Dispersion, spatial autocorrelation and mismatches between need and public transport supply in an intermediate city: San Juan de Pasto, Colombia

Jorge Luis Argoty Burbano[1], Daniel Mauricio Goyes Chaves[2], Ciro Jaramillo Molina[3]

[1]Doctoral Program in Urban and Territorial Management, Universidad del Valle, Cali, Colombia
jorge.argoty@correounivalle.edu.co https://orcid.org/0000-0001-6661-1398
[2]Universidad del Valle, Cali, Colombia. daniel.goyes@correounivalle.edu.co https://orcid.org/0000-0001-8923-8258
[3]School of Civil and Geomatic Engineering, Universidad del Valle, Cali, Colombia.
ciro.jaramillo@correounivalle.edu.co https://orcid.org/0000-0002-8820-2314

Abstract

The correspondence between public transport supply and the distribution of the population is usually assessed through a single reading of each variable and with correlations that ignore the neighborhood relationships between zones. This article characterizes the distribution, spatial autocorrelation and associations of nine socioeconomic, territorial and public transport variables in the 184 transport analysis zones (TAZs) of San Juan de Pasto, Colombia, an intermediate city. The 2018 census, the household origin–destination survey, the register of commercial establishments of the Chamber of Commerce and the scheduled supply of the Strategic Public Transport System (SETP) were integrated, and descriptive statistics, Spearman correlations and univariate and bivariate Moran statistics were applied. High dispersion does not imply spatial clustering: commercial establishments are the most dispersed variable (CV = 246.3%) but the least clustered (I = 0.070), whereas income, the least dispersed, is the most clustered (I = 0.455). Scheduled supply is associated with commercial establishments (ρ = 0.58) and income (ρ = 0.38), and is not correlated with households without a vehicle or with people seeking employment. However, the bivariate analysis shows that zones with more households in strata 1 and 2 are surrounded by lower supply (I = −0.235; −0.398 with household rates), a result that is robust to five specifications, and it locates a core of four zones and a northeastern sector where residential need coincides with lower route coverage. The results provide a baseline for assessing accessibility and territorial equity.

Keywords: intermediate cities; transport and land use; transport analysis zones;; spatial autocorrelation bivariate Moran; spatial mismatch; public transport supply; Colombia

INTRODUCTION

Public transport planning requires knowing how the population, resources and urban opportunities are distributed in space, because accessibility depends on the location of activities, the impedance of travel and the capabilities of individuals (Hansen, 1959; Geurs & van Wee, 2004; Páez et al., 2012), and sustainable mobility links living conditions with the available alternatives for travel (Banister, 2008). Socio-spatial characterization therefore has value in its own right as a diagnosis.

Mobility needs vary with income, vehicle availability, employment status and disability, and

their territorial representation is the basis of social need diagnoses (Currie, 2010; Foth et al., 2013). In Colombia, the spatial equity of accessibility has been assessed in Cali (Delmelle & Casas, 2012) and Bogotá (Bocarejo & Oviedo, 2012), and inequalities in access to work, education and fares have been documented (Guzman et al., 2017; Guzman & Oviedo, 2018; Oviedo & Guzman, 2020; Arellana et al., 2021); in the region, it has been examined in Montevideo (Hernández, 2018) and Santiago de Chile (Tiznado-Aitken et al., 2018), together with the relationship between daily mobility and residential segregation (Rodríguez Vignoli, 2008). This evidence comes mainly from large metropolitan areas (Vecchio et al., 2020), and less attention has been paid to the joint reading of the statistical distribution and the spatial structure of the variables in intermediate cities. Moreover, the correspondence between need and supply is usually assessed with non-spatial correlations, which compare zones without considering their neighborhood, even though the population also uses the transport services of contiguous zones. The notion of spatial mismatch, originally formulated for the separation between residence and employment (Kain, 1968), has been transferred to public transport in studies of gaps between need and supply (Currie, 2010; Jiao & Dillivan, 2013; Fransen et al., 2015), which compare each territorial unit through gap or accessibility indices but rarely assess whether the zones with the greatest need are surrounded by lower supply.

San Juan de Pasto, the administrative, commercial and service center of the department of Nariño, is an intermediate city, a category defined by its regional functions and socio-spatial attributes rather than by its size (Villegas Alzate, 2023). Studies of the city address urban expansion and land occupation models (González Bastidas, 2021; Mesías Rosas, 2024) or the social impact of the SETP from an organizational management perspective (Recalde Martínez, 2019). No characterization was found that integrates, on a single transport zoning, residential composition, commercial establishments and SETP supply through descriptive and spatial statistics, or that contrasts need with the supply of the surrounding area through bivariate spatial statistics. This is the gap addressed by the present study.

The objective of the article is to characterize the distribution, spatial autocorrelation and associations of nine socioeconomic, territorial and public transport variables in the 184 TAZs of Pasto, to examine the extent to which statistical dispersion and spatial clustering offer different readings of the territory, and to establish whether the correspondence between need and supply changes when the surroundings of each zone are considered. To this end, three questions are posed. How are residential, land-use and transport supply variables distributed across zones, and how clustered are they in space? What associations do these variables show with one another, and do they hold when the neighborhood between zones is considered? How are the patterns of the resident population, commercial establishments and service related, and where are the mismatches between need and supply located?

The study contributes an integrated zonal baseline for an intermediate Colombian city and evidence, robust to five specifications, that non-spatial correlation conceals mismatches between need and supply that bivariate spatial statistics make it possible to locate. Its scope

is descriptive; effective accessibility and distributive equity are not evaluated here.

2. ANALYTICAL FRAMEWORK

2.1 Residents, destinations and transport supply

Four dimensions are distinguished: population and land occupation; socioeconomic conditions and potential mobility constraints; located commercial establishments; and scheduled public transport supply. Each answers a different question, and commercial establishments and service represent territorial resources. From the perspective of transport and land use, density and commercial establishments correspond to the land-use component (Ewing & Cervero, 2010), and the scheduled service supply to the transport component. Lack of a vehicle characterizes modal availability, whereas disability introduces specific functional barriers (Casas, 2007; Lucas, 2012; Park & Chowdhury, 2022).

Accessibility studies combine land-use, transport, temporal and individual components (Geurs & van Wee, 2004; Páez et al., 2012); social need studies compare supply with the places where need is concentrated (Currie, 2010; Foth et al., 2013; Lucas, 2012), and equity assessments rely on zonal comparisons between supply and population (Karner & Niemeier, 2013; Boisjoly & El-Geneidy, 2017) or on Lorenz curves (Delbosc & Currie, 2011). This article remains at the level of diagnosis: an observed territorial difference does not in itself constitute an equity judgment (Lucas et al., 2016; Pereira et al., 2017), and the location of resources is not equivalent to their effective use (Grengs, 2015).

2.2 Zonal measurement and spatial structure

Zone-based aggregate analysis is conditioned by the scale and delimitation of the polygons (Openshaw, 1984; Fotheringham & Wong, 1991; Javanmard et al., 2023), and associations between aggregates do not describe individuals (Robinson, 1950). Zonal totals locate volumes and may covary simply because both grow with zone size (Pearson, 1897; Kronmal, 1993), whereas density or mean income describe intensities or levels; furthermore, a variable may be dominated by a few extreme zones and show little similarity between neighbors, or vice versa.

Spatial structure is examined with the global Moran's I and local indicators of spatial association (LISA), which assess the similarity between neighboring zones (Moran, 1950; Getis & Ord, 1992; Anselin, 1995) and depend on the definition of neighborhood (Bivand & Wong, 2018). A skewed distribution, a high correlation and a local cluster are therefore complementary pieces of evidence. When the interest lies in the correspondence between two variables, the bivariate Moran's I and its local version relate the value of one variable in each zone to that of another variable in its neighbors (Lee, 2001; Anselin, Syabri, & Smirnov, 2002; Anselin, Syabri, & Kho, 2006). This perspective is relevant for public transport, because the service available to a zone includes the routes of contiguous zones, and it makes it possible to identify mismatches that non-spatial correlation does not detect. The term mismatch is used here in a descriptive sense, as the combination of high residential need and low supply in the surrounding area, without implying causality. There are three analytical expectations: (E1) the dispersion of a variable does not anticipate its

spatial clustering; (E2) zones with more households in strata 1 and 2 will have lower mean income and may have fewer territorial resources, a possibility that is examined rather than assumed; and (E3) the absence of non-spatial correlation between supply and social conditions does not imply the absence of spatial mismatch.

3. DATA AND METHODS

3.1 Study area, spatial unit and sources

The study covers the 184 TAZs of San Juan de Pasto, including discontinuous peripheral units; in intermediate cities, the transport zoning is the common support for census, survey and operational information, and it is therefore the unit of analysis here. The 2018 National Population and Housing Census and its cartography (Departamento Administrativo Nacional de Estadística [DANE], 2018), the household origin–destination survey (EODH), the commercial establishments registered with the Pasto Chamber of Commerce as of January 2024, geocoded from the information provided by this entity (hereafter, commercial establishments), and the SETP routes, stops and service headway table (Table 1) were integrated, with the Land Use Plan as the framework (Concejo Municipal de Pasto, 2015). Processing was carried out in R. Because the sources refer to different periods, the database is an integrated territorial characterization rather than a simultaneous measurement. The nine variables were assigned to the same 184 polygons in all analyses (Figure 1).

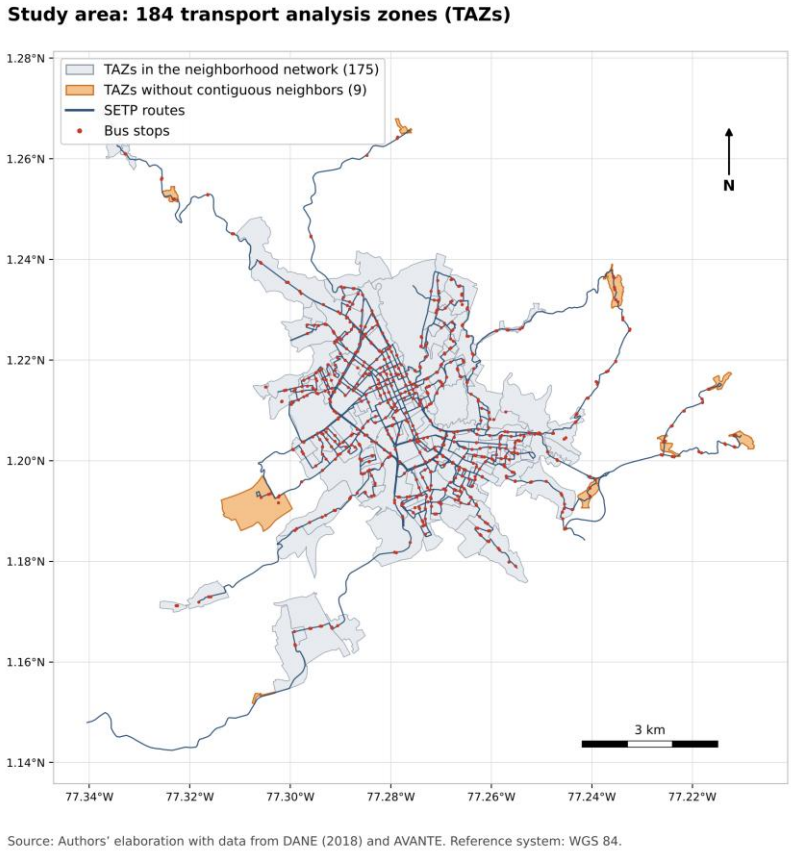


Figure 1. Study area: TAZs of San Juan de Pasto, zones without contiguous neighbors, and SETP route and stop network. Source: Authors' elaboration with data from DANE (2018) and AVANTE.

Table 1. Information sources used to build the zonal database.

Source	Information used	Period, coverage and sample
2018 National Population and Housing Census and its cartography (DANE)	Population, households, socioeconomic strata and zone boundaries	2018; entire municipality
Household origin–destination survey (EODH)	Vehicle availability, disability, job seeking and household income	2017; 4,509 households and 16,047 persons assigned to the TAZs
Commercial register of the Pasto Chamber of Commerce	Registered commercial establishments, geocoded and assigned to the TAZs	Register as of January 2024
SETP route and stop layers and service headway table (AVANTE)	Routes, stops and service headway per route (minutes between buses)	2020 operational design (Consorcio Pasto 2020, 2020); weekday, 05:30–21:30

3.2 Selection and measurement of variables

The nine variables were selected for their conceptual relevance, their availability and their assignment to the TAZs (Table 2). They describe the resident population, understood as potential SETP users, since the survey does not make it possible to isolate those who actually use the system. The population aged 19 to 65, labeled working population in the database, is used as a measure of the resident working-age population, without certifying actual employment; commercial establishments correspond to the registered establishments, aggregated by zone into a field with fractional values; and job seeking is the activity declared in the survey, not an official unemployment rate. Households in strata 1 and 2 are those located in the two lowest of the six socioeconomic strata that Colombia assigns to residential properties, a residential classification rather than a direct measure of poverty.

