Cities 158 (2025) 105655

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

Trends and drivers of pedestrian mobility in Barcelona: A fine-grained study across its commercial tissue

Clément Rames [a,e], Daniel Rhoads [a], Antoni Meseguer-Artola [b], Sergi Lozano [c,d], Javier Borge-Holthoefer [a], Albert Solé-Ribalta [a,*]

[a] Internet Interdisciplinary Institute, Universitat Oberta de Catalunya, Spain
[b] Faculty of Economics and Business, Universitat Oberta de Catalunya, Spain
[c] Faculty of Economics and Business, Universitat de Barcelona, Spain
[d] University of Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Spain
[e] Urban Sociology Laboratory, EPFL, Lausanne, Switzerland

ARTICLE INFO

Keywords:
Walkability
Urban design
Pedestrian mobility
Accessibility

ABSTRACT

Identifying factors that promote active mobility, especially walking, is essential for designing resilient and livable cities and promoting sustainable urban mobility. In spite of recent advances in this direction, available data often remains too spatially and temporally coarse, which constrains analysis. This paper leverages high resolution data from over 200 pedestrian count sensors, placed along Barcelona's commercial areas, providing a detailed understanding of how walking volume has evolved over the past five years, how it varies across neighborhoods, and which socioeconomic and urban attributes influence it. We find that while overall pedestrian traffic has increased, a neighborhood-scale analysis reveals a nuanced picture of fluctuations, including increases, declines, and periodic patterns. The use of global regression models allows us to identify seven key urban factors that shape pedestrian mobility. Subsequently moving the analysis to spatially-aware regression models, we identify the spatial non-stationarity of these factors across the city, indicating the presence of distinct behavioral groups within the urban population. The detailed spatial resolution of our findings provides municipal decision-makers with insights for implementing precise interventions and continually evaluating their effects. Moreover, monitoring pedestrian traffic before and after urban initiatives, while adjusting for seasonal, daily, and time-of-day variations, can yield critical insights for developing pedestrian-oriented urban environments.

1. Introduction

In response to the growing threat of climate change and its impact on urban areas, cities across the globe have increasingly recognized the need to incorporate climate considerations into their policy and planning decisions. On the mobility front, increasing active travel, and walking in particular, is widely accepted as critical to achieving sustainable cities. This is reflected in global and European policies, from the United Nations' Sustainable Development Goals for 2030 (SDG 3, 11 and 13) to the EU's "Climate Neutral and Smart Cities" mission for the period 2030–2050. Thus, the debates around pedestrian mobility incentives and walkable city planning have become central in the political, social and scientific discourse.

Recognizing the importance of walking and its incentives in urban settings is crucial to promote it and to understand their implications. Walking is a singular mode of transport highly influenced by the anatomical and physiological traits of individuals (Gonzalez, Hidalgo, & Barabasi, 2008). This involves elements such as a low average walking speed (Bosina & Weidmann, 2017), which is biologically optimized to minimize metabolic energy expenditure (Henderson, 1971). This inherent limitation in speed restricts the feasible distance that can be traversed within an acceptable timeframe (Berrett, Leake, May, & Whelan, 1988; Daniels & Mulley, 2013) and leads to a long-tailed distribution in the observed range of walking distances (Alessandretti, Aslak, & Lehmann, 2020; Bongiorno et al., 2021). Predominantly, shorter walking trips are more frequent (Marquet & Miralles-Guasch, 2014), resonating with urban design principles such as the "15-min city" (Moreno, Allam, Chabaud, Gall, & Pratlong, 2021). Conversely, longer walking journeys are rarer and typically serve to distinct purposes (Bongiorno et al., 2021). On the other side, from a practical

* Corresponding author at: Rambla del Poblenou, 156, 08018 Barcelona, Spain.
E-mail address: asolerib@uoc.edu (A. Solé-Ribalta).

https://doi.org/10.1016/j.cities.2024.105655
Received 10 May 2023; Received in revised form 3 December 2024; Accepted 9 December 2024
Available online 29 December 2024

standpoint, walking emerges as a pragmatic mode choice for numerous city dwellers, facilitating the completion of their daily tasks and bridging the first/last mile of many trips (Cerin et al., 2022; Park, Farb, & Chen, 2021). Widely regarded as one of the most inclusive forms of transport (Claris & Scopelliti, 2016), walking also holds significant implications for its social, economic, and environmental benefits (Claris & Scopelliti, 2016; Gössling, Choi, Dekker, & Metzler, 2019). Therefore, gaining insight into the factors that motivate city dwellers to walk, reflecting the varying urban needs and individual choices, is essential.

Among the several research lines exploring this topic, the scientific community has made substantial progress in understanding how built environment influence pedestrian behavior and enhance walkability (Dovey & Pafka, 2020; Ewing & Cervero, 2010; Sevtsuk et al., 2024; Van Dyck, Cardon, Deforche, & De Bourdeaudhuij, 2011). This body of work provides a direct avenue for addressing the intricate challenges associated with promoting pedestrian-friendly urban spaces. The literature encompasses a broad spectrum of factors related to urban structure—such as urban design, land use, availability of recreational opportunities, and accessibility to public transportation (Ewing & Cervero, 2010; Gascon et al., 2019)—in addition to examining social and demographic influences (Gascon et al., 2019). Collectively, these studies emphasize the intricate interplay between these elements and highlight the critical role of trip purpose in understanding pedestrian behavior. Broadly speaking, daily tasks, such as grocery shopping, are usually done in nearby locations and completed by short walks (Marquet & Miralles-Guasch, 2014; Perchoux et al., 2019), with the exception of commuting to work which presents an exception. Only 10 to 15 % of these trips being completed on foot, likely due to the constraints related to distance and time. Indeed, factors associated with reduced walking levels, such as car ownership, appear to become more salient as time and distance constraints increase (Perchoux et al., 2019). In contrast, for activities with more flexible time constraints, such as leisurely walks (Perchoux et al., 2019), visiting relatives, or accompanying others (e.g., taking children to school) (Perchoux et al., 2019; Yang & Diez-Roux, 2012), the impact of distance on the decision to walk is minimal to nonexistent. Moreover, these drivers, attitudes, and perceptions toward walking exhibit variations based on socio-cultural contexts and geographic locations (Boakye et al., 2023; Hatamzadeh, Habibian, & Khodaii, 2014). For instance, recent analysis indicates that street pedestrianization primarily impacts shopping districts with dining establishments, while its influence on local commerce remains negligible (Yoshimura et al., 2022).

Adding a further layer of complexity to understand pedestrian attitudes, Anselin's principle of spatial heterogeneity (Anselin, 2013; Goodchild, 2004) challenges the assumption of stationarity in these results, suggesting that the statistical properties of a system might not remain constant across space. The diversity of urban amenities (Hidalgo, Castañer, & Sevtsuk, 2020), socio-economic aspects (Nilforoshan et al., 2023), varying pedestrian activities (Sevtsuk, Basu, & Chancey, 2021), and other spatial dynamics (Triguero-Mas et al., 2022) all suggest complex interactions that could influence the factors affecting pedestrian mobility across spatial and temporal dimensions in urban landscapes. Recognizing and effectively modeling these spatial heterogeneity is crucial for urban planning, enabling precise analysis and avoiding a one-size-fits-all approach across diverse spatial scales. This understanding may significantly influence urban development and governance, and how different areas can accommodate the community's varied needs. In a similar vein, several studies have investigated the temporal variability of factors in pedestrian volume estimations (Lu, Mondschein, Buehler, & Hankey, 2018). Interestingly, despite schedule variations, the influence of urban attributes (Basu & Sevtsuk, 2022; Lu et al., 2018) and activities (Sevtsuk et al., 2021) on pedestrian volumes remains quite consistent over time. This uniformity implies a certain constancy in pedestrian behavior, even amid changing temporal patterns. Given these findings, the key element to focus on seems to be spatially.

The exploration of these heterogeneities has been facilitated by a variety of analytical tools (Anselin, 2013; Brunsdon, Fotheringham, & Charlton, 1996), underscoring the crucial need to account for spatial variations in urban planning and pedestrian mobility research. Relying on these tools, several authors (Chen, Lu, Ye, Xiao, & Yang, 2022; Feuillet et al., 2015; Feuillet et al., 2018; Lu, Sullivan, & Troy, 2012; Pfiester, Thompson, & Zhang, 2021) made important contributions to the state of the art. But, in general, they tend to focus on either broad or confined spatial scales, indicating the necessity for additional, more detailed investigations into spatial heterogeneities within urban environments.

At the national level, Feuillet et al. (Feuillet et al., 2018) identified the differing factors influencing walking behaviors in urban and non-urban areas in France, considering activity type (such as walking for leisure or errands). At a lower scale, Lu et al. (Lu et al., 2012) demonstrate varying degrees of positive correlations between building density and intersection density, affirming a greater propensity for non-motorized travel in urban compared to rural areas of Chittenden County, Vermont (USA). At the regional level, Feuillet et al. (Feuillet et al., 2015) also observed significant variability of active commuting attitudes in relation to the built and socio-economic environment in Paris and its immediate suburbs. At a finer scale, it is worth highlighting the work of Cheng et al. (Cheng, Shi, De Vos, Cao, & Witlox, 2021) and Graells-Garrido et al. (Graells-Garrido, Serra-Burriel, Rowe, Cucchietti, & Reyes, 2021), probably the closest in spatial framing to our study. Cheng et al (Cheng et al., 2021), focusing on older adults, detect, among others, interesting variations, from positive to negative, in the impact of population density and land use mix on pedestrian mobility in Nanjing (China). Graells-Garrido et al. (Graells-Garrido et al., 2021), despite focussing on a different but similar problem, identifies several spatial heterogeneities in the factors that motivates resident population to daily move between different areas of the city. Finally, focusing on a much lower scale, Pfiester et al. (Pfiester et al., 2021) investigated the downtown of Melbourne (Australia), obtaining higher goodness-of-fit for models considering spatial heterogeneities with respect to global models. However, the examination of spatial determinants' variations remains constrained, primarily showing only some heterogeneity concerning the proximity to emergency facilities. Similar findings regarding the spatial variation of determinants have been noted in analogous contexts, such as cycling mobility (Lu et al., 2018; Munira & Sener, 2020) or in other travel modes (Lu et al., 2012). In conclusion, there appears to be a specific scale at which spatial heterogeneities emerge, and this scale seems to be the urban scale which remains mainly unexplored.

In this study, our objective is to address the identified gap in spatial-scale analysis by examining the variability of pedestrian mobility drivers in Barcelona's commercially salient areas. We achieve this by leveraging footfall data—pedestrian counts continuously gathered in commercial areas over five years (2017–2022) by TC Group (TC Group Solutions, 2022)—in a prototypical European city: Barcelona, located in Catalonia, Spain. Barcelona constitutes a particularly interesting case study, because of the prevalence of pedestrian mobility in the city. In the context of European cities, statistics show that the modal share for walking urban trips typically range from 10 % to 50 % (Buehler & Pucher, 2023), reaching an average of 41 % (Ajuntament de Barcelona, 2019) in Barcelona for working days. Notably, such modal choice is higher than that of private or public motorized transport. This may be explained by the plethora of actions to incentivize sustainable mobility, including pedestrian friendly commercial streets, city-wide speed reduction zones, low emission zone or the flagship projects the *Superilles* ("Superblocks") and *Eixos Verds* ("Green Corridors") at city-wide scale. These policies, combined with the city's national and international scientific interest, have already been the object of several studies related to the topic of mobility (Gascon et al., 2019; Graells-Garrido et al., 2021; Marquet & Miralles-Guasch, 2014; Vich, Marquet, & Miralles-Guasch, 2019).

