Review

Sidewalk networks: Review and outlook

Daniel Rhoads [a], Clément Rames [a], Albert Solé-Ribalta [a], Marta C. González [b,c], Michael Szell [d,e,f], Javier Borge-Holthoefer [a,*]

[a] *Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain*
[b] *Department of City and Regional Planning, University of California, Berkeley, CA 94720, USA*
[c] *Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[d] *IT University of Copenhagen, Copenhagen, Denmark*
[e] *ISI Foundation, Turin, Italy*
[f] *Complexity Science Hub, Vienna, Austria*

ARTICLE INFO

ABSTRACT

From a transport perspective, increasing active travel –and walking in particular– is crucial for the future of sustainable cities, as reflected in global decarbonisation policies and agendas. Further, walking is much more than a mere mode of transport: it provides a fundamental social function, fostering vibrant cohesive communities. Arguably, walking and its associated infrastructure –sidewalks– should rank among the highest priorities for planning authorities. However, efficiency- and speed-driven urbanisation has gradually reallocated street space to private cars, leading to automobiles being the prioritised mode of transport today. Empirical research has generally followed suit, and a systemic understanding of walking as a phenomenon is largely missing, *i.e.*, questions like how connected, resilient, accessible, or socially equitable is the pedestrian infrastructure of whole neighbourhoods and cities. Such relative neglect of sidewalk network research is, first and foremost, the consequence of a generalised lack of publicly available data on sidewalk infrastructure worldwide. A second reason might be its apparent lack of interest from a systemic standpoint: pedestrian mobility does not produce coordination challenges on the scale that cars do. In this work, we confront this perception by showing that there is ample research potential in the study of system-wide sidewalk networks, with both structural and dynamical challenges which might be critical to pursue the latest aspirations towards sustainable mobility in cities.

1. Introduction

Walking[1] is arguably the most fundamental mode of human travel. It is cheap and sustainable, and provides wide-ranging transport, health, and environmental benefits to cities (Mueller et al., 2015; World Health Organization, 2022), some of which can be calculated in economic terms. For example, one recent cost-benefit analysis estimated that each kilometer walked in the European Union provides a net positive economic benefit of €0.37 to society, while each kilometer driven by car incurs a cost of €0.11 (Gössling, Choi, Dekker, & Metzler, 2019) due to pollution, traffic congestion, and health impacts, among other things. For these reasons and more, the promotion of pedestrian mobility, along with other active modes of travel like cycling, is increasingly recognized as a priority for sustainable urban development, and it is widely recommended that cities should become more "walkable" (Lo, 2009; Speck,

2013; Katz, Scully, & Bressi, 1994).

However, walking is much more than a mere mode of transport. With its human scale, it provides a fundamental social function, fostering lively, cohesive communities, as well as local commerce. Already many decades ago, sidewalks were recognised as "uniquely vital and irreplaceable organs of city safety, public life and child rearing" (Jacobs, 1961). Many of the positive ideas we associate with the "city" – the palpable energy, the diverse street scenes, the capacity to surprise – have their roots in the vibrant pedestrian life enabled by high-quality, walkable urban design.

Throughout most of human history, all cities were walking cities, and any discussion of the benefits of walking would have seemed unnecessary. Since the beginning of the 20th century, however, most cities around the planet have gradually reallocated increasing amounts of public urban street space exclusively to a new, modern form of

---

[1] Throughout this work, the terms "walking" and "walk" are used generally for any form of pedestrian mobility, including manual or powered wheelchair use.

transportation: private cars. Throughout what Peter D. Norton has called the "Motor Age" (Norton, 2011), most new developments have prioritized the needs of private cars, and existing urban forms have been retrofitted to accommodate them, at the expense of all other modes of transit. This car-oriented development has led to widespread "car dependency" in cities throughout the world (Newman & Kenworthy, 1989; Saeidizand, Fransen, & Boussauw, 2022; Haustein & Nielsen, 2016), and has contributed to a host of social problems including wide-spread traffic congestion, driver and pedestrian deaths from car crashes, and air and noise pollution (Solé-Ribalta, Gómez, & Arenas, 2018; Gössling et al., 2019; Rifaat, Tay, & De Barros, 2011).

Because of these developments, there is a growing push among planners, researchers, and the general public, to shift back towards a human-scale model of urban (re) development (Cervero & Radisch, 1996; Resch & Szell, 2019; Nieuwenhuijsen & Khreis, 2016), reclaiming the city for pedestrians. As a consequence, walking has received significant attention in the transportation literature over the past few years in terms of localized physical design, for example Pedestrian Level of Service (PLOS) models (inspired by vehicular LOS models) that operate on a small spatial scale, *e.g.* at the level of intersections and mid-blocks (Nag, Goswami, Gupta, & Sen, 2020). On the other hand, a plethora of walkability indices have been developed (Vale, Saraiva, & Pereira, 2016) that incorporate both characteristics of infrastructure quality, as well as proximity of daily wants and needs.

In the case of both PLOS models and walkability indices, walkability is measured on a local-scale (even if averaged across larger areas such as neighborhoods or cities). Meanwhile, the systemic and social perspective is largely missing. How connected, resilient, accessible, or socially equitable (Pereira, Schwanen, & Banister, 2017; Gössling, 2016) is the pedestrian infrastructure of whole neighborhoods or cities, and how can it be improved? Such large-scale understanding can be achieved well with the tools of network science: the large number of intersections, links, and connected elements in the system of sidewalks, crosswalks, and paths are naturally formalized as spatial networks (Barthelemy, 2022). Indeed, for road networks many systemic questions have already been answered this way (Barthelemy, 2016; Batty, 2012), but active transportation modes –cycling, and especially walking– remain crucially understudied (Olmos et al., 2020; Rhoads, Solé-Ribalta, González, & Borge-Holthoefer, 2021; Vybornova, Cunha, Gühnemann, & Szell, 2022).

What factors might explain the relative neglect that sidewalk network research has experienced? The most evident reason is the generalized lack of publicly available data on sidewalk infrastructure worldwide. Many municipal open data portals lack clearly available sidewalk data entirely. When available, they are not standardised to the same extent as road network data. More subtly, a second reason for a scarce research tradition in city-wide sidewalk networks is their apparent lack of interest from a systemic standpoint: despite some coordination challenges for pedestrian mobility especially in high-density cities, from pavement maintenance/safety to accessibility and potential sidewalk congestion (Corazza, Di Mascio, & Moretti, 2016; Aghaabbasi, Moeinaddini, Asadi-Shekari, & Shah, 2019; Feliciani & Nishinari, 2018), pedestrian mobility does typically not appear to produce the pressing coordination challenges on the scale that private cars do, whether they be structural (large infrastructural investment, ecosystem fragmentation, etc.) or dynamical (congestion, pollution, etc.). This imbalanced perception can be due to many reasons, such as marginalisation or visibility differences (Colville-Andersen, 2018; Szell, 2018; Rhoads et al., 2021), the much higher potential to congestion of vehicular traffic, or the more self-organising nature of pedestrian mobility (Helbing, Molnár, Farkas, & Bolay, 2001; Feliciani & Nishinari, 2018).

In this work, we hope to challenge this perception by showing that there is ample research potential in the study of system-wide sidewalk networks: from understanding and predicting pedestrian flows, to estimating origin–destination matrices, to assessing walking-distance allocation of services, *i.e.* the "15-min city" (Moreno, Allam, Chabaud, Gall,

& Pratlong, 2021; Xu, Olmos, Abbar, & González, 2020). To effectively address these issues, we make the case that sidewalk networks need to be able to closely approximate the structure and behavior of the physical pedestrian infrastructure systems they model. Researchers have called such close virtual analogues of physical systems "digital twins" (Ferré-Bigorra, Casals, & Gangolells, 2022). Achieving this means thinking of sidewalk networks on their own terms: not as a derivative of road networks, but as a unique class of system, with unique properties and applications. The necessity of having available and using sidewalk network data sets is underscored by a plethora of applications –relevant beyond pedestrian mobility itself– such as the last mile problem (Park, Farb, & Chen, 2021; Tzouras et al., 2023; Ha, Ki, Lee, & Ko, 2023), multimodal mobility (Alessandretti, Natera Orozco, Battiston, Saberi, & Szell, 2022), machine learning and network approaches to pedestrian safety (Osama & Sayed, 2017; Bustos et al., 2021; Ghomi & Hussein, 2023), or sidewalk interactions between humans and autonomous delivery robots (Jennings & Figliozzi, 2019; Gehrke, Phair, Russo, & Smaglik, 2023).

Before outlining the contents of this review, let us first clarify which topics are *not* dealt with here. This work does not engage with individual or localized pedestrian behavior, *i.e.*, evacuation dynamics, human crowd management, or micro-scale pedestrian navigation decisions. Such phenomena, often studied with a blend of empirical data and agent-based models, depend on the perception of the immediate environment (is a car coming my way? Are there obstacles that should be avoided?) (Seneviratne & Morrall, 1985; Brown, Werner, Amburgey, & Szalay, 2007; Kretz, Grünebohm, Kaufman, Mazur, & Schreckenberg, 2006), and system-level planning is unnecessary. Instead, we are concerned with a broad spectrum of questions, ranging from centrality measures, to network robustness, to multilayer transportation networks, that consider the walkable city as a complete system that can be studied, quantified, assessed, and optimized for its purpose of facilitating and encouraging walking as a modal choice.

To start with, Sections 2 and 3 provide the minimal ingredients for a sidewalk network to be constructed. With the relevant object at hand, the next step is to study its structural characteristics (Section 4) under the general framework of spatial networks. Section 5 provides an overview of some open problems in pedestrian networked dynamics. Opening to a broader perspective, Section 6 places the pedestrian network research in the context of multimodal transportation networks (Alessandretti et al., 2022). Finally, Section 7 unfolds some open problems and future challenges that urban analytics, data science and transportation research may address in the near future, in the interests of promoting a healthier, more sustainable and inclusive mobility.

2. Defining pedestrian networks

Many empirical physical and social systems can be usefully represented as networks. A network can be formally defined as a graph $G = (V, E)$ where V is a set of elements called vertices or nodes, and E is a set of ordered or unordered pairs of nodes, called links or edges, which represent relationships between those nodes. In the case of urban transportation networks, nodes represent discrete locations, which could be origins, destinations, or intermediate points on a journey, while edges represent navigable connections between those locations. In some cases, the translation from real-world system to network model can be intuitive and straightforward. The nodes of a subway system, for example, can be naturally defined as the stations or stops along the system's lines, with the edges indicating the rail connections between one station and another.

Sometimes, however, the mapping from empirical system to network model is not so immediately obvious. Consider the example of network representations of urban road systems. The most common approach to model such systems is the so-called "primal" approach (Porta, Crucitti, & Latora, 2006; Lin & Ban, 2013), wherein network nodes represent the intersections of different road segments, and each road segment between two intersections is considered as a link. Essentially, nodes serve as

decision points where a driver can choose to either continue down the same street, or change its direction. This matches our intuitive understanding of how roads are used, but there is a problem: the model implies that drivers begin and end their journeys at intersections. Attempting to resolve this by, say, inverting the mapping such that road segments are the nodes (thus also origins and destinations), and intersections are the edges, has some advantages, but this "dual" approach (Porta, Crucitti, & Latora, 2006) (see Fig. 1A) conflicts with our understanding of how these systems are actually navigated. Accordingly, for most use cases, the primal approach is preferred, despite the information loss it incurs by abstracting away from exact origins and destinations. In cases where precise point-to-point routing is needed, dynamic segmentation (Fischer, 2006; Dueker & Vrana, 1992) of primal

networks can allow for the placement of temporary nodes along the network corresponding to actual origin and destination points, without overloading the size of the underlying network data structure.

The point here is that the translation from an empirical system to a network model has to be done intentionally and carefully, and the choices made in the process should respond to the needs of the application, *i.e.* be fit for purpose. The network that emerges will not be a one-to-one reflection of reality, but instead an incomplete yet useful abstraction. For our purposes, this means that we need to make decisions as to what a pedestrian network should look like –how should the model be designed to best reflect reality and achieve its practical purposes?