Table 2. Variables analyzed, source, analytical meaning and transport equity dimension.

Variable	Source and operation	Interpretation	Equity dimension (references)
Populatio n aged 19 to 65	Integrated layer by TAZ (census)	Volume of resident population; does not certify actual employment	Potential demand (Geurs & van Wee, 2004)

Households in strata 1 and 2	Integrated census layer	Volume of households by socioeconomic stratum	Vertical equity: socioeconomic condition (Lucas et al., 2016; Pereira et al., 2017)
Population density (inhab./ha)	Census and cartographic layer	Intensity of land occupation	Land-use component (Ewing & Cervero, 2010)
Commercial establishments	Commercial establishments registered with the Pasto Chamber of Commerce, geocoded	Magnitude of registered commercial activity; does not measure size or capacity and does not include informal or unregistered establishments	Destinations and opportunities (Geurs & van Wee, 2004; Páez et al., 2012)
Households without a vehicle	EODH; expansion and imputation where necessary	Limited modal availability; not equivalent to bus use	Modal constraint and social need (Currie, 2010; Grengs, 2015)
Persons with disabilities	EODH; expansion and imputation where necessary	Functional vulnerability	Functional barrier (Casas, 2007; Park & Chowdhury, 2022)
Persons seeking employment	EODH; expansion and imputation where necessary	Declared employment status; not an official rate	Social exclusion and access to employment (Lucas, 2012)
Estimated monthly household income	EODH; category midpoints and imputation	Approximate economic capacity	Vertical equity: economic capacity (Pereira et al., 2017; Guzman & Oviedo, 2018)
Service frequency (SETP)	Routes, stops and service headway per route, in minutes between buses (AVANTE)	Scheduled supply (sum of headways); does not measure waiting times or effective accessibility	Transport supply versus need (Currie, 2010; Delbosc & Currie, 2011)

Note. Source: Authors' elaboration based on the census, the household survey, the register of commercial establishments of the Pasto Chamber of Commerce and public transport records.

3.3 Linkage, expansion and imputation of the survey

The household, person and vehicle modules were linked through the questionnaire and folio identifiers and assigned to the TAZs. Expansion used census factors, with a minimum of five surveyed households for a direct estimate. Household income was estimated by assigning representative values to the individual income categories, summing them by household and averaging them by TAZ; the no-income option was coded as zero, the non-fixed-income option was excluded and the open upper category was represented by its lower bound.

TAZs without the minimum sample received imputations of rates and means from at least two donor zones sharing a common boundary, up to third-order neighborhood, or, in their absence, from the three nearest observed zones. Eighteen TAZs were imputed (eight with immediate neighbors, four with second- or third-order neighbors and six by proximity), only for the four survey variables (households without a vehicle, disability, job seeking and income). These zones are retained in the descriptive statistics (Table 3), the maps and the main spatial analysis (Table 4 and Figures 6 to 9), but are excluded from the Spearman correlations (Figure 11), because imputation may introduce artificial similarity between neighbors; their influence on Moran's I is examined in Section 4.5 (Table 5).

3.4 Construction of service supply

Stops were cleaned (0.1 m tolerance) and linked to routes within 50 m and to TAZs within 25 m. The aggregate frequency of each TAZ adds, only once, the service headway, in minutes between buses, of each distinct route serving it, according to the weekday headway table (Consorcio Pasto 2020, 2020). By construction, it increases with the number of routes and with the length of their headways, so it is interpreted as scheduled route supply and not as buses per hour; hereafter, aggregate frequency and scheduled supply refer to the same variable. Nine TAZs recorded zero supply, two of them (213011 and 203013) also isolated, and 17 showed a possibly incomplete correspondence between stops and routes.

3.5 Statistical and spatial analysis

For each variable, the mean, median, standard deviation, coefficient of variation (CV), quartiles, interquartile range (IQR), skewness and excess kurtosis were calculated, and values outside [Q1 − 1.5 IQR; Q3 + 1.5 IQR] were identified as outliers; kurtosis describes the weight of the extremes and not geographic clustering (NIST/SEMATECH, n.d.).

Six variables express volumes per TAZ (population aged 19 to 65, households in strata 1 and 2, households without a vehicle, disability, job seeking and scheduled supply), which indicate where potential demand is concentrated; their maps and LISA are interpreted as concentrations of persons or households, not as a higher proportion of the condition.

After a normality diagnosis (Lilliefors, 1967), the associations between pairs of variables were described with the Spearman correlation on the 166 observed zones, with two-sided p-values

adjusted by the Holm (1979) method for the 36 pairs.

Spatial autocorrelation was assessed with the global Moran's I and LISA, using first-order queen contiguity and row-standardized weights. Nine TAZs, peripheral units sharing no edge or vertex with any other zone, have no neighbors and were excluded from the network of 175 zones (Figure 1); this condition is geometric and distinct from imputation (Section 3.3). A total of 9,999 permutations were used, conditional for LISA, and local results with pseudo-p ≤ 0.05 were classified as High-High (HH), Low-Low (LL), Low-High (LH) or High-Low (HL); their significance is nominal and exploratory. As a sensitivity analysis, Moran's I was recalculated without the 18 imputed TAZs (166 zones, 161 connected).

The bivariate Moran's I and the bivariate LISA (Anselin, Syabri, & Smirnov, 2002; Anselin, Syabri, & Kho, 2006) relate the standardized value of a variable x in each zone to the average of another variable y in its neighbors, that is, each zone to its surroundings (Lee, 2001). Eight pairs were analyzed: the three linking territorial resources (commercial establishments, income and supply), households in strata 1 and 2 with income, and the four social conditions with supply, using the same network, matrix and number of permutations. The calculation was implemented in Python (NumPy) with the same weight matrix, which exactly reproduces the univariate Moran's I values obtained in R.

Robustness was assessed with four variants: without the 18 imputed TAZs (161 zones); household rates instead of counts, to control for zone size; a matrix of the six nearest neighbors (184 zones); and the reverse direction, with supply in the zone and the condition in the surroundings. Zones classified as High-Low in at least three of the four specifications that preserve the direction of the pair were considered a robust core of mismatch.

4. RESULTS

4.1 Overall statistical profile

Table 3 summarizes the nine variables. All have positive skewness and a mean greater than the median, but the weight of their extremes differs. Commercial establishments are the most extreme variable (CV = 246.3%; skewness = 10.66; kurtosis = 129.42; 11 outliers): their maximum, 704.1 in TAZ 101152, is about 56 times the median. Households without a vehicle have the second-highest kurtosis (9.82), and the population aged 19 to 65 is the most regular (kurtosis = 0.20; two outliers).

Table 3. Location, dispersion, extremes and shape of the nine variables in the 184 TAZs.

Variable	Min.	Q1	Median	Mean	Q3	Max.	CV (%)	Skew.	Kurt.	Outl.
Population aged 19 to 65	22.0	610.5	1077.0	1137.7	1537.8	3591.0	59.3	0.63	0.20	2
Households in strata 1 and 2	5.0	60.5	306.0	356.1	555.5	1710.0	89.0	1.06	1.19	1
Density	0.6	9.0	16.5	18.	25.	74.4	68.	1.	2.1	4

(inhab./ha)				4	3		8	24	6	
Commercial establishments	0.0	2.2	12.6	22.2	26.6	704.1	246.3	10.66	129.42	11
Households without a vehicle	4.0	110.0	200.5	225.4	279.3	1353.0	75.4	2.20	9.82	7
Persons with disabilities	0.0	18.0	51.5	64.4	85.5	314.0	95.6	1.64	2.94	12
Persons seeking employment	0.0	14.0	35.0	51.7	69.3	269.0	104.7	1.64	2.81	12
Household income (thousand COP)	369	1233	1617	1756	2083	4601	41.9	1.14	1.78	9
Service frequency (min)	0.0	12.0	26.0	30.3	41.3	112.0	73.7	1.13	1.35	5

Note. Units specific to each variable; income in thousands of Colombian pesos (COP) per month, rounded. CV: coefficient of variation; Skew.: skewness; Kurt.: excess kurtosis; Outl.: outliers. Source: integrated database, n = 184 TAZs, with expansion and imputation where applicable; includes the 18 TAZs imputed in the four survey-derived variables.

Kurtosis measures the intensity of the extremes, not their number: disability has 12 outliers with a much lower kurtosis than commercial establishments. Income has the lowest CV (41.9%), although nine outliers reveal an upper tail. Zeros indicate that no positive value was recorded or estimated, not missing data.

4.2 Distribution and spatial structure of the nine variables

4.2.1 Population aged 19 to 65

The population aged 19 to 65 has a median of 1,077 persons (mean of 1,137.7; CV = 59.3%) and its maximum, 3,591, is in TAZ 110131. The largest volumes are in the north, northeast and south. Spatial dependence is positive ($I = 0.236$; $p = 0.0001$), with 13 HH zones, 11 LL zones and 19 contrasts (14 LH and 5 HL) (Figure 2), so residential concentration changes rapidly between nearby zones. It is associated with households without a vehicle ($\rho = 0.79$), households in strata 1 and 2 (0.70), disability and job seeking (0.54), but not with income (0.06); since these are zonal totals, part of this covariation reflects zone size.

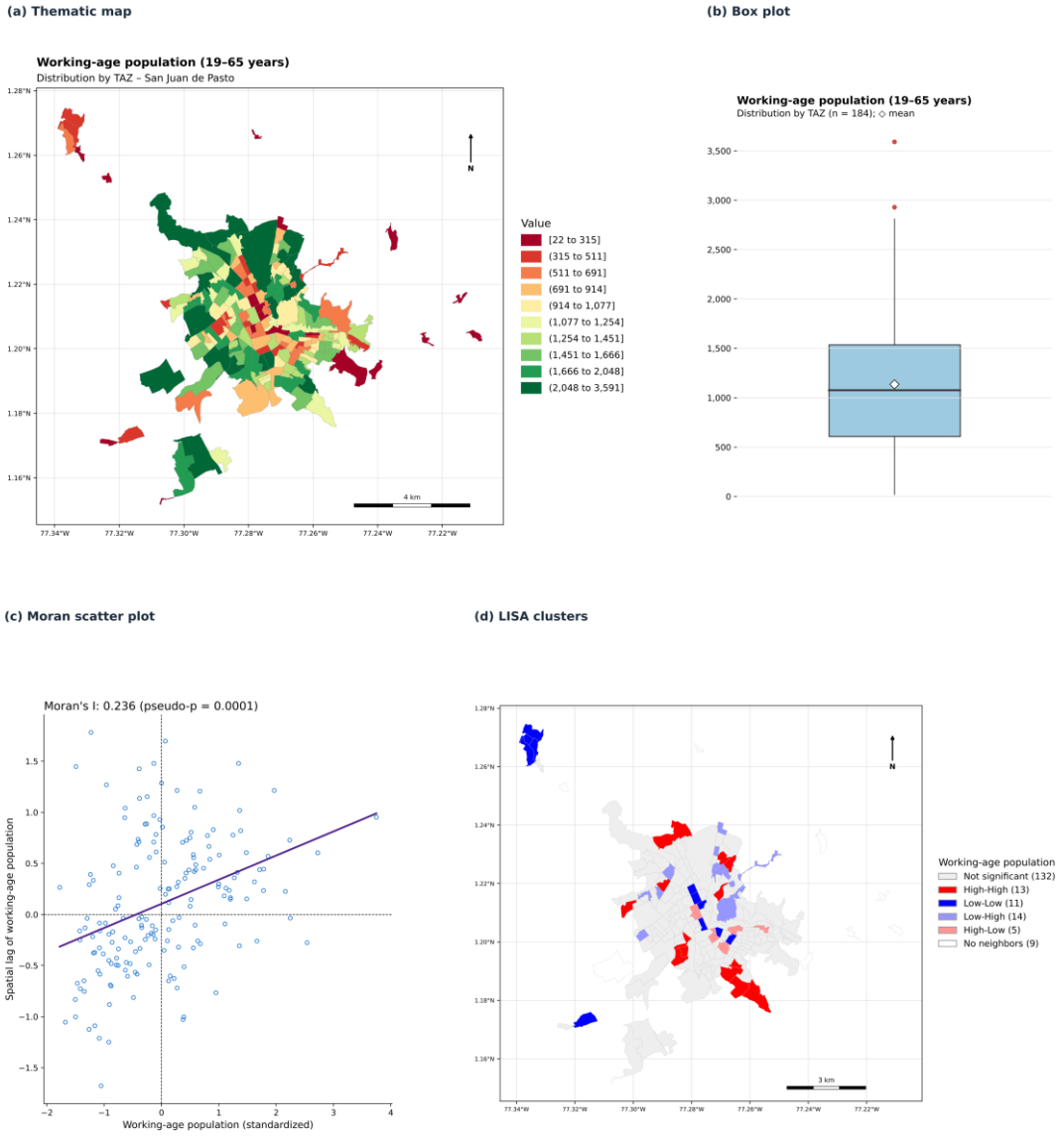


Figure 2. Population aged 19 to 65: (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.2 Households in strata 1 and 2

Households in strata 1 and 2 have a median of 306 (mean of 356.1; CV = 89.0%) and a maximum of 1,710 in TAZ 110131, with concentrations in the northeast and south. Their autocorrelation is among the highest (I = 0.402; p = 0.0001), with 19 HH, 30 LL, 8 LH and 2 HL zones (Figure 3): they form contiguous groups rather than an isolated maximum. They are negatively associated with income ($\rho = -0.42$), and their correlation with commercial establishments (−0.05) and supply (−0.16) is not significant; Section 4.6 shows, however, that these zones tend to be surrounded by lower supply.