In this context, our research provides a comprehensive analysis of the factors influencing pedestrian movement patterns in Barcelona, particularly focusing on the spatial heterogeneity that characterize these factors. We aim the insights from our study serve as valuable inputs for urban planning and mobility strategies, not just for Barcelona but for cities globally that encounter parallel challenges and opportunities facilitating pedestrian infrastructure and promoting walking as a sustainable transportation option.

The paper is organized as follows. First, we examine the evolution of pedestrian traffic across the city over the last five years, aiming to uncover both global trends and local particularities in the changing volumes of pedestrian flows on streets. Next, zooming in on a single year - 2019, the year before the pandemic lock-downs - we relate the observed volumes to several urban variables. A two-step regression analysis enables us to pinpoint the determinants of pedestrian activity and explore their variation throughout the city. This approach accounts for the spatially non-stationary regression models of urban factors associated with walking, allowing for a spatial understanding of how these influences differ across different areas. Finally, the article closes by providing a summary of the obtained results, their implications and limitations, and identifying future lines of research.

2. Material and methods

2.1. Acquiring and cleaning pedestrian count data

Although in recent years several analyses have been conducted to explore walking patterns at a global (Alessandretti & Jørgensen, 2021; Bongiorno et al., 2021; Ewing & Cervero, 2010; Gascon et al., 2019; Moro, Calacci, Dong, & Pentland, 2021; Sevtsuk et al., 2024; Simini, González, Maritan, & Barabási, 2012) or local (Chen et al., 2022; Feuillet et al., 2015; Feuillet et al., 2018; Lu et al., 2012; Pfiester et al., 2021) scale, publicly available pedestrian mobility data is scarce – especially in the European Union, where data protection regulations are quite strict. Additionally, such data is often proprietary, held by private entities.

Here, we exploit raw pedestrian count data provided by TC Group (TC Group Solutions, 2022). This retail intelligence company continuously captures pedestrian data from hundreds of sidewalks and shopping centers in Barcelona. Although the oldest sensors have been operating for >10 years, only since 2017 has there been a sufficient number of sensors for a robust analysis with sufficient spatial resolution. Fig. 1 shows the distribution of the subset of sensors used in the present study and their annual averaged hourly traffic. As shown, the sensors are primarily distributed in the city center and some residential neighborhoods. Despite TC Group's focus on retail companies, these areas in Barcelona encompass a rich mix of local commerce, retail hubs, and mixed-use zones characterized by dense demographic, touristic, and residential activity.

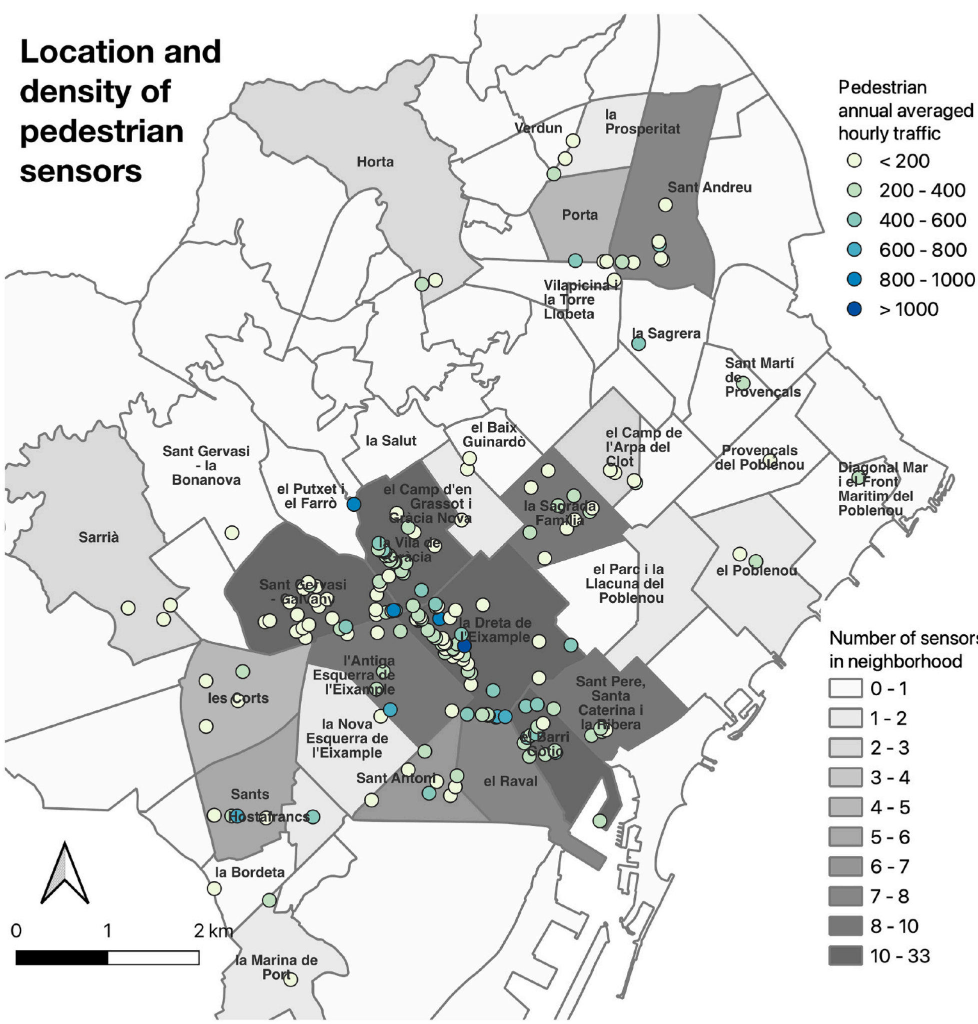

Fig. 1. The left panel shows the distribution of the 196 sensors used in the present study. The colour of each circle indicates average hourly pedestrian traffic.

Each sensor contains a photo-diode facing the outside of a retail store, which is able to count individual pedestrians with a minimum longitudinal separation of 25 cm, at a maximum distance of 3 to 5 m from the facade of the store. Notably, this kind of sensing device does not compromise privacy in any way. Based on these hardware characteristics, we identified three potential issues with data collection that need to be addressed: (1) sensors may not capture the full width of the sidewalk, (2) under-sampling of pedestrians walking in groups and (3) over-sampling of sensors located close to bar terraces. Thus, we have adjusted each sensor count by: scaling the counts by considering the range of the sensor and the current width of the sidewalk and the expected amount of groups walking in sidewalks (Harms, van Dijken, Brookhuis, & De Waard, 2019). Additionally, to compensate for (3), we manually discarded sensors close to (within a few meters of) locations susceptible to capture non-transiting pedestrians (e.g., bar terraces). After additionally discarding the sensors in indoor shopping malls and the ones susceptible to produce anomalous counts, we ended up with 115 sensors spanning several longitudinal periods (from 5 years to few months) up to March 2022 with a temporal resolution of 30 min.

2.2. Identifying tendencies in pedestrian count time series

As a preliminary analysis of the dataset, in Sec. 3.1, we will focus on determining trends within the pedestrian flow data, to examine the progression of pedestrian counts over time. To this objective, we rely on time series analyses to identify *break points*, defined as points (in time) where the time series experience significant changes in its mean or slope (Killick, Fearnhead, & Eckley, 2012). Since the method in (Killick et al., 2012) needs to be provided with the number of break points to be identified as a parameter, we use the elbow method heuristic (usually applied to k-means clustering algorithm) to find the optimal number of change points to consider. The elbow method, grounded in the principle of diminishing returns, enables us to determine the optimal number of break points by identifying the point of maximum curvature on the plot that contrasts the number of break points with the sum of residuals.

2.3. Urban factors shaping pedestrian mobility

The main objective of the study is to understand how pedestrian mobility is influenced by the characteristics of the surrounding area. Guided by the findings reported in the literature, we look for potential relationships between the observed pedestrian count data and the variables describing different aspects of the urban fabric and societal context: socio-demographic, comercial, economic, techno-infrastructural, environmental, and governance data. Considering the variables identified as significant in the literature (Hatamzadeh et al., 2014; Marquet & Miralles-Guasch, 2014; Perchoux et al., 2019; Yang & Diez-Roux, 2012), we assembled the list of descriptors summarized in Table 1. In the following sections, we explain the details of each of these variables and, where needed, their pre-processing.

2.3.1. Socio-demographic variables

Highly-resolved socio-demographic data was obtained from the city of Barcelona's open data portal (Ajuntament de Barcelona, 2022). Although socio-demographic data can be obtained for several years, we did not observe significant variations in their values. Thus, we focus on data from 2019 at the smallest available geographical scale: census tract. For reference, the city of Barcelona is divided into 10 districts, 73 neighborhoods, 233 basic statistical areas, and 1068 census tracts (regular shapes containing around 1000 voters). Since pedestrian count sensors capture flow at a very local scale, it is most adequate to consider data at the smallest possible geographical scale. Population density was calculated by dividing the total population by the area of each census tract. The median age was calculated using age pyramids in each census tract. The gender ratio refers to the female population divided by the male population. The education level variable refers to the percentage of the population with a university degree. Finally, motorization index (vehicle ownership rate per 1000 inhabitants) is calculated as the sum of automobiles, motorcycles, scooters and vans registered in the census tract, divided by the population.

2.3.2. Urban infrastructure variables

To quantify the impact of urban features and infrastructure, such as amenities, transportation systems, and green spaces, on pedestrian traffic, we concentrate on quantifying their volumes within a defined distance buffer (Lu et al., 2012; Pfiester et al., 2021). Different geographical scales were explored to conduct the analysis: neighborhood, census tract, and 500 m-buffers around sensors. Finally, so as to maximize the number of observations while maintaining a high geospatial granularity, we opted for aggregating the independent variables at 100 m-buffers around the sidewalk network where the sensor is installed and use the corresponding census-tract level data (the finest available) to incorporate the socio-demographic information. For this purpose, we use the sidewalk network for Barcelona provided in (Rhoads, Solé-Ribalta, González, & Borge-Holthoefer, 2021). The sidewalk segments on which sensors are located have a median length of 79 m, which translates to a median buffer area of 47,000 m^2.

Street intersections were obtained from OpenStreetMap data of the Barcelona road network, using the open-source Python package OSMnx (Boeing, 2017). The intersection density was calculated by counting the

Table 1
Summary of variables assembled for the regression analysis.

Category	Variable	Units	Resolution	Year
Dependent variable	Pedestrian counts	Hourly pedestrian counts	Point data	2012–22
Socio-demographics	Age median	Median age	Census tract	2019
	Gender ratio	# Female/# male	Census tract	2019
	Education level	% People with university degree	Census tract	2019
	Vehicle ownership	# Vehicles/1000 inhabitants	Census tract	2019
	Population density	# Residents/km^2	Census tract	2019
Urban infrastructure	Intersection density	# Road intersections/km^2	Point data	2022
	Floating population	# Hotel beds/km^2	Point data	2022
	% Green spaces	Green space area/total area	Polygon data	2019
	% Pedestrianized areas	Pedestrianized area/total area	Polygon data	2019
	Arrogance of space	Sidewalk area/roadbed area	Polygon data	2017
	Number of bus stops	# Bus stops/km^2	Point data	2022
	Number of metro stations	# Metro stops/km^2	Point data	2022
Access to services	Education and cultural facilities	# POIs/km^2	Point data	2019
	Health and related facilities	# POIs/km^2	Point data	2019
	Local neighborhood shops	# POIs/km^2	Point data	2019
	Restaurants, bars and hotels	# POIs/km^2	Point data	2019
	Offices and services businesses	# POIs/km^2	Point data	2022
	Touristic points of interest	# POIs/km^2	Point data	2022

number of intersections within each buffer, and dividing by the buffer area. Bus stops and metro stop locations were obtained from the Barcelona Open Data portal (Ajuntament de Barcelona, 2022). Note that the data include the exact location of all entrances to a given metro station, and bus stops on both sides of the street. Thus, the bus and metro stop densities are obtained by counting the number of bus stops and metro stops respectively within each buffer, and dividing by the buffer area.