Up to now, the answer in the literature has most often been to use some variation of the well-studied and standard primal model of the
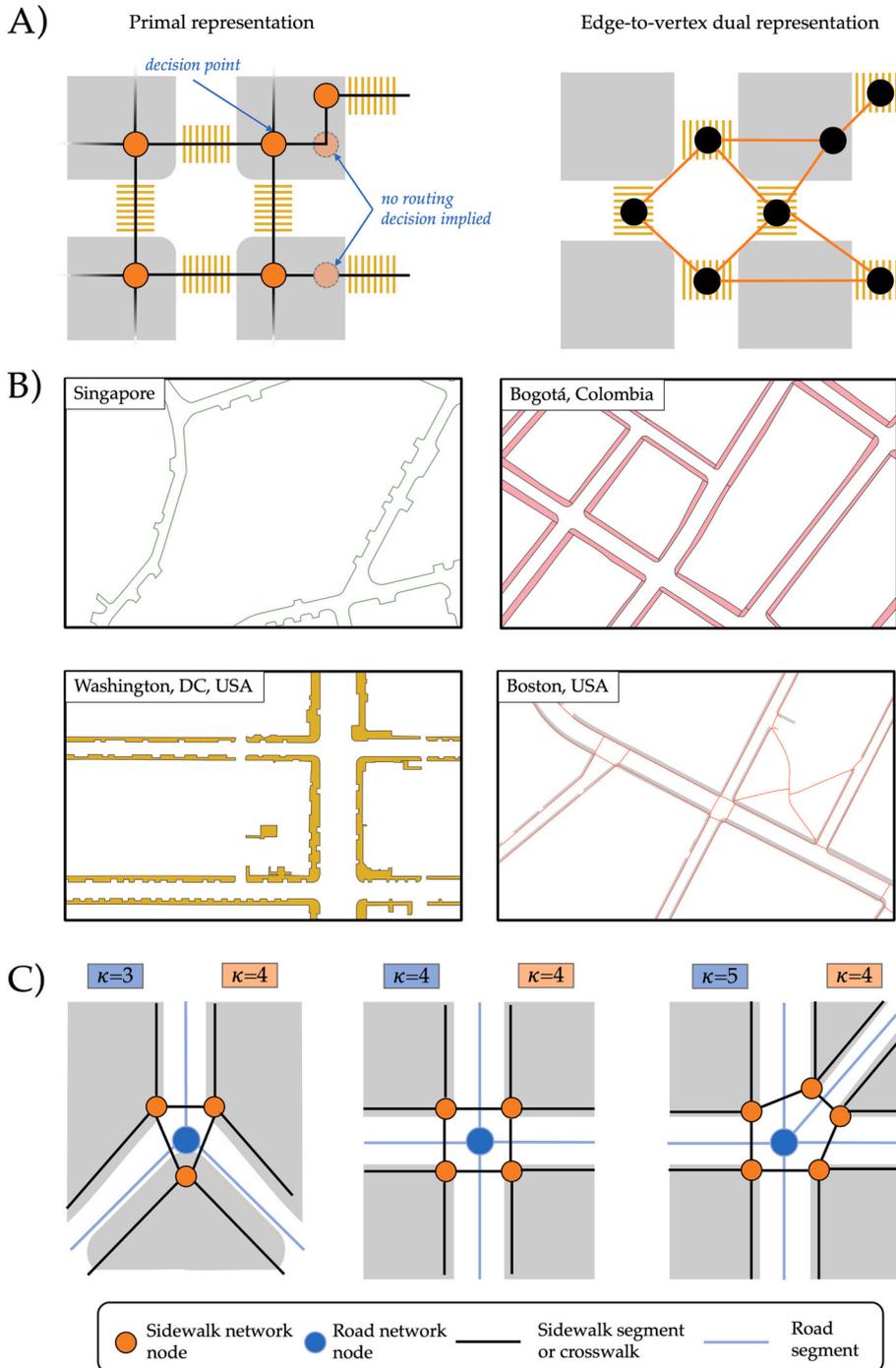
Fig. 1. (A) The "primal" and "dual" representations of real urban road infrastructure are illustrated. Neither of these representations is more "correct", and each can serve a purpose depending on context. "Decision" points are highlighted in the primal representation, and also "no routing decision" points where placing a node is not needed. Essentially, the condition to place a node boils down to the that node having more than two connections, $k > 2$. (B) Examples of existing sidewalk data in a variety of forms: curb lines (Singapore), sidewalk polygons (Bogota, Colombia and Washington, DC), and sidewalk center lines (Boston). (C) The primal model of sidewalk networks under common urban design patterns. Sidewalk nodes (orange) tend to have degree $k = 4$, regardless of the configuration and degree of the underlying street network (blue).

road network as a "good enough" proxy for a pedestrian network. After all, most urban street space is divided between carriageways for vehicles, and the sidewalks that run adjacent to them (Colville-Andersen, 2018; Szell, 2018; Rhoads et al., 2021). It follows that the structures of these spatially-embedded networks should be related. However, as explained next, there are several properties of pedestrian infrastructure networks that make them unique, and that call for dedicated study separate from other urban networks.

2.1. Moving away from road network analogies

For several years now in the planning literature, it has been known that reducing pedestrian networks to the topology of their adjacent road networks results in important miscalculations of network metrics relevant to pedestrian mobility. Chin, Van Niel, Giles-Corti, and Knuiman (2008) created a pedestrian network by supplementing the road network of a study area in Perth (Australia), with links corresponding to pedestrian-only paths and trails. Several of these links connected closed cul-de-sacs with adjacent roads, thus increasing the connectivity of the pedestrian network with respect to the road network. This same approach was taken up by Tal and Handy (2012), with similar results, in their study of walkability in Davis (California, US).

Building on such work, the authors of Cambra, Gonçalves, and Moura (2019) developed a useful typology of links on pedestrian networks, including those which are for exclusive pedestrian use. The authors distinguish between the formal and the informal parts of the network. The former is comprised of sidewalks and marked or legal crosswalks, while the latter includes paths through more amorphous spaces like empty lots, parking areas, green spaces, etc., as well as informal crossings. In the central neighborhood of Avenidas Novas (in Lisbon, Portugal), the share of the network taken up by informal links was as high as 25% (in length). While defining such informal links is difficult due to lack of data, it is clear that they cannot be simply discounted when they compose such a large part of the usable network.

At the very least, then, a pedestrian network must incorporate links that are unique to it, such as trails and pedestrian streets. This most basic model of the pedestrian network is already available at good quality and on a large scale through OpenStreetMap (OpenStreetMap contributors, 2017), and specifically through the OSMnx Python package (Boeing, 2017). Defining the pedestrian network as an augmented road network in this way is useful in some cases, but still abstracts away from important details of the pedestrian experience.

2.2. Sidewalk segments as fundamental units

Networks can be enriched with data far beyond the basic interconnection between nodes. The most detailed and realistic road network models, for example, include edge metadata related to the characteristics of each road segment. For the common purposes of route planning and traffic modelling, the most useful characteristics to include are some proxy for road segment capacity (*e.g.* number of lanes) and maximum speed at free flow. These characteristics help to determine whether or not a segment is navigable under given conditions, and whether or not it forms part of the shortest path between any two points (i, j) in the city. Likewise, sidewalk networks can integrate relevant data on individual sidewalk characteristics to allow for more realistic walkability and accessibility analysis. These characteristics cannot be assigned on a street-by-street basis, but must be determined for individual sidewalk segments. It may be the case that a sidewalk on one side of a street is wide, or of good quality, while its counterpart is narrow, or damaged, or otherwise unusable by all or some pedestrians. Additionally, a pedestrian's specific destination will always be on one side of the street or the other, and conflating the two can be problematic when no legal or safe crossing exists between them. All in all, it is clear that the sidewalks on each side of the street must be considered as separate, fundamental units of pedestrian network navigation.

Defining a model that accounts for this means moving towards a "primal" representation of sidewalk networks, adjacent but irreducible to the road network it shares space with. Following the decision-point criteria used in the construction of road networks, the nodes of a sidewalk network can be placed at any point where a pedestrian has to make a decision, either to turn and walk down another side of the current block, or to cross the roadbed to another block (Rhoads et al., 2021). In Fig. 1A, the idea of a "necessary decision" is highlighted in the primal representation. Note as well that, a block corner with no crosswalk does not imply a routing decision, and so there is no need to define a node. In practice, this can be achieved in several ways. In the best case, where geometric data on crosswalks and other legal crossings are available (see Section 3 for further discussion of data availability for network construction), the points where these crossings intersect with the sidewalk can be designated as nodes. Operationally, placing a node on the network demands that such node has more than two connections, that is, $k > 2$.

Under this model, sidewalks spanning each edge of each block are the edges of the network. This opens the door for the use of edge metadata to "tag" edges with sidewalk properties such as length, width, slope, or even pavement quality. This will be further discussed in later sections, particularly in the context of percolation on sidewalk networks. It bears mentioning that pedestrian-only paths such as park trails or stairs can be incorporated in this model under the same principle, with nodes placed at decision points. Network models of this type have begun to be the standard for city-wide studies of pedestrian networks (Rhoads et al., 2021; Bolten et al., 2015; Bolten, Mukherjee, Sipeeva, Tanweer, & Caspi, 2017; Bolten & Caspi, 2021; Bolten, 2020; Rhoads, Solé-Ribalta, & Borge-Holthoefer, 2023; Hennessy & Ai, 2023).

Having established the parameters for what a sidewalk network should look like, we can begin our discussion of how to use them in practical research. Unlike some more abstract classes of networks (*e.g.* random networks), sidewalk networks are singularly useful as representations of real world physical systems. Accordingly, constructing and studying sidewalk networks means starting with empirical data on urban pedestrian infrastructure –a topic which we will now take up.

3. Data standards and sources

Effective research on urban transportation networks requires reliable data. As of today, publicly-available sidewalk network data sets are nearly non-existent. As a consequence of the current lack of easily-available data, researchers interested in studying pedestrian networks as an entity structurally distinct from the road network (as described in the previous section), have often resorted to constructing networks by hand for select areas of a given study city (Chin et al., 2008; Tal & Handy, 2012; Cambra et al., 2019) –an approach which quickly becomes unfeasible with increasing scale. A more practical approach is to start with a pre-existing geographic data set from which network geometries can be extracted or inferred.

To construct a pedestrian network in the way laid out in Section 2, the main input is a set of polyline geometries representing the right-of-way of each one of the pedestrian paths (sidewalks, crossings, trails, etc.) of the target area (city, neighborhood, etc.). These polylines will be the links of the network. Any intersections between two or more polylines is designated as a node, indicating that a pedestrian navigating the network can choose to traverse a new link, or to finish their journey. Once such an input dataset is available, constructing the network is a fairly trivial GIS exercise. As it stands, acquiring and processing that input data set is the principal challenge of sidewalk network construction.

In the following section we will survey the various sources and types of sidewalk infrastructure data that are currently available, from official data sets to crowd-sourced alternatives, and we will discuss the potential for future data curation and management efforts.

3.1. Municipal data

For researchers hoping to make progress in the study of pedestrian networks, the most obvious point of departure is municipal data sources. City governments are the public entities best equipped to create and maintain data sets of pedestrian infrastructure. Considering the current state of affairs laid out so far in this review, it is not surprising that almost no ready-to-use sidewalk network data sets are available from municipal sources. Instead, the best approach is to construct the networks using available non-networked data on sidewalk and other pedestrian infrastructure that municipalities make publicly available. When available at all, data on urban sidewalks is presented in at least 3 different forms: curb lines, sidewalk polygons, and sidewalk center lines. Examples of all 3 are provided in Fig. 1B.

The lack of uniformity in data sources available from municipal portals may be attributable to the lack of clarity regarding the position of sidewalks in the division of public space. While urban streets are almost universally understood to be a public good to be maintained by the government, legal norms for sidewalk stewardship vary. In cities such as New York, Sao Paolo, and Brussels, property owners and residents are responsible for sidewalk installation and maintenance, while in Los Angeles, Washington DC, and London, that responsibility falls to municipal authorities.[2]

Curb lines are one form of sidewalk-related data that is likely maintained by most city governments, whether or not they are made available publicly, since they delineate the legal right-of-way for cars in public roadways. However, while all curb lines coincide with the edge of a roadway, not all of them will have a corresponding sidewalk. Instead, they may abut private lawns, walls, or non-navigable patches of green space. Furthermore, curb line data sets provide no information on sidewalk characteristics such as width.

Sidewalk polygon data overcomes these drawbacks. When available from municipal sources, these polygons are usually manually constructed with the use of aerial imagery as a guide. These datasets will produce no false positives (indicating the presence of a sidewalk where none exists), and they implicitly provide some information on sidewalk attributes, such as area and width. However, the data is geographic. Transforming such data into a network requires significant processing which can vary depending on the particular nature of the polygon data, the diversity of which is illustrated by the examples of data available from Bogotá, Colombia and Washington, DC (Fig. 1B).

Fewer cities offer data sets of sidewalk center line geometries. From the perspective of building a network, center lines are the easiest data type to manage, since they are similar in structure to the final network. While these data sets do not maintain the physical shape of the sidewalk in the same way that polygons do, lines can be tagged (manually or otherwise) with attributes like sidewalk width and area.

Adding to the difficulties of heterogeneity, it is not rare to find inaccurate, incomplete, incorrect or outdated data. Fig. 2 illustrates some typical situations that can be encountered when closely comparing high-resolution aerial imagery with polygon geometries supposedly reporting on the presence of sidewalks and/or crosswalks.

It is important to note that municipal data sets vary widely in their treatment of non-sidewalk elements of pedestrian infrastructure. Data sets on crosswalks may or may not be available. Where available, this data can come in the form of points, lines, or polygons. Likewise, formal non-sidewalk pedestrian facilities such as paths through parks, hiking trains, tunnels, etc. may or may not be available. As will be discussed later, future efforts to standardize sidewalk network data should aim to generate data sets that are as comprehensive as possible. In the meanwhile, alternative data sources such as the ones described below are often required for non-sidewalk elements.

3.2. Crowd-sourcing and OpenStreetMap

OpenStreetMap (OSM) is by far the best known and most used crowd-sourcing platform for open geographic data. As has already been mentioned, much of the data available on OSM is derived from data produced by public authorities. However, the open nature of the platform does allow for the possibility of crowd-sourcing new data sets that do not yet exist.

Like the EU's INSPIRE and the US Census Bureau's TIGER/Line data, OSM maintains its own data standards for the representation of different geographic entities. Currently, in a reflection of the broader state of the art, the standard for pedestrian infrastructure is uneven. In many cases, data on pedestrian infrastructure is limited to the road network level, where the field "sidewalk" on a street segment can be set to the values "left", "right", "both", and "no", indicating the presence or absence of sidewalks on either side of the street. There is also a separate "footway" entity that encapsulates several categories of pedestrian paths. The use of one or the other standard is dependent on data available in the study area. The OpenSidewalks[3] or SharedStreets[4] initiatives are examples of programs intending to give more uniformity to the standards around pedestrian infrastructure, in order to guide future contributions to OSM.

Major changes to the prevailing public policy surrounding pedestrian infrastructure and network data (*i.e.*, no policy) should be in sight. In the meantime, crowd-sourcing efforts do present opportunities, at least at a city-scale. One successful project (Saha et al., 2019) asked volunteers to assess street view imagery in order to enrich sidewalk data in Washington DC, indicating, for example, whether or not sidewalks were damaged, or had accessible curb ramps –work that might otherwise be delegated to field workers from the city government. Similar efforts exist for some European cities (Bartzokas-Tsiompras, Photis, Tsagkis, & Panagiotopoulos, 2021). The organization of "mapathons" dedicated to the manual construction of pedestrian networks for upload to OSM with the aid of high-quality aerial imagery is another promising route (Tanweer et al., 2017; Gaspari et al., 2021).

3.3. Towards abundant pedestrian infrastructure data

Establishing a standard for the representation of pedestrian infrastructure in network form, as was done in Section 2, facilitates future efforts to develop the pool of available sidewalk data. Standards of this kind already support most research on motorized transportation networks. Many works in the study of road networks rely on the excellent OSMnx Python package (Boeing, 2017) to download urban road networks in a standardized format with a few lines of code, but it should not be forgotten that OSM data relies critically on government data sources, whether at the national or municipal level. The US Census Bureau's TIGER/Line GeoData standard, often the basis for OSM line data in the US, presents road network data at a national scale. No standard exists for sidewalks, much less pedestrian networks, in the TIGER/Line data. The same goes for the European Union's INSPIRE standard, which has specifications for air, road, and rail networks, but not for pedestrian networks. In the face of such a lack of consensus around data formats, and the limitations of crowd-sourced platforms, it seems clear that researchers might need to tap alternative resources which should not depend on institutional initiatives.

