

# Obtaining The Origin-Destination Matrix In Euclidean Distances By Means Of Haversine's Law For The Transport Analysis Zones (Tazs) Of The City Of San Juan De Pasto

**Jorge Luis Argoty Burbano**

<https://orcid.org/0000-0001-6661-1398>

*Universidad de Nariño, Facultad de Ingeniería, Programa de Ingeniería Civil, Pasto, Colombia*

*Autor de correspondencia: [jlargoty@udenar.edu.co](mailto:jlargoty@udenar.edu.co)*

## ABSTRACT

Intermediate Andean cities require updated and georeferenced origin-destination (O-D) matrices to support evidence-based transport planning. San Juan de Pasto, with 184 Transport Analysis Zones (ZAT) and complex topography, lacks a spatially explicit and methodologically rigorous open-access distance matrix. This research generates a 184×184 Euclidean-geodesic distance matrix through application of the Haversine law to geographic coordinates of centroids in the WGS84 system. The matrix calculates 33,856 origin-destination pairs using spherical trigonometry with Earth's mean radius of 6,371 km, producing distances with numerical tolerance below  $10^{-6}$  km. Statistical analysis of 33,672 interzonal distances reveals a mean distance of 5.47 km (standard deviation 2.15 km, coefficient of variation 39.3%), characterizing a compact intermediate city with evident spatial stratification: densely occupied central core (mean <2.4 km), medium urban periphery (3–7 km) containing 39.78% of pairs, and periurban zones (>8 km) where spatial friction acts as an accessibility barrier. The matrix exhibits perfect symmetry without outliers, validating mathematical consistency and computational integrity. Results provide a homogeneous and replicable technical baseline that enables calibration of gravity models for trip distribution, calculation of cumulative accessibility indicators, evaluation of optimization scenarios for the Strategic Public Transport System (SETP), and territorial equity analysis in access to urban services. The methodology is cost-effective, open-source, and replicable in intermediate Latin American cities lacking automatic fare collection systems, establishing rigorous technical precedent that transcends geometric input to become a strategic tool underpinning sustainable, equitable, and innovative urban mobility in Andean contexts.

**Keywords:** origin-destination matrix, Haversine law, geodesic distances, transport zoning, urban accessibility, intermediate cities.

## 1. Introduction

Urban mobility in intermediary cities requires robust quantitative information to understand spatial interactions between territories. The Origin-Destination (OD) matrix is fundamental in transport modeling, as it synthesizes for each pair of areas the number of trips or the spatial cost of interaction, allowing analysis of demand, accessibility, and evaluation of public policies (Ortúzar & Willumsen, 2011). Its construction based on transport analysis zones (TAZs) is suitable for medium-sized cities, as they constitute homogeneous spatial groupings that integrate

socioeconomic information, modal supply and mobility in an organized geographical representation (Lovelace et al., 2019).

San Juan de Pasto, an Andean intermediate city with compact morphology, lacks spatially explicit and updated OD matrices for its 184 TAZs, limiting the analytical capacity to evaluate interzonal accessibility and detect territorial imbalances. Unlike large cities with advanced smart card-based information systems, intermediary cities rely on incomplete data and outdated surveys (Toole et al., 2012; International Transport Forum/CPB, 2015), justifying the development of low-cost, replicable and technically sound methodologies that characterize the spatial cost between TAZs as a basic input.

Haversine geodetic distance is adopted as a standard metric to measure the spatial separation between ZAT centroids, offering consistency on the earth's surface without relying on limited information on road networks (Lovelace et al., 2019). Although Euclidean distances tend to overestimate with respect to network distances, both are highly correlated and function as reasonable proxies in aggregate analyses of proximity structure (Jeon et al., 2013). At the same time, methods based on big data have generated O-D matrices for hundreds of cities (Toole et al., 2012); however, even these approaches require an explicit zoning system, so the calculation of Haversine distance matrices between ZAT centroids remains an essential structural element for any internationally comparable and reproducible mobility modeling (Ortúzar & Willumsen, 2011; Lovelace et al., 2019).

This research generates an OD matrix of geodetic distances for the 184 TAZs of Pasto through the practical application of geocomputing: transformation of the official shapefile into a manipulable data object, calculation of centroids in the WGS84 system compatible with Haversine, and construction of a symmetrical matrix  $184 \times 184$ . The main utility lies in having a homogeneous, technically transparent and replicable spatial baseline, which allows the subsequent addition of observed flow matrices, socioeconomic attributes and transport supply indicators for integrated analyses of accessibility, territorial equity and efficiency in an urban system of an Andean medium-sized city. The matrix is not intended to replace estimates with big data or surveys, but rather to constitute the minimum structural component required for mobility modeling schemes that are reproducible, comparable, and aligned with international best practices (Ortúzar & Willumsen, 2011; Lovelace et al., 2019).

