

Stepwise multiple linear regression applied to study the influence of pedestrian volumes according to environmental characteristics

Juan Carlos Gómez-Sánchez, María Victoria de-la-Fuente-Aragón,
Lorenzo Ros-McDonnell

Recibido: 12 de Enero de 2021
Aceptado: 15 de Marzo de 2021

<https://doi.org/10.37610/dyo.v0i75.609>

Abstract

Since the start of the 1990s, the evolution of socio-economic, local and regional factors in western society has placed pedestrians at the focal point of urban mobility plans until the present-day. This comes over in the many research works conducted in this area, which evidence the need to lay the foundations of the relation between urban planning and people's health. Despite many methodologies being proposed to analyse this relation, those that strike a balance between the pedestrian volume in an area and its environmental characteristics provide the best results, but start with multivariable models that are not easy to calculate. This article aims to collect and reduce those variables that significantly affect population movements depending on the study area's commercial level. Of all the different existing mathematical models, stepwise linear regression was chosen to do this. Finally, the present work concludes that connectivity and urban furnishing in prime areas, the presence of bars and offices in secondary areas and connectivity, residential density and mixed-use zones in tertiary areas are key factors that should be considered when planning the studied areas in the future.

Keywords

Multiple linear regression analysis, spatial data, commercial characteristics, pedestrian volumes

1. Introduction

In the 1960s, very few authors believed that walkability was a forgotten means of transport because all studies then centred on analysing the relation between accessibility to cities and vehicle use (Handy, 2002). The first reference made to the importance of pedestrians when developing cities appears in an article by Jacobs (1961). This author believed that cities' vitality depended on population density, mixed land uses, green areas, connectivity and presence of buildings.

Nevertheless, not much attention was paid to these pieces of evidence until the 1990s when the New Urbanist movement came into being (Calthorpe, 1993). This current managed to change the city planning concept it placed pedestrians at the centre of attention in new urban plans. Today's need to reduce pollution in cities means having to maintain the research that commenced in the 1960s, but by placing an emphasis on seeking relations between the urban environment and pedestrian mobility (Sung, 2015). The number of variables influencing the behaviour of those who walk around streets is high and hard to predict, and also presents a varied typology. External regional variables can, however, significantly affect those people who co-exist in an area, which makes the task of detecting those significant variables even more difficult. Consequently, existing relations with the urban environment are better explained by emphasising from the local point of view (Feuillet, 2016).

Since 2003, the tendency followed to establish correlations between urban variables consists in using GIS-based spatial models. This technology allows a large quantity of data to be handled that can be used in urban planning, and to predict how people behave (Anselin, 2006) with a time series. However, not all variables can be employed in these models because collinearity exists in some variables (Feuillet, 2016). Therefore, multiple linear regression and principal component analyses are carried out to reduce significant variables (Fang, 2019). In these studies, the variance inflation factor is used as an indicator of collinearity among variables, and any over a high value are ruled out.

✉ María Victoria de la Fuente Aragón *
marivi.fuente@upct.es

 ORCID: 0000-0002-5283-8003

Juan Carlos Gómez Sánchez *
j.carlosgomez@upct.es

 ORCID: 0000-0002-2828-483X

Lorenzo Brian Ros McDonnell *
lorenzo.ros@upct.es

 ORCID: 0000-0002-1736-7525

* Grupo Investigación "Ingeniería de Organización", ETSII.
Universidad Politécnica de Cartagena

The literature includes only a few studies that relate the influence between streets' commercial level and the number of pedestrians walking around them. A common fact appears in them all: the more pedestrians, the more sales (Kumar, 2006). So it is necessary to look into this study in-depth to establish which variables significantly affect the activity of commercial areas in cities.

2. Determining the variables that affect pedestrian mobility

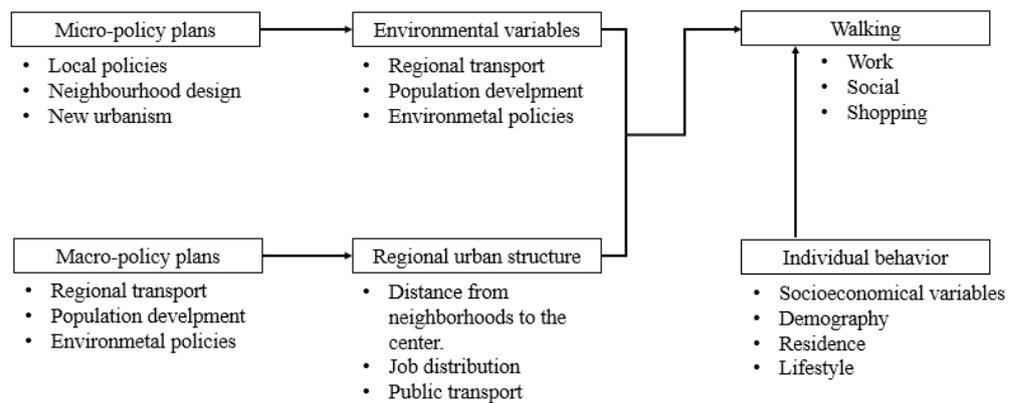
Cervero (1997) introduced the 3Ds concept (diversity, density and demography) to explain people's intentions when taking certain routes to walk or drive along. This author states that these variables are not easy to measure because they depend on the space typology, urban patterns, the interconnection of communication networks, intersections among streets, etc., and had to introduce new variables like accessibility and distance to the final trip distance. The first three variables can be locally characterised, but not the last two, which are regionally characterised. Giles-Corti (2002)