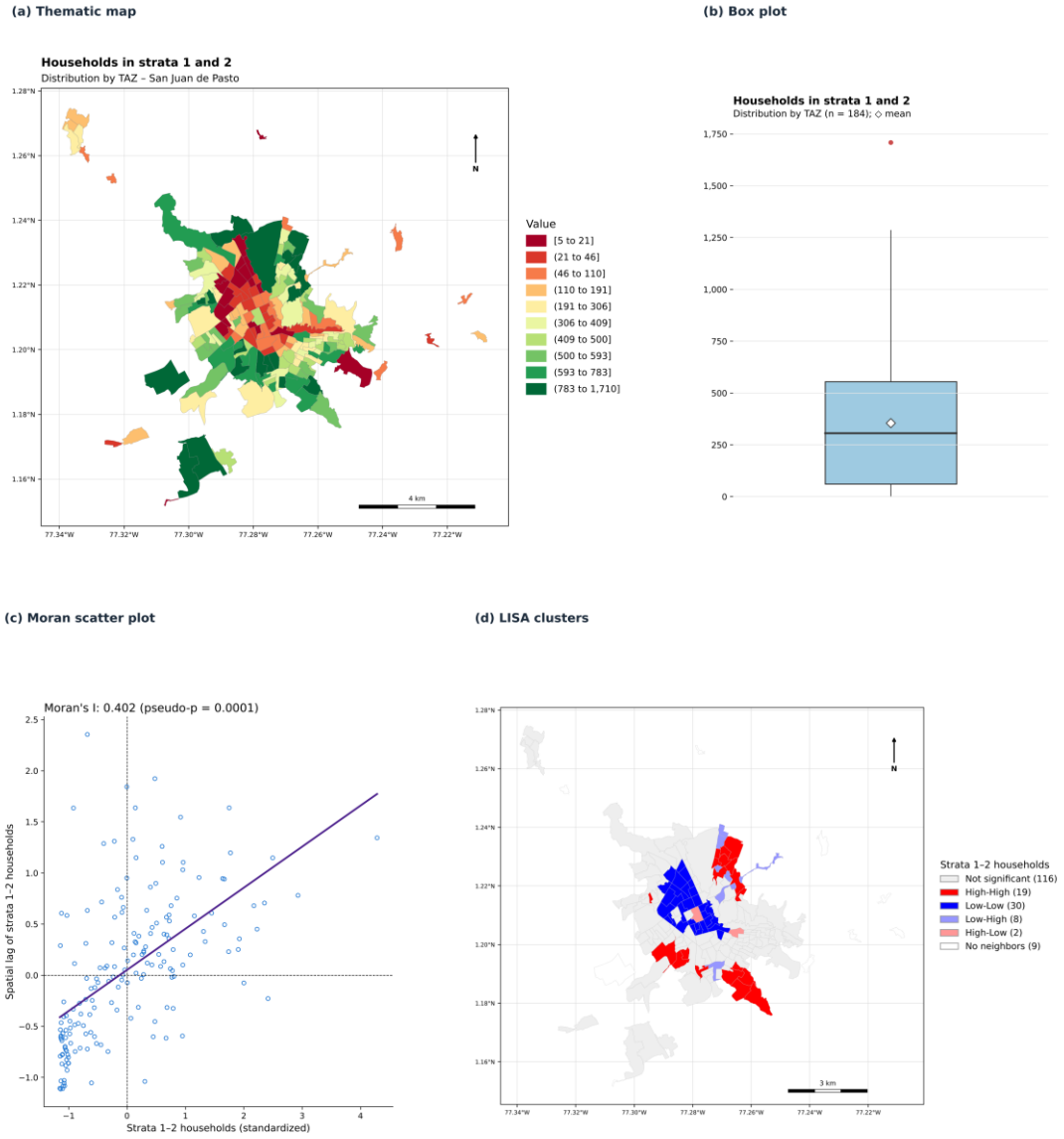


Figure 3. Households in strata 1 and 2: (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.3 Population density

Density has a median of 16.5 inhab./ha (mean of 18.4; CV = 68.8%) and a maximum of 74.4 in TAZ 112115. Consolidated and eastern foci contrast with extensive low-density peripheral units. Moran's I is 0.344 (p = 0.0001), with 11 HH and 21 LL zones (Figure 4), and its geometry differs from that of residential volume, because a small zone can be dense without concentrating many residents. It is moderately associated with households in strata 1 and 2 (ρ = 0.36), population (0.29) and commercial establishments (0.25), and not with supply (0.02).

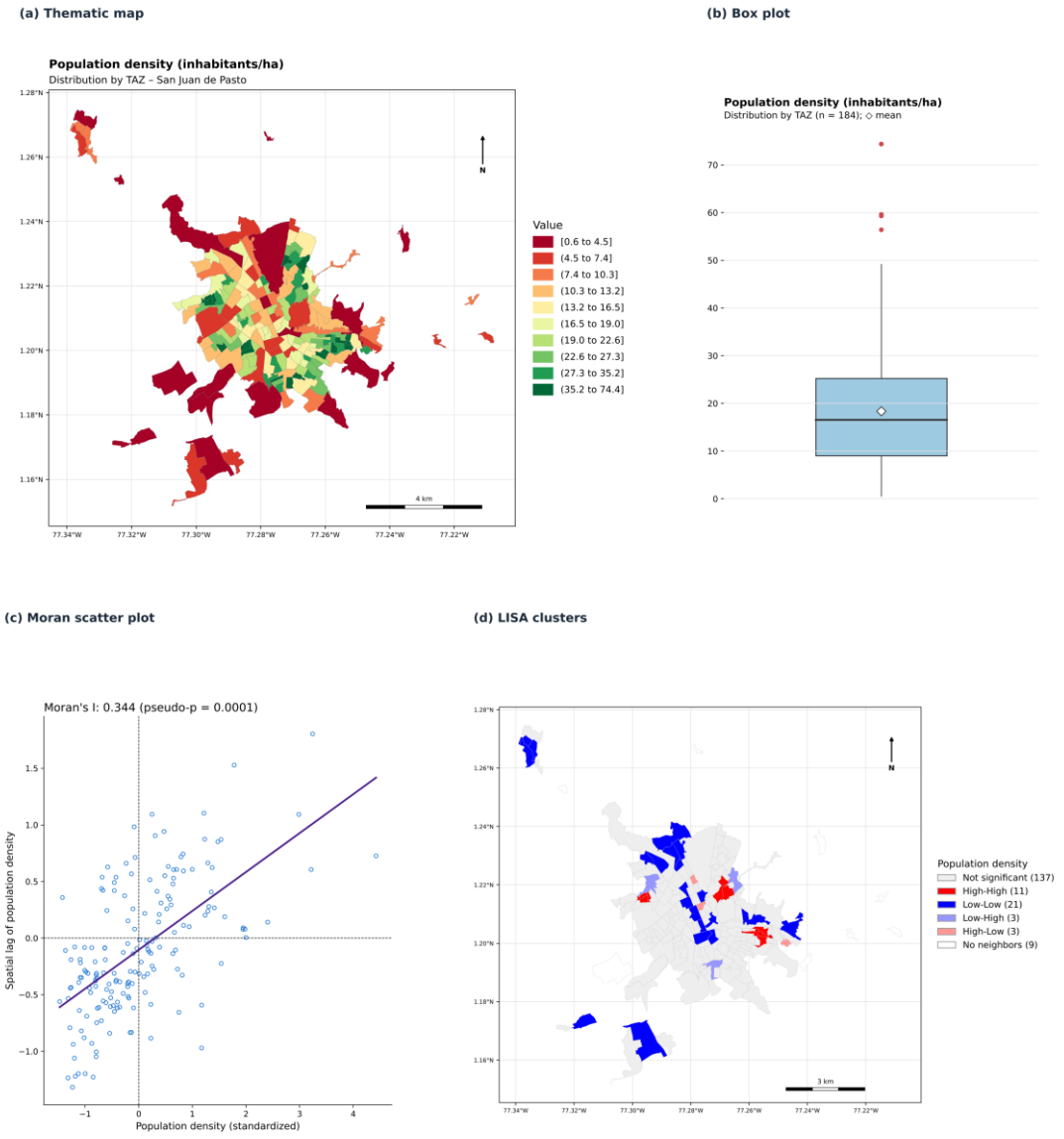


Figure 4. Population density (inhab./ha): (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.4 Commercial establishments

Commercial establishments have a median of 12.6 and a mean of 22.2, pulled up by the maximum. Their statistical concentration coexists with a low Moran's I (I = 0.070; p = 0.0105): LISA identifies 10 HH and 14 LL zones, and 145 of the 175 zones are not significant (Figure 5), because the central maximum is not reproduced in its neighbors. They are associated with supply (ρ = 0.58), income (0.42) and population (0.32), and not with households in strata 1 and 2 (−0.05). The field expresses the registered magnitude, not the capacity of the establishments.

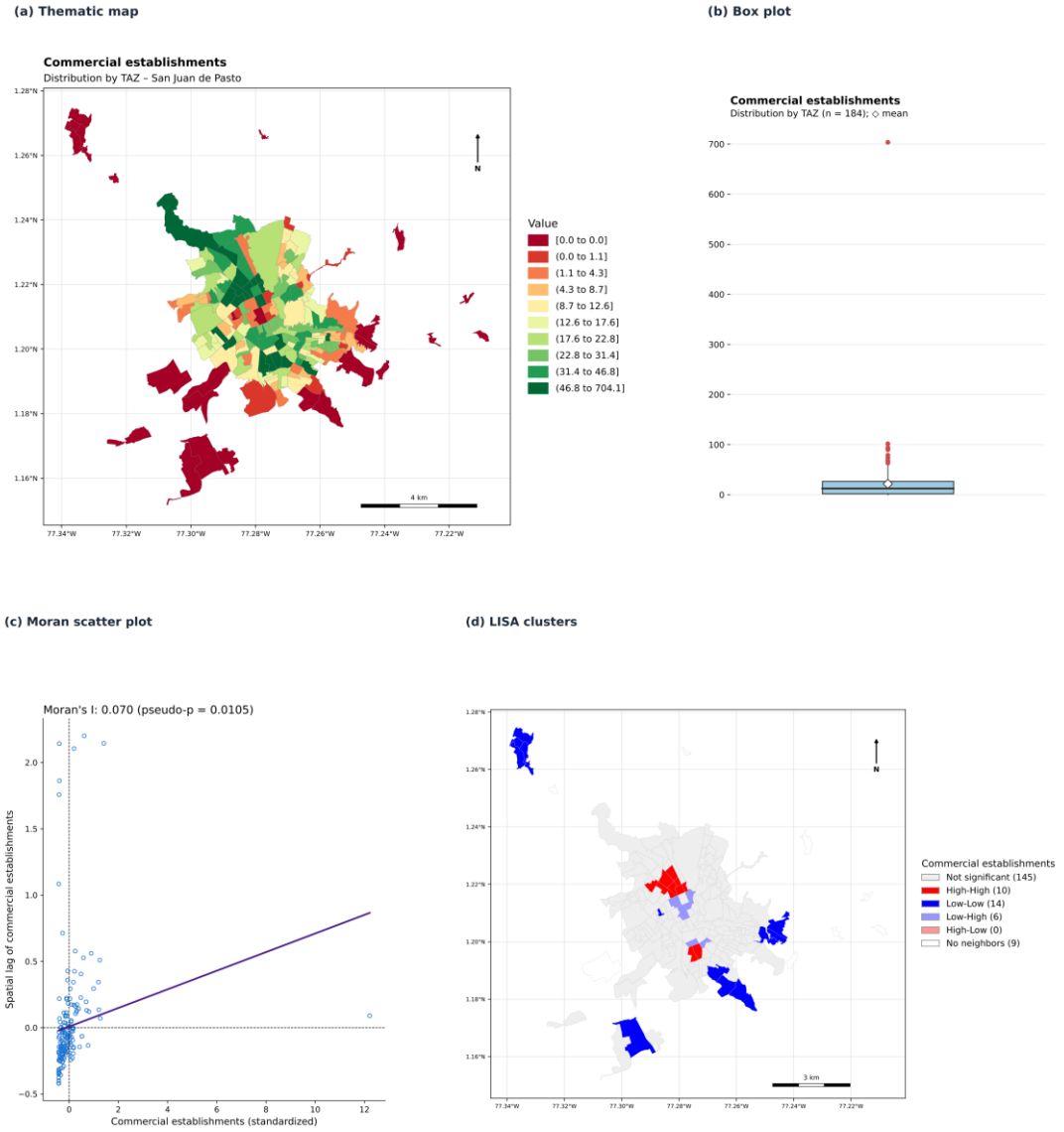


Figure 5. Commercial establishments: (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.5 Households without a vehicle