Barcelona is among the most touristic European cities, attracting over 23 million visitors in 2019 (Barcelona Regional, 2019). In certain neighborhoods, and during the peak season, many more tourists than local residents can be seen walking. Therefore, some measure of touristic activity must be included in our analysis. The Barcelona Tourism Activity Report (Observatori del Turisme a Barcelona, 2020) uses the number of beds in hotels and other tourist accommodation to quantify tourist activity. We used a census of all registered hotels and tourist accommodations in Barcelona to retrieve their address and the corresponding number of beds. We then geocoded these addresses to obtain point coordinates of tourist accommodations, using open-source service Nominatim (Boeing, 2017). Finally, we counted the number of accommodations in each buffer, summed the number of beds, and divided by the buffer area.

Whereas the above data can be represented as points, other features of urban fabric are better represented as polygons: green spaces, pedestrianized streets and pedestrian areas. The city of Barcelona's Guia Urbana offers a rich set of polygon data (Ajuntament de Barcelona, 2022). We perform a spatial overlay between the green spaces polygons and the buffers to obtain the area of green space contained within each buffer. We then divide the green area within each buffer by its total area to obtain the % green area. We repeat this process to obtain the % pedestrian area within the buffers. Note that pedestrian area here refers to pedestrianized streets and public spaces, not sidewalks of regular streets.

The last urban feature we accounted for is the "Arrogance of Space", coined by the urban designer Colville-Andersen (Colville-Andersen, 2018) to highlight how much urban street space was occupied by cars and car-related infrastructure, compared with other street users (pedestrians, cyclists, public transport users). Building upon this concept, we calculated here the ratio of sidewalk (pedestrian) space to roadbed (car) space, at the geographical resolution of our buffers. Sidewalk and roadbed polygons were constructed in-house using the Cartographic and

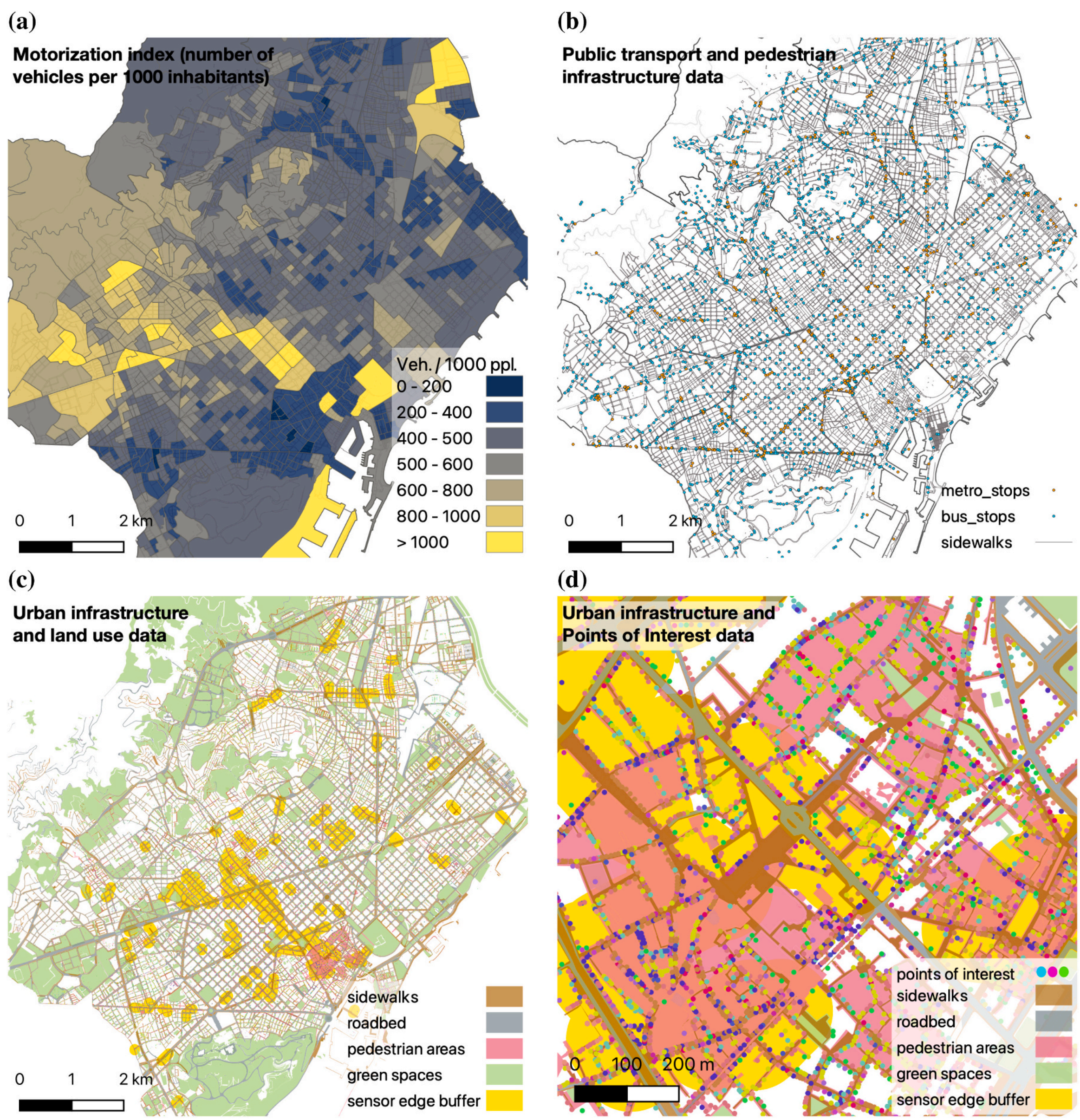

Fig. 2. Series of maps illustrating the variables used in the different regression models.

Geological Institute of Catalonia (ICGC)'s 1:1000 topographic maps (Rhoads et al., 2023). We then repeated the same spatial overlay process as for green space and pedestrian space, and simply divided sidewalk area by road area in each buffer.

2.3.3. Points of interest data

Barcelona's Open Data portal (Ajuntament de Barcelona, 2022) features a rich dataset of points of interest (POI), consisting of over 80,000 points, split between 18 categories, and 83 activities. We have grouped these POIs into 5 typologies quantifying similar services, using the metadata provided in the portal:

- POI 1: Educational and cultural facilities
- POI 2: Health facilities
- POI 3: Daily, neighborhood shops (food, pharmacy, clothing, tobacco, newspapers, etc.)
- POI 4: Hotels, bars and restaurants
- POI 5: Other businesses (hair salons, veterinaries, travel agencies, fitness centers, etc.)

The density of each POI typology is calculated by counting the number of POI in each buffer, and dividing by its area.

2.3.4. Spatial analysis of the urban descriptors

To complement the textual definitions above, we provide a more visual description of the variables considered in our study in Fig. 2. Panel (a) shows the motorization index (number of vehicles per 1000 inhabitants) which varies widely across Barcelona. Lower-income, dense neighborhoods - whether inner city or peripheral - can have under one vehicle for every five people, while affluent neighborhoods around Passeig de Gràcia or above Avinguda Diagonal have more than one vehicle for each inhabitant. It is worth noting that industrial zones in the south and north edges of the city feature also a high number of corporate vehicles registered with few residents, which explains their high motorization index.

As can be seen in Panel (b), public transport infrastructure is fairly homogeneously distributed across the city (with the exception of the hilly areas in the northwest). The bus network of Barcelona was recently redesigned to maximize transfer efficiency, while the metro network has been extended to the city's immediate suburbs.

Panel (c) shows the variables describing the physical layout and land use of the city. Barcelona notoriously lacks access to green spaces, which are concentrated on the Montjuïc hill in the south of the city, and along the northwestern border of the city, in the Collserola hills. The ratio of sidewalks to roadbed (i.e. the reverse of Colville-Andersen's "arrogance of space") is higher in the historic core of the city, as can be seen on Panel (d), as well as the pedestrianized cores of more residential neighborhoods, and by the seaside. In addition, neighborhoods such as Gòtic, Raval or Sant Pere, Santa Caterina i la Ribera in the center but also Vila de Gràcia, Sants or Sant Andreu in the periphery feature a high percentage of pedestrianized areas, as can be seen in Fig. 2 panels (c) and (d).

Finally, the over 80,000 points of interest inscribed in the city of Barcelona's database are heterogeneously distributed across the territory. The highest density of educational and cultural facilities are located in the city center, Eixample and Gràcia districts. Hospitals, clinics and health facilities seem to be clustered in certain peripheral districts: les Corts, Sarrià-Sant Gervasi and Gràcia. Local, neighborhood shops are found all over the city, yet higher densities can be observed in areas with stronger local life: Sant Antoni, Vila de Gràcia, Hostafrancs, etc. Unsurprisingly, hotels, bars and restaurants are mostly concentrated in the central districts (Ciutat Vella and Eixample), the most touristic area of the city. Finally, offices and other services are less present in the center and more evenly distributed across the periphery.

2.4. Regression modeling

2.4.1. Generalized linear models

Linear regression models are a simple yet insightful tool to investigate the nature of relationships between variables (Dobson & Barnett, 2018; Nelder & Wedderburn, 1972). As such it is one of the first tools used to determine the relationship between the influential factors of walkability, pedestrian volumes and urban vibrancy (Boakye et al., 2023; Chen et al., 2022; Zhao & Wan, 2020). The most simple regression models assume a linear relationship between the dependent variable, y, and the p independent ones, x. Each observation i of the dependent variable, $y(i)$, can be modeled as:

$$y\left(i\right) = \beta_0 + \sum_k \beta_k x_{ik} + \epsilon_i \tag{1}$$

where β's are the coefficients relating the independent variables to the dependent ones. In the equation, index i refers to the different observations and x_{ik} to the value of the $k \in \{1, \ldots, p\}$ predictor variable associated to observation i. Given the different observations, one can find the β_k coefficients that minimize the least squares between the observations and the predicted value. Once adjusted, the model accuracy is usually assessed in terms of the coefficient of determination, R^2 (or Adjusted R^2), describing the proportion of the variance of the dependent variable that the model is able to explain. Eq. 1 can be generalized to different assumptions regarding the type and distribution of dependent variables, assuming $g(y(i))$. Function g is the link function describing the relationship between the linear predictors and the mean of the distribution of the dependent variable. In this work, we will focus on the use of linear regression models, where g is the identity function, and Poisson regression models, where g is the logarithmic function.

Often, adding too many independent variables can reduce the accuracy of the model (due to potentially insignificant relationships) and affect the reliability of parameter estimates, which pivots on the ratio of observations to independent variables. In these scenarios, the standard strategy involves selecting a good subset of independent variables. In this sense, Backward elimination (Thompson, 1978) consists in the recursive elimination of independent variables that present non-statistically significant relations with the dependent variable. In contrast, forward selection (Thompson, 1978) consists in iteratively adding independent variables to the model until the overall performance of the model no longer improves.