reasserted this idea, but added that regional parameters lead to instability, while Cho (2015) specified that population distribution depends on where job posts are located. If we bear in mind the reference bibliography, all these variables can be divided into three main groups:

1. Regional urban environment variables: connectivity with means of transport, closeness to other urban centres and the employment rate stand out (Cervero, 1997).
2. Local urban environment variables: including commercial and residential density, land uses, street connectivity and obstacles, width and length, among others (Cerin, 2013).
3. Individual behaviour: the socio-economic level, demographic factors, culture, etc. (Frank, 2006).

Figure 1 shows the relation that Cho (2015) proposes among these groups:

Figure 1 Relation between local and regional variables.
 Source: Cho (2015)



2.1. Regional urban environment variables

Authors like Oakes (2007) analysed populational movements in accordance with the size of population nuclei and realised that, if a given area was very dense, pedestrian volumes moving towards would increase. The same idea had been previously stated by Saelens (2003), who believed that it was much simpler to observe people's movements in denser areas than places further away from city centres.

A city centre's activity depends on its connection with other nearby centres, and their good connectivity is a key factor for their development (Cervero, 1997). Cho (2014) analysed which regional variables most affected the way people from certain cities behaved, and noticed that distance to public transport stations (train, bus or boat), distance to job posts and distance to public transport stops were relevant. Handy (1993) developed the concept of regional accessibility to city centres (by emphasising job posts) versus local accessibility.

Understanding the relation between regional and local variables is fundamental because they are closely related (Naess, 2005). From the macroscale point of view, jointly analysing the location of districts in relation to city centres, regional employment distribution, road connections, regional transport plans and environmental policies with socio-economic factors, such as age, gender, academic training, level of income, occupational activity, is an extremely complex task (Cho, 2015), which was why those that can be measured locally were previously analysed and collected: distance to institutions, number of bus lines in the area, distance to green areas, intersections, employment rate, number of inhabitants, distance to the nearest bus station and number of bus stops. Handy (1993) also stated that providing certain areas of a region with accessibility will involve more journeys being made, and not only going back and to from work. Therefore, the authorities should encourage the aforementioned variables.

2.2 Local urban environment variables and their relation to intention to walk

The easiness of taking measurements locally has encouraged many authors to analyse the relation linking environmental characteristics and pedestrian volumes. Nonetheless, the complexity of relating local variables with people's behaviour has led existing research lines to jointly deal with them. A close relation exists between local variables and density level, existing transport and intention to walk (Oakes, 2007). On the first of these, Saelens (2003) analysed the relation between pedestrian volumes in accordance with residential density, mixed land uses, street connectivity, walking facilities, aesthetic aspects, safety to move around and crime, and observed that some of these variables are subjective. This is why he emphasises socio-economically characterising the population according to its age, gender, ethnic group, cultural level, etc.

Until 2009, works measured pedestrian volumes according to building height, street length and width, footpaths and their dimensions, and street typology (Ewing, 2009). However, it is not easy knowing people's intentions, and especially how they perceive the environment. Handy (1996) and Ewing (2001) were responsible for the first references, whose works included architecture, parking and visual perceptions, but they quantified qualitative perceptions based on a panel of experts. Imageability, visual enclosure, human scale, transparency, complexity, legibility, linkage and coherence are specified as the eight dimensions defining human perceptions when walking, of which the first five are fundamental. Cerin (2013) analysed the way people perceived environmental variables when walking along a certain street. Their results indicated that the following were the most important: presence of people, facilities, urban furnishing, slope, commercial density and residential density. The variables to be quantified in the five above-cited dimensions were specified by Ewing (2013) as follows:

- Imageability: pedestrian volume, presence of historic buildings, green areas, restaurant terraces, buildings without corners, noise level, views, architecture
- Visual enclosure: left footpath occupancy, right footpath occupancy, view ahead, sky view, signals/lighting furnishing
- Human scale: urban furnishing present, ground-floor premises with windows, height of buildings, trees present and architecture
- Transparency: once again ground-floor premises with windows, level of abandonment, level of occupied walls on the same side where pedestrians walk
- Complexity: building density, colour of buildings, public art, restaurant terraces present

Street design and presence of buildings in an urban area are fundamental factors that influence people's intention to walk down a given street (Handy, 2008). Yet these perceptions imply considerable variability in the models that are not easy to explain (Ewing, 2013). Lee (2017) state that people's behaviours and, therefore, pedestrian volumes, differ in residential and commercial areas, while urban elements most strongly influence commercial activity. Revitalising streets involves favouring populational density, mixed land uses and maintaining historic buildings (Park, 2012).

In short, by taking into account many local research works, we conclude that commercial density, employment rate, maintaining urban furnishings and building density (Jacobs, 1961; Salom, 2012; Alho, 2015), along with improvements to mobility and accessibility (Ewing, 2010), positively influence the pedestrian volumes moving around a street. Although few regional studies are found because measuring this type of variables involves using large databases, they all justify that the analysis scale strongly impacts the results, which means that regional variables must also be considered in the results (Giles-Corti, 2002). Among these variables, employment rate, which is indirectly measured with commercial activity, distance to public transport and distance to train stations (Boarnet, 2008) are the main variables to bear in mind.

Finally, it is worth stressing that most of the literature has focused on studying people's individual behaviour and how they perceive the environment, which means that the results of local studies are well-verified. We conclude that intention to walk along one street or another depends, according to Reisi (2019), on: availability of crossings and safety, street width, accessibility, cultural aspects, variety of businesses, urban furnishings and public transport. These variables relate to those considered by regional studies.

3. Measuring spatial data in urban environments

In order to contribute specific measures to allow the efficient planning of land uses in cities, the research team proposes a methodology that allows the minimum number of variables that significantly affect pedestrian volumes in a given area to be obtained. To this end, a multivariable spatial regression model was considered (Anselin, 2006) to combine the calculation methodologies put forward by Song (2007), Lee (2017) and Fang (2019), and the variables indicated in points 2 to 4 in the articles by Jacobs (1961), Cervero (1997), Frank (2006), Cerin (2013), Ewing (2013), Sung (2015) and Hamidi (2019).

People's intention to walk is considered to depend on the commercial street typology, urban characteristics, and social, morphological and economic factors (Salom, 2012), of which the first is interesting to study its influence on pedestrian

volumes (Gómez, 2019; Palomares, 2017). It favoured the model proposed by our research team to distinguish among the different existing commercial areas.

The selected zone that was included in the study area contains more than 14,000 cadastral references, over 750 commercial premises and a mean population density of 384 inhab/km². It also has plenty of economic activity that leads to pedestrian volumes coming from outside the city inwardly, and vice versa. Hence such a large quantity of data cannot be arbitrarily processed, which was why a methodology that works with spatial data was proposed (Lee, 2017).

3.1. Characterisation of commercial zones. The model's hypothesis

If we consider the articles of Palomares (2017) and Sastre (2013), different commercial areas can be classified into prime, secondary or tertiary commercial areas. Therefore, three groups were defined to include: existing commercial diversity; administration activities and services rendering; the appeal of the area. The first considered commercial areas are prime areas with a high pedestrian volume, many varied commercial premises, and offices that attend to all types of public. Their drawbacks tend to be high noise levels, constant congested traffic at peak times and high pollution rates, which result from the former. The premises they cover are: high diversity of shops covering maximum requirements, marked attraction with a large population or a tourist attraction, administration services rendering, children's playgrounds, public transport stations and areas with leisure activities or events. These areas are those that public administrations ultimately pedestrianise (Kumar, 2006).

Secondary areas are located near prime areas, but are less frequented. So shops are not as varied and there is less demand, so they tend to have lower employment rates. The restrictions in such areas are the same as those in prime areas. Finally, tertiary areas lie further away from prime areas, are not frequented so much and tend to be consolidated residential places. These places offer no commercial variety and come with these premises: not very appealing, not a

wide range of services and low pedestrian volumes.

By bearing all this in mind, the research team proposed answering these hypotheses:

1. Which independent variables are to be considered in the model?
2. Which variables significantly affect prime, secondary and tertiary areas?
3. Do these variables exactly explain the model?

3.2. Study area delimitation

The city of Cartagena lies in southeast Spain, 48 km from the city of Murcia and has one of the best ports of the Mediterranean Sea. On 1 January 2020, it had 216,920 registered inhabitants according to its City Council, and it covers 558 km². The city's main activities include those related to leisure, commerce, administration and businesses.

