigated in the future by using image processing on traffic videos to detect bicycle flows. The difficulty with this approach is that, to have a minimal coverage of a city, this would require analyzing hundreds, or even thousands, of video cameras in real-time to detect and count bike flows. To the best of our knowledge, this has not yet been done on a scale that would allow any city-wide analysis of mobility flows. Another alternative would be to use the location service of smart phones to detect bike (and walking) trips (e.g., based on speed). This would require tracking the location of millions of people, which has strong privacy implications but, in fact, is already performed by cell phone carriers and mobile OS vendors.

The pedestrian dataset has some limitations. First, the data was collected from an activity-oriented mobile phone application, which makes the sample non-representative of the total population. Second, we do not have a way to verify whether users activated the app for all trips they made. Third, to anonymize the data, a distance of 100 m was erased from the start and end of each trajectory to avoid identification, which makes the specific origin and destination a little less accurate.

7. Conclusions and future work

Non-motorized modes of transportation are an essential element for the quality of life in cities. They have a direct impact on people's health and help reducing pollution, mainly when replacing car trips. In this paper, we analyzed two mobility datasets to identify similarities and differences between walking and cycling characteristics. This was the first such study analyzing large datasets both from pedestrian and cycling trips in a combined form, contributing to the understanding of the dynamics of active transportation. These two modes of non-motorized mobility have a good potential not only to ameliorate the health of citizens but also to improve the environment of contemporary cities, making it less noisy, less polluted, and more comfortable for humans.

This study points out that pedestrian trips and bike trips serve different purposes, complementing each other. Walking is less affected by weather conditions, possibly due to the superior quality of the walking infrastructure. Potentially, these findings have a strong spatial impact for a variety of constituents, including policy makers, planners, and traffic managers. Higher exposure to car traffic and to environmental conditions are probably some of the most influential elements to inhibit bike mobility: a matter that indicates the need of redesigning streets to facilitate bike mobility. In particular, dedicated bike lanes, covered by trees or artificial canopies as well as safely lit and detached from motorized mobility, are measures that urban planners and policy makers could propose for increasing the sense of safety and thus making the use of bikes more attractive. These interventions have a strong impact on the design of street sections and could be combined with specific regulations such as speed and access limitations for motorized vehicles. These interventions should also take into account the spatial distribution of the focal nodes of certain neighborhoods, where walkability and transit availability aspects, such as density of retail destinations, density of recreational open space, intersection density, residential density and density of subway stops are higher (Duncan et al., 2013). Also, expert recommendations and educational campaigns about using bikes during cold and rainy seasons as well as about appropriate equipment could expand its use during these periods of the year.

As future work, it would be interesting to analyze the evolution of pedestrian and bike sharing mobility over longer periods of time (e.g., 10 years), and relate it to changes in the city socio-economic variables. It would also be valuable to extend the analysis to other cities and to assess to which extent the results can be generalized. In addition, it would be worth investigating what portion of car trips would be amenable to a switch towards non-motorized mobility and what actions could be taken by policy makers to enable that change. For informing policy making processes, it would be essential to highlight the complementarity of non-motorized transportation modes and the relation to the bike infrastructure and to the street section, in particular with regard to safety, which might be a major issue in facilitating bicycle use.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgements

We are grateful for the feedback from our colleagues of the MIT Senseable City Lab. Fabio Kon, a professor from the University of São Paulo, participated in this research while a visiting professor at the MIT Senseable City Lab funded by a Fulbright Visiting Professor Fellowship. This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq proc. 465446/2014-0, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, FAPESP proc. 14/50937-1, and FAPESP proc. 15/24485-9.