Examples include the use of user-volunteered GPS traces to track pedestrian activity (Hunter et al., 2021; Mobasheri, Huang, Degrossi, & Zipf, 2018), or inference from mobile phone traces (Jiang et al., 2013) based on call detailed records (CDRs). Still, these alternatives have their own shortcomings. Neither GPS traces nor CDRs are generally available, and, when they are, data belong to very specific locations and time ranges. Also, it is uncertain whether a given trace belongs to a

---

A)

B)

C)

D)

E)

F)
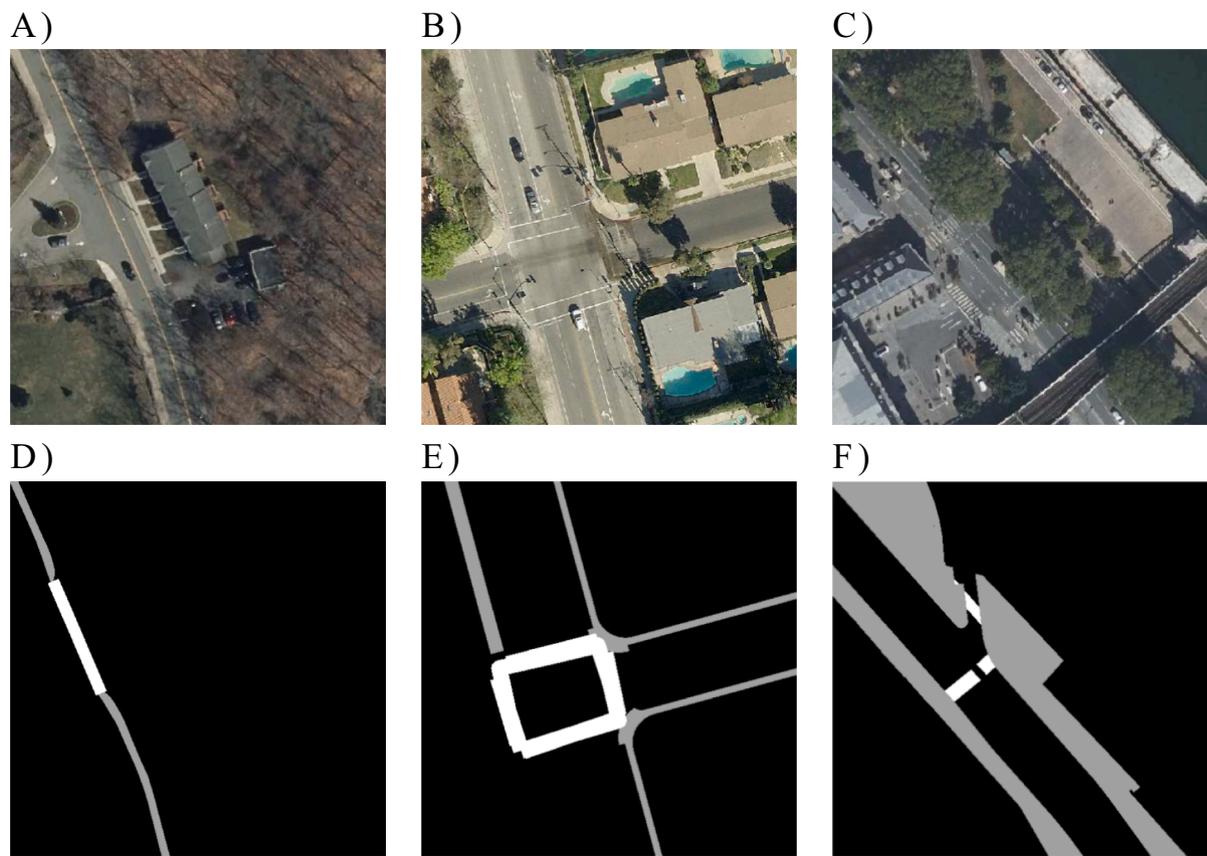
Fig. 2. Illustration of some typical errors in reported sidewalks and crosswalks polygon geometries. Top row displays the aerial imagery, bottom row the polygon geometries. Panels A and D show an example in Berlin (Germany) where two precarious sidewalks (grey) are reported to connect via a crosswalk (white), but the aerial image shows no sign of the existence of such connection. Panels B and E show an example in Los Angeles (US). There, the lower-left corner of the intersection is missing, such that the crosswalk (white) leading to that area becomes a dead-end. Further, the left-most portion of sidewalk does not connect to its corresponding crosswalk –an automatic script to build a network representation from this information would leave these two objects disconnected. Finally, panels C and F show and example in Paris (France) where we observe two disconnected crosswalk segments (white), and the two remaining crosswalks are simply absent in the polygon geometries. We also observe a large portion of missing sidewalk (grey) information nearby the river, where the image shows people gathering.

pedestrian, a cyclist, or a driver, and attempts to disambiguate this are not straightforward.

All in all, the most promising direction toward a pipeline that leads to reliable data on sidewalk infrastructure is remote sensing. In particular, computer vision for applications in remote sensing can support and resolve challenges exploiting, *e.g.*, large satellite or aerial image data sets to collect and identify features in an environment with accuracy and speed. In the context of road transportation, there are successful applications for the identification and mapping of road networks (Cheng et al., 2017; Máttyus, Luo, & Urtasun, 2017; Bastani et al., 2018), with a combination of pixel classification, image segmentation, and graph extraction techniques. Also, the use of street-level imagery has proved useful to assess different urban features (Zünd & Bettencourt, 2021; Bustos et al., 2021). These ideas have been put to work in the pedestrian infrastructure context in Li et al. (2018, 2020, 2022, 2023), and may be the path to solving the data scarcity problem around walking infrastructure.

4. Structure of sidewalk networks

Interestingly, the study of networks had unofficial origins in an urban setting. In 1736, the Swiss mathematician Leonhard Euler published his famous solution to the Königsberg bridge problem (that is, finding a round trip that crossed each of the bridges of Königsberg exactly once). Relevant to this section, the inaugural Königsberg bridge problem is not only urban in setting, but it relates directly to the structure of the network in question. Following the spirit of the Königsberg bridge

problem, the next subsections review some features and measures that broadly characterise the topology of real sidewalk networks. The focus of network-level analysis is on properties of networks as a whole. These may reflect typical or atypical traits relative to a specific application domain, or similarities occurring in networks of entirely different origin. In this sense, the following topics aim at connecting the interests and potentialities of complex network researchers –for whom the network's architecture is intimately coupled to its functional and dynamical aspects, such as growth, robustness, response to external perturbations, or the onset of congestion from the free-flow state– to those of urban science.

We are aware that, by taking this perspective, we momentarily overlook the fact that pedestrian dynamics take place mostly at a local scale, with heavy constraints on, for example, the length of a typical on-foot trip: pedestrian travel decays with distance, and trips beyond 2 kilometres seldom occur. However, global features are nonetheless relevant: the apparent discrepancy is cleared up as soon as we understand that all the scales in a network (the micro and macro levels) are intertwined –macro emerges from micro, and micro behaves constrained by macro.

4.1. Planarity and its constraints on graph descriptors

Despite different historical, political, geographical, or financial circumstances, at a coarse-grained level, infrastructure networks of very different cities share quantitative and structural similarities. This is the case because such networks are not only spatial (*i.e.*, nodes are located in

physical space), but often also planar or quasi-planar (Barthélemy, 2011). Planar graphs are those that can be drawn on a 2-dimensional euclidean plane such that none of the edges of the graph intersect.

Pedestrian networks are no exception: like other urban transportation networks, they are embedded in geographic space. Further, they are approximately planar, and this property alone heavily constrains certain graph properties (Barthélemy, 2011; Barthélemy, 2018). For example, the degree distribution of a planar network will always be somewhat narrow (in contrast to some social networks, no power-law degree distributions exist for planar graphs), with an average degree no greater than 6. Underpasses, overpasses, stairs, bridges, and even passageways through buildings are examples of situations where pedestrian networks might break planarity, but these cases are the exception rather than the rule.

Table 1, reproduced from Rhoads et al. (2021), reports the characteristics of the pedestrian networks of 3 cities, compared to their corresponding road networks. Road network geometries were extracted from OpenStreetMap (OSM) using the OSMnx Python package (Boeing, 2017), which provides a simple interface for querying OSM data. The package was used to extract the edges and nodes of the "Drive" network of each city. On the other hand, pedestrian networks were constructed by hand, in the sense that vector data on sidewalks were processed *ad hoc,* depending on the format in which they were offered (seeFig. 1B and accompanying text in the previous section). Those datasets are publicly available, as detailed in the main text and Supplementary Information of Ref. Rhoads et al. (2021).

It is immediately clear from the Table that the pedestrian networks are much larger than the road networks in terms of the number of nodes (N) –about 3 times larger in the case of New York City, and about 4 times larger in the case of Paris and Boston. This can be explained by the nature of the primal model to build pedestrian networks, as described in the previous section. Likewise, the average degree ($\langle k \rangle$, the number of edges that a node in the network is connected to) of the networks hovers between 3.11 (Paris) and 4.13 (Boston). The tendency to converge towards an average degree of 4 is unsurprising. Assuming the presence of all possible legal crossings and all possible sidewalks, under common street patterns, a node placed along a block's sidewalk at the closest point to a given street intersection will tend to have a degree of 4, independently of the degree of the street intersection (see Fig. 1C). Nodes with $k < 4$ may be cul-de-sacs, or end-points of incomplete sidewalks, or else areas where a theoretically-possible legal crossing is unavailable. Nodes with $k > 4$ most of the time form part of a pedestrian-only thoroughfare, which replicates the primal model of road

networks. It is also interesting to observe the differences in efficiency and diameter between the sidewalk and road networks. Sidewalk networks demonstrate greater efficiency, likely due to their higher density, although their diameter may be larger, as seen in Paris. Remarkably, the statistics presented in Table 1 highlight the significance of realism in terms of distances, especially when it comes to representing the sidewalk network, as pedestrians are particularly sensitive to distances.

4.2. Betweenness centrality

In network science, centrality refers to the various measures used to gauge the importance of individual network elements (nodes and edges) to the network as a whole. There are dozens of centrality measures, each useful in different applied contexts. In the case of transportation networks such as sidewalk networks, it can be especially useful to understand an individual element's contribution to transit, or flow, across the network –a feature well-described by betweenness centrality.

Initially introduced in the social sciences (Freeman, 1977), betweenness is a centrality measure that quantifies the importance of a node (or an edge) in terms of the number of paths crossing it. The shortest-path betweenness (B_i) considers only the least costly paths (typically length or traversal time in spatial networks) between locations and is defined, for a given node i, as

$$B_i = \frac{1}{\mathcal{N}} \sum_{o \neq d} \frac{\sigma_{od}(i)}{\sigma_{od}} \tag{1}$$

where σ_{od} is the number of shortest paths going from origin o to destination d, while $\sigma_{od}(i)$ is the number of these paths crossing i. Factor $\mathcal{N}$ represents a normalisation constant which may be different ($(N-1)(N-2), N^2, N, 1$) depending on the application. Note that this equation can be easily adapted to the particular constraints of pedestrian mobility –first and foremost, limitations on the typical on-foot trip. This is explicitly considered in Section 5 on pedestrian dynamics. In the remainder of this subsection, instead, we focus on the unrestricted (*i.e.,* paths from any node to any other node) centrality distribution as a case study to place the accent on the network as an object of study itself.

Betweenness centrality is implicitly related to the concept of path which, in turn, depends on the routes that elements take while traversing the network. This explains why, despite its origins in sociology, betweenness stands out as one of the most prominent descriptors in the analysis of traffic, routing, and congestion (*i.e.*, road networks and the vehicles traversing them). While congestion is not –generally– a problem in on-foot mobility, betweenness, and centrality measures in general, might constitute an insightful framework when considering the social dimension of sidewalk networks as a space for meeting and interaction, *e.g.* to extend the notion of "social interaction potential", as discussed in Farber, Neutens, Miller, and Li (2013), to networked systems where the concept of joint accessibility might be more clearly delineated (Zhang & Thill, 2017). Centrality measures might prove useful as well when determining the optimal location of public services (Xu et al., 2020).

Edge betweenness has a direct relation to pedestrian dynamics, and is discussed in length in Section 5. Here, instead, we focus on node betweenness and its distribution (Kirkley, Barbosa, Barthelemy, & Ghoshal, 2018) on a sidewalk network. Keeping in mind the above-mentioned features –planarity, and a narrow degree distribution around a well defined average: a quasi-regular network– we can foresee that a sidewalk network's betweenness will behave close to the expected values in a grid. Noteworthy, betweenness is a relatively costly descriptor to compute (Brandes, 2001), but analytical derivations are possible for a reduced –but relevant here– family of graphs, such as regular ones (Lampo, Borge-Holthoefer, Gómez, & Solé-Ribalta, 2021; Verbavatz & Barthelemy, 2022).

Table 1

This table compares the sidewalk and road networks of 3 cities according to 5 important network characteristics: number of nodes and edges (N and E), average degree ($\langle k \rangle$), average efficiency ($\langle Eff \rangle$), and network diameter (D). From Rhoads et al. (2021) with permission.