Households without a vehicle have a median of 200.5 (mean of 225.4; CV = 75.4%) and a maximum of 1,353, again in TAZ 110131. Spatial dependence is positive but lower than that of stratum and income ($I = 0.153$; $p = 0.0026$), with 7 HH zones, 12 LL zones and 146 non-significant zones (Figure 6); the foci in the northeast and south do not form a continuous band. They are associated with population ($\rho = 0.79$), households in strata 1 and 2 (0.60), disability (0.58) and job seeking (0.46), and not with supply (0.02) or income (−0.08).

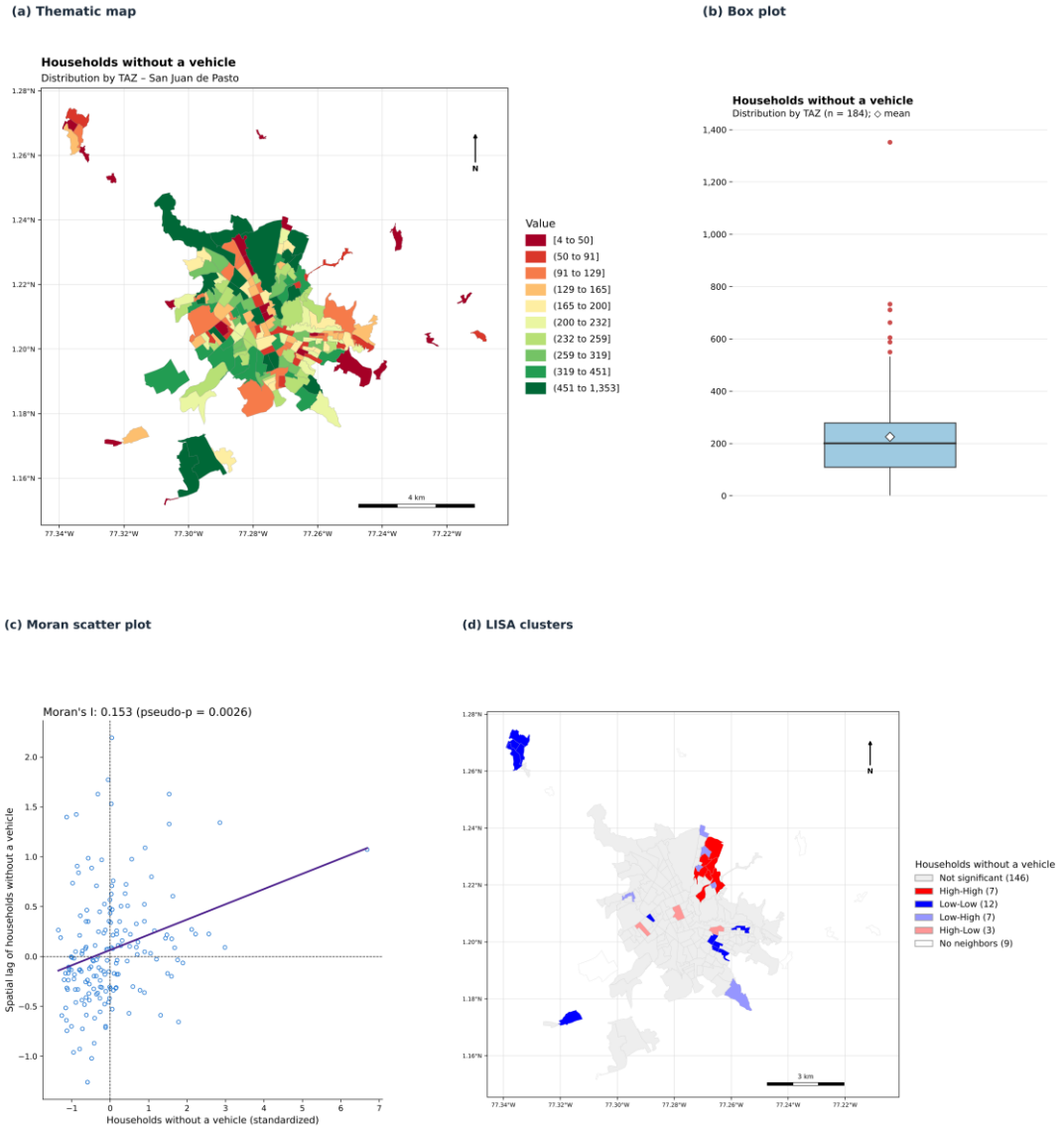


Figure 6. Households without a vehicle: (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.6 Persons with disabilities

Disability has a median of 51.5 persons (mean of 64.4; CV = 95.6%), a maximum of 314 and 12 outliers. Its global I is not significant at 5% (I = 0.077; p = 0.0524); only two zones are HH and six LL, and 160 are not significant (Figure 7). It is associated with households without a vehicle (ρ = 0.58), population (0.54), households in strata 1 and 2 (0.48) and job seeking (0.40), and not with supply (0.01). It contributes a functional dimension that is not substituted by income or service (Casas, 2007; Park & Chowdhury, 2022).

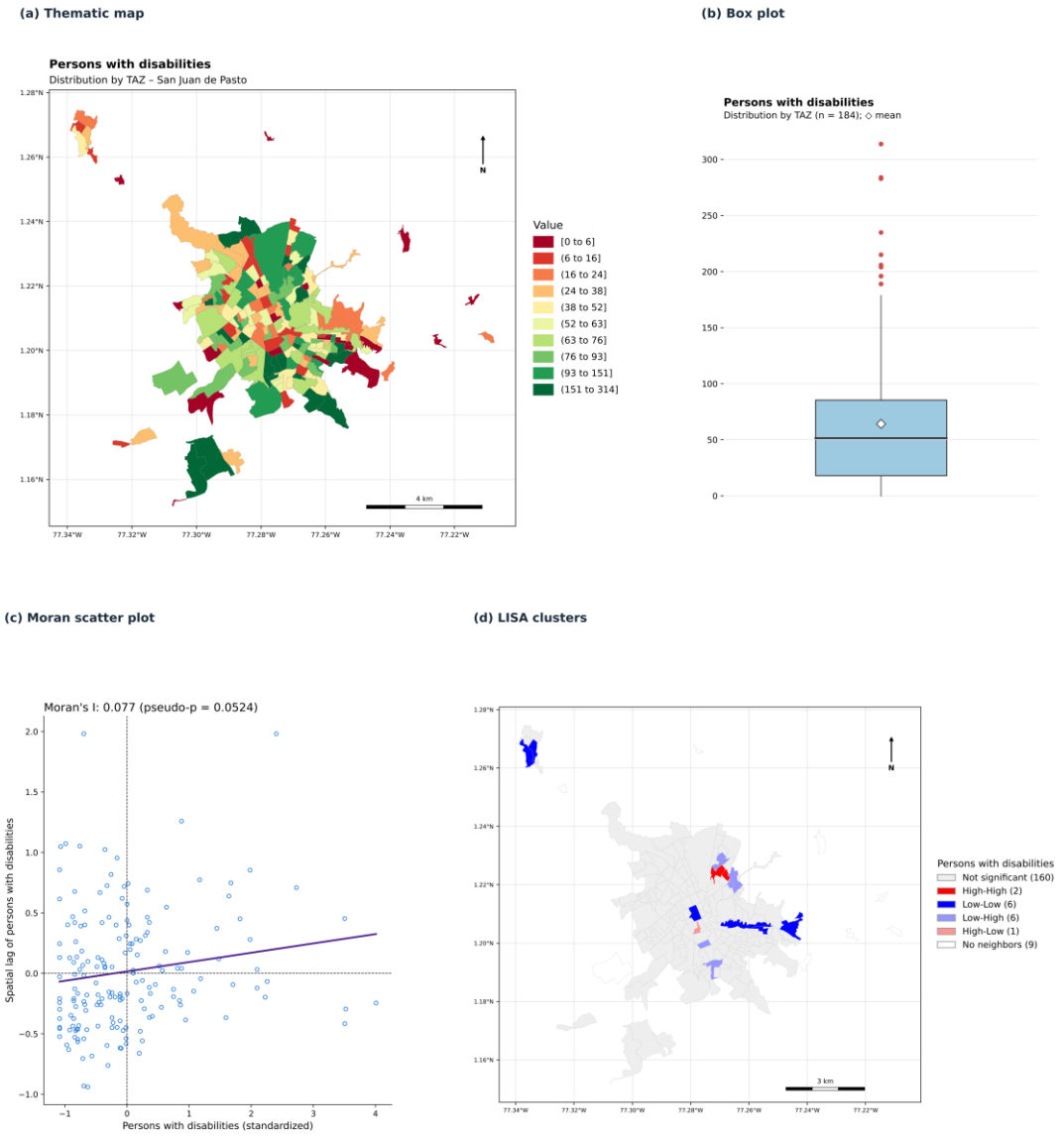


Figure 7. Persons with disabilities: (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.7 Persons seeking employment

Job seeking has a median of 35 persons (mean of 51.7; CV = 104.7%) and a maximum of 269, with concentrations in the west, northeast and southeast. Its autocorrelation is limited ($I = 0.165$; $p = 0.0020$), with 12 HH zones, 11 LL zones, ten contrasts and 142 non-significant zones (Figure 8). It is associated with population ($\rho = 0.54$), households in strata 1 and 2 (0.51), households without a vehicle (0.46) and disability (0.40), and not with supply (−0.05) or income (−0.16).