Given its maritime port, Cartagena is an eminently tourist city and its whole urban area is built around this port. In fact the city's historic centre is very close to the port, where its main tourist monuments stand. Consequently, the main administration buildings and finance offices are found there. However, the city suburbs are less urbanistic and the population density is much lower. This was why the city centre was chosen as the study area (see Figure 2).

Although the work team initially proposed 100 points corresponding to one measurement per street, they were increased to 130 after analysing the results. Of these, 30 corresponded to prime areas, 45 to secondary areas and 63 to tertiary areas. To help to easily locate these measurement points, in Figure 2 red, orange and green denote prime, secondary and tertiary areas, respectively. In the last two area types, only one measurement per street was made. In prime areas, the number of measurements was increased from 11 to 30 (every 300 m) because of the high collinearity among variables. This is explained below in more detail.

Figure 2 Study area with the selected measuring points. Source: the authors



3.3 Multiple linear regression analysis. Data collection

In order to analyse the relation among the variables, a multiple linear regression model was proposed in which the dependent and independent variables had to be firstly identified. To select the independent variables, the articles by Cervero (1997), Cerin (2013), Frank (2006), Jacobs (1961) and Sung (2015) were taken into account, which offer very different results and many variables. Nonetheless, collinearity exists among many of them, which means having to perform a VIF analysis (Lee, 2017) to rule out those variables with an index over 10.

It was necessary to introduce the commercial sector type into the independent variables to know the influence that pedestrian volume had on them. To do so, we used the shop classification contemplated by Muñozuri (2016) as follows: Sector 1 with chemists and perfume shops; Sector 2 with technical services; Sector 3 with offices and restaurants; Sector 4 with hotels, clothes shops, jewellers and warehouses; Sector 5 with food.

The difficulty with this analysis lies in data collection because it depends on the type of variable considered (qualitative or quantitative). This means having to measure pedestrian volumes and street characteristics manually from

Monday to Friday between 8 am and 6 pm. Nonetheless, commercial density, the connectivity group variables and street characteristics require following other reliable measuring techniques that indicate up-to-date data. So we employed GIS-based spatial data. Bearing this in mind, the work team decided to use the Goolzoom website with access to the Ministry of the Exchequer database. By considering the defined study area, more than 14,000 urban references were obtained from this application. They provide many urban data, including those herein defined. All these data were grouped into streets to reduce the number of independent variables. The urban references of those streets with more than one measuring point were divided into blocks of numbers and were averaged to obtain the measuring points considered in the mathematical model. Table 1 includes all the variables and those in bold were valid for this analysis.

We can see that not all the variables were quantitative (crossings, urban furnishing, slopes, footpaths), so it was necessary to use Likert scales to measure them. We opted to conduct street surveys at all the measuring points. If there were no pedestrians at some measuring points, research team members completed surveys themselves. Scores from 1 to 4 were given to each variable, but scales from 1 to 5 could have been used. Part of these qualitative data was taken from former works conducted in Cartagena (Ros, 2016).

Table 1 The considered independent variables

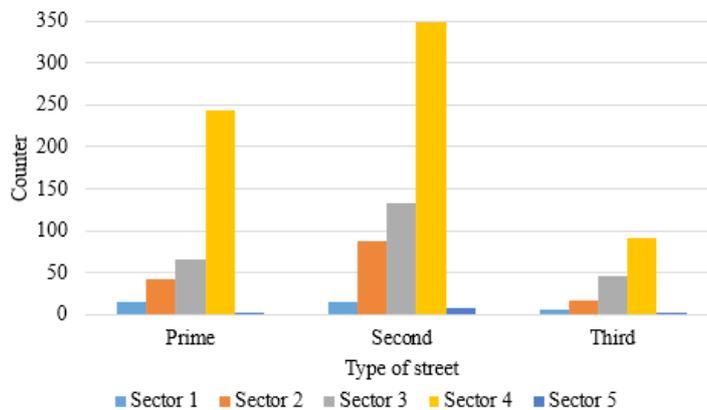
Socio-economic	Density	Street design	Connectivity	Construction
No. of residential buildings	Closed businesses	Length	Distance to institutions	No. buildings
Inhabitants	Sector 1	Width	No. bus lines	No. floors
Employment	Sector 2	Obstacles	Distance to green areas	Mean area
Pedestrian volume	Sector 3	Crossings	Distance to public transport	
	Sector 4	Urban furnishing	Distance to bus station	
	Sector 5	Slope	No. bus stops	
		Footpaths	Intersections	

The other quantitative variables were obtained from the Goolzoom website, as previously mentioned. We chose this tool because it provides lots of data simply that can be exported to any spreadsheet in a short time. Goolzoom was selected in the study area, and measurement points were placed by downloading a series of numerical data that had to be filtered and adapted to the model. The downloaded database distinguished between urban and commercial cadastral references or, within them, the commerce typology. They have related variables like surface area, number of floors, age, number of dwellings, registered inhabitants,

economic level, level of education, etc.