City	Metric	Sidewalk network	Road network
Boston	N	46,547	11,133
	E	96,022	25,455
	$\langle k \rangle$	4.13	2.3
	$\langle Eff \rangle$	9.8×10^{-5}	1.9×10^{-4}
	D	23,823 m	25,769 m
New York	N	129,520	46,007
	E	253,594	116,723
	$\langle k \rangle$	3.91	2.53
	$\langle Eff \rangle$	5.4×10^{-5}	9.3×10^{-5}
	D	39,226 m	51,404 m
Paris	N	40,095	10,073
	E	62,414	19,651
	$\langle k \rangle$	3.11	1.95
	$\langle Eff \rangle$	1.1×10^{-4}	2.2×10^{-4}
	D	22,427 m	19,533 m

4.2.1. Betweenness derivation in regular grids

Let us first consider a 4-regular grid with N_r number of rows and N_c number of columns ($N = N_r \times N_c$), where each node can be identified by its coordinates in that grid. Given a node with coordinates (x, y), i.e., located at x horizontal and y vertical steps from lower-left node $(0, 0)$ of the grid, we can trivially generalise Eq. (9) in Ref. Lampo et al. (2021) such that the (non-normalized) node betweenness can be measured as

$$B_{(x,y)} = \sum_{a=1}^{N_c} \sum_{b=1}^{N_r} \sum_{c=1}^{N_c} \sum_{d=1}^{N_r} \frac{\pi_{|c-x|,|d-y|} \, \pi_{|a-x|,|b-y|}}{\pi_{|a-c|,|b-d|}} \qquad (2)$$

where (a, b) and (c, d) are origin and destination nodes, respectively, and at least one shortest path between those goes through (x, y). On the other hand, $\pi_{x,y}$ is the number of different paths in the grid involving x horizontal and y vertical steps:

$$\pi_{x,y} = \binom{x+y}{x} = \frac{(x+y)!}{x! \, y!}. \qquad (3)$$

Fig. 3 provides a visual support for the explanation of the terms in Eq. (2). For this example, we aim to find out the number of shortest paths traversing node (x, y), such that they start at (a, b) and end at (c, d); as well as the total number of shortest paths between (a, b) and (c, d). To do so, we consider first the paths starting at (a, b) leading to (x, y) (yellow shade): there are 4 (see the legend on the right of the image). Then, we calculate the combinatorics of the paths leaving from (x, y) and leading to (c, d) (blue shade). Trivially then, there are 4×10 shortest paths between (a, b) and (c, d) traversing (x, y). Finally, the grey shadow circumscribes *all* possible shortest paths between (a, b) and (c, d). The relation between these two quantities quantifies how central (x, y) is, with respect to (a, b) and (c, d).

4.2.2. Comparison with empirical data

Fig. 4 sheds some light on the hypothesis above. In it, the nodes of Barcelona's sidewalk network are coloured according to their betweenness value (except for terminal nodes with betweenness 0, which amount to a small number). Interestingly, nodes in the periphery have indeed cooler colours that those in the centre. This is also true if we take some of its districts sub-graphs: see the insets of the figure, corresponding to sidewalk network nodes in the Eixample, Ciutat Vella and Gràcia, where again they are coloured according to their betweenness values.

However, unexpected anomalies are visible to the naked eye as well. These deviations from the idealised betweenness distribution of a standard grid reveal particularities of the sidewalk infrastructure of a given city as a networked system. For example, a glance at Fig. 4 reveals concentrations of high-betweenness nodes along several of the city's large diagonal thoroughfares (e.g., Avinguda Diagonal, Avinguda Meridiana). These high-betweenness axes run from the centre of the city to the periphery, and indicate the relative efficiency of such diagonal "shortcuts" through an otherwise orthogonal grid structure. Likewise, small pockets of low betweenness nodes can be identified throughout the city, generally indicating areas of low connectivity, or at least locations where the particular geometry of the local pedestrian infrastructure makes transit less efficient. By considering all possible origin–destination pairs, we can thus see how betweenness centrality might help to identify locations for interventions to make neighbourhoods more physically walkable.

4.3. Network accessibility and fragmentation

Global aggregate measures like those reported in Table 1 provide us with some insights into the characteristics of a given pedestrian network. However, despite their descriptive power, these quantities tend to obscure the network's functionality as a whole. An example of such obscured features is connectivity and its associated concepts. When thinking of pedestrians, the connectivity of the network, that is, the fact that an agent can travel from any node to any other node, is crucial to the wider concept of walkability, and to understanding whether the city facilitates (or hampers) active travel, social interaction, and other relevant features. This is directly related to the pressing issue of urban fragmentation: the existence of massive infrastructure that divide neighbourhoods otherwise connected. The presence of highways, expressways, and large industrial zones enable and seek long-range connections, while paradoxically creating disconnection at the community level. A relevant question is then how different connection patterns affect network connectivity, and the closely related characteristics of robustness and accessibility. A way to study this question –how easily will a connected network become a set of isolated subgraphs?– is by applying percolation theory. Percolation is the progressive removal of nodes or links in a network to signal their inaccessibility, and it can be envisioned as a process that allows system planners to anticipate structural vulnerabilities (which segments on the street patterns should
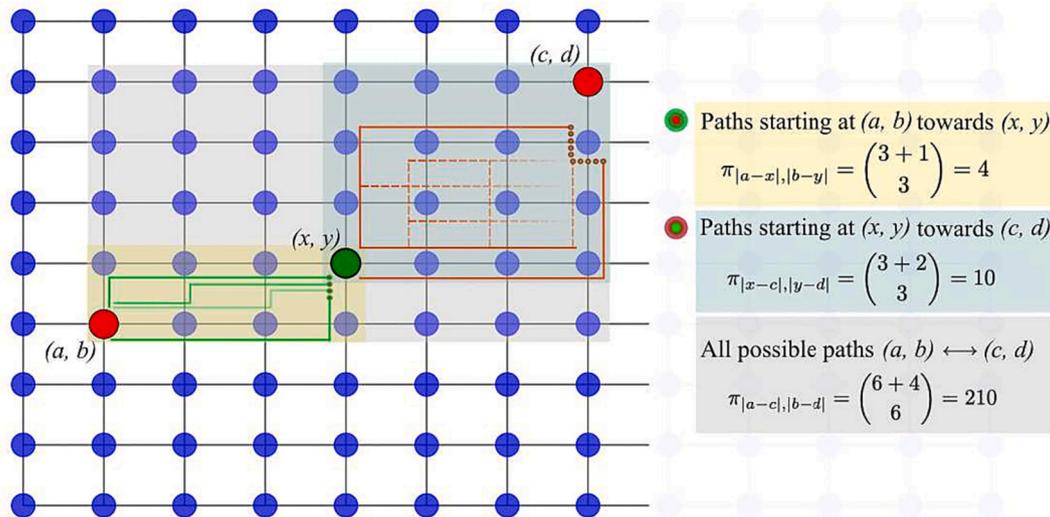
Fig. 3. Diagram of a regular grid, showing an example of how shortest paths through node (x, y) (green) are counted. In the illustration, we consider all those paths departing from node (a, b) and leading to node (c, d) (both in red). The centrality of the green node depends on three ingredients: the amount of shortest paths connecting (a, b) and (x, y) (yellow shadow); the amount of shortest paths connecting (x, y) and (c, d) (blueish shadow); and the total amount of shortest paths connecting (a, b) and (c, d) (grey shadow), regardless of their passing through (x, y) or not (see right legends).
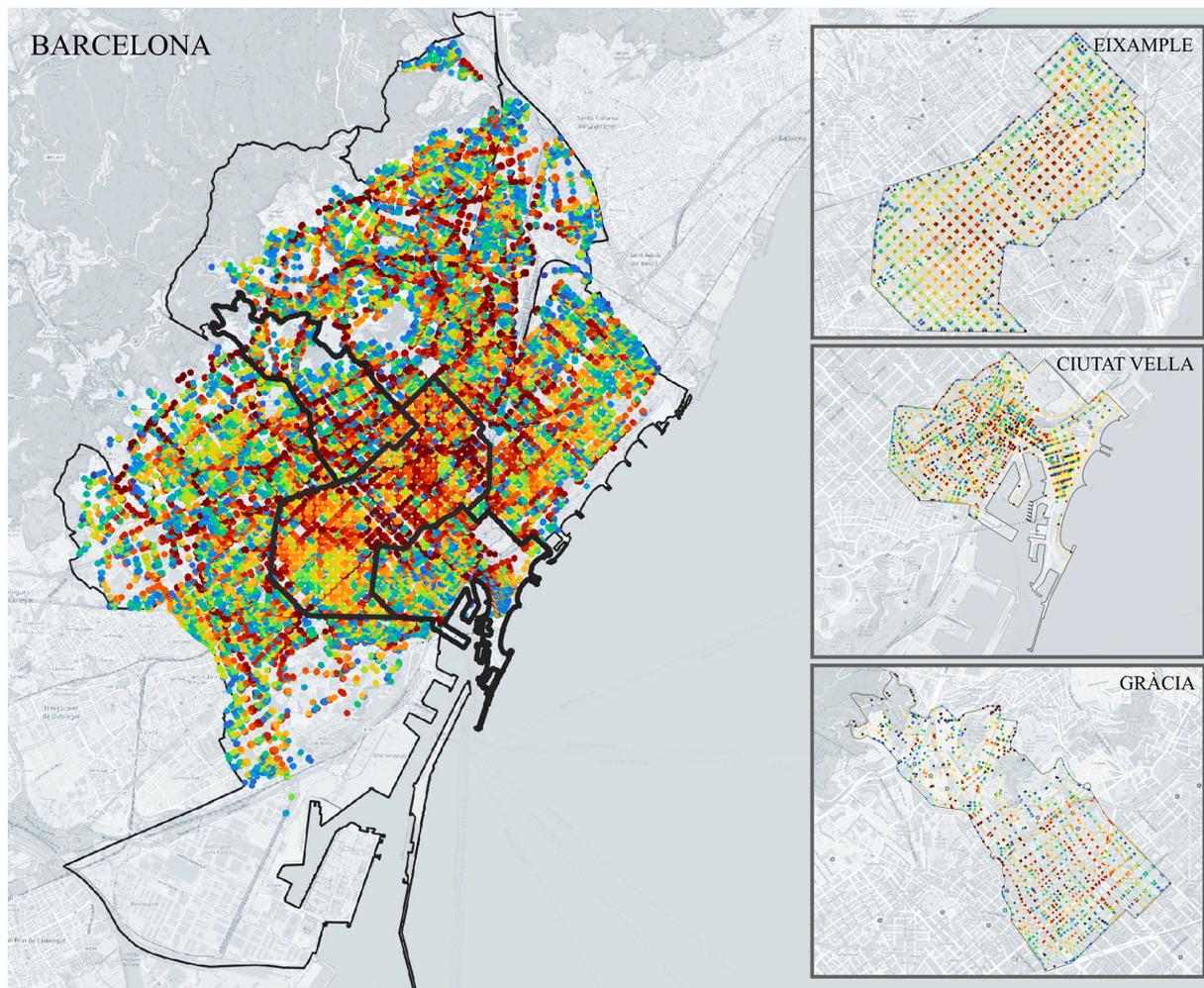
Fig. 4. Barcelona's sidewalk network ($N = 36728$), in which nodes are coloured according to their betweenness (as measured from the unweighted version of the graph). Cold colours (blue, green) correspond to lower betweenness values, as opposed to hot colours (orange, red). Nodes in the city outskirts mostly show cooler colours –as one would expect from peripheral nodes in a regular grid. For the sake of visibility, three district sub-graphs are shown as insets (Eixample, Ciutat Vella and Gràcia), where the trend is clear as well. Note that terminal nodes (cul-de-sacs) with betweenness 0 are not represented.

be secured to prevent the disconnection of a large portion of the city) or plan optimal interventions (where should a segment be placed to decrease the probability of disconnection). This approach has proved useful in several urban settings, although predominantly so far in relation to road networks (Li et al., 2015; Arcaute et al., 2016; Abbar, Zanouda, & Borge-Holthoefer, 2018; Serok, Levy, Havlin, & Blumenfeld-Lieberthal, 2019).

The inaccessibility of a node or a link can be thought of as the consequence of an exogenous and unpredictable event (*e.g.*, a natural disaster that might literally (physically) affect any part of the network Abbar et al., 2018), or by some intrinsic feature of the nodes/links that make them unavailable (*e.g.*, an exceedingly narrow sidewalk for a wheelchair user Rhoads et al., 2023). Translated to percolation processes, the former corresponds to a "random failure" (which node/link is about to become inaccessible can not be known in advance), while the latter can be simulated as a deterministic "attack" on the network (narrow sidewalks become inaccessible first). We illustrate these ideas with examples on random percolation processes. Fig. 5 (adapted from Rhoads et al. (2021)) shows the results of a series of random percolation processes on 10 sidewalk networks from cities in Europe, and North and South America. A single random attack simulation proceeds by sequentially eliminating links from the network at random, and monitoring the size of the two largest connected components of the network, the so-called *giant* and *second giant* components. The giant component

will decrease in size monotonically as links are removed, while the second giant commonly peaks sharply at a point known as the critical threshold $p = p^*$, where p is the percentage of network links that have been removed.

Different ideal classes of graphs exhibit different characteristic critical thresholds. Comparing these canonical percolation thresholds on abstract graph types, with the results of percolation simulations on empirical networks, can provide clues as to the underlying structure of dense systems. The results in Fig. 5 show that the percolation thresholds for most pedestrian networks lie somewhere between that of a triangular and a square lattice (thresholds $p = 0.347$ and $p = 0.5$ respectively Li et al., 2021).