## **2. Location and description of the study area**

San Juan de Pasto, capital of the department of Nariño, is located in southwestern Colombia in the Atriz valley at 2,500 meters above sea level, configuring itself as an articulating node of economic, administrative, and service flows in the southern Andean region and an international corridor to Ecuador (Cedre-Universidad de Nariño, 2024; National Planning Department, 2013). The city concentrates its demand for mobility in a consolidated urban area of approximately 26 km<sup>2</sup>, complemented by extensive rural territories (Cedre-Universidad de Nariño, 2024).

The urban morphology is based on a strongly broken topography, conditioned by the presence of the Galeras volcano, which generates marked altitudinal gradients, restrictions on urban expansion and winding road layouts that increase costs and travel times (National Planning Department, 2013). This configuration is critical for distance modeling and public transportation planning. The city administratively structures 12 urban communes and 17 rural townships, with more than 400 neighborhoods differentiated by socioeconomic status, generating marked residential segregation and pendulum travel patterns from low-income peripheries to employment, commerce, education and services centralities located in the historic center; (Ciudad Humana, 2013).

The Strategic Public Transport System (SETP) operates on a predominantly radial network with routes that converge towards the urban center, a fleet of more than 300 vehicles and commercial speeds of 16-17 km/h at rush hour, evidencing levels of congestion typical of intermediate Andean cities with accelerated urban growth (Ministry of Transport, 2024). For this research, the territory is disaggregated into 184 Transport Analysis Zones (TAZs), defined with criteria of socioeconomic homogeneity, spatial continuity, land use and functionality in the generation-attraction of trips, allowing to represent with a high level of detail the structure of mobility

demand differentiating between densely occupied central areas, low-density peripheral areas and rural sectors (National Planning Department, 2013; Cedre-Universidad de Nariño, 2024).

Pasto shares characteristics with other Andean intermediate cities, where the combination of rugged relief, perimeter expansion, and fragmented transport supply makes it essential to use rigorous distance models and origin-destination matrices to support investment decisions in strategic transport systems in contexts of sustainable and equitable mobility (CAF, 2019). The city is an ideal territorial laboratory to apply the calculation of Euclidean distances by means of Haversine's Law between ZAT centroids, integrating high-resolution geospatial information to accurately capture the spatial friction that conditions the behavior of public transport users and to provide a technical basis for the restructuring of routes, prioritization of corridors and improvement of accessibility for vulnerable populations (Ministry of Transport, 2024; Ciudad Humana, 2013).

### **3. Scope and delimitation of the research**

The research is limited to the exhaustive calculation of an origin-destination matrix in geodetic Euclidean distances among the 184 TAZs of Pasto using Haversine's Law in the WGS84 system, based on data from the Mobility Master Plan (2012-2013), building a matrix  $184 \times 184$  of 33,856 pairs that represents pure spatial costs without estimation of observed demand (National Planning Department, 2008). The spatial scope is limited to the administrative territory of Pasto without intermunicipal flows; methodologically, travel times, monetary costs, service levels, real road network models, and dynamic factors such as time congestion, seasonal variability, and real-time behavioral changes are deliberately excluded, isolating the purely geometric component of spatial separation (Abrahamsson, 1998; Zhao et al., 2024). It temporarily adopts a cross-sectional approach without exploring diachronic changes in spatial structure, representing a structural photograph of spatial friction from the period 2012-2013 that can be reused in subsequent modeling (Departamento Nacional de Planeación, 2008). The research is positioned as a basic spatial input for transport modeling in Andean intermediate cities, where precision in measuring distances is essential to evaluate sustainable transport scenarios, recognizing that incorporation of travel volumes, road network allocation, dynamic indicators and modal equity analysis constitutes a subsequent complementary agenda (Zhao et al., 2024).