Further tests in this line may also consider the Akaike information criterion (AIC), a metric that compares the fit of several regression models while taking into account the number of parameters included.

2.4.2. Spatially-aware regression models

A persistent issue with applying generalized linear models to spatially embedded data is their inability to account for potential spatial heterogeneities in the relationship between dependent and independent variables (Anselin, 2013; Goodchild, 2004). This limitation can be evaluated by analyzing the Moran's index of the model residuals, offering insights into the spatial distribution of these discrepancies (Pfiester et al., 2021). One viable approach to address this issue involves accommodating the spatial non-stationarity of the regression model's coefficients. This means assuming that observations in close proximity should exert similar effects on the explanatory variables, while this may not hold true for those farther apart. Considering Eq. 1, on one side we may expect that β coefficients vary across the study area, while close observations share similar β coefficients. Geographically weighted regression models address these two issues (Brunsdon et al., 1996). First, independent coefficients, β, are incorporated for each location i:

$$y\left(i\right) = \beta_0\left(i\right) + \sum_k \beta_k\left(i\right) x_{ik} + \epsilon_i, \tag{2}$$

which allow to accounts for spatial heterogeneity. Then, $\beta_k(i)$ coefficients are obtained following the weighted least squares approach:

$$arg\ min_{\beta_{1..p}(i)} \sum_j \omega\left(i,j\right)\left[y\left(j\right) - \left(\beta_0\left(i\right) + \sum_k \beta_k\left(i\right) x_{jk}\right)\right]^2 \tag{3}$$

where $\omega(i,j)$ is a function of the distance between observations i and j, that includes a spatial decay given by a kernel function governed by a bandwidth parameter. These weights force coefficients $\beta_k(i)$ to adjust not only to observation i, but also include the influence of nearby observations. This model can be generalized to consider other distributions of the data (Fotheringham & Brunsdon, 2002), such as Poisson or Negative Binomial. The current state-of-the-art also provides more elaborate methods to include other spatial and temporal influences over the dependent variable if necessary, such as the Multiscale Geographically Weighted Regression (Fotheringham, Yang, & Kang, 2017; Oshan, Li, Kang, Wolf, & Fotheringham, 2019) that accounts for different bandwidths for each independent variable, or Spatio-Temporal Weighted Regression (Que, Ma, Ma, & Chen, 2020) which also incorporates the temporal dimension within the regression model.

In this work, we designate spatially-aware regression models as "local regression models" to emphasize their ability to discern spatial intricacies. In contrast, models lacking spatial consideration will be jointly labeled as "global regression models," indicating their application across undifferentiated areas. This nomenclature distinguishes between the models based on their capacity to address local versus global characteristics within the data.

3. Results

In the following subsections, we aim to identify the significant factors outlined in Table 1 that determine pedestrian volumes on urban streets and examine their fluctuation across the urban landscape. Our methodology is designed as a two-step analysis, similar to the analysis pipeline of (Munira & Sener, 2020; Pfiester et al., 2021). We begin with a global regression model to identify which variables have a meaningful impact on the dependent variable (pedestrian counts). After that, we proceed with spatially-aware regression models to test our hypothesis regarding the spatial heterogeneity of factors and to gain a comprehensive insight into the variables that shape pedestrian mobility.

Building upon the insights presented in the introduction regarding the socio-economic and urban factors influencing pedestrian mobility, and recognizing that these factors vary within a city, as demonstrated for Barcelona in Fig. 1 and in (Graells-Garrido et al., 2021), we first extend this analysis to the dependent variables, considering their evolution over time and across various neighborhoods, admittedly with less spatial detail. This temporal analysis serves a dual purpose. Firstly, it helps determining whether distinct pedestrian activity profiles exist, which could suggest that a uniform city-wide model might overlook specific local variations. Secondly, it aids in selecting the most appropriate time frame for conducting the regression analysis in subsequent sections.

3.1. Temporal evolution of pedestrian traffic across Barcelona

To analyze the temporal evolution of the pedestrian counts, it is fundamental to consider the turnover of sensors in our dataset. Simply averaging counts over time and incorporating new sensors (which may have different volumetric characteristics) into each time window without adjustment could yield inaccurate results.

To gain an initial insight into the dataset, Fig. 3 presents weekly averages obtained from a selected subset of sensors, out of the 196 available, that were consistently operational during the entire analysis period from January 2017 to March 2022. Results show fluctuations and no evident general trend, but statistical analysis with the Augmented Dickey-Fuller test (Dickey & Fuller, 1979) indicates non-stationarity (p-value of 0.49) between 2017 and 2020, and an increment of 0.3168 pedestrians per hour every week. This tendency is more evident considering the full set of sensors during the period (see inset of Fig. 3), which displays a slightly positive trend up to the end of 2019. The sharp decrease in the Spring of 2020 corresponds to the first government-mandated Covid-19 lockdown in Spain, with a gradual recovery of pedestrian traffic as the pandemic-induced restrictions ceased. It is worth noting that, 2 years after the initial outbreak of the pandemic, pedestrian traffic has not yet recovered to pre-pandemic levels.

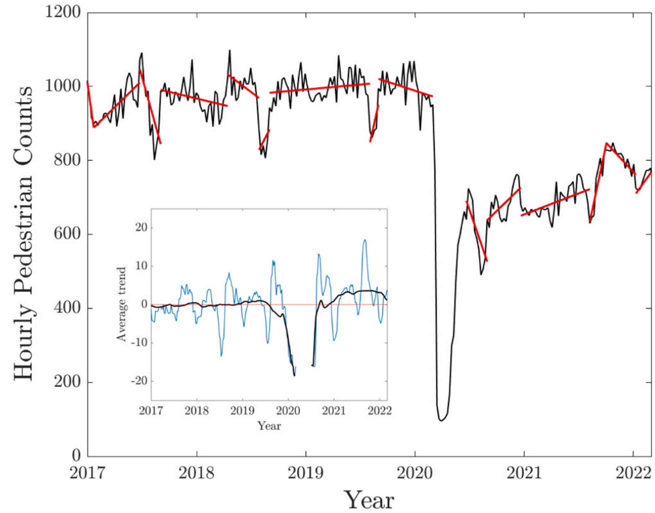

Fig. 3. Evolution of pedestrian counts across Barcelona, between 2017 and 2022. In the main plot, the black line represents raw average hourly traffic considering all sensors available for the full period 2017–2022. Sensors with missing data were discarded. The average hourly traffic is obtained by averaging the counts obtained on weekdays during business hours (10-14 h and 17 to 20 h) for each sensor, over all available sensors. The red lines represent the different trends obtained as described in Section Data and methods. The inset represents the average increment (pedestrians per week) obtained considering all sensors available at that time. Specifically, for each sensor we obtained the individual trend summary line (the equivalent of the red line in the main plot but considering each sensor independently) and then report the average trend of all sensors available for each week in the 2017–2022 period. The black line represents the moving average smoothing (60 days time window), displayed to ease interpretation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Narrowing our focus to the neighborhood level allows us to uncover the hypothesized localized trends. Fig. 4 (a) displays the weekly hourly averages of pedestrian traffic across six representative neighborhoods. A technical focus on the Augmented Dickey-Fuller test indicates again non-stationary trends for all these neighborhoods throughout the period 2017 to 2020 (p-values in the range of 0.3 to 0.5). Only two neighborhoods show very slight steady negative trends: el Barri Gòtic (slope $= -0.3$) and Sant Andreu (slope $= -0.7$), with strong regular seasonal variations. Interestingly, the sensors in these two neighborhoods represent paradigmatic examples of commercial areas characterized by predominantly touristic and residential demographics, respectively. Heterogeneous trends are observed in the neighborhood of Sant Pere, Santa Caterina i La Ribera, which appears to increase up until mid 2019, and then starts decreasing. Analysis of the weekly gradients for the period confirm these findings, see Fig. 4 panel (b). Neighborhoods Dreta de l'Eixample and Vila de Gràcia show positive slopes with a quite steady increase in them. The touristy, central neighborhoods of Gòtic or Sant Pere, Santa Caterina i La Ribera, display more heterogeneous trends. It is also important to highlight the drop associated to the first lockdown in practically all neighborhoods. Even though we exclude the lockdown period in the calculation of the trend line (black line), it seems that there were clear pre-lockdown effects.

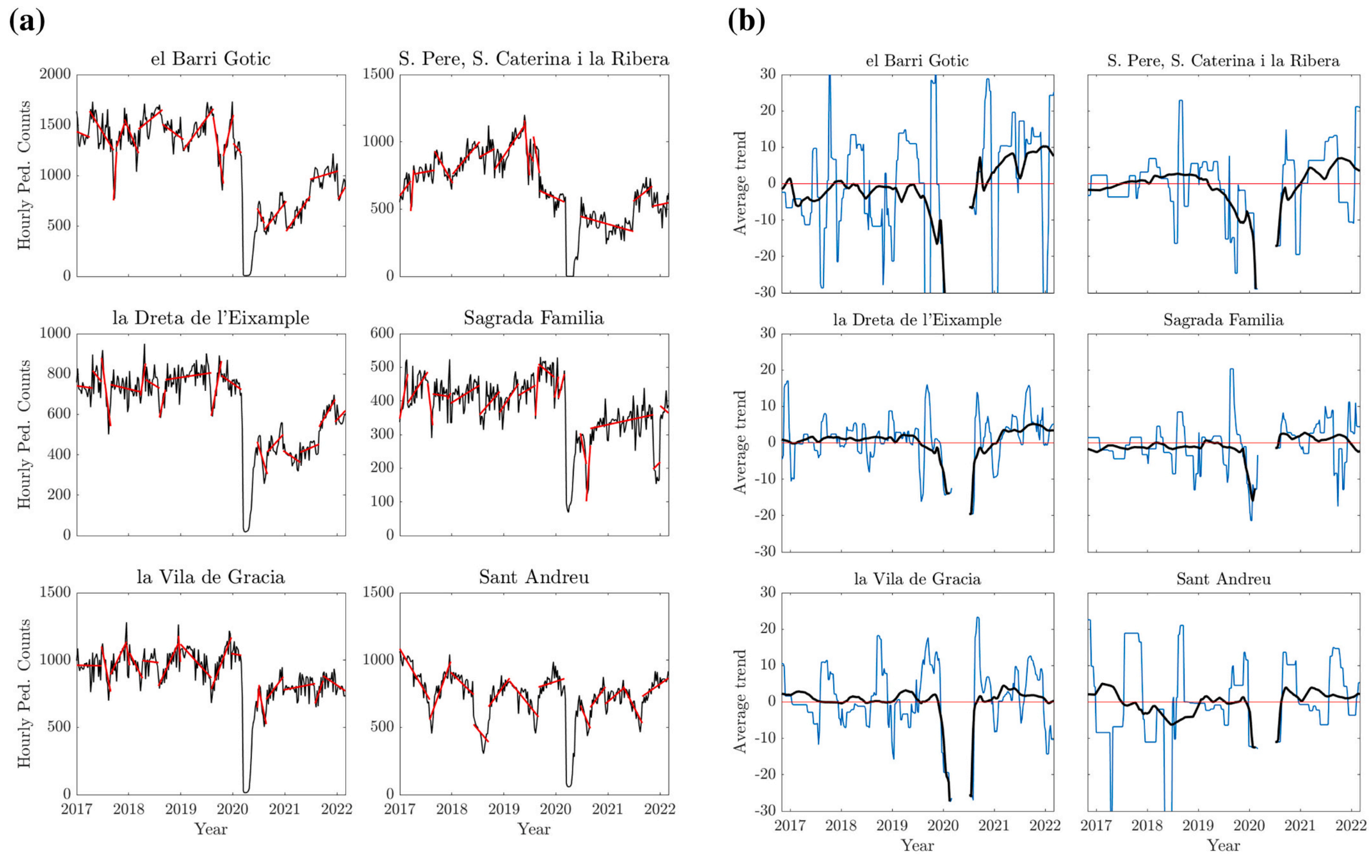

Fig. 4. Panel (a). Evolution of pedestrian counts in 6 neighborhoods of Barcelona between 2017 and 2022. The black line represents raw average hourly traffic for sensors available for the full period 2017–2022. Sensors with missing data within the period were discarded. The average hourly traffic is obtained using the counts obtained during business hours (10 to 14 h and 17 to 20 h) for each sensor, averaged over each neighborhood. The red line represents a trend line obtained as described in Section Data and methods. Panel (b). Evolution of pedestrian counts in the same 6 neighborhoods. The blue line represents the average weekly gradient from all sensors available at that time. For each sensor, we computed the individual trend line (the equivalent of the red line in the left panel but for each sensor independently) and report the average trend of all sensors available in the given neighborhood for each week in the 2017–2022 period. The black line represents a moving average smoothing (60 days time window) of the blue line, intended to ease interpretation. Units of increment represent pedestrians per week. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Contextual attributes influencing pedestrian traffic