The variables length, slopes, width, distance to green areas or public means of transport were manually applied, as were pedestrian volume, number of closed businesses and the classification of the commerces from Sectors 1 to 5 (Gómez, 2019). Finally, the employment rate was obtained with the data obtained from the Cartagena City Council, which were indicated per street. Figure 3 illustrates the measurements taken of premises according to street type.

Figure 3 Measurement of businesses per sector. Source: the authors



Of all the valid variables, pedestrian volume was selected as the dependent variable, just as most reference articles confirmed (Lee, 2017; Cho, 2015; Pikora, 2003; Cerin, 2013; Feuillet, 2016), and also because its measurement is easy in practice. Before running the regression analysis, it was necessary to check that all the model's variables met the linearity, normality, homogeneity, homoscedasticity and independence hypotheses. We had to perform a descriptive analysis with them all to decide if it was necessary to apply numerical transformation or not and to reduce asymmetry. The dependent variable, number of buildings and number of dwellings required logarithmic transformation.

as pedestrians' behaviour is hard to predict. Feuillet (2016) considered that the effects between the environmental and pedestrian volume variables can be better seen on the local scale as well-established behaviour patterns exist locally. So in the literature, a tendency comes over to use multivariable spatial models where local variables predominate (Anselin, 2006). Stepwise (Lee, 2017) or principle component analysis methods (Song, 2007) are employed to reduce the number of variables. Our research team opted for the stepwise method.

The three stepwise linear regression methods include:

- Ascending method: it groups those variables with the lowest p-value. Once aggregated, the variable is never eliminated. The process is repeated until the correlation coefficient improves or remains constant

The main difficulty in selecting a mathematical model is because a vast amount of data must be processed, and correlating subjective and objective variables is not easy

- Descending method: this method is followed in the opposite way to the previous method. It begins with all the variables and those with the highest p-value are eliminated. No variable can be re-introduced into the model once eliminated. The process is repeated until the correlation coefficient improves or remains constant
- Ascending-descending method: it combines both methods, and introduces into or eliminates variables from the model. To do so, how the model's correlation coefficient varies is analysed. The Akaike information criterion values related to the model's exactness are also analysed.

By considering all this, the selected multivariable regression model was the following (Anselín, 2016; Lee, 2017):

$$\text{Log (pedestrian volume)} = \beta_0 + \beta_1 (\text{socio-economic}) + \beta_2 (\text{density}) + \beta_3 (\text{street design}) + \beta_4 (\text{connectivity}) + \beta_5 (\text{building}) \quad (1)$$

4. Results

The proposed methodology gives very different results depending on the hour of day in which the pedestrian flow measurement is made, so the measurements have been made in those time periods where greater pedestrian traffic has been observed. Also, we can remark that in previous studies (Ros et al., 2016) the main activity occurs in the mornings.

To measure those variables that cannot be traced by the GoolZoom app (pedestrian flow, number of bus stops, number of closed shops, crossings, street furniture, slopes, distance to institutions, distance to green areas and distance to public transport) a mobile app based on Android has been used and has been used by the research team in previous studies called ContaThings (Gómez et al, 2020). This simple application has been customized for new measurement requirements and after measurements have been made the results have been exported to a csv file.

Having applied the proposed methodology, the results obtained after running a stepwise linear regression method are provided in Table 2 (Lee, 2017), as well as the R² value for the model's fit. To show the model's exactness, the Snedecor F-value, their significance, the Akaike information criterion and correlation coefficient values were collected as some authors point out (Song, 2007). So, the research team has chosen to use these coefficients in order to explain the model's significance.

Table 2 Results

Variable	β (prime)	Sig.	B (secondary)	Sig.	B (tertiary)	Sig.
Constant	197.004	0.011	3.168	0.001	2.080	0.025
Mean 7nal						
No. Bus stops						
No. Closed businesses			-0.015	0.053		
Intersections			-0.320	0.028		
Pedestrian crossings	-2.390	0.226				
Obstacles	-76.898	0.018	0.470	0.020		
Distance to institutions			0.012	0.004		
No. Bus lines	197.004	0.011				
LN number of buildings					2.735	0.035
LN number of residential buildings					3.312	0.044
Length	-0.027	0.018				
Number of floors	0.667	0.019			0.543	0.027
Distance to green areas	2.415	0.010	0.013	0.003	0.076	0.018
Distance to public transport	1.238	0.006	0.009	0.003		
Sector 1						
Sector 2						
Sector 3						
Sector 4			0.047	0.030		
Sector 5			0.841	0.008		
Distance to bus station						
Urban furnishings	119.585	0.014	-0.447	0.022		
Width			0.075	0.000	0.300	0.024
N	11		45		63	
R ²	0.677		0.529		0.683	
Akaike's information criterion	25.989		39.547		61.807	
ANOVA	F = 3.09	0.042	F = 3.814	0.037	F = 2.181	0.016