### **4. Identification and description of the problem.**

San Juan de Pasto faces progressive deterioration in urban mobility conditions associated with spatial expansion, residential segregation and operation of a SETP with suboptimal levels of efficiency and accessibility for the lower-income population (National Planning Department, 2008); although there are mobility diagnoses, there is no updated, georeferenced origin-destination (O-D) matrix built with homogeneous methodological criteria from the 184 ZATs, limiting the capacity of the administration to design scenarios for the restructuring of routes and inclusive accessibility policies, since the planning has been based on partial and indirectly inferred matrices without robust matrix input that incorporates the spatial structure of the city and its complex topography, despite the fact that the international literature shows that the quality of the O-D matrix is the main determinant of reliability of urban demand models (Amézquita, LL, 2016; Galliani et al., 2024). Although methods based on big data (smart cards, GPS) have advanced globally, their adoption in Andean intermediary cities is incipient due to significant gaps in basic information in rigorous measurement of distances and spatial costs. The absence of a spatially explicit O-D matrix generates difficulties in quantifying interaction between areas, identifying accessibility bottlenecks, and evaluating the redistributive impact of transport interventions on vulnerable populations, compromising progress towards sustainable and inclusive mobility (Espacio & GIZ, 2019; CAF, 2019). The broken topography of Pasto with Galeras volcano generates a significant gap between approximate flat distances and real geodetic distances, introducing biases through flat Euclidean metrics; in view of this, the calculation of an O-D matrix in geodetic distances using Haversine's Law is proposed as an essential methodological link that accurately captures the effective spatial separation between zones and serves as the basis for subsequent models of demand, allocation and accessibility (LocationIQ, 2017; Zhao et al., 2024). The research seeks to close a critical gap between regulatory frameworks and qualitative

diagnoses of the transport system and the lack of detailed quantitative representation of spatial interactions between zones, a problem especially visible in Latin American intermediary cities where technical capacity and data inputs continue to be structural constraints for informed decision-making in urban mobility (CAF, 2019; Despacio & GIZ, 2019).

## 5. Objective

Calculate and develop an origin-destination matrix in geodetic Euclidean distances for the 184 Transport Analysis Zones (TAZs) of the municipality of Pasto, applying Haversine's Law on georeferenced coordinates of centroids, with the purpose of building a spatially and methodologically robust input that facilitates the characterization of the effective distance friction between zones, allows the evaluation of accessibility patterns and supports the planning and optimization processes of the urban public transport system in contexts of intermediary cities (Zhao et al., 2024; Galliani et al., 2024).

## 6. Theoretical Framework

Haversine's Law is a formulation derived from spherical trigonometry that allows calculating the orthodromic distance (great-circle distance) between two points on the surface of a sphere, using only their latitude and longitude coordinates expressed in radians. Historically, the term "haversine" comes from the obsolete trigonometric function "half versed sine", where  $\text{hav}(\theta) = \sin^2(\theta/2)$ , a formulation that was rediscovered and popularized in modern navigation contexts due to its numerical superiority over the spherical cosine law for short and medium distance calculations (Community ESRI, 2021; LocationIQ, 2017). From a mathematical point of view, the geodetic distance  $d$  between two points A and B on the Earth's surface, defined by their latitudes  $\phi_A, \phi_B$  and longitudes  $\lambda_A, \lambda_B$ , is obtained by the sequence of equations:

$$a = \sin^2 \left( \frac{\phi_B - \phi_A}{2} \right) + \cos(\phi_A) \cdot \cos(\phi_B) \cdot \sin^2 \left( \frac{\lambda_B - \lambda_A}{2} \right)$$

$$c = 2 \cdot \arctan 2(\sqrt{a}, \sqrt{1-a})$$

$$d = R \cdot c$$

where corresponds to the mean radius of the Earth ( $R$ ), is the central angle between the two points and the resulting geodesic distance. This formulation is algebraically equivalent to the spherical cosine law, but presents better numerical stability for short distances, avoiding cancellations associated with the direct use of the cosine for small angles, a typical situation in intra-urban distances such as those modeled between centroids of the 184 Pasto TAZs (Community ESRI, 2021; Liu & Zhu, 2004). Recent comparative studies show that, although more complex models such as Vincenty's formulas on the ellipsoid WGS84 can achieve accuracies in the order of millimeters, the Haversine offers an optimal compromise between accuracy and computational efficiency in transport planning applications where the relative error of less than 0.5% is acceptable (Mapbox, 2016; Engström, 2023).  $RR \approx 6,371cd$

From transportation theory, the origin-destination matrix is defined as a square matrix of order  $n$ , where each element represents the number of trips that originate in the zone and have the zone as their destination during a given time interval, usually an average working day. Formally:  $T_{ij}$

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}$$

where the diagonal terms include intra-zonal trips and the non-diagonal terms include inter-zonal trips. In contemporary literature, a clear distinction is made between flow matrices (trips) and cost or impedance matrices, in which each element represents the "resistance" to travel between and , expressed as distance, time, monetary cost, or weighted combinations (Zhao et al., 2024;