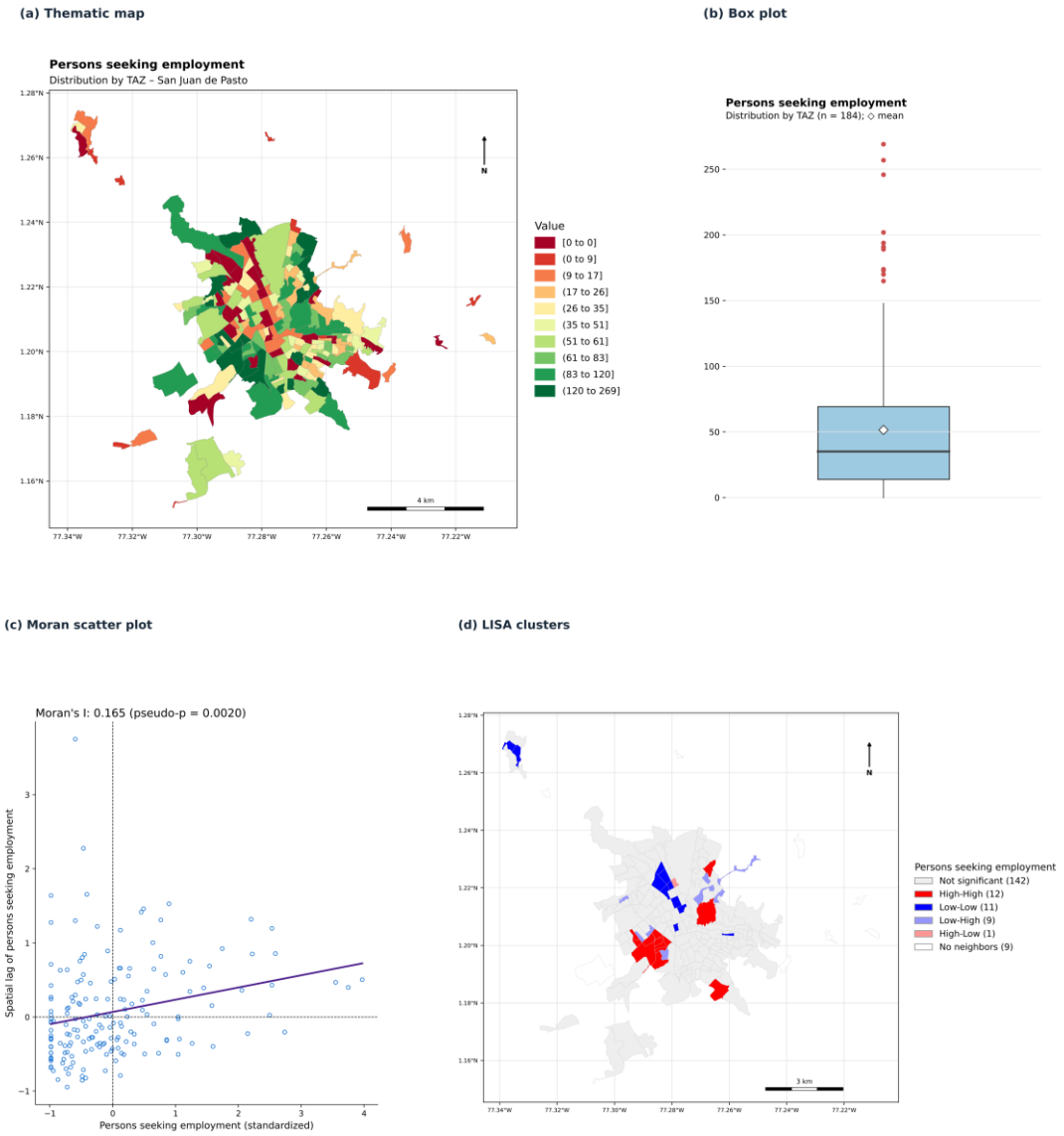


Figure 8. Persons seeking employment: (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.8 Estimated monthly household income

Estimated monthly income has a median of COP 1,617,153 and a mean of COP 1,756,215 across TAZs; its CV, of 41.9%, is the lowest, and its maximum, COP 4,601,023, is in TAZ 109098. It has the highest Moran's I (I = 0.455; p = 0.0001), with 19 HH and 21 LL zones (Figure 9): high values cluster in the center-west and northwest, and low values in the south and northeast. It is negatively associated with households in strata 1 and 2 ($\rho = -0.42$) and positively with commercial establishments (0.42) and supply (0.38). As a zonal mean, it does not express inequality between households.

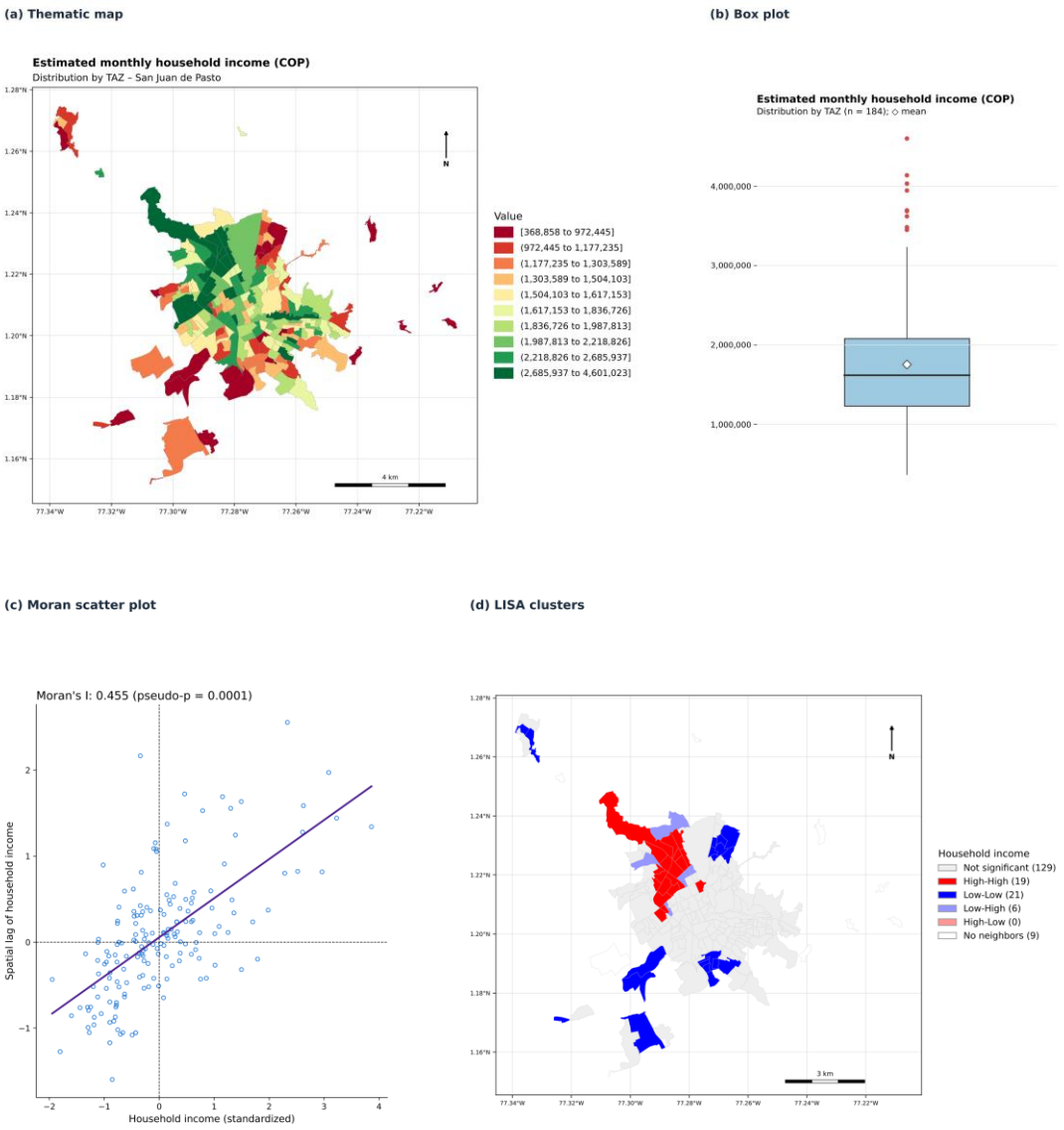


Figure 9. Estimated monthly household income (COP): (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.2.9 Aggregate service frequency (SETP)

Scheduled supply, expressed as a sum of headways in minutes, has a median of 26 (mean of 30.3; CV = 73.7%) and a maximum of 112 in TAZs 101154 and 101157; nine zones record zero. Its Moran's I is 0.448 (p = 0.0001), with 28 HH zones, the largest number among the variables, 10 LL and 6 LH (Figure 10): it clusters high values in the center and south, and low values in the north and periphery. It is associated with commercial establishments (ρ = 0.58) and income (0.38), and not with households without a vehicle (0.02), disability (0.01), job seeking (−0.05) or households in strata 1 and 2 (−0.16). The 17 zones with incomplete correspondence limit its precision.

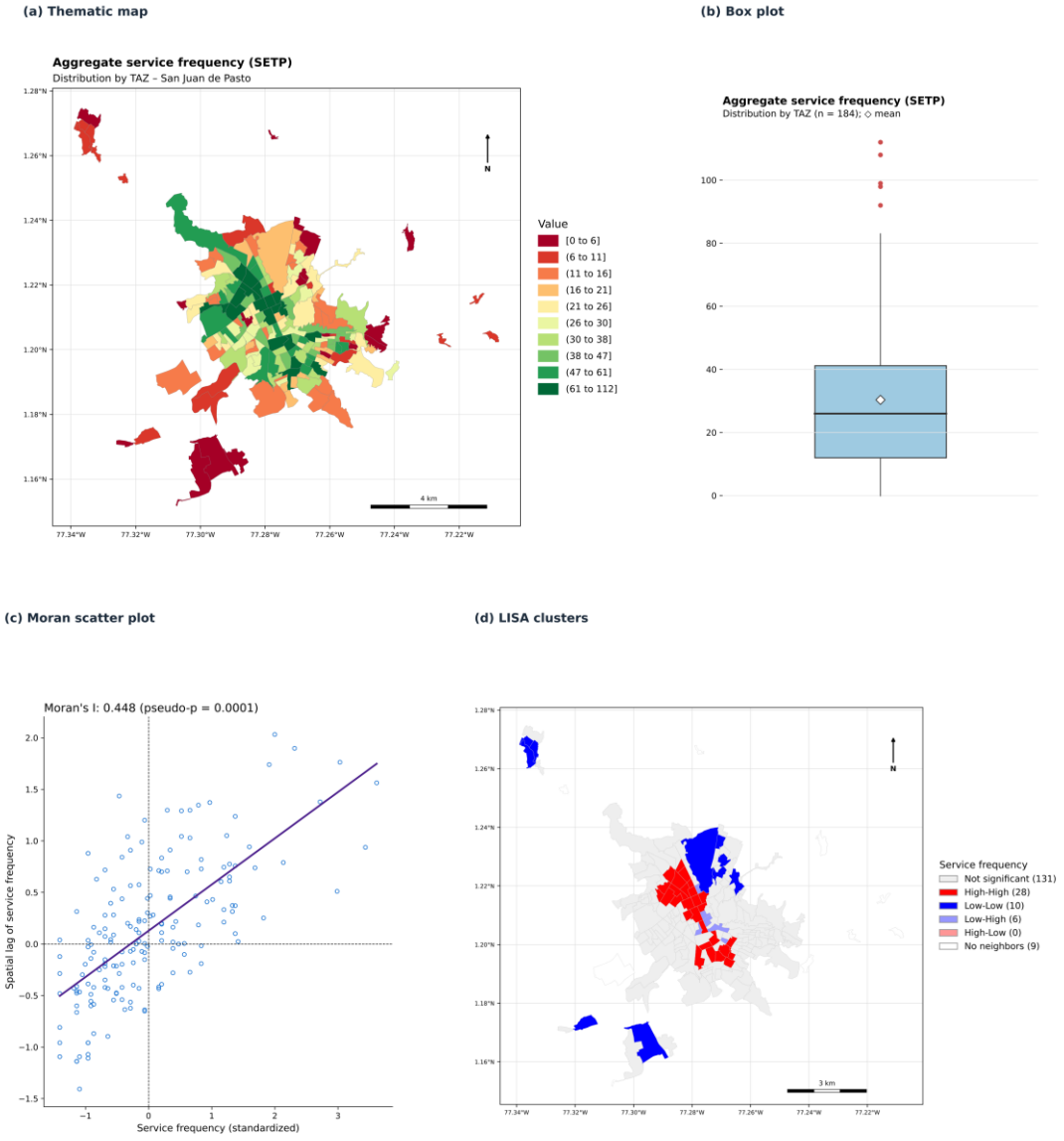


Figure 10. Aggregate service frequency (SETP): (a) thematic map, (b) box plot, (c) Moran scatter plot and (d) LISA clusters.

4.3 Associations between variables

Figure 11 presents the Spearman matrix computed with the observed data. The residential volume variables are positively and significantly associated with one another (ρ between 0.40 and 0.79), a covariation consistent with the shared size of the zones that does not allow the conclusion that the same persons combine these conditions (Pearson, 1897; Robinson, 1950). Territorial resources (supply, commercial establishments and income) are also associated with one another (ρ = 0.58, 0.42 and 0.38).