5. Analysing the regression model results

It is essential to reduce those variables that do not significantly influence in pedestrian flows, as reflected in the reference literature. Having focused on the quantitative variables' analysis, the high level of results' redundance determine that the models obtained are imprecise and very difficult to implement (Anselín, 2006; Song, 2007). The number of variables is high in all the articles analyzed, so the interpretation of the correlation coefficients gives results that are difficult to transmit to the administrations and, consequently, urban modifications are delayed. The proposed methodology allows obtaining the minimum number of variables that affect urban flows in a significant manner. And in a simple way and following a work procedure that is easy to reproduce by any technical team responsible in charge of urban planning of the city.

On the contrary, numerous authors have studied, from a qualitative point of view, those variables that encourage walking, gathering the people' needs and requirements. The qualitative models obtained are very difficult to implement due to the numerous dependent variables among them (Saelens, 2003). In no qualitative and quantitative analysis mentioned we can't find any reference to the commercial level of the streets, and the importance of these in the intention of walking. The methodology proposed by the research team deepens the developed research carried out at present by establishing a simple work procedure in order to obtain those variables that significantly influence pedestrian flows depending on the commercial level of the streets (prime, secondary and tertiary).

According to the obtained results, the stepwise linear regression model reduced the number of significant variables in all cases, which justified the redundance of the variables discussed earlier (Song, 2007). Of the three linear models proposed in all the commercial areas, the reliability of the results was, at least, 52.9%, and the tertiary area gave the best fit ($R^2 = 68.3\%$). Conversely, the secondary area obtained a worse fit because of the variability of the existing data. By analysing each model's degree of significance, that applied to the tertiary area gave the most reliable results because the associated p-value was the lowest of the three models. However, the prime area was the least reliable because its p-value came close to 0.05.

By individually analyzing the variables in each model, the relation between commercial and pedestrian activity was non-significant, unlike distance to public transport, distance to green areas, urban furnishings and number of bus lines/bus stops in the area because of their high degree of correlation, which confirms what Kumar investigated (2006). It was noteworthy how presence of obstacles and crossings negatively affected the model. Thus we emphasise modifying the linear regression model's parameters,

increasing the sample or using other models like a principle component analysis in prime areas (Song, 2007).

Our work team realised that when a single measurement point was taken in prime areas, it gave models whose significance exceeded 0.05 and, therefore, the model did not explain the dependent variable. However, the proposed methodology allowed data to be grouped into intervals of 100 m, 75 m and 50 m which, in turn, allowed the number of measurement points to increase. After several attempts, we reached the conclusion that a minimum 50-metre distance made the prime streets model explanatory.

The data variability in the secondary area was wide and, therefore, the model considered more variables. However, all the models' degrees of significance came close to 0, which means a strong influence on the dependent variable. We stress that the results revealed that urban furnishing, many crossroads and the presence of closed businesses were variables that negatively impacted the number of pedestrians on streets (Feuillet, 2016). Conversely, the sector with restaurants and offices, and administration buildings, were predominating factors for these streets' vitality.

Finally, the most marked reduction in variables took place in the tertiary area, where five variables were considered. These results confirm those reported by Cho (2015), who concluded that a positive relation appeared among residential density, transport and pedestrian volume. Thus improving public transport would be ideal to encourage pedestrian volume. Street dimensions are also important because areas with parking, green areas and services also positively influence pedestrians' intentions to walk along tertiary streets (Pikora, 2003). For the commercial level, as no significant relation was found with the dependent variable, we can state that these areas cannot guarantee some businesses being successful.

6. Conclusion

For more than 20 years, the centre of attention of urban planning has been pedestrians because of the rapid evolution of socio-economic, local and regional factors in western society. The literature has focused on analysing the existing relation between pedestrian volume and environmental variables from the objective point of view, but has neglected the subjective perspective. Since 2010, researchers have centred on explaining the subjective reasons that lead someone to have a given perception of the area they move through. If more people move along a given street, it will have a knock-on effect that will favour more people coming, which will imply greater commercial activity. Although many variables have been considered in the literature, it is necessary to reduce variables to leave only the most significant ones by multivariable or stepwise methods, or principal component analyses, to meet certain goals and to obtain feasible solutions.