To statistically discern which contextual attributes influence pedestrian mobility in the areas monitored by sensors, we employed two suitable regression global models: a classical linear regression model and a Poisson regression model. Although Poisson regression is generally more appropriate for discrete count data as it is restricted to non-negative integers, the observed averages of pedestrian counts in our study are sufficiently large to make the use of linear regression models feasible, which offer advantages in terms of interpretability. In both models, pedestrian counts serve as the dependent variable, while the independent variables include subsets of the socio-demographic, urban infrastructure and points of interest variables from Table 1.

Due to sensor turnover in our dataset, a direct and comprehensive spatio-temporal analysis is not feasible. To address potential temporal variations, we explored different aggregation windows for each regression model, following the strategy outlined in (Pfiester et al., 2021). This method allowed us to assess the impacts of seasons, days of the week, and times of day by selecting sensor data subsets from various time frames. This approach contrasts with the alternative method of incorporating season, day of the week, and time of day as categorical variables in the analysis, as seen in some studies (Lu et al., 2018).

We also examined the results from different years separately (2018 to 2022), in an attempt to tease out the evolution of walking trends in Barcelona. Data from 2020 were excluded due to the pandemic-induced consecutive lock-downs which paralyzed the city. Data from 2021 and 2022 overall displayed low R^2 values, suggesting complex mobility patterns in the post-pandemic era that our independent variables do not enable us to adequately explain. Finally, we opted to focus our attention on the year 2019 for four reasons: 1) the time series analysis shows us that is the most prolific period; 2) it is the most recent complete year we have before the pandemic; 3) the number of sensors active during the whole period is sufficient to provide a large-enough number of observations for our analysis; 4) most of the independent variables we gathered refer to this year. After some data cleaning, we are left with 115 sensors for 2019 distributed across 98 sidewalk network segments, split evenly between the central districts of *Ciutat Vella* and *Eixample* (49 observations) and the peripheral ones (49 observations).

In this scenario, we applied forward selection for both global regression models to identify the significant subset of variables from the initial 18 considered. This approach led to the consistent identification of 7 variables statistically significant across different temporal scenarios. The robustness of the subset selection process was cross-validated by proceeding with backward elimination from the initial subset, which yielded to the equivalent selection. We note that the ratio of the number of observations to independent variables is 14, which is considered to be acceptable (Hair, Black, Babin, & Anderson, 2009).

The demographic variables of age and gender were found to be non-statistically significant here in agreement with some previous studies (Gascon et al., 2019; Vich et al., 2019) and in contrast with other ones (Charreire et al., 2021; Marquet & Miralles-Guasch, 2014). Further, neither population density, education level, nor intersection density were found to be statistically significant here, unlike (Gascon et al., 2019; Gómez-Varo, Delclòs-Alió, & Miralles-Guasch, 2022; Rames et al., 2021; Vich et al., 2019). The stability of census data in Barcelona, showing minimal variation over the past decade, poses challenges in correlating temporal patterns with changes in socioeconomic variables known to influence walking levels (Gascon et al., 2019). Thus, determining correlation in these cases remains challenging.

The density of hotel beds, which we used as a proxy for *floating population*, did not appear to be statistically significant – neither did the density of touristic attractions. The ratio of sidewalk area to roadbed area and the density of metro stations also turned out to not be statistically significant in our analysis. Finally, among the points of interest, neither the *educational and cultural facilities*, nor the *health and related facilities* showed a significant relationship with pedestrian volumes.

The forward selection method identified the following variables consistently significant: number of vehicles/1000 inhabitants, density of bus stops per km^2, % of green spaces, % of pedestrianized areas and density of offices and other service businesses per km^2 (POI 5). The findings related to the density of *Daily use neighborhood shops* (POI3) and *Hotels, bars, and restaurants* (POI4) per km^2 presented contrasting outcomes based on the regression method applied. Specifically, only the Poisson regression method flagged POI3 and POI4 as statistically significant.

We also tested the final selected subset of variables for multicollinearity, and they appear to be mostly uncorrelated - with the exception of the POI 3 and POI 4 categories which have a 0.5 Pearson correlation coefficient. We did not consider this Pearson correlation significant enough to discard the variables. In our analysis of the residuals from various regression experiments, we observed homoskedasticity, indicating uniform variance across the predicted values. Additionally, our examination of the Q-Q plots for these residuals revealed that they largely conform to a normal distribution, though there are minor deviations at the distribution's tails.

Overall, the global regression models captures between 38 % and 51 % of the variance observed in the pedestrian counts (Table 2). This suggests that, while these environmental factors clearly influence walkability, pedestrian mobility in urban spaces is a very complex phenomenon with a high inherent variability.

Table 3 presents the results of the regression analysis. As expected, pedestrianized areas show a positive correlation with walking activity across all analyzed areas. A slightly counterintuitive finding is that the proximity of green spaces is negatively correlated with pedestrian counts. This may be explained by the fact that green spaces are primarily spaces for leisure rather than transiting through. Another possible explanation could lie in the particular distribution of green spaces in Barcelona, which are primarily found in the outskirts of the city or within building areas, with the exception of *Parc de la Ciutadella*. Thus, almost 40 % of sensors have zero green area within their respective buffers. Finally, offices and other businesses (POI 5) are negatively correlated with pedestrian volumes, potentially suggesting that white collar workers may prefer other modes for their professional mobility needs.

Seasonal variations are fairly minor, as in (Basu & Sevtsuk, 2022). Overall, the trends observed and detailed above are more pronounced during the warm weather months (April to September) than during the rest of the year, but patterns remain similar. Pedestrian volumes are higher in the Summer and on weekends, which seems to affect the explanatory power of the models as R^2 is generally higher in these periods. Finally, differences between morning and afternoon are quite negligible, with very similar patterns overall, perhaps slightly more pronounced in the afternoon. A similar results can be seen in (Lu et al., 2018).

3.3. *Spatial heterogeneity in pedestrian traffic determinants*

Global regression models have been instrumental in identifying variables with a statistically significant impact on pedestrian traffic, but they only capture a portion of the observed variance. Crucially, these models are insufficient to analyze the spatial heterogeneity of the drivers for mobility that likely characterizes different areas of the city. A simple test to evaluate whether a spatially aware model is required is to examine the spatial autocorrelation of the model residuals (Chen, 2016; Pfiester et al., 2021). Computing Moran's I over the global model residuals reveals statistically significant spatial autocorrelation of residuals in 14 out of the 16 scenarios examined (p-value <0.1), with Moran I values ranging from 0.476 to 0.748, as seen in Table 2. This numerically confirms the presence of substantial spatial auto-correlation in our data. Therefore, a spatially aware model is appropriate here, whether using a linear or Poisson regression model.

To examine the spatial variations of the influence of the independent variables, we applied Geographically Weighted Regression (GWR) adaptations of the initially considered global regression models. This analysis was conducted across the same eight temporal windows—distinguished by season (winter/summer), type of day (weekday/weekend), and time of day (morning/afternoon)—and took into account the same seven independent variables.

These analysis were conducted with the MGWR software package (Oshan et al., 2019) which allowed us to compute the optimal bandwidth associated with each independent variable. Yet, we obtained very similar bandwidth values across the different variables. Thus, the

Table 2
Geographically Weighted Linear (GWLR) and Poisson (GWPR) Regression model metrics for different combinations of seasons, days, hours and geographical scope, using 2019 data. Column names stand for: "Obs." (number of observations, sensors), "AIC" (Akaike information criterion), "BW" (optimal bandwidth of the regression models). Symbol'***' indicates p-value<0.01,'**', p-value<0.05 and'*' p-value<0.1. GWPR values are displayed in italics for ease of reading.

Season	Days	Hours	Obs.	Model	R^2	Adjusted R^2	AIC	BW	Moran's I
Oct-Mar.	w/day	a.m.	99	GWLR	0.517 (0.389)	0.414 (0.342)	1466 (1468)	82	0.507**
				GWPR	*0.515 (0.377)*	*0.413 (0.329)*	*11,510 (15605)*	*82*	*0.703*****
		p.m.	99	GWLR	0.626 (0.399)	0.486 (0.353)	1486 (1494)	64	0.502**
				GWPR	*0.632 (0.393)*	*0.499 (0.346)*	*10,016 (17330)*	*64*	*0.632**
	w/end	a.m.	99	GWLR	0.589 (0.427)	0.472 (0.382)	1453 (1456)	74	0.460
				GWPR	*0.597 (0.417)*	*0.488 (0.372)*	*10,184 (15540)*	*74*	*0.748*****
		p.m.	94	GWLR	0.728 (0.497)	0.609 (0.456)	1415 (1430)	57	0.490**
				GWPR	*0.743 (0.492)*	*0.636 (0.451)*	*9896 (20249)*	*57*	*0.656*****
Apr-Sept.	w/day	a.m.	99	GWLR	0.564 (0.485)	0.480 (0.445)	1457 (1456)	85	0.476*
				GWPR	*0.559 (0.475)*	*0.477 (0.435)*	*10,984 (13912)*	*85*	*0.742**
		p.m.	99	GWLR	0.671 (0.507)	0.562 (0.469)	1480 (1485)	68	0.480*
				GWPR	*0.684 (0.514)*	*0.585 (0.476)*	*9545 (15496)*	*68*	*0.685*****
	w/end	a.m.	99	GWLR	0.586 (0.505)	0.507 (0.467)	1440 (1440)	85	0.433
				GWPR	*0.579 (0.483)*	*0.501 (0.444)*	*11,084 (14389)*	*85*	*0.748******
		p.m.	94	GWLR	0.747 (0.507)	0.637 (0.469)	1407 (1426)	57	0.519**
				GWPR	*0.764 (0.486)*	*0.666 (0.444)*	*9336 (21041)*	*57*	*0.705******

Table 3
Linear and Poisson regression results for different combinations of seasons, days, hours and geographical scope, using 2019 data. LR stands for Linear Regression and PR for Poisson Regression, PR values are displayed in italics to facilitate visualization. Star nomenclature has the following meaning: * p-value<0.05, ** p-value<0.01, *** p-value<0.001. Poisson regression coefficients are scaled by 10^3.