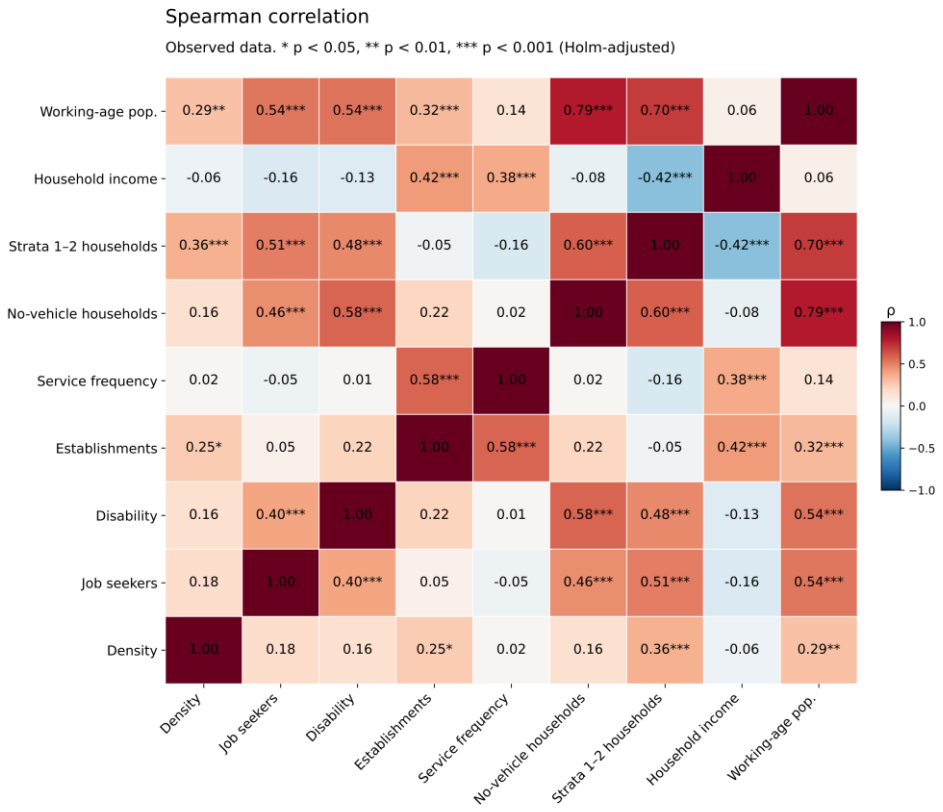


Figure 11. Spearman correlations between the nine variables, computed with the observed data (without imputed zones). Colors represent ρ; * p < 0.05; ** p < 0.01; *** p < 0.001, with Holm adjustment. Working-age pop. corresponds to the population aged 19 to 65; Strata 1–2 households, to households in strata 1 and 2; No-vehicle households, to households without a vehicle; Job seekers, to persons seeking employment; and Service frequency, to the aggregate service frequency (scheduled supply).

Between the two groups, supply is not associated with households without a vehicle, disability or job seeking (ρ between −0.05 and 0.02), and income is associated, negatively, only with households in strata 1 and 2 (−0.42); density is associated with the volumes and with commercial establishments, but not with service. These bivariate relationships are zonal and do not identify individual effects. In the zone-by-zone comparison, the operational pattern does not reproduce the distribution of households with mobility constraints; Section 4.6 examines this relationship taking into account the surroundings of each zone.

4.4 Global and local autocorrelation across variables

Table 4. Global autocorrelation and distribution of LISA clusters.

Variable	Moran's I	Pseudo-p	HH	LL	LH	HL
Population aged 19 to 65	0.236	0.0001	13	11	14	5
Households in strata 1 and 2	0.402	0.0001	19	30	8	2
Density (inhab./ha)	0.344	0.0001	11	21	3	3
Commercial establishments	0.070	0.0105	10	14	6	0

Households without a vehicle	0.153	0.0026	7	12	7	3
Persons with disabilities	0.077	0.0524	2	6	6	1
Persons seeking employment	0.165	0.0020	12	11	9	1
Household income	0.455	0.0001	19	21	6	0
Service frequency (SETP)	0.448	0.0001	28	10	6	0

Note. HH: High-High; LL: Low-Low; LH: Low-High; HL: High-Low. Queen contiguity and 9,999 permutations; nominal local significance at 5%. The nine isolated TAZs (without contiguous neighbors) are excluded from the network of 175 zones. Results include the 18 imputed TAZs; their exclusion is examined in Table 5.

Eight variables show positive and significant global autocorrelation; disability is the exception (Table 4). Income, supply and households in strata 1 and 2 have the highest indices, whereas commercial establishments combine the highest CV and the highest kurtosis with the lowest significant I: the weight of the extremes does not anticipate similarity between neighbors. LISA counts describe classified zones, not independent clusters, and, being univariate, they do not test the overlap between social and operational patterns, which is examined in Section 4.6.

4.5 Exploratory sensitivity to the imputed zones

Without the imputed zones, income goes from I = 0.455 to 0.437, supply from 0.448 to 0.441 and households in strata 1 and 2 from 0.402 to 0.412 (Table 5): the three retain the highest indices, although supply and income swap first place. Disability shows the largest reduction (from 0.077 to 0.038). The comparison does not isolate the effect of imputation, because the network also changes from 175 to 161 zones.

Table 5. Descriptive sensitivity of Moran's I to the exclusion of the imputed zones.

Variable	I with 184 TAZs (175 in the network)	I with 166 TAZs without imputation (161 in the network)	Difference
Households in strata 1 and 2	0.402	0.412	0.010
Persons with disabilities	0.077	0.038	−0.039
Household income	0.455	0.437	−0.018
Service frequency (SETP)	0.448	0.441	−0.007

Note. The three variables with the highest indices and disability, the only variable without global significance at 5%, are presented. Differences calculated from the coefficients presented, without new permutation tests.

4.6 Bivariate spatial association between social conditions, resources and supply

The relationships between resources hold when neighborhood is incorporated (Table 6):

zones with more commercial establishments and higher income are surrounded by more supply (I = 0.248 and 0.289; p = 0.0001), and income is associated with neighboring commercial establishments (I = 0.148; p = 0.0136). The negative relationship between households in strata 1 and 2 and income is also spatial (I = −0.261; p = 0.0001): 15 HL zones concentrate these households in a low-income environment, and 22 LH zones contain few of them in a high-income environment (Figure 12a).

The most relevant result appears when social conditions are contrasted with supply. Although the Spearman correlation between households in strata 1 and 2 and supply is not significant (ρ = −0.16), their bivariate spatial association is negative and significant (I = −0.235; p = 0.0001): zones with more households in these strata tend to be surrounded by lower supply. The same occurs, with lower intensity, with job seeking (I = −0.101; p = 0.0050), whereas households without a vehicle and disability show no significant global association (I = −0.048 and −0.047).

Table 6. Bivariate Moran and bivariate LISA clusters for selected pairs.

Pair (x in the zone – lag of y)	Bivariate I	Pseudo-p	HH	LL	LH	HL
Commercial establishments – scheduled supply	0.248	0.0001	22	10	13	1
Income – commercial establishments	0.148	0.0136	10	11	5	3
Income – scheduled supply	0.289	0.0001	22	9	13	1
Households in strata 1 and 2 – income	−0.261	0.0001	2	7	22	15
Households in strata 1 and 2 – scheduled supply	−0.235	0.0001	6	6	29	5
Households without a vehicle – scheduled supply	−0.048	0.1237	13	6	22	5
Job seeking – scheduled supply	−0.101	0.0050	7	8	28	3
Disability – scheduled supply	−0.047	0.1218	14	8	21	3

Note. x: variable in the zone; y: variable averaged over the neighboring zones. HH: high x and surroundings with high y; LL: low x and surroundings with low y; LH: low x and surroundings with high y; HL: high x and surroundings with low y. Network of 175 zones, queen contiguity, 9,999 permutations (conditional for LISA) and nominal local significance at 5%; includes the 18 imputed TAZs.

At the local level, five zones (109096, 110102, 110108, 112109 and 213014) are High-Low both for households in strata 1 and 2 and for households without a vehicle, and three of

them also for job seeking (Figures 12c and 12d): they concentrate these conditions in an environment of significantly low supply. Four are in the north and northeast, between 1.6 and 2.2 km from the core of highest supply, and the fifth is on the southern periphery; none is imputed. With 2.9% of the zones in the network, they contain 6.8% of the households in strata 1 and 2 and 5.1% of the households without a vehicle. The 13 HH zones of households without a vehicle, in the center-west and the south, combine instead concentration and high supply in the surroundings.

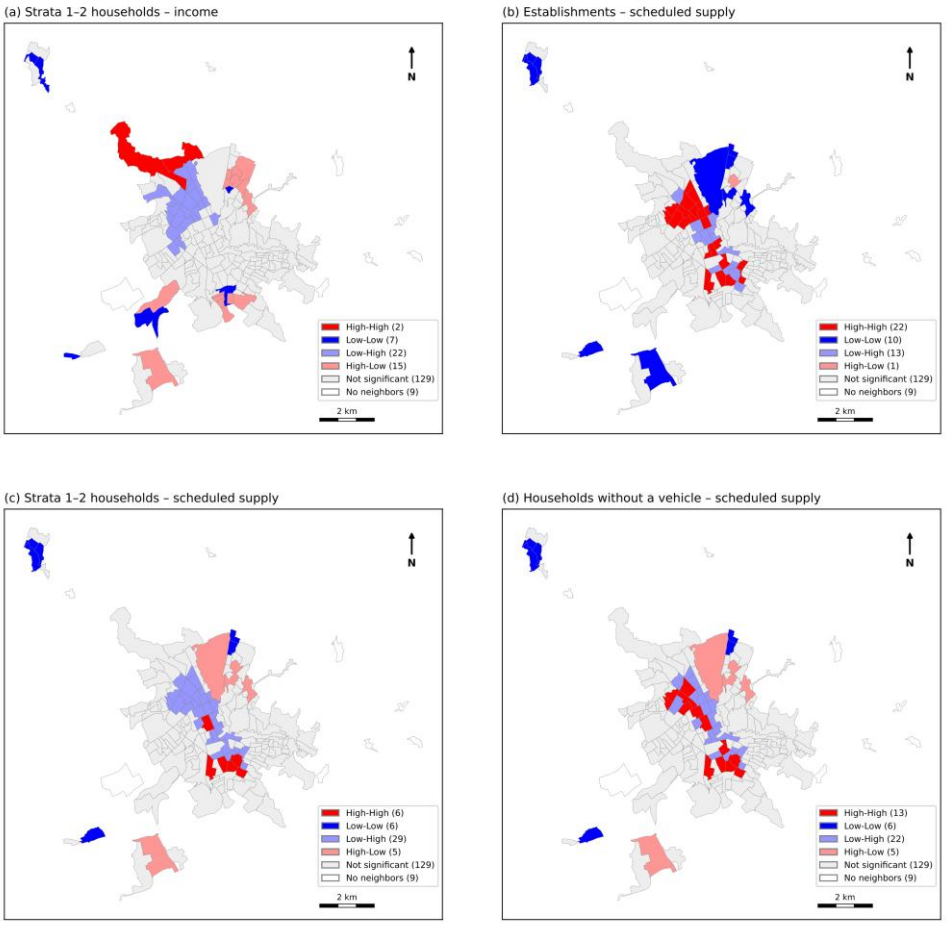


Figure 12. Bivariate LISA: (a) households in strata 1 and 2 and spatial lag of income; (b) commercial establishments and lag of scheduled supply; (c) households in strata 1 and 2 and lag of scheduled supply; (d) households without a vehicle and lag of scheduled supply. The first variable refers to the zone and the second to the average of its neighbors.

These results compare each zone with its surroundings and depend on the neighborhood used; since supply is strongly associated with the number of routes ($\rho = 0.96$), an environment of low supply indicates, above all, few routes in the neighboring zones.

4.7 Robustness of the bivariate analysis and mismatch zones

Table 7 compares the bivariate I across five specifications. The negative association between households in strata 1 and 2 and the supply of the surroundings is significant in all of them (I between -0.182 and -0.398; $p \leq 0.0001$) and is stronger with household rates ($I = -0.398$), indicating that the mismatch is not a mere effect of zone size. The relationship with job seeking is also negative and significant in all five (I between -0.081 and -0.140).

Households without a vehicle show a significant negative association only without the imputed zones and with rates (I = −0.089 and −0.087), and disability only without the imputed zones (I = −0.083), so their relationship with the supply of the surroundings is weak. The relationships between resources and between strata 1 and 2 and income hold in all variants, and the reverse direction reproduces the signs with similar magnitudes.

Table 7. Robustness of the bivariate Moran's I to five specifications.

Pair (x in the zone – lag of y)	Base	Without imputed	Household rates	6 neighbors	Reverse direction
Households in strata 1 and 2 – supply	−0.235***	−0.281***	−0.398***	−0.182***	−0.284***
Job seeking – supply	−0.101**	−0.140***	−0.082*	−0.081*	−0.138**
Households without a vehicle – supply	−0.048	−0.089*	−0.087*	−0.017	−0.070
Disability – supply	−0.047	−0.083*	−0.032	−0.016	−0.065
Commercial establishments – supply	0.248***	0.239***	—	0.229***	0.247***
Income – supply	0.289***	0.264***	—	0.274***	0.291***
Households in strata 1 and 2 – income	−0.261**	−0.291***	−0.463**	−0.246**	−0.305**

Note. Base: counts, queen contiguity and 175 zones (Table 6). Without imputed: 161 connected zones. Household rates: conditions divided by the total number of census households in the zone; not applicable (—) to commercial establishments and income. 6 neighbors: six nearest neighbors, 184 zones. Reverse direction: the second variable in the zone and the first in the surroundings. 9,999 permutations; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Because the local classification is more sensitive than the global coefficient, a robust core was defined with the zones that are High-Low, for households in strata 1 and 2 or households without a vehicle, in at least three of the four specifications that preserve the direction of the pair: 110102, 110108, 112109 and 213014 (Table 8 and Figure 13). Zones 109096 and 110097 appear in two specifications and are considered sensitive; none of the six is imputed.