Rodríguez-Rueda et al., 2021). In this work, a matrix of spatial impedances of dimensions is constructed, where each element corresponds to the geodesic distance calculated by Haversine's Law between the centroids of the TAZ and . By geometric definition, this matrix is symmetrical ( $D_{ij} = D_{ji}$ ), with representing the intrazonal distance, which can be approximated to zero or an average radius of the area depending on the level of detail required (Liu & Zhu, 2004; Zhao et al., 2024). This distance matrix is the basic input for the specification of distance friction functions and for the calibration of trip distribution models.  $t_{ij} = c_{ij} D_{ij}^{184} \times 184 d_{ij} d_{ij} = d_{ij} d_{ij}$

Travel distribution models, particularly gravitational models, rely on the analogy with Newton's law of gravitation, postulating that the flow between two zones is proportional to the "weight" of those zones (their generation and attraction) and inversely proportional to an increasing friction function of the distance between them. A canonical formulation of the gravitational model is:  $T_{ij} = O_i D_j$

$$T_{ij} = k \cdot O_i \cdot D_j \cdot f(d_{ij})$$

where  $k$  is an adjustment factor and usually takes forms such as  $k = \alpha \beta$  or  $k = \alpha \beta d_{ij}^{-\alpha}$ , with empirically calibrated parameters (Zhao et al., 2024). In this framework, the distance matrix calculated with Haversine is not a mere geometric component, but a direct determinant of the shape and slope of distance friction, affecting the allocation of flows between pairs of zones and the representation of phenomena such as the preference for short-distance travel or the existence of strongly concentrated poles of attraction. In intermediary cities with marked territorial inequality, the systematic underestimation or overestimation of the distances between peripheral and central areas can lead to erroneous diagnoses of the intensity of spatial interaction and accessibility gaps (CAF, 2019; Despacio & GIZ, 2019).  $k f(d_{ij}) = e^{-\beta d_{ij}} f(d_{ij}) = d_{ij}^{-\alpha} \alpha \beta$

Zoning in Transport Analysis Zones (TAZs) is, in turn, a structural component of the theoretical framework. The specialized literature on transport planning indicates that the choice of zoning shapes the spatial "granularity" of all subsequent models, affecting the sensitivity of the results to local changes and the possibility of identifying intra-urban inequalities (Liu & Zhu, 2004). Each ZAT is defined as a relatively homogeneous spatial unit in terms of socio-economic conditions, land use and mobility patterns, and is associated with a centroid that synthesizes its location. From a geometric point of view, the centroid of the area can be defined as the weighted average of its spatial coordinates:  $i$

$$\phi_i^c = \frac{\sum_{p \in A_i} \phi_p \cdot w_p}{\sum_{p \in A_i} w_p}, \lambda_i^c = \frac{\sum_{p \in A_i} \lambda_p \cdot w_p}{\sum_{p \in A_i} w_p}$$

where  $w_p$  is a weight that may represent population, employment or another variable of interest within the area. In Pasto, the definition of 184 TAZs within the framework of the Mobility Master Plan and the Strategic Public Transport System (SETP) responds precisely to these criteria, allowing the heterogeneity between densely urbanized areas, areas of peripheral expansion and rural townships articulated to the urban system to be captured (Cedre-Universidad de Nariño, 2024; Ministry of Transport, 2024). The methodological decision to calculate distances between ZAT centroids using Haversine allows for the coherent integration of the local zonal structure with international best practices in geospatial analysis.  $w_p A_i$

Internationally, the estimation and use of O–D matrices for public transport has evolved from classic approaches based exclusively on mobility surveys and manual gauging to hybrid and machine learning methods that integrate massive data sources such as automatic collection systems (AFCs), fleet GPS records and mobile phone data. Zhao et al. (2024) propose a weighted multimodal network approach to estimating public transport O–D matrices, showing that the combination of smart card data and detailed network structures significantly improves the accuracy of the estimated matrices compared to traditional methods. Galliani et al. (2024) develop a variational method for dynamic estimation of O–D matrices in multimodal networks, integrating

disaggregated flow measurements and optimization techniques in graphs, with applications in high-demand European networks. Yang et al. (2020) formulate a deep optimization framework for estimating dynamic O–D matrices in metro systems using AFC data, achieving fine time forecasts of demand and demonstrating the potential of these inputs for advanced supply management (Yang et al., 2020). Studies such as those by Rodríguez-Rueda et al. (2021) and Kuusela (2020) explore methodologies based on conditional statistical models and machine learning to estimate or update O–D matrices from socioeconomic variables and traffic counts, reducing costs compared to extensive surveys and showing how the underlying spatial structure including the distance matrix is key to controlling the quality of estimates (Rodríguez-Rueda et al., 2021; Kuusela, 2020). These works, although focused on different networks and scales, reinforce the premise that a rigorous geometric representation of the distances between zones is an indispensable prerequisite for the adoption of advanced O–D estimation methodologies.