Season	Days	Hours	Model	Vehicles	Bus st.	Green sp.	Ped. sp.	POI 3	POI 4	POI 5
Oct.-Mar.	w/day	a.m.	LR	−0.54***	3.62***	−16.42*	8.22***	−0.06	−0.04	−0.63***
			PR	−0.59***	3.54***	−18.60***	7.67***	−0.05***	−0.01	−0.67***
		p.m.	LR	−0.65***	3.59***	−21.41**	10.07***	−0.07	0.23	−0.70***
			PR	−0.61***	3.27***	−22.62***	8.21***	−0.06***	0.17***	−0.63***
	w/end	a.m.	LR	−0.66***	2.39**	−20.69***	9.06***	−0.07	0.14	−0.51**
			PR	−0.83***	3.08***	−30.48***	8.77***	−0.02***	0.03	−0.57***
		p.m.	LR	−0.86***	2.43*	−35.24***	13.81***	−0.05	0.29	−0.96***
			PR	−0.89***	2.97***	−52.86***	11.17***	−0.06***	0.20***	−0.94***
Apr.-Sep.	w/day	a.m.	LR	−0.53***	3.41***	−28.58***	10.33***	−0.06	0.17	−0.64***
			PR	−0.55***	3.59***	−37.59***	9.34***	−0.06***	0.14***	−0.65***
		p.m.	LR	−0.65***	3.66***	−34.39***	12.81***	−0.06	0.30	−0.72***
			PR	−0.59***	3.60***	−42.67***	10.08***	−0.06	0.23	−0.63***
	w/end	a.m.	LR	−0.61***	2.31**	−22.90***	10.36***	−0.01	0.16	−0.53***
			PR	−0.77***	3.30***	−38.54***	10.40***	−0.03***	0.14***	−0.61***
		p.m.	LR	−0.77***	2.97**	−29.45***	15.41***	−0.06	0.36	−0.83***
			PR	−0.81***	3.73***	−45.06***	13.19***	−0.07***	0.30***	−0.83***

influence of scale was not found to be significant within the city of Barcelona. Therefore, we report only the results of the GWR analysis, configured to use the bi-square distance kernel (see Eq. 3).

The spatial variation in the coefficients from the Geographically Weighted Linear Regression (GWLR) model, illustrated in Figs. 5 and 6, offers valuable insights into the city's heterogeneous relationship with pedestrian mobility. The detailed results for the GWPR model, which are qualitatively consistent with the Geographically Weighted Poisson Regression (GWPR) results, can be found in the supplementary materials, as shown in Figs. S1 and S2.

While motor vehicle ownership rates are always negatively correlated with walking, this effect is much stronger in wealthier, peripheral neighborhoods north to the Diagonal avenue. Conversely, the presence of bus stops is much more strongly positively correlated with walking in the center of Barcelona, where public transit infrastructure is dense and parking availability is scarce. Proximity to green spaces appears to be negatively associated with walking volumes [?]. However, in some residential working class districts (Nou Barris, Sant Andreu or Sant Martí) they are positively correlated with walking. One may contrast the negative city-wide correlation between pedestrian activity and proximity to green areas is notable, but the correlation may be misleading. Zhang et al. (Zhang, Melbourne, Sarkar, Chiaradia, & Webster, 2020) conclude that, in London, park density has no significance, and one needs to look around areas where parks are located to find positive

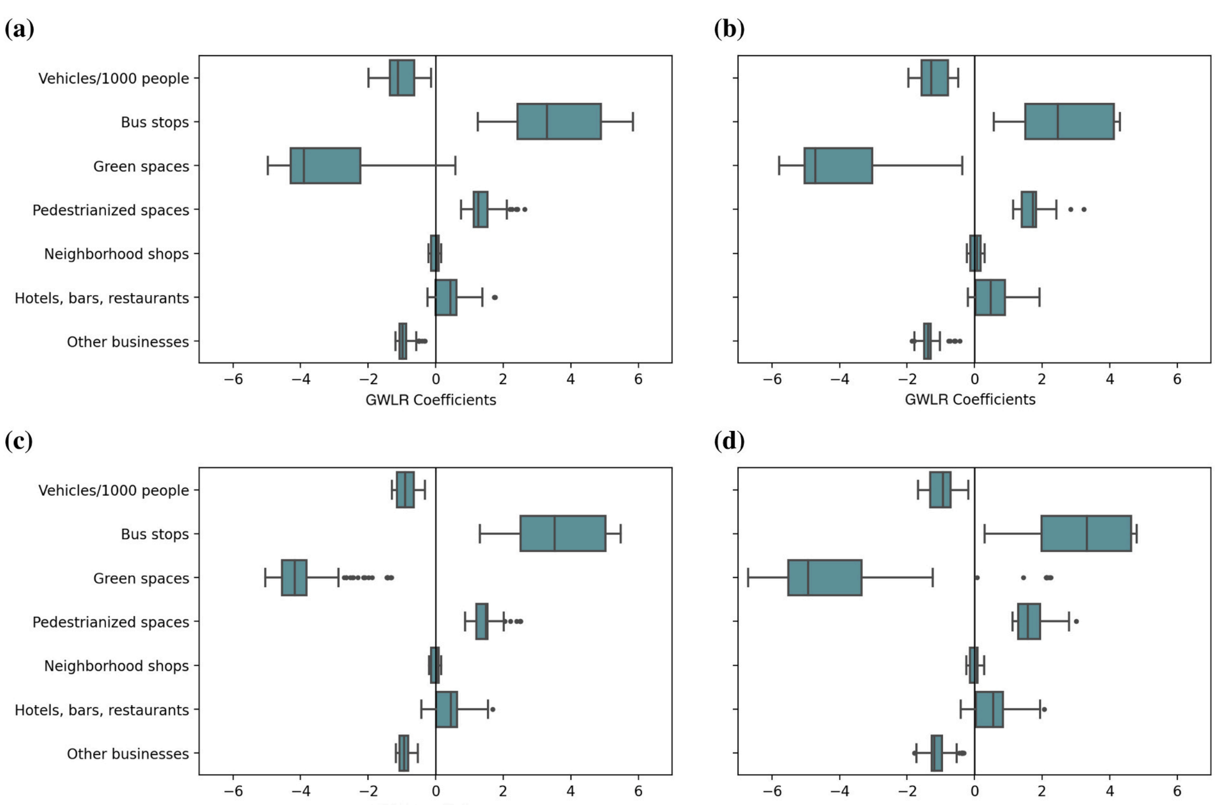

Fig. 5. Box plot of the distribution of GWLR coefficients for all dependent variables. a) winter, weekday; b)winter, weekend; c) summer, weekday; d) summer, weekend. Note that the coefficients for green space and pedestrianized space were divided by 10 so as to fit within the same order of magnitude as the other variables. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

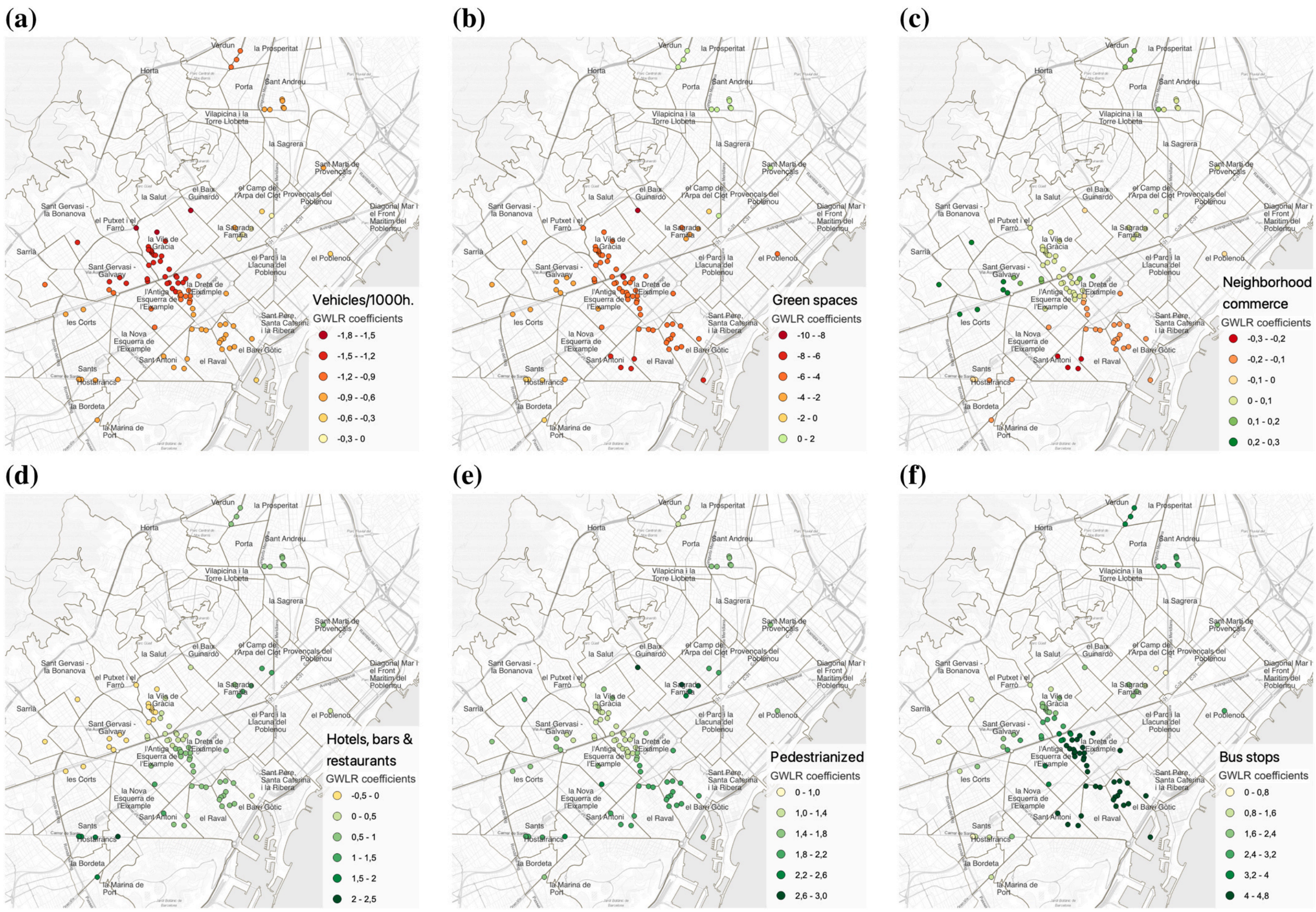

Fig. 6. Maps showing the spatial distribution of the geographically weighted regression coefficients for the following dependent variables: a) vehicles/1000 people; b) green spaces; c) neighborhood commerce; d) hotels, bars and restaurants; e) pedestrianized spaces; f) bus stops. Maps are sorted by negative to positive influence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

significant relations with pedestrian activity. Pedestrianized streets appear to be most strongly correlated with walking in the vicinity of touristic attractions (such as Sagrada Família, Park Güell or the historic city center).

Neighborhood shops are negatively related to pedestrian traffic in gentrified and touristified neighborhoods (Sant Antoni, Barri Gòtic, and Sant Pere, Santa Caterina i la Ribera). In contrast, they are positively correlated with pedestrian traffic in residential, peripheral districts, whether affluent (Les Corts, Sarrià-Sant Gervasi) or working class (Sant Andreu, Nou Barris). This may denote that residents of peripheral districts tend to walk to their local shops, while in the center a large portion of the pedestrian volume is made up of profiles of pedestrians who do not use these kinds of shops, such as tourists.