This article aimed to collect the main variables that significantly affect population movements and to analyse which ones need to be promoted according to the study area's commercial level. The difficulty of this analysis lies in data collection because the considered variables are both qualitative and quantitative. This means that some involve conducting surveys, while others are manually applied. Nonetheless, commercial density, connectivity group variables and street characteristics require other reliable measurement techniques that indicate up-to-date data, which is why GIS-based spatial data were employed. We herein opted to employ Goolzoom with access to the Ministry of the Exchequer database. When considering the defined study area, more than 14,000 urban references with many urban data were obtained, including those herein defined.

To make the subsequent data analysis easier, all the data were grouped into streets to work with fewer independent variables. If streets had more than one measurement point, urban references were divided into blocks of numbers and were averaged to obtain the measurement points contemplated in the mathematical model. After generating the mathematical model's generic expression, stepwise multiple linear regression was applied to not only obtain this model's coefficients, but also its individual and overall significance values.

According to our results, we can state that the stepwise linear regression model reduced the number of significant variables in all cases. Of the three linear models proposed in each commercial area, that obtained for the tertiary area offered the best fit ($R^2 = 68.3\%$). Conversely, the same cannot be stated of the secondary area given the variability of the existing data. Calculation problems arose in the prime area because, if only one measurement point was taken in prime areas, the model gave a significant result above 0.05 and did not explain the dependent variable. However, the proposed methodology allowed data to be grouped into intervals of 100 m, 75 m and 50 m which, in turn, allowed the number of measuring points to be increased. After a few attempts, we concluded that a minimum 50-metre distance made the prime streets model explanatory. Finally, we concluded that connectivity and urban furnishings in prime areas, the presence of bars and offices in secondary areas, and connectivity, residential density and mixed land uses in tertiary areas are key factors that should be taken into account when planning the studied areas in the future.

7. References

- Alho, A., De-Abreu, J. (2015). Utilizing urban form characteristics in urban logistics analysis: a case study in Lisbon, Portugal. *Journal of Transport Geography*, 42, 57-71.
- Anselin, L. (2006). *Spatial Regression*. University of Illinois Press. Nueva York.
- Boarnet, M.G., Greenwald, M., McMillan, T.E. (2008). Walking, Urban Design, and Health Toward a Cost-benefit. Analysis Framework. *Journal of Planning Education and Research*, 27, 341-358.
- Calthorpe, P. (1993), *The Next American Metropolis: Ecology, Community, and the American Dream*. Princeton Architectural Press.
- Cerin, E., Macfarlane, D., Sit, C.H. (2013). Effects of Built Environment on Walking among Hong Kong Older Adults. *Hong Kong Medical Journal*, 19(4), 39-41.
- Cervero, R., Kockelman, K. (1997). Travel Demand and the 3Ds: Density, Diversity and Design. *Transportation Research D*, 2(3), 199-219.
- Cho, G.H., Rodríguez, D.A. (2014). Location or design? Associations between neighborhood location, built environment and walking. *Urban Studies*, 1 (20).
- Cho, G.H., Rodríguez, D.A. (2015). Neighborhood design, neighborhood location, and three types of walking: results from the Washington DC area. *Environment and Planning*, 42.
- Ewing, R., Greenwald, M.J., Zhang, M. (2009). *Measuring the Impact of Urban Form and Transit Access on Mixed Use Site Trip Generation Rates - Portland Pilot Study*. Washington, DC: U.S. Environmental Protection Agency.
- Ewing, R., Cervero, R. (2001). Travel and the built environment: A synthesis. *J. Transp. Res. Board*, 1780, 87-114.
- Ewing, R., Cervero, R. (2010). Travel and the built environment: a meta-analysis. *Journal of the American Planning Association*, 76, 265-294.
- Ewing, R., Clemente, O. (2013). *Measuring Urban Design: Metrics for Livable Places*; Island Press: Washington, DC, USA.
- Fang, K., Wang, X., Chen, L., Zhang, Z., Furuya, N. (2019). Research on the correlation between pedestrian density and street spatial characteristics of commercial blocks in downtown area: a case study on Shanghai Tianzifang. *Journal of Asian Architecture and Building Engineering*, 18 (3), 233-246.
- Feuillet, T., Salze, P., Charreire, H. (2016). Built Environment in Local Relation with Walking: Why Here and Not There?. *Journal of Transport & Health*, 12(4).
- Frank, L.D., Sallis, J.F., Conway, T.L. (2006). Many Pathways to Health: Associations between Neighborhood Walkability and Active Transportation. *Journal of the American Planning Association*, 72, 75-87.