Moving past the commonalities, though, a city-by-city analysis can reveal interesting facts. For example, the results for Buenos Aires show a clear peak at almost exactly 0.5, in line with several studies of the city's street and block configuration that identify it as exceptionally regular and grid-like in form (Louf & Barthelemy, 2014; Boeing, 2019). While the behaviour of the second giant component across 100 simulations for Buenos Aires, Paris, and, to some extent, Barcelona and Brussels is quite uniform, cities such as Denver, Bogotá, and New York present various paths to network breakdown from iteration to iteration, indicating lower redundancy and higher vulnerability. The peak of Montreal's second giant component is in fact the most glaring exception to the rule, peaking at around $p = 0.1$ –a sign of deep systemic vulnerability that
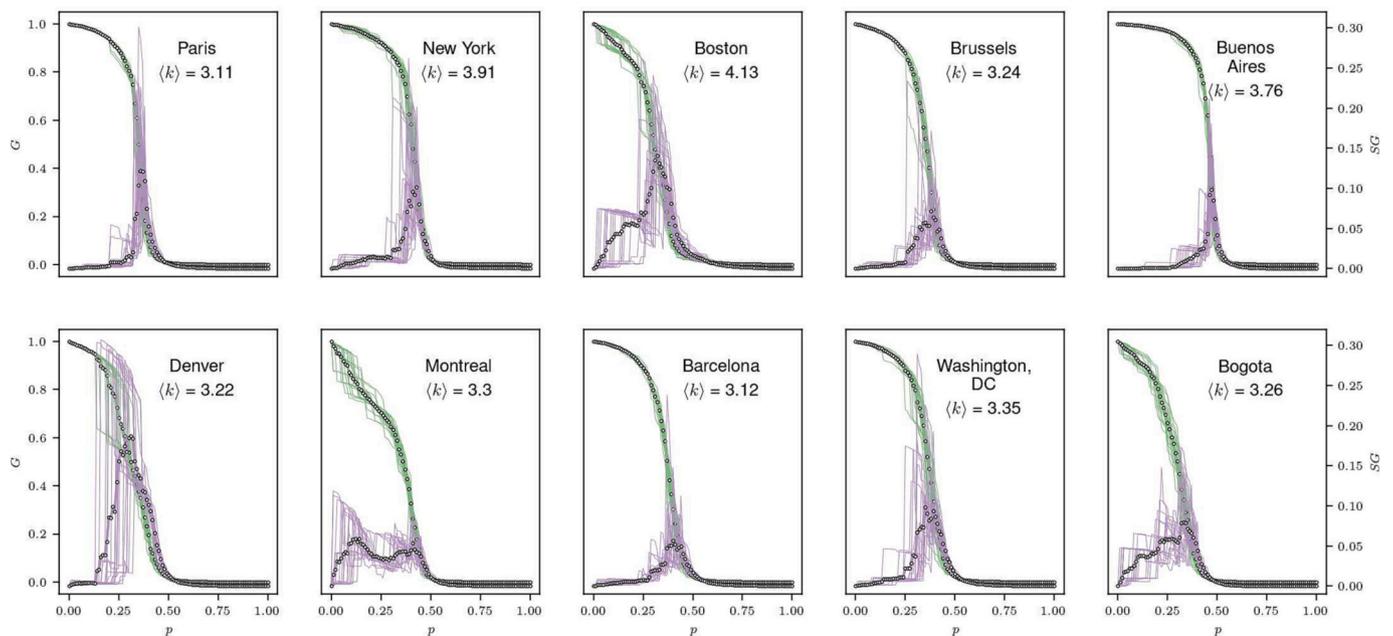
Fig. 5. The panels for different cities illustrate the breakdown in network connectivity for 10 empirical pedestrian networks under 100 iterations of a random percolation process. The green lines indicate the size of the largest connected network component as a fraction of network size *N*, and the purple lines show the same for the second largest component. Individual iterations are indicated with thin lines, while the thicker lines with points show the average behaviour. The critical threshold occurs when the red lines spike to their highest point, indicating system failure. With some variation, breakdown tends to occur in the area $0.347 < p < 0.5$, between the canonical percolation thresholds for bond percolation on triangular and square lattices (Li et al., 2021). Adapted from Rhoads et al. (2021) with permission.

reflects a low density of sidewalks in the southern portions of the city.

These inter-city differences in network robustness, ranging from the subtle to the extreme, suggest that some systems of pedestrian infrastructure work better than others. The question of assessing network quality, and its variation from neighbourhood to neighbourhood and city to city, is addressed in the following section.

4.4. Network completeness

When assessing the quality of an urban transportation network, the most fundamental question might be: can a resident use the network to navigate between any two given points *a* and *b* in the city? If this question is posed from the perspective of car users, the answer, in modern cities, is almost always yes. Conversely, much like the well-known phenomenon of public transit deserts (Jiang et al., 2013), there also exist pedestrian deserts: areas of the city where pedestrian infrastructure is either of low quality or lacking entirely. This makes the basic assessment of network completeness more relevant and pressing in the context of pedestrian networks.

Generally, incompleteness in pedestrian networks can take two forms: absence of sidewalks, and lack of sufficient crossing infrastructure. We discuss both below.

4.4.1. Sidewalk coverage

Sidewalk availability has been positively associated with walking in a variety of contexts and across a wide range of demographics, including among the elderly in Taiwan (Chen, Hsueh, Rutherford, Park, & Liao, 2019), university students and faculty in North Carolina (Rodríguez & Joo, 2004) and residents of Portland, Oregon (Ewing & Cervero, 2010). This seems logical because, on the one hand, walking in the street exposes pedestrians to the (sometimes extreme) risk of being hit by a car, and because low sidewalk availability can lead to longer, more tortuous, less efficient, and more cognitively costly paths (more on that later).

The availability of sidewalks varies within and between different cities. One recent study (Coppola & Marshall, 2020) of four US cities

found a wide range of sidewalk coverage levels, from 27.2% and 29.1% in Austin, TX and Raleigh, NC on the lower end, to 49% and 58% for Portland, OR and Denver, CO on the other. Even those higher numbers indicate a stark gap in sidewalk provision with respect to road infrastructure. Inequalities can also exist between neighbourhoods of the same city. A study of Cascade, a low-income, predominantly black neighbourhood in Atlanta (Gaither et al., 2016), gave statistical backing to the complaints of residents about the lack of adequate sidewalk provision: sidewalk acreage per square mile was half that of the city as a whole. Meanwhile, in Washington, DC, a crowd-sourced sidewalk mapping project (mentioned in previous sections) found over 40,000 points in the city where sidewalks were either missing or discontinuous.

The results presented in the preceding paragraph are statistical and treat sidewalks as individual entities rather than as a networked whole. One recent paper (Bolten & Caspi, 2021) moves in the direction of a network analysis of sidewalk coverage, systematically comparing the "reach" of the sidewalk and road networks for journeys beginning at a given point *i*. This measure, called normalised sidewalk reach, can be defined as:

$$f(i) = \frac{\sum_{j} l_{sidewalk}(r_{ij})}{2\sum_{j} l_{road}(r_{ij})} \qquad (4)$$

where *j* is the set of locations reachable from i with given restrictions (e. g., only paths of less than 400 meters, less than 800 meters, etc.), $l_o(r_{ij})$ is the length of the path *r* from *i* to *j*. Constraints of this type are known variously as pedsheds, walksheds, and egohoods. The ratio of pedestrian to road network reach is divided by 2 under the assumption that each road segment will have at most 2 parallel sidewalk segments.

One promising avenue in this regard, so far unexplored to our knowledge, would be to establish a baseline model for the sidewalk network of a given city. In terms of coverage, for example, an ideal scenario would consider that every possible sidewalk (on both sides of each street) exists and is of good quality. This would allow for a kind of reverse vulnerability analysis, in which the construction of new

sidewalks in underserved areas could be prioritized. Such approach could further be extended to crosswalks, which we will discuss below.

4.4.2. Crosswalk placement

The structure of most cities, which isolate sidewalks between car right-of-ways, leads to a further complication in the design of sidewalk networks –the placement of crosswalks (a.k.a. pedestrian crossings, zebra crossings). Legal crosswalks may be marked or unmarked, depending on the regulation of the particular locality. They may or may not be signalled with a light requiring vehicles to stop at certain intervals. In many cases, what might be a useful crossing point for pedestrians is rendered unusable because of the conditions of the infrastructure, together with the flow of vehicles on the roadway to be crossed. This inability to cross, whether structural or dynamically caused, translates to a link removal on the network.

Most commonly, local laws permit pedestrians to cross from one block to another only at street segment intersections. Further, with very few exceptions (*e.g.*, the famous Shibuya Crossing in Tokyo), pedestrians must only cross a single street when moving from one block to another –in other words, crossing diagonally through the intersection is prohibited. Fig. 1C provides examples of common sidewalk crossing configurations. It is important to note that these are transportation design choices, and not necessities. Pedestrians often follow desire lines and cross at non-designated, irregular points, according to well-studied preferences that take into account factors like street width, traffic flow, and distance to the nearest legal crossing (Cherry, Donlon, Yan, Moore, & Xiong, 2012; Pawar & Patil, 2016; Sisiopiku & Akin, 2003). It follows that the placement of designated crossing points can have an influence on crossing behavior and, in turn, pedestrian safety (Sisiopiku & Akin, 2003). This phenomenon has been explored through agent-based and traditional modeling approaches (Sargoni & Manley, 2020). However, these models tend to operate at the scale of the city block and do not take into account the acute or cumulative effects of the presence or absence of legal crossings on system performance in terms of path lengths or connectivity.

5. Dynamics on sidewalk networks

Often echoed in the literature, the central *mantra* of complex network theory is the interplay between structure and dynamics: just as the structure constrains how dynamics behave across the network substrate, dynamics can, as they unfold in time, strengthen or erode parts of the network. In transport engineering, this fundamental entanglement is most evident when studying vehicle congestion dynamics, which are completely driven by a structural feature –betweenness (Solé-Ribalta, Gómez, & Arenas, 2016; Barthelemy, 2016). More generally, we encounter a very rich and long tradition that deals with micro-, meso- and macro-scale vehicle dynamics, from car-following models, to congestion-driven dynamic percolation, to origin–destination matrix calculation. For studies at the meso- and macro-levels, a precise representation of structure is essential.

Turning from vehicle to pedestrian mobility, a comparable wealth of quantitative knowledge and modelling capacity is replicated in a limited fashion only at the micro-scale, with agent-based modelling of pedestrian flows at the segment or intersection levels, and the development of the fundamental diagram for pedestrians (Seyfried, Steffen, Klingsch, & Boltes, 2005). The closest we get to a system-wide account of pedestrian dynamics is the concept of walkability, which at least delineates the conditions under which such dynamics take place.

Walkability is a multifaceted concept, depending both on subjective (*e.g.*, perception of enjoyment, attractiveness, etc.) and objective features (Arellana, Saltarín, Larrañaga, Alvarez, & Henao, 2020) such as the physical condition of sidewalks. Presently, walkability indices either rely on precise measurements of sidewalks in small areas (Zhao, Sun, & Webster, 2021), or simply resort altogether to road network metrics. Therefore, when developing pedestrian dynamics models at the

meso- and macro-scales, it is crucial to first have a dependable representation of the physical walking infrastructure, which includes the sidewalk network. Then, it needs to consider the particularities of pedestrian mobility as well, which are obviously different from those of drivers, cyclists, and transit riders. First and foremost, pedestrians move at a low and relatively uniform speed. Second, walking trips tend to be limited to short distances, and the probability of choosing walking as a transportation mode has been shown to fall exponentially with the distance to destination (Marquet & Miralles-Guasch, 2014).

5.1. Crowd dynamics at the mesoscale

The walking speed of pedestrians at free-flow follows a Gaussian distribution with $\mu = 1.34 ms^{-1}$ with $\sigma = 0.26$, which is known to minimise the metabolic energy (Henderson, 1971), and each pedestrian occupies a minimum space of $0.15 m^2$ when standing still, an area that increases when they are walking. As the density of pedestrians increases, the available area per pedestrian is reduced and interactions between pedestrians take an important role, especially with regard to walking speed. This motivated the description of the fundamental diagram for pedestrians (Seyfried et al., 2005), relating explicitly sidewalk density and walking speed. The classical work of Weidmann (1993) models the fundamental diagram for uni-directional flows, as a function of the free flow pedestrian velocity minus an exponential decay that depends on the density of pedestrians

$$v(D) = v_f(1 - e^{-\gamma(1/D - 1/D_{max})}) \tag{5}$$

where γ is a free parameter, v_f stands for the velocity at the free flow condition, D is the current density and D_{max} is the maximum density that allows some flow to exist. Although global parameters γv_f and D_{max} can be found for the general case (Martinez-Gil, Lozano, García-Fernández, & Fernández, 2017), the fundamental diagram is highly dependent of other factors, including the bi-directionality of the flow (Zhang, Klingsch, Schadschneider, & Seyfried, 2012) characteristic of sidewalk segments. Studies of the influence to the fundamental diagram of merging pedestrian flows in T-junctions are also relevant (Zhang, Klingsch, Schadschneider, & Seyfried, 2013).

Road traffic analysis has shown us that the fundamental diagram is a basic actor to understand system-scale phenomena, such as congestion, traffic assignment and system optimality, etc. (Ambühl, Menendez, & González, 2023). From the perspective of pedestrians, these implications are still to be explored. Their fundamental diagram has been only studied, up to now, independently of the networked structure of sidewalks. Clearly, pedestrian congestion does not have the same definition as for cars, but the concept still relates to the level of service that a sidewalk network can offer. Gehl Architects' studies indicate that sidewalks exceeding more than 13 pedestrians per width-meter and minute are not comfortable for pedestrians (Gehl et al., 2008). Discerning whether there is an influence of the fundamental diagram at system scale is central to having a complete understanding of sidewalk network use and their design.

5.2. Estimation of pedestrian flows at the city scale

As when modelling road traffic, the estimation of pedestrian flows is mainly based upon the consideration of two elements: origin––destination matrices (OD-matrix) and pedestrian route choice models. In some situations, these two factors may be independently calculated, while in others there is a formal dependence between them, in the form of route choice assumptions, or the type of data at hand. The following two sections will focus on the estimation of these two components, emphasizing these dependencies.

Other techniques based on time series analysis and machine learning models exist for pedestrian and road traffic prediction (Jiang, Ma, & Li, 2022; Cohen & Dalyot, 2020). In general, these methods do not directly

consider the underlying network structure of the sidewalk or road network, and consequently are not considered in the following sections.

5.2.1. Origin–destination matrices

An origin–destination (OD) matrix T is two dimensional array of M rows (origins) and N columns (destination), where its elements t_{ij} encode travel demand between location i and destination j in whatever temporal resolution needed, *e.g.*, annual average daily traffic (AADT), rate per minute, etc. Further, depending on the problem and on the required scale of analysis, pedestrian OD-matrices may depend on a non-trivial combination of time of day, socioeconomic status, scale (which depends on zone definition), and other factors such as multimodality. It is important to note that the common integration of walking trips into multimodal trips –where walking often serves as the mode of transportation for the critical first and last miles– means that realistic pedestrian OD estimation should take into account the interface between pedestrian networks and other transportation networks such as the bus, metro, train, and even road. Such a multimodal approach to urban networks is discussed further in Section 6.

Despite the presence of surveys, the collection of demand estimates for pedestrian transportation is considerably smaller compared to that for motorized transportation, with only a few scattered works available for pedestrian OD-matrix estimations (Buehler & Pucher, 2012; Miller, 2006). Within those, the particularities of pedestrian mobility are generally overlooked, *i.e.*, approximately constant speed, distance constraints, specific fundamental diagram, diversion from shortest paths, mobility restrictions, bidirectional sidewalks, existence of crosswalks, etc.