References

Alderighi, M., Cento, A., Nijkamp, P., Rietveld, P., 2005. Network competitionthe coexistence of hub-and-spoke and point-to-point systems. J. Air Transp. Manag. 11 (5), 328–334.
Austwick, M.Z., OBrien, O., Strano, E., Viana, M., 2013. The structure of spatial networks and communities in bicycle sharing systems. PLoS One 8 (9), e74685.
Bagler, G., 2008. Analysis of the airport network of India as a complex weighted network. Phys. A Statistical Mech. Appl. 387 (12), 2972–2980.
Bill, E.M., Baker, G., Ferguson, N.S., Drinkwater, D., Mutrie, N., 2015. Representing active travel: a formative evaluation of a computer visualization tool demonstrating a new walking and cycling route. Environ. Plan. B Plan. Des. 42 (3), 450–467.
Cervero, R., Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. Transp. Res. Part D: Transp. Environ. 2 (3), 199–219.
Corburn, J., 2007. Reconnecting with our roots: American urban planning and public health in the twenty-first century. Urban Aff. Rev. 42 (5), 688–713.
Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., Mateo-Babiano, D., 2014. Spatio-temporal patterns of a public bicycle sharing program: the effect of weather and calendar events. J. Transp. Geogr. 41, 292–305.
Crane, R., Crepeau, R., 1998. Does neighborhood design influence travel?: a behavioral analysis of travel diary and GIS data. Transp. Res. Part D: Transp. Environ. 3 (4), 225–238.
Crucitti, P., Latora, V., Porta, S., 2006. Centrality measures in spatial networks of urban streets. Phys. Rev. E 73 (3), 036125.
De Nazelle, A., et al., 2011. Improving health through policies that promote active travel: a review of evidence to support integrated health impact assessment. Environ. Int. 37 (4), 766–777.
Dehejia, R.H., Wahba, S., 1999. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. J. Am. Stat. Assoc. 94 (448), 1053–1062.
DiCiccio, T.J., Efron, B., 1996. Bootstrap confidence intervals. Stat. Sci. 11 (3), 189–212.
Duncan, D.T., Aldstadt, J., Whalen, J., Melly, S.J., 2013. Validation of walk scores and transit scores for estimating neighborhood walkability and transit availability: a small-area analysis. GeoJournal 78 (2), 407–416.
Faghih-Imani, A., Eluru, N., El-Geneidy, A.M., Rabbat, M., Haq, U., 2014. How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. J. Transp. Geogr. 41, 306–314.
Forsyth, A., 2015. What is a walkable place? The walkability debate in urban design. Urban Des. Int. 20 (4), 274–292.
Forsyth, A., Hearst, M., Oakes, J.M., Schmitz, K.H., 2008. Design and destinations: factors influencing walking and total physical activity. Urban Stud. 45 (9), 1973–1996.
Gini, C., 1936. On the measure of concentration with special reference to income and statistics. Colorado College Publ. 208 (9), 73–79.
Griffin, G., Nordback, K., Götschi, T., Stolz, E., Kothuri, S., March 2014. Monitoring Bicyclist and Pedestrian Travel and Behavior: Current Research and Practice. Transportation Research Circular E-C183.
Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., 2002. How the built environment affects physical activity - views from urban planning. Am. J. Prev. Med. 23 (2), 64–73.
Kang, hD, 2017. Measuring the effects of street network configurations on walking in Seoul, Korea. Cities 71 (September 2016), 30–40.
Krizek, K.J., Handy, S.L., Forsyth, A., 2009. Explaining changes in walking and bicycling behavior: challenges for transportation research. Environ. Plan. B Plan. Des. 36 (4), 725–740.
LAB, 2018. The Bicycle Friendly America Award Database. Tech. Rep. The League of American Byciclists URL. www.bikeleague.org/sites/default/files/bfareportcards/BFC_Fall_2017_ReportCard_Boston_MA.pdf.
Lee, C., Moudon, A.V., 2006. The 3Ds + R: quantifying land use and urban form correlates of walking. Transp. Res. Part D: Transp. Environ. 11 (3), 204–215.
Litman, T., 2015. Evaluating Active Transportation Benefits and Costs. Tech. Rep. Victoria Transportation Policy Institute, Victoria URL. www.vtpi.org/nmt-tdm.pdf.
Malleson, N., Vanky, A., Hashemian, B., Santi, P., Verma, S.K., Courtney, T.K., Ratti, C., 2018. The characteristics of asymmetric pedestrian behavior: a preliminary study using passive smartphone location data. Trans. GIS 22 (2), 616–634.
Marchetti, C., 1994. Anthropological invariants in travel behavior. Technol. Forecast. Soc. Chang. 47 (1), 75–88.
McKenzie, G., 2019. Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, D.C. J. Transp. Geogr. 78, 19–28.
Newman, M.E., 2002. Assortative mixing in networks. Phys. Rev. Lett. 89 (20), 208701.
Newman, M., 2018. Networks. Oxford University Press, pp. 168–234.
Newson, P., Krumm, J., 2009. Hidden Markov Map Matching through Noise and Sparseness. Proc. 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '09. pp. 336.
O'Brien, O., Cheshire, J., Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. J. Transp. Geogr. 34, 262–273.
O'Kelly, M.E., Miller, H.J., 1994. The hub network design problem: a review and synthesis. J. Transp. Geogr. 2 (1), 31–40.
Page, L., Brin, S., Motwani, R., Winograd, T., 1999. The Pagerank Citation Ranking: Bringing Order to the Web. Tech. Rep. Stanford InfoLab URL. http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf.
Pan, W., Bai, H., 2015. Propensity Score Analysis: Fundamentals and Developments. Guilford Publications.
Pastor-Satorras, R., V'azquez, A., Vespignani, A., 2001. Dynamical and correlation properties of the internet. Phys. Rev. Lett. 87 (25), 258701.
Pikora, T., Giles-Corti, B., Bull, F., Jamrozik, K., Donovan, R., 2003. Developing a framework for assessment of the environmental determinants of walking and cycling.