Finally, hotels, bars and restaurants are positively associated with walking volumes in most of the city, with the exception of the residential upper Diagonal districts of Gràcia and Sarrià-Sant Gervasi, where their impact is less negative.

4. Discussion and conclusions

By utilizing count point data and geospatial regression models, our research uncovers significant spatial heterogeneity in the factors influencing pedestrian behavior in urban environments. Our work contributes to the state of the art in two main ways. Firstly, we identify nuanced spatial patterns in the drivers of urban pedestrian mobility within cities. Secondly, we demonstrate the valuable role of simple automatic pedestrian counting technologies in addressing these issues, highlighting their multiple benefits. In the rest of the section, we elaborate further on our contributions, detailing the significance and implications of our findings within the broader context of urban pedestrian mobility.

4.1. Spatial heterogeneity in urban pedestrian mobility factors

The exploration of pedestrian mobility drivers has garnered substantial attention worldwide, including across-city comparisons (Boakye et al., 2023; Gascon et al., 2019) and case studies like Beijing (Zhao & Wan, 2020), Shanghai (Chen et al., 2022), Melbourne (Pfiester et al., 2021), Paris (Feuillet et al., 2015), Beirut (Sevtsuk et al., 2024), Tel-Aviv (Angel & Plaut, 2024), Blacksburg (Lu et al., 2018) or Barcelona (Graells-Garrido et al., 2021; Marquet & Miralles-Guasch, 2014; Vich et al., 2019). This extensive corpus of research has significantly helped advance the field of urban planning, generating insights to design urban environments that meet the needs of their inhabitants. In general, these studies aim to establish how different factors influence pedestrian activity, frequently reducing these complex interactions to common cause-and-effect relationships in the studied population. This approach might inaccurately ascribe modeling inaccuracies to factors not considered in the analysis or to noise. As emphasized in (Anselin, 2013; Feuillet et al., 2018; Goodchild, 2004), the existence of spatial heterogeneities is a critical consideration that must be addressed.

Our spatially-aware regression analysis has revealed that significant heterogeneities become apparent when examining the entire population of cities (including resident and floating population) as several factors exhibit pronounced spatial variations across the city, from positive to negative impacts. Our spatially-aware models achieved superior explanatory power, with R^2 values ranging from 0.52 to 0.75, compared with global regression models, which reported R^2 values between 0.38 and 0.51. The improved R^2 values, coupled with the Moran's I analysis of the residuals from global models, highlight the importance of accounting for spatial variations in urban studies. This consideration extends beyond merely adjusting the magnitude of influence, as certain factors may even exhibit different direction of impact (sign of the coefficients) in different areas.

The spatial interpretation of the local regression coefficients reveals the emergence of city-wide trends and distinct spatial patterns within the analyzed commercial areas. Pedestrian mobility is found to be driven by different factors in central and peripheral districts, which is intuitively influenced by their activities, resident population, and potentially gentrified areas in the center of the city. In this sense, our results differently align with the state-of-the-art within city regions. Walking is negatively associated with car ownership (Zhao & Wan, 2020), most strongly in residential neighborhoods. Likewise, walking is positively related with public transport (Chen et al., 2022;Lu et al., 2018; Pfiester et al., 2021) and services (Gascon et al., 2019) in the city center. By conducting a direct comparison of our results with those from other studies focused on urban areas near city centers, we are able to pinpoint shared determinants of pedestrian traffic, especially concerning public transportation (Pfiester et al., 2021) and pedestrianized space (Lu et al., 2018). However, we also find differences, such as the role of neighborhood shops in Melbourne, which are especially important for residential neighborhoods in Barcelona, suggesting distinct dynamics in the city centers of these two cities. We also find find several non-conclusive relations with residents' age and gender (where literature point in different directions (Charreire et al., 2021; Graells-Garrido et al., 2021; Marquet & Miralles-Guasch, 2014; Vich et al., 2019)), household income (Lu et al., 2018), or education level (Gascon et al., 2019). Indeed the different results found in the literature linked to pedestrian traffic reinforce the idea of variability in each local community within cities, as our results have shown and reflecting the Anselin's concept of spatial heterogeneity (Goodchild, 2004).

Beyond this general contextualization within the state-of-the-art, our results need to be carefully contrasted to the research conducted by Cheng et al. (Cheng et al., 2021) and Graells-Garrido et al. (Graells-Garrido et al., 2021), who address similar issues at the same urban scale. Both papers focus on population subgroups within a city: older adults (60+ years old) in Nanjing and resident population in Barcelona; both representing a fraction of the total population in the corresponding cities. Barcelona's distinctive combination of residential, commercial, and tourist areas, coupled with its considerable internal and external mobility — characterized by approximately 150,000 tourists per day (Barcelona Regional, 2019) and over 2 million daily commuting trips (Ajuntament de Barcelona, 2019) —, fosters particular pedestrian mobility patterns that can only be studied including a broad cross-section of the population, in our case everyone walking in the streets. This diversity is particularly noteworthy when comparing our results with the study by Graells-Garrido et al. (Graells-Garrido et al., 2021), which was also conducted in Barcelona to understand mobility between neighborhoods in terms of origin-destination pairs without disregarding by transportation mode. Our findings distinctly highlight the variation in influencing factors between touristic and residential zones, offering a contrast to the outcomes presented in (Graells-Garrido et al., 2021), which focused solely on the residential population. Qualitatively, a similar influence attributed to public transport can be inferred, while other factors do not exhibit (at least in our interpretation) comparable patterns. Cheng et al. (Cheng et al., 2021) focusing specifically on the factors affecting walking time for older adults, also find large spatial heterogeneity in the factors they analyze. However, apart from the influence on the results due to the broader cohort we consider, Barcelona is considerably smaller in area (102 km^2) compared to the main city area of Nanjing (estimated to be 450 km^2) and has a much higher population density, with 16,000 inhabitants/km^2, versus Nanjing's 1400 inhabitants/km^2 (Nanjing city). This point is crucial to emphasize, as it allows us to deduce from the current state-of-the-art that heterogeneities may arise within certain spatial, comparing outcomes from studies conducted at smaller scales (e.g., (Pfiester et al., 2021)) with those at broader spatial scales (e.g., (Feuillet et al., 2018; Lu et al., 2012)). Therefore, our findings underscore the essential need to account for these heterogeneities also in more compact cities.

We finally emphasize the importance of recognizing the aggregation of these spatial patterns in the context of the Modifiable Areal Unit Problem (Fotheringham & Rogerson, 2008), to determine whether they constitute distinct spatial clusters or represent a gradual continuum of changes and how they compare among different cities. This understanding is crucial for informing urban planning and policy-making, impacting infrastructure development, transportation planning, and the provisioning of community services, as well as determining areas where behavioral constraints might limit the effectiveness of these initiatives.

4.2. Use of count point data to determine pedestrian mobility factors

While surveys and manual pedestrian counts (Sevtsuk et al., 2024) remain a valuable tool for analyzing pedestrian mobility and attitudes, the study of the complex interactions driving pedestrian movement faces a fundamental, longstanding challenge regarding data acquisition and accessibility. Generally costly, these classical data sources often limit the spatial and temporal resolution of analysis, providing only aggregated results. In a complementary way, the availability and accessibility of extensive data sources (Boeing et al., 2022; Graells-Garrido et al., 2021; Moro et al., 2021) has shifted the paradigm, currently boosting urban science and data-driven urban planning (Bibri, 2019; Bibri & Krogstie, 2020; Kandt & Batty, 2021). Detailed GPS trajectories (Bongiorno et al., 2021), street-view images (Chen et al., 2022), pedestrian sensor count data (Lu et al., 2018; Pfiester et al., 2021), and other digital sources (Yoshimura et al., 2022) are continuously being incorporated in the literature. While each has their pros and cons, these new data sources have already proven valuable to study pedestrian mobility, for example to identify the social and cultural determinants of pedestrian behavior, route choice (Bongiorno et al., 2021; Gallotti, Porter, & Barthelemy, 2016; Quercia, Schifanella, & Aiello, 2014), urban attractors (Hidalgo et al., 2020), etc. Despite their potential, the applicability to study pedestrian mobility with detailed spatial and temporal resolution is still challenging in many cities, due to sparse sensorization and lack of openly accessible data.

The analysis pipeline developed here could be easily replicated in many other cities at low cost, without compromising privacy requirements. In particular, understanding the commonalities, differences and how they relate to the urban planning strategies adopted by cities would be crucial to help establish context-specific guidelines for designing more walkable cities (Claris & Scopelliti, 2016).

4.3. Policy implications

Our results show that count point data may offer relevant spatial and temporal information for data driven targeted urban policies at city scale. Highly spatially resolved insights may enable municipal decision-makers to make more geographically focused interventions.

For instance, the strongly negative influence of vehicle ownership on walking in the upper diagonal part of the city hints at the necessity of addressing car-dependence in these areas. This insight is corroborated by the lower impact of pedestrianized streets and bus stops in these areas, which highlights the opportunity for localized infrastructural interventions. Identifying negative relation of walking with local commerce in combination with positive relation with bars and restaurants, as we see in the touristic center of Barcelona, may be clear signs of gentrification and touristification processes. In these areas, the decoupling of local commerce and walking volumes may signal to the need for preserving vitality and proximity for long-time residents in the context of urban transformations, such as *Superilles* and *Eixos Verds*.

Findings in Barcelona, a Mediterranean city shaped by centuries of history and urban transformations, in some sense, may not be generalized to all cities worldwide. However, the rich spatially-resolved methodology used here may help shed-light on the nuanced, intra-city variations at play in all cities. We hope that the significance of our results shows to the municipalities the potential of sensorizing cities to offer open data for the scientific community to support the transition to sustainable cities.

CRediT authorship contribution statement

Clément Rames: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Daniel Rhoads:** Writing – review & editing. **Antoni Meseguer-Artola:** Writing – review & editing, Supervision, Methodology. **Sergi Lozano:** Writing – review & editing, Conceptualization. **Javier Borge-Holthoefer:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Albert Solé-Ribalta:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

The research presented here is funded by “la Caixa” Foundation and Barcelona City Council throughout the project LIKE-BCN: sociaL accessIble, and walKablE Barcelona (21S09383-001), which is part of the Barcelona Science Plan 2020-2023. Additionally, A.S.-R., A. M-A and J. B-H. acknowledge the support of the Spanish MICINN project PID2021-128966NB-I00 and J.B-H. acknowledges financial support from the Spanish MICINN through the Ramón y Cajal grant RYC2020-030609-I.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cities.2024.105655.

Data availability

The data that has been used is confidential.

References

Ajuntament de Barcelona. (2019). Enquesta de Mobilitat en Dia Feiner 2019. https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/119315/1/EMEF_2019_Informe_Resum_Executiu.pdf [Online; accessed 28-July-2022].

Ajuntament de Barcelona. (2022). *Open Data BCN, barcelona.cat/opendata, [Online; accessed 28-July-2022]*.

Alessandretti, L., & Jørgensen, S. L. (2021). Law of human travel uncovered. *Nature, 593*(7860), 515–516.

Alessandretti, L., Aslak, U., & Lehmann, S. (2020). The scales of human mobility. *Nature, 587*(7834), 402–407.

Angel, A., & Plaut, P. (2024). Tempo-spatial analysis of pedestrian movement in the built environment based on crowdsourced big data. *Cities, 149*, Article 104917.