We classify the different approaches depending on the type of input data and the theoretical basis (and assumptions) of the algorithms. Data sources and methods can be quite broad, but a minimum list includes the following:

- OD Surveys are the most common, intuitive, and direct method to obtain pedestrian OD-matrices. Although costly and time-intensive, surveys are valuable for traffic flow estimation and to understand the factors that promote pedestrian mobility. Yet, they are generally agnostic to the concept of sidewalk network and, consequently, they cannot provide information about its use.
- A much finer detail is provided by mobility traces captured with passive sensing. These new methods for data collection are currently displacing the survey approach (Bonnel & Munizaga, 2018). Mobility traces are defined as a sequence of locations where an individual has been at a given time scale. Nowadays, traces are collected via cell-phone devices (Friedrich, Immisch, Jehlicka, Otterstätter, & Schlaich, 2010), including their apps (*e.g.*, Cuebiq (Moro, Calacci, Dong, & Pentland, 2021)) and service providers (Google, Apple, Orange, etc.), or the use of social networks (Hawelka et al., 2014). Cáceres, Romero, and Benítez (2020) show that their performance to estimate vehicle OD matrices is similar to surveys, and advantageous for low-populated areas.

  Regarding the use of such data to study the use of sidewalk networks, the main problem is to identify if a given trace actually corresponds to a walking trip. State-of-the-art techniques (Wu et al., 2022; Roy, Fuller, Nelson, & Kedron, 2022) apply machine learning algorithms to feature vectors composed of metrics provided by the GPS positioning (speed, turn angles, etc.), combined with descriptors of the geographic context of the trace. The structure and coverage of sidewalk network and their attributes (*e.g.*, width) may play an important role in improving the accuracy of these methods.

  Other than this challenge, the estimation of origin–destination matrices from pedestrian traces is not difficult. Unfortunately, the availability of such type of data is very limited and generally sold as Data as a Service (DaaS).
- Different from the previous two direct measurements of the pedestrian OD-matrix, most remaining methods rely on partial

information, which is often easier to obtain in terms of economic cost and privacy issues. This is the case of count points. We define a count point as a device that is able to enumerate how many individuals pass by a location. There exist a wide set of technologies that can be used to accomplish this. In public transport, data is generated from the counts obtained from turnstiles or gates, *e.g.*, the Oyster Dataset (De Domenico, Solé-Ribalta, Gómez, & Arenas, 2014) for citizens entering public transport in London. Other technologies may include image sensors, light-interruptor sensors, Bluetooth sensor readings (Sarkar et al., 2017), etc. Choosing the location of the sensors to maximise information value is an open problem (Gentili & Mirchandani, 2012).

5.2.2. Pedestrian route choice

The volume of pedestrians on a sidewalk segment is heavily dependent on structural and dynamic features of the surrounding environment. On the other hand, for pedestrians using public space to walk between locations, the efficiency (in terms of space or time) of the sidewalk network plays an important role. The combination of both factors implies that pedestrian route choice considers cost, but also that chosen routes may differ from the shortest path (Bongiorno et al., 2021; Tang & Levinson, 2018). Indeed, several studies on large-scale datasets of GPS traces report that 20% of trips do not follow the shortest path (Miranda, Fan, Duarte, & Ratti, 2021) and, on average, these paths are 20% longer than shortest route (Malleson et al., 2018). Based on these observations, several way-finding strategies (Bhowmick, Winter, & Stevenson, 2019) have been studied over the years to unveil the specific mechanisms guiding pedestrian mobility, putting special emphasis on strategies that reduce the pedestrian cognitive load such as the "longest-leg" strategy or the "fewest turns" strategies.

Finally, pedestrians often prefer one route to arrive to a destination, and a different one to return to their origin. This has been referred to as a hysteresis effect in route choice (Helbing et al., 2001), or simply as asymmetry (Bongiorno et al., 2021). Sidewalk networks could help explain the underlying reasons for these deviations –which may of course include a myriad of other factors, including physical constraints of pedestrians (Bolten et al., 2017; Bolten, 2020; Rhoads et al., 2023); (perceived) safety and attractiveness of the urban environment (Miranda et al., 2021); crime, pollution and noise avoidance (Knittel, Miller, & Sanders, 2016; De Nadai et al., 2016); etc. Independent of the method to construct the pedestrian OD-matrix, or the route choice assumption taken, the individual flow on a sidewalk network segment can be measured in the following way:

$$V_{ij} = \sum_{od} t_{od} p_{od}^{ij} \qquad (6)$$

where V_{ij} is the volume on the (i, j) segment, t_{od} is the entry of the OD matrix T for origin o and destination d, and p_{od}^{ij} is a variable describing the proportion of trips from o to d that pass along link i, j, *i.e.*, the route choice. Note that there is a clear parallelism between Eq. (6) and the non-normalised edge betweenness centrality of the sidewalk networks B_{ij}, which is obtained as $B_{ij} = \sum_{od} p_{od}^{ij}$. Indeed, V_{ij} can be considered a weighted version of the betweenness, and it is sometimes called "augmented betweenness" (Puzis et al., 2013) or "effective betweenness" (Guimerà, Díaz-Guilera, Vega-Redondo, Cabrales, & Arenas, 2002).

In this framework, the objective is to estimate the $M \times N$ entries of the OD matrix, t_{ij}, that are compatible with the set of data observations stated by Eq. (6). The problem is under-determined and relies on assumptions over the particular structure of the OD-matrix, *e.g.*, a gravity model, entropy maximisation, maximum likelihood, etc. For a general description of these models, we refer the reader to de Dios Ortúzar and Willumsen (2011). The specific application of these models to pedestrian OD matrix estimation is limited to regressing (scaling) gravity models to fit observed counts (Sevtsuk, Basu, & Chancey, 2021). The

validity of alternative approaches to estimate pedestrian OD-matrices, such as entropy maximisation, remains to be assessed. Inherent to such type of algorithms, it is important to understand the tight binding between the method and the route choice assumptions over the sidewalk network structure.

6. Pedestrian networks and multimodal urban transportation

From door to door, we begin and end practically every journey as a pedestrian (Walker, 2012). As a consequence, most trips are *multimodal*, *i.e.* they use multiple modes of transportation (Buehler & Hamre, 2015) –unless they are entirely pedestrian trips to begin with. Pedestrian infrastructure therefore forms an integral part of the wider fabric of urban transportation networks. From a systems perspective, this fabric can be understood as a multilayer or multiplex network, where each layer corresponds to one mode of transport (Alessandretti et al., 2022) –such as a city's sidewalk network, its rail/metro network, its bicycle network, or its street network for vehicular traffic, see Fig. 6.

One central challenge is the question of how to most effectively connect these layers to enable trips that are both short, and that use sustainable modes of transport (walking, cycling, and public transit). From a complexity perspective, this interplay between layers is well described via multilayer network metrics (Aleta, Meloni, & Moreno, 2017) which can be used to categorize cities by their interlayer connectedness, a measure which crucially can take the pedestrian network into account (Natera Orozco, Battiston, Iñiguez, & Szell, 2020). The efficient layer connection problem can be studied from different perspectives such as coupling (Morris & Barthelemy, 2012), temporality and time-preservation (Gallotti & Barthelemy, 2014), or navigability and information overload (Gallotti, Porter, & Barthelemy, 2016). Note that, despite the importance of the pedestrian layer, it is often simplified or neglected in the first step of multimodal analyses (Strano, Shai, Dobson, & Barthelemy, 2015; Chodrow, Al-Awwad, Jiang, & González, 2016).

In transportation planning, walking metrics such as the walk radius and the directness of the pedestrian network, are fundamental to deciding on the spacing of stops and stations (Walker, 2012). Similarly, the walk radius is useful for bicycle network planning (Szell, Mimar, Perlman, Ghoshal, & Sinatra, 2022).

Related to the effective connection of transport network layers and

the walk radius is the last mile problem (Park et al., 2021), where non-walking trips should end (or start) as close as possible to the destination (or origin) so that the last (or first) part is short enough to be walkable. How the last mile problem is handled has crucial implications on urban form and vice versa. For example, average availability of car parking closer than public transport stops prevents sustainable development (Knoflacher, 2006). Further, typical suburban cul-de-sac road network patterns can obstruct pedestrian or bicycle networks, and therefore need to be pierced with footpaths to decrease walking distance to transit stops (Walker, 2012). Spatial urban design and transport features also have a crucial impact on walkability and health (Cerin et al., 2022). In general, land use and transport policies should be designed to maximize accessibility to daily amenities and public transport via walking (Liu et al., 2022; Moreno, Allam, Chabaud, Gall, & Pratlong, 2021; World Health Organization, 2022).

7. Conclusions

In the last years, the alignment of the urban agenda to the urgencies of climate change has rendered a long trail of desiderata at the European and world levels. In relation to transportation, increasing active travel, and walking particularly, has been set as crucial for the future of sustainable cities. This is reflected in global and EU policies: EU Green Deal, Flagship Mission and Strategy of the EU for "Climate Neutral and Smart Cities", or the EU Urban Mobility Framework; and of course several of the United Nations' Sustainable Development Goals for 2030.

However, the car-centric developments of the 20th century have a lingering impact on current developments towards sustainable mobility. Society still tends to follow a prioritisation of mobility strategies that are highly skewed against walking, favoring efficiency and modal shift to other existing, yet more sustainable, vehicular modes of transport. This bias against walking is observed across various scales, ranging from the preference for micro-mobility vehicles for short-range trips to public transport for longer distances. Despite this general trend, projects to incentivise walking are slowly gaining social and political attention, as they offer broader societal benefits beyond just improving transit efficiency. These strategies (Claris & Scopelliti, 2016), which encompass social, economic, and political perspectives, are primarily formulated at the intracity level, *e.g.*, pedestrian-friendly streets in downtown areas, city-wide speed reduction zones, and the development of
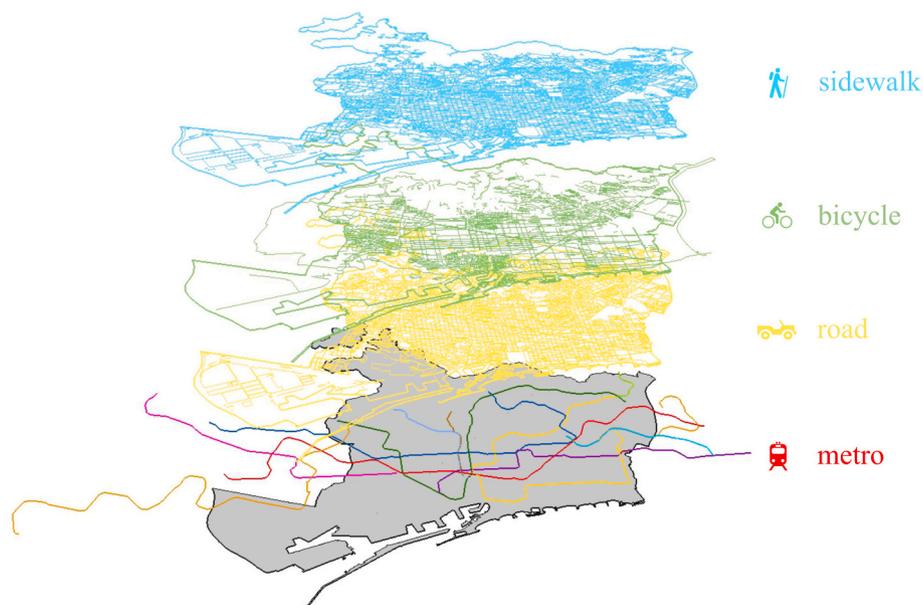
Fig. 6. Multimodal network representation of Barcelona, Spain, with four layers of transport infrastructure (pedestrian, bicycle, road, and metro networks), with data from the Open Data BCN portal for the sidewalk geometries, bicycle paths and metro lines, and OpenStreetMap for the road layer (via OSMnx Boeing, 2017).

neighbourhoods designed to facilitate access within a 15- to 20-min radius. Greater challenges arise when considering longer-scale mobility, such as commuting to work, as pedestrians face distance and time constraints. Some argue that the most effective approach in these cases involves not only prioritising the adoption of electric vehicles to decarbonise mobility (Henderson, 2020) but also focusing on reducing and localising overall mobility (Holden, Banister, Gössling, Gilpin, & Linnerud, 2020).

These trends are reflected in academic research as well. With a longer tradition and better-defined challenges, urban science has been almost exclusively motor vehicle-oriented when it comes to macro-scale, city-level mobility. In this paper, we have argued that the main and most obvious impediment to a quantitative and large-scale approach to walking is the lack of reliable and standardised data. Cartographic services do not consider footways as a category per se, but rather as an addendum to (or attribute of) the road layer. This is true for most public national/regional services, but also, critically, for today's most important open digital mapping platform, OpenStreetMap. Consequently, it is difficult to find reliable data on sidewalk geometries and walkable space. For most cities, it is simply impossible. Considering these circumstances, there is a lack of comprehensive academic research on pedestrian infrastructure at a large scale. For instance, there is a scarcity of studies that examine the connectivity and potential fragmentation of pedestrian mobility in entire neighbourhoods and even less so in whole cities.

To chart a path towards resolving these challenges for the rigorous study of active mobility in present and future cities, we began this paper with a review of existing approaches to collect data and formalise it as a network, considering the advantages and limitations of each approach. Then, we have identified some challenges in which a complex systems perspective might be helpful for better understanding and improving pedestrian infrastructure in cities. We have gathered these problems under the umbrellas of structure, dynamics of sidewalk networks, on one hand, and multilayer transportation networks, on the other. In doing so, a comprehensive account on existing efforts has been laid out, while highlighting possible gaps –research opportunities– that lay ahead.

We foresee that the changes in the nature and quality of pedestrian infrastructure due to social developments will also change the way scientists approach its modelling and analysis. While the further thinning of pedestrian spaces could lead to even less interest in sidewalk networks, the opposite development is also possible: from increased pedestrianization (Soni & Soni, 2016), to pedestrian "superblocks" (Mueller et al., 2020), to car-free cities (Nieuwenhuijsen & Khreis, 2016), the state-of-the-art approach to treating sidewalks as networks could evolve towards more general models of pedestrian spaces. Given how the latter are the most promising paths towards sustainable urban mobility, we believe further inquiry is warranted and a rapid growth of the field can be expected.

References

Abbar, S., Zanouda, T., & Borge-Holthoefer, J. (2018). Structural robustness and service reachability in urban settings. *Data Mining and Knowledge Discovery, 32*(3), 830–847.