Table 8. Profile of the zones of mismatch between need and supply.

Zone	Pop. aged 19 to 65	Households S1-S2 (%)	Households without a vehicl	Income (thousand	Supply (min; routes)	Specif. with HL

			e (%)	CO P)		
110102	2066	922 (91.1)	490 (48.4)	845	30; 4	3 of 4
110108	1014	398 (88.6)	238 (52.9)	112 8	30; 4	4 of 4
112109	2324	1105 (93.2)	385 (32.5)	178 0	22; 3	4 of 4
213014	2129	880 (87.0)	506 (50.0)	123 8	6; 1	4 of 4
109096	2812	1124 (90.4)	483 (38.9)	202 9	17; 2	2 of 4
110097	2617	1149 (88.5)	711 (54.7)	926	6; 1	2 of 4
Median of the 184 TAZs	1077	306 (83.6)	200.5 (40.4)	161 7	26; 3	—

Note. Households as a number and, in parentheses, as a percentage of the census households in the zone. Supply: sum of headways (minutes) and number of routes. Specif. with HL: specifications (base, without imputed, rates and 6 neighbors) in which the zone is High-Low for households in strata 1 and 2 or without a vehicle, with a permutation sampling independent of that of Table 6.

The core zones have between 1,014 and 2,324 persons aged 19 to 65, proportions of households in strata 1 and 2 between 87.0% and 93.2%, above the median (83.6%), and, except for 112109, an income below the median. They contain 5.0% of the households in strata 1 and 2 and 3.9% of the households without a vehicle in 2.3% of the zones. Their own supply is not always low: 110102 and 110108 exceed the median (30 versus 26), which shows that the mismatch lies in the surroundings, whereas 213014 has a single route.

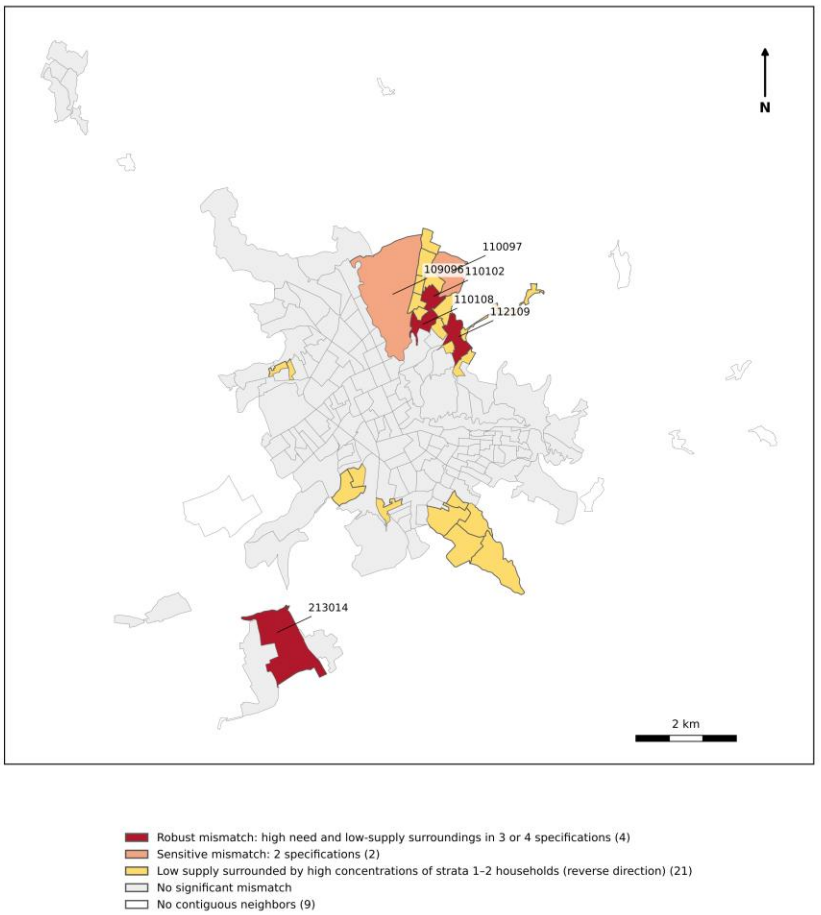


Figure 13. Synthesis of the mismatch between need and supply: robust core (High-Low in three or four specifications), sensitive zones (two specifications) and low-supply zones surrounded by a high concentration of households in strata 1 and 2 (reverse direction).

In the reverse direction, 25 zones with below-average supply are surrounded by high concentrations of households in strata 1 and 2; their median supply is 17, with two routes, compared with 26 and three in the city, and they are home to 24.0% of these households and 18.9% of the households without a vehicle. Most form a continuous sector in the northeast, which includes TAZ 110131, and a second group is located in the southeast (Figure 13): the mismatch shapes sectors of the city and not only isolated zones.

5. DISCUSSION

Regarding the first question, all variables have positive skewness, but the weight of their extremes and their clustering vary independently. Regarding the second, residential volumes and territorial resources covary within each group, also when neighborhood is considered. Regarding the third, the relationship between the two groups is weak in the non-spatial correlation, but when the surroundings are considered, the zones with more households in strata 1 and 2 and more persons seeking employment are surrounded by lower supply, in a core of four zones and a sector in the northeast.

5.1 Dispersion is not equivalent to clustering

The difference between dispersion and spatial continuity confirms that box plots and kurtosis describe the shape of a distribution, whereas Moran and LISA provide its geographic structure (Anselin, 1995; Bivand & Wong, 2018). Ranking zones by their

extremes points to individual outliers, such as TAZ 101152; detecting contiguous groups points to territories, such as those of income or supply. A diagnosis that reports only one of the two describes a different territory from one that reports both.

5.2 Residential volumes and territorial resources

The concentration of maxima in TAZ 110131 indicates an accumulation of volumes, not that the same persons combine these characteristics (Robinson, 1950), and the covariation between volumes may be due to zone size (Pearson, 1897; Kronmal, 1993). For this reason, the use of rates in the robustness test (Section 4.7) is relevant: the mismatch between strata 1 and 2 and supply becomes stronger with rates, and the negative relationship between these households and income, an average per household, does not depend on zone size.

Regarding the expectations, E1 is supported by the contrast between commercial establishments and income and by the lack of correspondence between the CV or kurtosis and Moran's I. E2 is partially met: households in strata 1 and 2 are negatively associated with income, but not with commercial establishments nor, in the non-spatial correlation, with supply, although the bivariate analysis shows that they are surrounded by lower supply. E3 is met: the correlation between these households and supply is not significant ($\rho = -0.16$), but their bivariate spatial association is negative and robust (I between -0.182 and -0.398).

5.3 Social conditions, destinations and transport supply

Modal, functional and labor constraints present different geometries: disability does not form significant clusters, so a selection of zones based only on large clusters would omit this heterogeneity, which supports distinguishing social profiles in transport diagnoses (Currie, 2010; Foth et al., 2013; Pereira et al., 2017) and a specific dimension for functional barriers (Park & Chowdhury, 2022).

The association between commercial establishments, supply and income indicates that scheduled supply follows the location of activities and income rather than that of mobility constraints, in line with Colombian evidence (Bocarejo & Oviedo, 2012; Arellana et al., 2021) and with studies in Bogotá, Cali, Montevideo and Santiago, where public transport accessibility partly reproduces the socio-spatial structure (Delmelle & Casas, 2012; Guzman et al., 2017; Hernández, 2018; Tiznado-Aitken et al., 2018). Establishing whether this responds to planning criteria or to network design, and assessing the actual benefits, requires measuring destinations, travel times and travel costs (Geurs & van Wee, 2004; Grengs, 2015).

The bivariate analysis qualifies this reading: the absence of correlation is not equivalent to spatial neutrality. The neighborhood structure, which the article shows to be relevant for each variable, is also relevant for pairs of variables, and it is what makes it possible to locate the mismatches. The result is consistent with studies of social need (Currie, 2010; Foth et al., 2013) and of public transit deserts (Jiao & Dillivan, 2013), and it specifies that, in an intermediate city, the gap may not be observed in the zone itself, but in the route coverage of its surroundings.

5.4 Implications for planning in intermediate cities

Five implications follow for the management of an intermediate city: interpreting volume maps as concentrations of persons or households, in line with the recommendation to incorporate explicit distributive measures into transport plans (Karner & Niemeier, 2013; Boisjoly & El-Geneidy, 2017); presenting dispersion statistics and spatial statistics together; giving specific treatment to functional conditions such as disability; cleaning supply data before ranking zones; and assessing the correspondence between need and supply with bivariate spatial statistics, which in Pasto points to the core of four zones and the northeastern sector as priority candidates for reviewing the layout and coverage of the SETP.

6. CONCLUSIONS

The study supports the following conclusions on the socio-spatial structure of San Juan de Pasto and the correspondence between residential need and the scheduled supply of the SETP.

Statistical dispersion and spatial clustering are distinct properties, and a diagnosis that reports only one of them describes an incomplete territory. Commercial establishments are the most unequal variable across zones (CV = 246.3%; kurtosis = 129.42), but the least clustered (I = 0.070): their map is dominated by a single peak in TAZ 101152, whose value is about 56 times the median, and 145 of the 175 zones show no local significance (Figure 5). Income, by contrast, is the least dispersed variable (CV = 41.9%) and the most clustered (I = 0.455), with contiguous blocks of high income in the center-west and northwest and of low income in the south and northeast (Figure 9).

The resident population and social conditions make up a geography of their own, concentrated in the north, the northeast and the south. The volumes of population aged 19 to 65, households in strata 1 and 2, households without a vehicle and persons seeking employment covary with one another (ρ between 0.40 and 0.79) and reach their maxima in TAZ 110131 (Figures 2, 3, 6 and 8). This geography is not homogeneous: population has the largest number of neighborhood contrasts (19 LH or HL zones), and households without a vehicle form foci in the northeast and the south that do not constitute a continuous band, so residential concentration changes rapidly between nearby zones.

Territorial resources make up a second geography, concentrated in the center. Scheduled supply has the largest number of High-High zones in the set (28), in the center and the south, and low values in the north and the periphery (Figure 10); it is associated with commercial establishments (ρ = 0.58) and income (ρ = 0.38), and the bivariate analysis confirms that zones with more commercial establishments and higher income are surrounded by more supply (I = 0.248 and 0.289; Figure 12b). The SETP network therefore follows the location of activities and income, whereas density, which clusters in eastern foci, is not associated with supply (ρ = 0.02; Figure 4): the densest zones do not necessarily receive more service.

The city shows a sharp socio-spatial polarization between income and stratum. Households in strata 1 and 2 are negatively associated with income both in the correlation ($\rho = -0.42$)

and in the bivariate analysis (I = −0.261; −0.463 with rates), and the bivariate map distinguishes 15 zones that concentrate these households in a low-income environment and 22 zones with few households in these strata in a high-income environment (Figure 12a). Inequality is therefore expressed not in isolated zones, but in contiguous blocks of the city.

Non-spatial correlation conceals a mismatch between need and supply that the bivariate analysis makes visible. The correlation between households in strata 1 and 2 and supply is not significant (ρ = −0.16), but the zones with more households in these strata are surrounded by lower supply (I = −0.235), a relationship that becomes stronger when the condition is expressed as a proportion of households (I = −0.398) and that holds without the imputed zones, with another definition of neighborhood and in the reverse direction (Table 7). Persons seeking employment show the same pattern, with lower intensity, in all five specifications. The third expectation of the study is confirmed: the absence of correlation is not equivalent to the absence of spatial mismatch.

The mismatch has a precise location and shapes sectors of the city (Figure 13). Four zones (110102, 110108, 112109 and 213014) combine, in at least three of four specifications, high concentrations of households in strata 1 and 2 and of households without a vehicle with an environment of significantly low supply; with 2.3% of the zones in the network, they contain 5.0% of the households in strata 1 and 2 (Table 8). In the reverse direction, 25 zones with below-average supply, surrounded by high concentrations of these households, form a continuous sector in the northeast, which includes TAZ 110131, and a second group in the southeast; they are home to 24.0% of the households in strata 1 and 2 and 18.9% of the households without a vehicle, with a median supply of 17 and two routes, compared with 26 and three routes in the city as a whole.