Anselin, L. (2013). *Spatial econometrics: Methods and models* (Vol. 4). Springer Science & Business Media.

Barcelona Regional. (2019). Redacció de l'Estratègia de mobilitat turística de Barcelona. https://www.bcnregional.com/wp-content/uploads/2019/10/MO02_MOBILITAT_TURISTICA_V1.pdf.

Basu, R., & Sevtsuk, A. (2022). How do street attributes affect willingness-to-walk? City-wide pedestrian route choice analysis using big data from Boston and San Francisco. *Transportation Research Part A: Policy and Practice, 163*, 1–19.

Berrett, B., Leake, G., May, A., & Whelan, J. (1988). Ergonomic standards for pedestrian areas for disabled people: Literature review and consultations. *Technical Report*.

Bibri, S. E. (2019). On the sustainability of smart and smarter cities in the era of big data: An interdisciplinary and transdisciplinary literature review. *Journal of Big Data, 6*(1), 1–64.

Bibri, S. E., & Krogstie, J. (2020). The emerging data–driven smart city and its innovative applied solutions for sustainability: The cases of London and Barcelona. *Energy Informatics, 3*(1), 1–42.

Boakye, K., Bovbjerg, M., Schuna, J., Jr., Branscum, A., Mat-Nasir, N., Bahonar, A., , … Seron, P., et al. (2023). Perceived built environment characteristics associated with walking and cycling across 355 communities in 21 countries. *Cities, 132*, Article 104102.

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., Higgs, C., Liu, S., Giles-Corti, B., Sallis, J. F., Cerin, E., , … Moudon, A. V., et al. (2022). Using open data and open-source software to develop spatial indicators of urban design and transport features for achieving healthy and sustainable cities. *The Lancet Global Health, 10*(6), e907–e918.

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

Bosina, E., & Weidmann, U. (2017). Estimating pedestrian speed using aggregated literature data. *Physica A: Statistical Mechanics and its Applications, 468*, 1–29.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28*(4), 281–298.

Buehler, R., & Pucher, J. (2023). Overview of walking rates, walking safety, and government policies to encourage more and safer walking in europe and north america. *Sustainability, 15*(7), 5719.

Cerin, E., Sallis, J. F., Salvo, D., Hinckson, E., Conway, T. L., Owen, N., , … Moudon, A. V., 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.

Charreire, H., Roda, C., Feuillet, T., Piombini, A., Bardos, H., Rutter, H., … Oppert, J.-M. (2021). Walking, cycling, and public transport for commuting and non-commuting travels across 5 european urban regions: Modal choice correlates and motivations. *Journal of Transport Geography, 96*, Article 103196.

Chen, L., Lu, Y., Ye, Y., Xiao, Y., & Yang, L. (2022). Examining the association between the built environment and pedestrian volume using street view images. *Cities, 127*, Article 103734.

Chen, Y. (2016). Spatial autocorrelation approaches to testing residuals from least squares regression. *PLoS One, 11*(1), Article e0146865.

Cheng, L., Shi, K., De Vos, J., Cao, M., & Witlox, F. (2021). Examining the spatially heterogeneous effects of the built environment on walking among older adults. *Transport Policy, 100*, 21–30.

Claris, S., & Scopelliti, D. (2016). Cities alive: Towards a walking world. *Technical Report*.

Colville-Andersen, M. (2018). *Copenhagenize, Island Press, Washington, DC* (Ch. The Arrogance of Space).

Daniels, R., & Mulley, C. (2013). Explaining walking distance to public transport: The dominance of public transport supply. *Journal of Transport and Land Use, 6*(2), 5–20.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(366a), 427–431.

Dobson, A. J., & Barnett, A. G. (2018). *An introduction to generalized linear models*. Chapman and Hall/CRC.

Dovey, K., & Pafka, E. (2020). What is walkability? The urban dma. *Urban Studies, 57*(1), 93–108.

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

Feuillet, T., Charreire, H., Menai, M., Salze, P., Simon, C., Dugas, J., , … Weber, C., et al. (2015). Spatial heterogeneity of the relationships between environmental characteristics and active commuting: Towards a locally varying social ecological model. *International Journal of Health Geographics, 14*(1), 1–14.

Feuillet, T., Commenges, H., Menai, M., Salze, P., Perchoux, C., Reuillon, R., , … Hercberg, S., et al. (2018). A massive geographically weighted regression model of walking-environment relationships. *Journal of Transport Geography, 68*, 118–129.

Fotheringham, A. S., & Rogerson, P. A. (2008). The SAGE handbook of spatial analysis. *Sage*.

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers, 107*(6), 1247–1265.

Fotheringham, C., & Brunsdon, M. (2002). *Charlton*. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, Wiley.

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.

Gascon, M., Götschi, T., de Nazelle, A., Gracia, E., Ambròs, A., Márquez, S., , … Iacorossi, F., et al. (2019). Correlates of walking for travel in seven european cities: The pasta project. *Environmental Health Perspectives, 127*(9), Article 097003.

Gómez-Varo, I., Delclòs-Alió, X., & Miralles-Guasch, C. (2022). Jane jacobs reloaded: A contemporary operationalization of urban vitality in a district in Barcelona. *Cities, 123*, Article 103565.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A.-L. (2008). *Understanding individual human mobility patterns, nature, 453*(7196), 779–782.
Goodchild, M. F. (2004). Giscience, geography, form, and process. *Annals of the Association of American Geographers, 94*(4), 709–714.
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.
Graells-Garrido, E., Serra-Burriel, F., Rowe, F., Cucchietti, F. M., & Reyes, P. (2021). A city of cities: Measuring how 15-minutes urban accessibility shapes human mobility in Barcelona. *PLoS One, 16*(5), Article e0250080.
Hair, J., Black, W., Babin, B., & Anderson, R. (2009). *Multivariate data analysis 7th edition Pearson prentice hall*.
Harms, I. M., van Dijken, J. H., Brookhuis, K. A., & De Waard, D. (2019). Walking without awareness. *Frontiers in Psychology, 10*, 1846.
Hatamzadeh, Y., Habibian, M., & Khodaii, A. (2014). Walking behaviors by trip purposes. *Transportation Research Record, 2464*(1), 118–125.
Henderson, L. (1971). The statistics of crowd fluids. *Nature, 229*(5284), 381–383.
Hidalgo, C. A., Castañer, E., & Sevtsuk, A. (2020). The amenity mix of urban neighborhoods. *Habitat International, 106*, Article 102205.
Kandt, J., & Batty, M. (2021). Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities, 109*, Article 102992.
Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association, 107*(500), 1590–1598.
Lu, G. X., Sullivan, J., & Troy, A. (2012). Impact of ambient built-environment attributes on sustainable travel modes: A spatial analysis in chittenden county, Vermont. In *, Vol. 21. 91st annual meeting of the Transportation Research Board. Washington, DC*.
Lu, T., Mondschein, A., Buehler, R., & Hankey, S. (2018). Adding temporal information to direct-demand models: Hourly estimation of bicycle and pedestrian traffic in Blacksburg, va. *Transportation Research Part D: Transport and Environment, 63*, 244–260.
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.
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.
Munira, S., & Sener, I. N. (2020). A geographically weighted regression model to examine the spatial variation of the socioeconomic and land-use factors associated with strava bike activity in Austin, Texas. *Journal of Transport Geography, 88*, Article 102865.
Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 135*(3), 370–384.
Nilforoshan, H., Looi, W., Pierson, E., Villanueva, B., Fishman, N., Chen, Y., Sholar, J., Redbird, B., Grusky, D., & Leskovec, J. (2023). Human mobility networks reveal increased segregation in large cities. *Nature*, 1–7.
Observatori del Turisme a Barcelona. (2020). Barcelona tourism activity report 2019. https://www.observatoriturisme.barcelona/sites/default/files/IAT19.pdf [Online; accessed 05-May-2023].
Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). Mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information, 8*(6), 269.
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.
Perchoux, C., Brondeel, R., Wasfi, R., Klein, O., Caruso, G., Vallée, J., , … Chaix, B., et al. (2019). Walking, trip purpose, and exposure to multiple environments: A case study of older adults in Luxembourg. *Journal of Transport & Health, 13*, 170–184.
Pfiester, L. M., Thompson, R. G., & Zhang, L. (2021). Spatiotemporal exploration of Melbourne pedestrian demand. *Journal of Transport Geography, 95*, Article 103151.
Que, X., Ma, X., Ma, C., & Chen, Q. (2020). A spatiotemporal weighted regression model (stwr v1. 0) for analyzing local nonstationarity in space and time. *Geoscientific Model Development, 13*(12), 6149–6164.
Quercia, D., Schifanella, R., & Aiello, L. M. (2014). *The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city, in: Proceedings of the 25th ACM conference on Hypertext and social media, ACM* (pp. 116–125).
Rames, C., Wilson, A., Zimny-Schmitt, D., Neri, C., Sperling, J., & Romero-Lankao, P. (2021). A data-driven mobility–energy typology framework for New York state. *Environment and Planning B: Urban Analytics and City Science, 48*(8), 2254–2271.
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.
Rhoads, D., Rames, C., Solé-Ribalta, A., González, M. C., Szell, M., & Borge-Holthoefer, J. (2023). Sidewalk networks: Review and outlook. *Computers, Environment and Urban Systems, 106*, Article 102031.
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.
Sevtsuk, A., Kollar, J., Pratama, D., Basu, R., Haddad, J., Alhassan, A., … Abou-Zeid, M. (2024). Pedestrian-oriented development in Beirut: A framework for estimating urban design impacts on pedestrian flows through modeling, participatory design, and scenario analysis. *Cities, 149*, Article 104927.
Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature, 484*(7392), 96–100.
TC Group Solutions. (2022). https://www.tcgroupsolutions.com/ [Online; accessed 28-July-2022].
Thompson, M. L. (1978). Selection of variables in multiple regression: Part i. a review and evaluation. *International Statistical Review/Revue Internationale de Statistique*, 1–19.
Triguero-Mas, M., Anguelovski, I., Connolly, J. J., Martin, N., Matheney, A., Cole, H. V., , … Argüelles, L., et al. (2022). Exploring green gentrification in 28 global north cities: The role of urban parks and other types of greenspaces. *Environmental Research Letters, 17*(10), Article 104035.
Van Dyck, D., Cardon, G., Deforche, B., & De Bourdeaudhuij, I. (2011). Do adults like living in high-walkable neighborhoods? Associations of walkability parameters with neighborhood satisfaction and possible mediators. *Health & Place, 17*(4), 971–977.
Vich, G., Marquet, O., & Miralles-Guasch, C. (2019). Green streetscape and walking: Exploring active mobility patterns in dense and compact cities. *Journal of Transport & Health, 12*, 50–59.
Yang, Y., & Diez-Roux, A. V. (2012). Walking distance by trip purpose and population subgroups. *American Journal of Preventive Medicine, 43*(1), 11–19.
Yoshimura, Y., Kumakoshi, Y., Fan, Y., Milardo, S., Koizumi, H., Santi, P., … Ratti, C. (2022). Street pedestrianization in urban districts: Economic impacts in spanish cities. *Cities, 120*, Article 103468.
Zhang, X., Melbourne, S., Sarkar, C., Chiaradia, A., & Webster, C. (2020). Effects of green space on walking: Does size, shape and density matter? *Urban Studies, 57*(16), 3402–3420.
Zhao, P., & Wan, J. (2020). Examining the effects of neighbourhood design on walking in growing megacity. *Transportation Research Part D: Transport and Environment, 86*, Article 102417.