Aghaabbasi, M., Moeinaddini, M., Asadi-Shekari, Z., & Shah, M. Z. (2019). The equitable use concept in sidewalk design. *Cities, 88*, 181–190.

Alessandretti, L., Natera Orozco, L. G., Battiston, F., Saberi, M., & Szell, M. (2022). Multimodal urban mobility and multilayer transport networks. *Environment and Planning B: Urban Analytics and City Science*, Article 23998083221108190.

Aleta, A., Meloni, S., & Moreno, Y. (2017). A multilayer perspective for the analysis of urban transportation systems. *Scientific Reports, 7*(1), 1–9.

Ambühl, L., Menendez, M., & González, M. C. (2023). Understanding congestion propagation by combining percolation theory with the macroscopic fundamental diagram. *Communications Physics, 6*(1), 26.

Arcaute, E., Molinero, C., Hatna, E., Murcio, R., Vargas-Ruiz, C., Masucci, A. P., et al. (2016). Cities and regions in britain through hierarchical percolation. *Royal Society Open Science, 3*(4), Article 150691.

Arellana, J., Saltarín, M., Larrañaga, A. M., Alvarez, V., & Henao, C. A. (2020). Urban walkability considering pedestrians' perceptions of the built environment: A 10-year review and a case study in a medium-sized city in Latin America. *Transport Reviews, 40*(2), 183–203.

Barthélemy, M. (2011). Spatial networks. *Physics Reports, 499*(1–3), 1–101.

Barthelemy, M. (2016). *The structure and dynamics of cities*. Cambridge University Press.

Barthelemy, M. (2018). *Morphogenesis of spatial networks*. Springer.

Barthelemy, M. (2022). *Spatial Networks: A Complete Introduction. From Graph Theory and Statistical Physics to Real-World Applications*. Springer Nature.

Bartzokas-Tsiompras, A., Photis, Y. N., Tsagkis, P., & Panagiotopoulos, G. (2021). Microscale walkability indicators for fifty-nine european central urban areas: An open-access tabular dataset and a geospatial web-based platform. *Data in Brief, 36*, Article 107048.

Bastani, F., He, S., Abbar, S., Alizadeh, M., Balakrishnan, H., Chawla, S., et al. (2018). Roadtracer: Automatic extraction of road networks from aerial images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4720–4728).

Batty, M. (2012). Building a science of cities. *Cities, 29*, S9–S16.

Bhowmick, D., Winter, S., & Stevenson, M. (2019). Comparing the costs of pedestrian wayfinding heuristics across different urban network morphologies.

Boeing, G. (2017). Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems, 65*, 126–139.

Boeing, G. (2019). Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science, 4*(1), 1–19.

Bolten, N. (2020). *Equitable Network Modeling of Diverse Modes of Built Environment Pedestrian Navigation*. University of Washington.

Bolten, N., & Caspi, A. (2021). Towards routine, city-scale accessibility metrics: Graph theoretic interpretations of pedestrian access using personalized pedestrian network analysis. *PLoS One, 16*(3), Article e0248399.

Bolten, N., Amini, A., Hao, Y., Ravichandran, V., Stephens, A., & Caspi, A. (2015). Urban sidewalks: visualization and routing for individuals with limited mobility. In *Proceedings of the 1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics* (pp. 122–125).

Bolten, N., Mukherjee, S., Sipeeva, V., Tanweer, A., & Caspi, A. (2017). A pedestrian-centered data approach for equitable access to urban infrastructure environments. *IBM Journal of Research and Development, 61*(6), 10–1.

Bongiorno, C., Zhou, Y., Kryven, M., Theurel, D., Rizzo, A., Santi, P., et al. (2021). Vector-based pedestrian navigation in cities. *Nature Computational Science, 1*(10), 678–685.

Bonnel, P., & Munizaga, M. A. (2018). Transport survey methods-in the era of big data facing new and old challenges. *Transportation Research Procedia, 32*, 1–15.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology, 25*(2), 163–177.

Brown, B. B., Werner, C. M., Amburgey, J. W., & Szalay, C. (2007). Walkable route perceptions and physical features: Converging evidence for en route walking experiences. *Environment and Behavior, 39*(1), 34–61.

Buehler, R., & Hamre, A. (2015). The multimodal majority? Driving, walking, cycling, and public transportation use among American adults. *Transportation, 42*(6), 1081–1101.

Buehler, R., & Pucher, J. (2012). Walking and cycling in western europe and the united states: trends, policies, and lessons. *Tr News, 280*.

Bustos, C., Rhoads, D., Solé-Ribalta, A., Masip, D., Arenas, A., Lapedriza, A., et al. (2021). Explainable, automated urban interventions to improve pedestrian and vehicle safety. *Transportation Research Part C: Emerging Technologies, 125*, Article 103018.

Cáceres, N., Romero, L., & Benítez, F. G. (2020). Exploring strengths and weaknesses of mobility inference from mobile phone data vs. travel surveys. *Transportmetrica A: Transport Science, 16*(3), 574–601.

Cambra, P. J., Gonçalves, A., & Moura, F. (2019). The digital pedestrian network in complex urban contexts: A primer discussion on typological specifications. *Finisterra, 54*(110), 155–170.

Cerin, E., Sallis, J. F., Salvo, D., Hinckson, E., Conway, T. L., Owen, N., et al. (2022). Determining thresholds for spatial urban design and transport features that support walking to create healthy and sustainable cities: Findings from the IPEN Adult study. *The Lancet Global Health, 10*(6), e895–e906.

Cervero, R., & Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods. *Transport Policy, 3*(3), 127–141.

Cheng, G., Wang, Y., Xu, S., Wang, H., Xiang, S., & Pan, C. (2017). Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing, 55*(6), 3322–3337.

Chen, B.-I., Hsueh, M.-C., Rutherford, R., Park, J.-H., & Liao, Y. (2019). The associations between neighborhood walkability attributes and objectively measured physical activity in older adults. *PLoS One, 14*(9), Article e0222268.

Cherry, C., Donlon, B., Yan, X., Moore, S. E., & Xiong, J. (2012). Illegal mid-block pedestrian crossings in china: Gap acceptance, conflict and crossing path analysis. *International Journal of Injury Control and Safety Promotion, 19*(4), 320–330.

Chin, G. K., Van Niel, K. P., Giles-Corti, B., & Knuiman, M. (2008). Accessibility and connectivity in physical activity studies: The impact of missing pedestrian data. *Preventive Medicine, 46*(1), 41–45.

Chodrow, P. S., Al-Awwad, Z., Jiang, S., & González, M. C. (2016). Demand and congestion in multiplex transportation networks. *PLoS One, 11*(9), Article e0161738.

Claris, S., & Scopelliti, D. (2016). Cities alive: towards a walking world.

Cohen, A., & Dalyot, S. (2020). Machine-learning prediction models for pedestrian traffic flow levels: Towards optimizing walking routes for blind pedestrians. *Transactions in GIS, 24*(5), 1264–1279.

Colville-Andersen, M. (2018). The arrogance of space. In *Copenhagenize*. Washington, DC: Island Press.

Coppola, N. A., & Marshall, W. E. (2020). An evaluation of sidewalk availability and width: Analyzing municipal policy and equity disparities. In *International Conference on Transportation and Development 2020* (pp. 13–24). American Society of Civil Engineers.

Corazza, M. V., Di Mascio, P., & Moretti, L. (2016). Managing sidewalk pavement maintenance: A case study to increase pedestrian safety. *Journal of Traffic and Transportation Engineering (English Edition), 3*(3), 203–214.

de Dios Ortúzar, J., & Willumsen, L. G. (2011). *Modelling transport*. John Wiley & Sons.

De Domenico, M., Solé-Ribalta, A., Gómez, S., & Arenas, A. (2014). Navigability of interconnected networks under random failures. *Proceedings of the National Academy of Sciences, 111*(23), 8351–8356.

De Nadai, M., Vieriu, R. L., Zen, G., Dragicevic, S., Naik, N., Caraviello, M., et al. (2016). Are safer looking neighborhoods more lively? A multimodal investigation into urban life. In *Proceedings of the 24th ACM international conference on Multimedia* (pp. 1127–1135).

Dueker, K. J., & Vrana, R. (1992). Dynamic segmentation revisited: A milepoint linear data model. *Journal of the Urban and Regional Information Systems Association, 4*(2), 94–105.

Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American Planning Association, 76*(3), 265–294.

Farber, S., Neutens, T., Miller, H. J., & Li, X. (2013). The social interaction potential of metropolitan regions: A time-geographic measurement approach using joint accessibility. *Annals of the Association of American Geographers, 103*(3), 483–504.

Feliciani, C., & Nishinari, K. (2018). Measurement of congestion and intrinsic risk in pedestrian crowds. *Transportation Research Part C: Emerging Technologies, 91*, 124–155.

Ferré-Bigorra, J., Casals, M., & Gangolells, M. (2022). The adoption of urban digital twins. *Cities, 131*, Article 103905.

Fischer, M. M. (2006). Gis and network analysis. *Spatial Analysis and GeoComputation: Selected Essays*, 43–60.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 35–41.

Friedrich, M., Immisch, K., Jehlicka, P., Otterstätter, T., & Schlaich, J. (2010). Generating origin–destination matrices from mobile phone trajectories. *Transportation Research Record, 2196*(1), 93–101.

Gaither, C. J., Himmelfarb, D., Hitchner, S., Schelhas, J., Shepherd, J. M., & Binita, K. (2016). "Where the sidewalk ends": Sustainable mobility in Atlanta's Cascade community. *City & Society, 28*(2), 174–197.

Gallotti, R., & Barthelemy, M. (2014). Anatomy and efficiency of urban multimodal mobility. *Scientific Reports, 4*(1), 1–9.

Gallotti, R., Porter, M. A., & Barthelemy, M. (2016). Lost in transportation: Information measures and cognitive limits in multilayer navigation. *Science Advances, 2*(2), Article e1500445.

Gaspari, F., Stucchi, L., Bratic, G., Jovanovic, D., Ponti, C., Biagi, L. G. A., et al. (2021). Innovation in teaching: The polimappers collaborative and humanitarian mapping course at Politecnico di Milano. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVI-4/W2-2021*, 63–69.

Gehl, J., Westermark, E., Rison, J., Mortensen, H., Adams, R., & Rorbech, J. (2008). World class streets: Remaking new york city's public realm, New York.

Gehrke, S. R., Phair, C. D., Russo, B. J., & Smaglik, E. J. (2023). Observed sidewalk autonomous delivery robot interactions with pedestrians and bicyclists. *Transportation Research Interdisciplinary Perspectives, 18*, Article 100789.

Gentili, M., & Mirchandani, P. B. (2012). Locating sensors on traffic networks: Models, challenges and research opportunities. *Transportation Research Part C: Emerging Technologies, 24*, 227–255.

Ghomi, H., & Hussein, M. (2023). Investigating the application of deep learning to identify pedestrian collision-prone zones. *Journal of Transportation Safety & Security*, 1–31.

Gössling, S. (2016). Urban transport justice. *Journal of Transport Geography, 54*, 1–9.

Gössling, S., Choi, A., Dekker, K., & Metzler, D. (2019). The social cost of automobility, cycling and walking in the european union. *Ecological Economics, 158*, 65–74.

Guimerà, R., Díaz-Guilera, A., Vega-Redondo, F., Cabrales, A., & Arenas, A. (2002). Optimal network topologies for local search with congestion. *Physical Review Letters, 89*(24), Article 248701.

Ha, J., Ki, D., Lee, S., & Ko, J. (2023). Mode choice and the first-/last-mile burden: The moderating effect of street-level walkability. *Transportation Research Part D: Transport and Environment, 116*, Article 103646.

Haustein, S., & Nielsen, T. A. S. (2016). European mobility cultures: A survey-based cluster analysis across 28 european countries. *Journal of Transport Geography, 54*, 173–180.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science, 41*(3), 260–271.

Helbing, D., Molnár, P., Farkas, I. J., & Bolay, K. (2001). Self-organizing pedestrian movement. *Environment and Planning B: Planning and Design, 28*(3), 361–383.

Henderson, L. (1971). The statistics of crowd fluids. *Nature, 229*(5284), 381–383.

Henderson, J. (2020). EVs are not the answer: A mobility justice critique of electric vehicle transitions. *Annals of the American Association of Geographers, 110*(6), 1993–2010.

Hennessy, E. R., & Ai, C. (2023). A comparative analysis of pedestrian network connectivity and accessibility using network approximation. *Journal of Transport Geography, 111*, Article 103637.

Holden, E., Banister, D., Gössling, S., Gilpin, G., & Linnerud, K. (2020). Grand narratives for sustainable mobility: A conceptual review. *Energy Research & Social Science, 65*, Article 101454.

Hosseini, M., Sevtsuk, A., Miranda, F., Cesar, R. M., Jr, & Silva, C. T. (2023). Mapping the walk: A scalable computer vision approach for generating sidewalk network datasets from aerial imagery. *Computers, Environment and Urban Systems, 101*, Article 101950.

Hou, Q., & Ai, C. (2020). A network-level sidewalk inventory method using mobile lidar and deep learning. *Transportation Research part C: Emerging Technologies, 119*, Article 102772.

Hunter, R. F., Garcia, L., de Sa, T. H., Zapata-Diomedi, B., Millett, C., Woodcock, J., et al. (2021). Effect of covid-19 response policies on walking behavior in us cities. *Nature Communications, 12*(1), 1–9.

Jacobs, J. (1961). *The death and life of great American cities*. New York: Random House.

Jennings, D., & Figliozzi, M. (2019). Study of sidewalk autonomous delivery robots and their potential impacts on freight efficiency and travel. *Transportation Research Record, 2673*(6), 317–326.

Jiang, S., Fiore, G. A., Yang, Y., Ferreira Jr., J., Frazzoli, E., González, M. C. (2013). A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In *Proceedings of the 2nd ACM SIGKDD international workshop on Urban Computing* (pp. 1–9).

Jiang, F., Ma, J., & Li, Z. (2022). Pedestrian volume prediction with high spatiotemporal granularity in urban areas by the enhanced learning model. *Sustainable Cities and Society, 79*, Article 103653.

Katz, P., Scully, V. J., & Bressi, T. W. (1994). *The new urbanism: Toward an architecture of community* (Vol. 10). New York: McGraw-Hill.