Soc. Sci. Med. 56 (8), 1693–1703.
Pucher, J., Komanoff, C., Schimek, P., 1999. Bicycling renaissance in North America? Recent trends and alternative policies to promote bicycling. Transp. Res. A Policy Pract. 33 (7–8), 625–654.
Pucher, J., Buehler, R., Merom, D., Bauman, A., 2011. Walking and cycling in the United States, 20012009: evidence from the National Household Travel Surveys. Am. J. Public Health 101 (SUPPL. 1).
Quercia, D., Schifanella, R., Aiello, L.M., 2014. The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media. ACM, pp. 116–125.
Rietveld, P., oct 2001. Biking and Walking: The Position of Non-Motorized Transport Modes in Transport Systems. Tinbergen Institute, Tech. Rep.
Rojas-Rueda, D., de Nazelle, A., Teixid'o, O., Nieuwenhuijsen, M.J., 2012. Replacing car trips by increasing bike and public transport in the greater Barcelona metropolitan area: A health impact assessment study. Environ. Int. 49, 100–109.
Rybarczyk, G., Wu, C., 2014. Examining the impact of urban morphology on bicycle mode choice. Environ. Plan. B Plan. Des. 41 (2), 272–288.
Saelens, B.E., Ph, D., Frank, L.D., Ph, D., Med, A.B., 2003. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Environ. Phys. Activity 25 (2), 80–91.
Santucci, D., Fugiglando, U., Li, X., Auer, T., Ratti, C., 2018. Methodological framework for evaluating liveability of urban spaces through a human centred approach. In: Windsor Conference - Rethink Comfort, pp. 789–797.
Sarkar, A., Lathia, N., Mascolo, C., 2015. Comparing cities' cycling patterns using online shared bicycle maps. Transportation 42 (4), 541–559.
Southworth, M., 2005. Designing the Walkable City. J. Urban Plan. Dev. 131 (4), 246–257.
Squartini, T., Picciolo, F., Ruzzenenti, F., Garlaschelli, D., 2013. Reciprocity of weighted networks. Sci. Rep. 3, 2729.
Stuart, E.A., 2010. Matching methods for causal inference: a review and a look forward. Statistical Sci. 25 (1), 1.
USDOT, 2018. U.S. Department of Transportation National Household Travel Survey. Available online at: https://nhts.ornl.gov/vehicle-trips (web archive link, 18 December 2020) and https://nhts.ornl.gov/person-trips, accessed: 2018-12-20. Washington, D.C.: USDOT.
Vanky, A.P., Verma, S.K., Courtney, T.K., Santi, P., Ratti, C., 2017. Effect of weather on pedestrian trip count and duration: City-scale evaluations using mobile phone application data. Prev. Med. Rep. 8, 30–37.
Wang, X., Lindsey, G., Schoner, J.E., Harrison, A., 2016. Modeling bike Share Station activity: effects of nearby businesses and jobs on trips to and from stations. J. Urban Plan. Dev. 142 (1), 04015001.
Winters, M., Brauer, M., Setton, E.M., Teschke, K., 2013. Mapping bikeability: a spatial tool to support sustainable travel. Environ. Plan. B Plan. Des. 40 (5), 865–883.
Zhang, Y., Brussel, M.J., Thomas, T., van Maarseveen, M.F., 2018. Mining bike-sharing travel behavior data: an investigation into trip chains and transition activities. Comput. Environ. Urban. Syst. 69, 39–50.
Zhou, X., 2015. Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. PLoS One 10 (10).