- Giles-Corti, B., Donovan, R.J. (2002). The relative influence of individual, social and physical environment determinants of physical activity. *Social Science & Medicine*, 54, 1793–1812.
- Gómez, J.C., de la Fuente, M.V., Ros, L. (2019). Characterization of pedestrian and mixed streets and its influence in freight activities. 15th Annual International Bata Conference for Ph.D.
- Gómez, J.C., de la Fuente, M.V., Ros, L. (2020). Análisis del comportamiento de repartidores logísticos en zonas peatonales. *Dirección y Organización*, 71, 14-30.
- Hamidi, S., Moazzeni, S. (2019). Examining the Relationship between Urban Design Qualities and Walking Behavior: Empirical Evidence from Dallas, TX. *Sustainability*, 11, 2720
- Handy, S. (1993). Regional Versus Local Accessibility: Implications for Non-work Travel. *Transportation Research Record*, 1400.
- Handy, S. L. (1996). Urban form and pedestrian choices: Study of Austin neighborhoods. *Transportation Research Record: Journal of the Transportation Research Board*, 1552, 135–144.
- Handy, S., Boarnet, M., Ewing, R., Killingsworth, M. (2002). How the Built Environment Affects Physical Activity. *Views from Urban Planning. American Journal of Preventive Medicine*, 23, 64-73.
- Handy, S., Mokhtarian, P. (2008). *Growing Cooler: The Evidence on Urban Development and Climate Change*. Urban Land Institute: Washington, DC, USA, 2008.
- Jacobs, J. (1961). *The Death and Life of Great American Cities*. Vintage. Nueva York.
- Kumar, K., Ross, W. (2006). Effects of pedestrianization on the commercial and retail areas: study in Khao San road, Bangkok. *World Transport Policy & Practice*, 13.
- Muñuzuri, J., Onieva, L., Cortes, P. y Guadix, J. (2016). Gestión pública de la logística urbana de mercancías. Una norma, múltiples necesidades. *Economía Industrial*, 400, 21–28.
- Næss, P. (2005). Residential location affects travel behavior – But how and why? The case of Copenhagen metropolitan area. *Progress in Planning*, 63(2), 167–257.
- Oakes, J.M., Forsyth, A., Schmitz, K.H. (2007). The effects of neighborhood density and street connectivity on walking behavior: The Twin Cities walking study. *Epidemiologic Perspectives & Innovations*, 4(16).
- Palomares, B. (2015). *Marketing en el punto de venta*. ESIC Editorial, Madrid.
- Park, D., Nakagawa, R., Matsunaka, T. (2012). Factor Structure Which Makes Streets Vibrant with People—Based on Field Survey of Kyoto, Seoul and Florence. *Journal of Japan Society of Civil Engineers Ser.* 68 (5).
- Pikora, T., Giles-Corti, B., Bull, F. (2003). Developing a framework for assessment of the environmental determinants of walking and cycling. *Social Science & Medicine*, 56(8), 1693-1703
- Reisi, M., Ahmadi, M., Aye, L. (2019). Local walkability index: assessing built environment influence on walking. *Bulletin of Geography. Socio-economic Series*, 46, 7–21.
- Ros-McDonnell, L., De-la-Fuente, M.V., Szander, N., Ros-McDonnell, D. (2016). A study of the walkability impact in a compact Mediterranean city. *International Joint Conference - CIO-ICIEOM-III-E-AIM (IJC 2016)*. San Sebastián, Spain.
- Saelens, B., Sallis, J., Black, J., Chen, D. (2003) Neighborhood-Based Differences in Physical Activity: An Environment Scale Evaluation. *American Journal of Public Health*, 93(9).
- Salom, J., Carrasco, J.M., Puebla, A. (2014). Delimitation and characterization of new urban spaces in Valencia. *Boletín de la Asociación de Geógrafos Españoles*, 64, 127-149.
- Sastreb, A., Gamo, A.M., Gaztelu, T. (2013). Economic Impact of Pedestrianisation in Historic Urban Centre, the Valdemoro Case – Study (Spain). *Procedia. Social and Behavioral Sciences*, 104(2013),737 – 745.
- Song, Y., Knaap, G. (2007). “Quantitative classification of neighborhoods: the neighbourhoods of new single-family homes in the Portland Metropolitan Area”. *Journal of Urban Design*, 12.
- Sugie, L., Hyungun, S., Ayoung, W. (2009). The Spatial Variations of Relationship between Built Environment and Pedestrian Volume: Focused on the 2009 Seoul Pedestrian Flow Survey in Korea. *Journal of Asian Architecture and Building Engineering*, 16(1).
- Sung, H., Lee, S. (2015), Residential Built Environment and Walking Activity: Empirical Evidence of Jane Jacobs' Urban Vitality, *Transportation Research: Part D*, 41, 318-329.