Kirkley, A., Barbosa, H., Barthelemy, M., & Ghoshal, G. (2018). From the betweenness centrality in street networks to structural invariants in random planar graphs. *Nature Communications, 9*(1), 1–12.

Knittel, C. R., Miller, D. L., & Sanders, N. J. (2016). Caution, drivers! Children present: Traffic, pollution, and infant health. *Review of Economics and Statistics, 98*(2), 350–366.

Knoflacher, H. (2006). A new way to organize parking: The key to a successful sustainable transport system for the future. *Environment and Urbanization, 18*(2), 387–400.

Kretz, T., Grünebohm, A., Kaufman, M., Mazur, F., & Schreckenberg, M. (2006). Experimental study of pedestrian counterflow in a corridor. *Journal of Statistical Mechanics: Theory and Experiment, 2006*(10), P10001.

Lampo, A., Borge-Holthoefer, J., Gómez, S., & Solé-Ribalta, A. (2021). Multiple abrupt phase transitions in urban transport congestion. *Physical Review Research, 3*(1), Article 013267.

Li, H., Cebe, J., Khoeini, S., Xu, Y. A., Dyess, C., & Guensler, R. (2018). A semi-automated method to generate gis-based sidewalk networks for asset management and pedestrian accessibility assessment. *Transportation Research Record, 2672*(44), 1–9.

Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H. E., et al. (2015). Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proceedings of the National Academy of Sciences, 112*(3), 669–672.

Li, M., Liu, R.-R., Lü, L., Hu, M.-B., Xu, S., & Zhang, Y.-C. (2021). Percolation on complex networks: Theory and application. *Physics Reports*.

Lin, J., & Ban, Y. (2013). Complex network topology of transportation systems. *Transport Reviews, 33*(6), 658–685.

Liu, S., Higgs, C., Arundel, J., Boeing, G., Cerdera, N., Moctezuma, D., et al. (2022). A generalized framework for measuring pedestrian accessibility around the world using open data. *Geographical Analysis, 54*(3), 559–582.

Lo, R. H. (2009). Walkability: What is it? *Journal of Urbanism, 2*(2), 145–166.

Louf, R., & Barthelemy, M. (2014). A typology of street patterns. *Journal of The Royal Society Interface, 11*(101), 20140924.

Malleson, N., Vanky, A., Hashemian, B., Santi, P., Verma, S. K., Courtney, T. K., et al. (2018). The characteristics of asymmetric pedestrian behavior: A preliminary study using passive smartphone location data. *Transactions in GIS, 22*(2), 616–634.

Marquet, O., & Miralles-Guasch, C. (2014). Walking short distances. The socioeconomic drivers for the use of proximity in everyday mobility in barcelona. *Transportation Research Part A: Policy and Practice, 70*, 210–222.

Martinez-Gil, F., Lozano, M., García-Fernández, I., & Fernández, F. (2017). Modeling, evaluation, and scale on artificial pedestrians: A literature review. *ACM Computing Surveys (CSUR), 50*(5), 1–35.

Máttyus, G., Luo, W., & Urtasun, R. (2017). Deeproadmapper: Extracting road topology from aerial images. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3438–3446).

Miller, H. J. (2006). Modelling accessibility using space-time prism concepts within geographical information systems: fourteen years on. *Classics from IJGIS: Twenty Years of the International Journal of Geographical Information Science and Systems, 5*(3), 175–179.

Miranda, A. S., Fan, Z., Duarte, F., & Ratti, C. (2021). Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Computers, Environment and Urban Systems, 86*, Article 101563.

Mobasheri, A., Huang, H., Degrossi, L. C., & Zipf, A. (2018). Enrichment of openstreetmap data completeness with sidewalk geometries using data mining techniques. *Sensors, 18*(2), 509.

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-minute city": Sustainability, resilience and place identity in future post-pandemic cities. *Smart Cities, 4*(1), 93–111.

Moro, E., Calacci, D., Dong, X., & Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large us cities. *Nature Communications, 12*(1), 1–10.

Morris, R. G., & Barthelemy, M. (2012). Transport on coupled spatial networks. *Physical Review Letters, 109*(12), Article 128703.

Mueller, N., Rojas-Rueda, D., Cole-Hunter, T., de Nazelle, A., Dons, E., Gerike, R., et al. (2015). Health impact assessment of active transportation: A systematic review. *Preventive Medicine, 76*, 103–114.

Mueller, N., Rojas-Rueda, D., Khreis, H., Cirach, M., Andrés, D., Ballester, J., et al. (2020). Changing the urban design of cities for health: The superblock model. *Environment International, 134*, Article 105132.

Nag, D., Goswami, A. K., Gupta, A., & Sen, J. (2020). Assessing urban sidewalk networks based on three constructs: A synthesis of pedestrian level of service literature. *Transport Reviews, 40*(2), 204–240.

Natera Orozco, L. G., Battiston, F., Iñiguez, G., & Szell, M. (2020). Extracting the multimodal fingerprint of urban transportation networks. *Transport Findings, 13171*.

Newman, P. G., & Kenworthy, J. R. (1989). *Cities and automobile dependence: An international sourcebook*. Aldershot, England: Gower Publishing.

Nieuwenhuijsen, M. J., & Khreis, H. (2016). Car free cities: Pathway to healthy urban living. *Environment International, 94*, 251–262.

Ning, H., Li, Z., Wang, C., Hodgson, M. E., Huang, X., & Li, X. (2022). Converting street view images to land cover maps for metric mapping: A case study on sidewalk network extraction for the wheelchair users. *Computers, Environment and Urban Systems, 95*, Article 101808.

Norton, P. D. (2011). *Fighting traffic: The dawn of the motor age in the American city*. MIT Press.

Olmos, L. E., Tadeo, M. S., Vlachogiannis, D., Alhasoun, F., Alegre, X. E., Ochoa, C., et al. (2020). A data science framework for planning the growth of bicycle infrastructures. *Transportation Research Part C: Emerging Technologies, 115*, Article 102640.

OpenStreetMap contributors (2017). Planet dump retrieved from https://planet.osm.org, https://www.openstreetmap.org.

Osama, A., & Sayed, T. (2017). Evaluating the impact of connectivity, continuity, and topography of sidewalk network on pedestrian safety. *Accident Analysis & Prevention, 107*, 117–125.

Park, K., Farb, A., & Chen, S. (2021). First-/last-mile experience matters: The influence of the built environment on satisfaction and loyalty among public transit riders. *Transport Policy, 112*, 32–42.

Pawar, D. S., & Patil, G. R. (2016). Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials. *Safety Science, 86*, 295–303.

Pereira, R. H., Schwanen, T., & Banister, D. (2017). Distributive justice and equity in transportation. *Transport Reviews, 37*(2), 170–191.

Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: A primal approach. *Environment and Planning B: Planning and Design, 33*(5), 705–725.

Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: A dual approach. *Physica A: Statistical Mechanics and its Applications, 369*(2), 853–866.

Puzis, R., Altshuler, Y., Elovici, Y., Bekhor, S., Shiftan, Y., & Pentland, A. (2013). Augmented betweenness centrality for environmentally aware traffic monitoring in transportation networks. *Journal of Intelligent Transportation Systems, 17*(1), 91–105.

Resch, B., & Szell, M. (2019). Human-centric data science for urban studies. *ISPRS International Journal of Geo-Information, 8*(584).

Rhoads, D., Solé-Ribalta, A., & Borge-Holthoefer, J. (2023). The inclusive 15-minute city: Walkability analysis with sidewalk networks. *Computers, Environment and Urban Systems, 100*, Article 101936.

Rhoads, D., Solé-Ribalta, A., González, M. C., & Borge-Holthoefer, J. (2021). A sustainable strategy for open streets in (post) pandemic cities. *Communications Physics, 4*(1), 1–12.

Rifaat, S. M., Tay, R., & De Barros, A. (2011). Effect of street pattern on the severity of crashes involving vulnerable road users. *Accident Analysis & Prevention, 43*(1), 276–283.

Rodríguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment, 9*(2), 151–173.

Roy, A., Fuller, D., Nelson, T., & Kedron, P. (2022). Assessing the role of geographic context in transportation mode detection from GPS data. *Journal of Transport Geography, 100*, Article 103330.

Saeidizand, P., Fransen, K., & Boussauw, K. (2022). Revisiting car dependency: A worldwide analysis of car travel in global metropolitan areas. *Cities, 120*, Article 103467.

Saha, M., Saugstad, M., Maddali, H. T., Zeng, A., Holland, R., Bower, S. A., et al. (2019). Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

Sargoni, O.T., & Manley, E. (2020). A sequential sampling model of pedestrian road crossing choice. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation* (pp. 10–19).

Sarkar, S., Chawla, S., Ahmad, S., Srivastava, J., Hammady, H., Filali, F., et al. (2017). Effective urban structure inference from traffic flow dynamics. *IEEE Transactions on Big Data, 3*(2), 181–193.

Seneviratne, P. N., & Morrall, J. F. (1985). Analysis of factors affecting the choice of route of pedestrians. *Transportation Planning and Technology, 10*(2), 147–159.

Serok, N., Levy, O., Havlin, S., & Blumenfeld-Lieberthal, E. (2019). Unveiling the inter-relations between the urban streets network and its dynamic traffic flows: Planning implication. *Environment and Planning B: Urban Analytics and City Science, 46*(7), 1362–1376.

Sevtsuk, A., Basu, R., & Chancey, B. (2021). We shape our buildings, but do they then shape us? A longitudinal analysis of pedestrian flows and development activity in melbourne. *PloS One, 16*(9), Article e0257534.

Seyfried, A., Steffen, B., Klingsch, W., & Boltes, M. (2005). The fundamental diagram of pedestrian movement revisited. *Journal of Statistical Mechanics: Theory and Experiment, 2005*(10), P10002.

Sisiopiku, V. P., & Akin, D. (2003). Pedestrian behaviors at and perceptions towards various pedestrian facilities: An examination based on observation and survey data. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(4), 249–274.

Solé-Ribalta, A., Gómez, S., & Arenas, A. (2016). A model to identify urban traffic congestion hotspots in complex networks. *Royal Society Open Science, 3*(10), Article 160098.

Solé-Ribalta, A., Gómez, S., & Arenas, A. (2018). Decongestion of urban areas with hotspot pricing. *Networks and Spatial Economics, 18*(1), 33–50.

Soni, N., & Soni, N. (2016). Benefits of pedestrianization and warrants to pedestrianize an area. *Land Use Policy, 57*, 139–150.

Speck, J. (2013). *Walkable city: How downtown can save America, one step at a time*. Macmillan.

Strano, E., Shai, S., Dobson, S., & Barthelemy, M. (2015). Multiplex networks in metropolitan areas: Generic features and local effects. *Journal of the Royal Society Interface, 12*(111), 20150651.

Szell, M. (2018). Crowdsourced quantification and visualization of urban mobility space inequality. *Urban Planning, 3*(1), 1–20.

Szell, M., Mimar, S., Perlman, T., Ghoshal, G., & Sinatra, R. (2022). Growing urban bicycle networks. *Scientific Reports, 12*(6765).

Tal, G., & Handy, S. (2012). Measuring nonmotorized accessibility and connectivity in a robust pedestrian network. *Transportation Research Record, 2299*(1), 48–56.

Tang, W., & Levinson, D. M. (2018). Deviation between actual and shortest travel time paths for commuters. *Journal of Transportation Engineering, Part A: Systems, 144*(8), 04018042.

Tanweer, A., Bolten, N., Drouhard, M., Hamilton, J., Caspi, A., Fiore-Gartland, B., et al. (2017). *Mapping for accessibility: A case study of ethics in data science for social good, Bloomberg Data for Good Exchange Conference. New York City, NY, USA*.

Tzouras, P. G., Mitropoulos, L., Koliou, K., Stavropoulou, E., Karolemeas, C., Antoniou, E., et al. (2023). Describing micro-mobility first/last-mile routing behavior in urban road networks through a novel modeling approach. *Sustainability, 15*(4), 3095.

Vale, D. S., Saraiva, M., & Pereira, M. (2016). Active accessibility: A review of operational measures of walking and cycling accessibility. *Journal of Transport and Land Use, 9*(1), 209–235.

Verbavatz, V., & Barthelemy, M. (2022). Betweenness centrality in dense spatial networks. *Physical Review E, 105*(5), Article 054303.

Vybornova, A., Cunha, T., Gühnemann, A., & Szell, M. (2022). Automated detection of missing links in bicycle networks. *Geographical Analysis*, 1–29.

Walker, J. (2012). *Human transit: How clearer thinking about public transit can enrich our communities and our lives*. Island Press.

Weidmann, U. (1993). Transporttechnik der fußgänger: Transporttechnische eigenschaften des fußgängerverkehrs, literaturauswertung. *IVT Schriftenreihe, 90*.

World Health Organization. (2022). *Walking and cycling: Latest evidence to support policy-making and practice*.

World Health Organization. (2022). Walking and cycling: latest evidence to support policy-making and practice. *Regional Office for Europe*.

Wu, H., Jin, H., Xu, Z., Liu, C., Li, J., & Huang, W. (2022). Travel mode classification based on GNSS trajectories and open geospatial data. *Transactions in GIS, 26*(6), 2598–2620.

Xu, Y., Olmos, L. E., Abbar, S., & González, M. C. (2020). Deconstructing laws of accessibility and facility distribution in cities. *Science Advances, 6*(37), Article eabb4112.

Zhang, J., Klingsch, W., Schadschneider, A., & Seyfried, A. (2012). Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram. *Journal of Statistical Mechanics: Theory and Experiment, 2012*(02), P02002.

Zhang, J., Klingsch, W., Schadschneider, A., & Seyfried, A. (2013). Experimental study of pedestrian flow through a t-junction. In *Traffic and Granular Flow'11* (pp. 241–249). Springer.

Zhang, W., & Thill, J.-C. (2017). Detecting and visualizing cohesive activity-travel patterns: A network analysis approach. *Computers, Environment and Urban Systems, 66*, 117–129.

Zhao, J., Sun, G., & Webster, C. (2021). Walkability scoring: Why and how does a three-dimensional pedestrian network matter? *Environment and Planning B: Urban Analytics and City Science, 48*(8), 2418–2435.

Zünd, D., & Bettencourt, L. (2021). Street view imaging for automated assessments of urban infrastructure and services. In *Urban Informatics* (pp. 29–40). Springer.