The problem is, above all, one of coverage of the surroundings and not only of the zone. Two of the core zones, 110102 and 110108, have their own supply above the median, but are surrounded by zones with few routes, whereas 213014, on the southern periphery, depends on a single route. Since scheduled supply is closely associated with the number of routes (ρ = 0.96), improving the correspondence between need and supply in these sectors requires reviewing the extent and connectivity of the network, and not only the frequency of the existing routes.

Not all mobility constraints behave in the same way. Disability does not form significant clusters (I = 0.077; p = 0.0524), it is the variable most sensitive to imputation and its relationship with the supply of the surroundings is weak (Figure 7 and Table 7). Since it is not concentrated in defined sectors, addressing it cannot rely on territorial targeting and requires universal accessibility criteria applied to the network as a whole. Households without a vehicle, for their part, show a mismatch with the supply of the surroundings only in some specifications, which suggests treating them as a complementary condition rather than as the main indicator of need.

The results are stable with respect to the main measurement decisions. When the 18 imputed zones are excluded, scheduled supply, income and households in strata 1 and 2

retain the three highest Moran indices (Table 5), and the mismatch between strata 1 and 2 and supply is significant in all five specifications of the bivariate analysis. Supply, measured as a sum of scheduled headways, should be read as route coverage, and the nine zones without assigned service and the 17 with incomplete correspondence between stops and routes should be cleaned before any ranking of zones.

The study offers a methodological lesson and a planning tool. The way variables are read defines the territory that a diagnosis makes visible: reporting dispersion, univariate clustering and bivariate spatial association together prevents a single dimension from being taken for the whole and makes it possible to move from stating that there is no correlation to identifying where the underserved zones are. For the transport authority and municipal planning, the core of four zones and the northeastern and southeastern sectors are the priority candidates for reviewing the layout and coverage of the SETP, and the zonal database constructed constitutes the baseline for measuring, in later stages, accessibility to employment centers and territorial equity in Pasto.

REFERENCES

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis, 27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical Analysis, 38*(1), 5–22. https://doi.org/10.1111/j.0016-7363.2005.00671.x

Anselin, L., Syabri, I., & Smirnov, O. (2002). Visualizing multivariate spatial correlation with dynamically linked windows. In L. Anselin & S. Rey (Eds.), *New tools for spatial data analysis: Proceedings of the Specialist Meeting*. Center for Spatially Integrated Social Science, University of California, Santa Barbara.

Arellana, J., Oviedo, D., Guzman, L. A., & Alvarez, V. (2021). Urban transport planning and access inequalities: A tale of two Colombian cities. *Research in Transportation Business & Management, 40*, Article 100554. https://doi.org/10.1016/j.rtbm.2020.100554

Banister, D. (2008). The sustainable mobility paradigm. *Transport Policy, 15*(2), 73–80. https://doi.org/10.1016/j.tranpol.2007.10.005

Bivand, R. S., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST, 27*(3), 716–748. https://doi.org/10.1007/s11749-018-0599-x

Bocarejo S., J. P., & Oviedo H., D. R. (2012). Transport accessibility and social inequities: A tool for identification of mobility needs and evaluation of transport investments. *Journal of Transport Geography, 24*, 142–154. https://doi.org/10.1016/j.jtrangeo.2011.12.004

Boisjoly, G., & El-Geneidy, A. M. (2017). How to get there? A critical assessment of accessibility objectives and indicators in metropolitan transportation plans. *Transport Policy, 55*, 38–50. https://doi.org/10.1016/j.tranpol.2016.12.011

Casas, I. (2007). Social exclusion and the disabled: An accessibility approach. *The Professional Geographer, 59*(4), 463–477. https://doi.org/10.1111/j.1467-9272.2007.00635.x

Concejo Municipal de Pasto. (2015). *Acuerdo 004 del 14 de abril de 2015: Por el cual se adopta el Plan de Ordenamiento Territorial del Municipio de Pasto 2015–2027, Pasto Territorio Con-Sentido* [Agreement 004 of April 14, 2015: Adopting the Land Use Plan of the Municipality of Pasto 2015–2027]. https://www.pasto.gov.co/index.php/pot/396-pot-2015-2027

Consorcio Pasto 2020. (2020). *Diseño operacional del SETP de Pasto: Informe ejecutivo* [Operational design of the Pasto SETP: Executive report; Technical report, December 2020, prepared for the Unidad Administrativa Especial AVANTE SETP]. Movilidad Sostenible.

Currie, G. (2010). Quantifying spatial gaps in public transport supply based on social needs. *Journal of Transport Geography, 18*(1), 31–41. https://doi.org/10.1016/j.jtrangeo.2008.12.002

Delbosc, A., & Currie, G. (2011). Using Lorenz curves to assess public transport equity. *Journal of Transport Geography, 19*(6), 1252–1259. https://doi.org/10.1016/j.jtrangeo.2011.02.008

Delmelle, E. C., & Casas, I. (2012). Evaluating the spatial equity of bus rapid transit-based accessibility patterns in a developing country: The case of Cali, Colombia. *Transport Policy, 20*, 36–46. https://doi.org/10.1016/j.tranpol.2011.12.001

Departamento Administrativo Nacional de Estadística [DANE]. (2018). *Censo Nacional de Población y Vivienda 2018 (CNPV 2018)* [2018 National Population and Housing Census]. https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivienda-2018

Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American Planning Association, 76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Foth, N., Manaugh, K., & El-Geneidy, A. M. (2013). Towards equitable transit: Examining transit accessibility and social need in Toronto, Canada, 1996–2006. *Journal of Transport Geography, 29*, 1–10. https://doi.org/10.1016/j.jtrangeo.2012.12.008

Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A, 23*(7), 1025–1044. https://doi.org/10.1068/a231025

Fransen, K., Neutens, T., Farber, S., De Maeyer, P., Deruyter, G., & Witlox, F. (2015). Identifying public transport gaps using time-dependent accessibility levels. *Journal of Transport Geography, 48*, 176–187. https://doi.org/10.1016/j.jtrangeo.2015.09.008

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography, 12*(2), 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005

González Bastidas, A. M. (2021). *Territorio y ciudad. Del paisaje agrícola al paisaje urbano en los modelos de ocupación de Pasto 2000-2020* [Territory and city: From the agricultural landscape to the urban landscape in the land occupation models of Pasto 2000–2020; Master's thesis, Universidad Nacional de Colombia]. https://repositorio.unal.edu.co/handle/unal/80898

Grengs, J. (2015). Nonwork accessibility as a social equity indicator. *International Journal of Sustainable Transportation, 9*(1), 1–14. https://doi.org/10.1080/15568318.2012.719582

Guzman, L. A., & Oviedo, D. (2018). Accessibility, affordability and equity: Assessing 'pro-poor' public transport subsidies in Bogotá. *Transport Policy, 68*, 37–51. https://doi.org/10.1016/j.tranpol.2018.04.012

Guzman, L. A., Oviedo, D., & Rivera, C. (2017). Assessing equity in transport accessibility to work and study: The Bogotá region. *Journal of Transport Geography, 58*, 236–246. https://doi.org/10.1016/j.jtrangeo.2016.12.016

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners, 25*(2), 73–76. https://doi.org/10.1080/01944365908978307

Hernández, D. (2018). Uneven mobilities, uneven opportunities: Social distribution of public

transport accessibility to jobs and education in Montevideo. *Journal of Transport Geography, 67*, 119–125. https://doi.org/10.1016/j.jtrangeo.2017.08.017

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70. https://www.jstor.org/stable/4615733

Javanmard, R., Lee, J., Kim, J., Liu, L., & Diab, E. (2023). The impacts of the modifiable areal unit problem (MAUP) on social equity analysis of public transit reliability. *Journal of Transport Geography, 106*, Article 103500. https://doi.org/10.1016/j.jtrangeo.2022.103500

Jiao, J., & Dillivan, M. (2013). Transit deserts: The gap between demand and supply. *Journal of Public Transportation, 16*(3), 23–39. https://doi.org/10.5038/2375-0901.16.3.2

Kain, J. F. (1968). Housing segregation, Negro employment, and metropolitan decentralization. *The Quarterly Journal of Economics, 82*(2), 175–197. https://doi.org/10.2307/1885893

Karner, A., & Niemeier, D. (2013). Civil rights guidance and equity analysis methods for regional transportation plans: A critical review of literature and practice. *Journal of Transport Geography, 33*, 126–134. https://doi.org/10.1016/j.jtrangeo.2013.09.017

Kronmal, R. A. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society: Series A, 156*(3), 379–392. https://doi.org/10.2307/2983064

Lee, S.-I. (2001). Developing a bivariate spatial association measure: An integration of Pearson's r and Moran's I. *Journal of Geographical Systems, 3*(4), 369–385. https://doi.org/10.1007/s101090100064

Lilliefors, H. W. (1967). On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association, 62*(318), 399–402. https://doi.org/10.1080/01621459.1967.10482916

Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy, 20*, 105–113. https://doi.org/10.1016/j.tranpol.2012.01.013

Lucas, K., van Wee, B., & Maat, K. (2016). A method to evaluate equitable accessibility: Combining ethical theories and accessibility-based approaches. *Transportation, 43*(3), 473–490. https://doi.org/10.1007/s11116-015-9585-2

Mesías Rosas, O. (2024). Nodos de desarrollo rururbanos: una alternativa para la sostenibilidad territorial. Caso de estudio: cabeceras corregimentales del municipio de Pasto, Colombia [Rural–urban development nodes: An alternative for territorial sustainability. Case study: Township seats of the municipality of Pasto, Colombia]. *Seminario Internacional de Investigación en Urbanismo*. https://doi.org/10.5821/siiu.13079

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*(1–2), 17–23. https://doi.org/10.1093/biomet/37.1-2.17

NIST/SEMATECH. (n.d.). Measures of skewness and kurtosis. In *e-Handbook of Statistical Methods*. https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm

Openshaw, S. (1984). *The modifiable areal unit problem* (Concepts and Techniques in Modern Geography, No. 38). Geo Books.

Oviedo, D., & Guzman, L. A. (2020). Revisiting accessibility in a context of sustainable transport: Capabilities and inequalities in Bogotá. *Sustainability, 12*(11), Article 4464. https://doi.org/10.3390/su12114464

Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: Positive and normative implementations of various accessibility indicators. *Journal of Transport Geography, 25*, 141–153. https://doi.org/10.1016/j.jtrangeo.2012.03.016

Park, J., & Chowdhury, S. (2022). Towards an enabled journey: Barriers encountered by public

transport riders with disabilities for the whole journey chain. *Transport Reviews, 42*(2), 181–203. https://doi.org/10.1080/01441647.2021.1955035

Pearson, K. (1897). Mathematical contributions to the theory of evolution.—On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London, 60*, 489–498. https://doi.org/10.1098/rspl.1896.0076

Pereira, R. H. M., Schwanen, T., & Banister, D. (2017). Distributive justice and equity in transportation. *Transport Reviews, 37*(2), 170–191. https://doi.org/10.1080/01441647.2016.1257660

Recalde Martínez, J. H. (2019). *Análisis de la incidencia social para los usuarios del Sistema Estratégico de Transporte Público de la ciudad de Pasto prestado por el operador Unión Temporal Ciudad Sorpresa, en los años 2016–2017* [Analysis of the social impact on users of the Strategic Public Transport System of the city of Pasto provided by the operator Unión Temporal Ciudad Sorpresa, 2016–2017; Master's thesis, Universidad Nacional Abierta y a Distancia]. https://repository.unad.edu.co/handle/10596/27989

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*(3), 351–357. https://doi.org/10.2307/2087176

Rodríguez Vignoli, J. (2008). Movilidad cotidiana, desigualdad social y segregación residencial en cuatro metrópolis de América Latina [Daily mobility, social inequality and residential segregation in four Latin American metropolises]. *EURE, 34*(103), 49–71. http://www.eure.cl/index.php/eure/article/view/1364

Tiznado-Aitken, I., Muñoz, J. C., & Hurtubia, R. (2018). The role of accessibility to public transport and quality of walking environment on urban equity: The case of Santiago de Chile. *Transportation Research Record, 2672*(35), 129–138. https://doi.org/10.1177/0361198118782036

Vecchio, G., Tiznado-Aitken, I., & Hurtubia, R. (2020). Transport and equity in Latin America: A critical review of socially oriented accessibility assessments. *Transport Reviews, 40*(3), 354–381. https://doi.org/10.1080/01441647.2020.1711828

Villegas Alzate, J. G. (2023). Consideraciones teórico-metodológicas para el estudio de ciudades intermedias en sus atributos socioespaciales [Theoretical and methodological considerations for the study of intermediate cities in their socio-spatial attributes]. *Jangwa Pana, 22*(1), 59–76. https://doi.org/10.21676/16574923.4736