The concept of spatial accessibility occupies a central place in the theoretical framework, particularly in studies that link mobility, equity and urban rights. Liu and Zhu (2004) develop an integrated GIS tool for accessibility analysis in urban transport planning, distinguishing between potential accessibility, cumulative accessibility, and measures based on travel time, and show that the way in which travel distances and costs are modeled directly influences territorial equity outcomes. More recently, Shi (2021) assesses the accessibility and equity of urban transport in Kunshan, China, using weighted average accessibility indicators and the Theil index applied to both public transport and the car, concluding that public transport has less absolute accessibility but greater spatial equity than private vehicles, and proposing network adjustments to reduce "accessibility poverty" (Shi, 2021). These studies show that the combination of O–D matrices, distance matrices, and accessibility models provides a powerful analytical framework to evaluate the performance of transportation systems in terms of equity and efficiency, a particularly relevant objective in intermediate cities such as Pasto, where peripheral urban expansion and the concentration of services in the central area generate potential accessibility gaps (Cedre-Universidad de Nariño, 2024; CAF, 2019).

In Latin America, the discussion on mobility and sustainability in intermediary cities has emphasized the need for analysis tools that integrate geospatial data and transport models. CAF (2019) stresses that the implementation of integrated and strategic transport systems in medium-sized Colombian cities requires overcoming information deficits on mobility and accessibility, proposing the use of indicators of spatial coverage, travel times, and accessibility to prioritize corridors and projects. Despacio & GIZ (2019) highlight that, in cities in the region, inequality in access to opportunities manifests itself in longer and more expensive travel patterns for lower-income peripheral populations, and that accessibility analysis based on distances and travel times is key to designing policies that "leave no one behind". The research in Pasto is explicitly situated in this line, by proposing the construction of a matrix of Euclidean geodesic distances between the 184 TAZs as a baseline for future accessibility analyses, evaluation of SETP routes and territorial equity studies in access to public transport and urban services (Ministry of Transport, 2024).

Finally, from the methodological perspective, the use of Haversine's Law for the calculation of the matrix of distances between TAZs in Pasto aligns with recent developments in network analysis and computational geoscience. Technical documentation such as the "distance" package or Haversine implementations in Python and R geospatial libraries highlight the ubiquity of this formula in geographic analysis workflows and its compatibility with big data architectures (Distancia, 2023; LocationIQ, 2017). Studies such as that of Engström (2023) highlight that, while more complex formulas such as Vincenty may be preferable for high-precision scientific applications, Haversine remains a robust and computationally efficient solution for transport analysis on an urban and regional scale, especially when it comes to calculating dense matrices of distances between hundreds or thousands of points (Engström, 2023). In short, the theoretical framework presented here justifies the relevance of integrating Haversine's Law, the theory of origin-destination matrices and the approaches of accessibility and spatial equity into a single

analytical device: the O–D matrix of Euclidean geodesic distances between the 184 TAZs of Pasto, conceived as an essential technical input for the modeling of mobility, the planning of the Strategic Public Transport System and the formulation of sustainable and equitable urban mobility policies in an Andean intermediate city.

### 7. Calculations.

The determination of the calculations of the variables and the computational procedure to construct the origin-destination matrix in geodetic Euclidean distances by means of Haversine's Law for the 184 Transport Analysis Zones (TAZs) of Pasto is based on a methodological sequence that integrates GIS geoprocessing techniques, rigorous mathematical formulation and matrix calculation procedures implemented in scientific programming languages such as R or Python. aligned with the best practices reported in the international literature on estimation of O–D matrices and analysis of accessibility in urban transport networks (Liu & Zhu, 2004; Cheng et al., 2022). In the first phase, a shapefile is based on a vector shapefile containing the polygonal delimitation of the 184 TAZs of the municipality, which is loaded into the spatial analysis environment (for example, using the sf or sp packages in R, or geopandas in Python), verifying and, if necessary, reprojecting the coordinate reference system to the WGS84 geodetic system (EPSG:4326). an indispensable condition to ensure that the latitude and longitude coordinates used in Haversine's Law are consistent with standard geodetic conventions (Liu & Zhu, 2004; Community ESRI, 2021). Once the topology of the polygons has been validated (without invalid geometries, undue overlaps or unwanted gaps), the centroids of each ZAT are calculated, understood as representative points of the spatial location of each area; In geometric terms, if the area is composed of a set of vertices that define the polygon, the geometric centroid can be approximated by:

$$\phi_i^c = \frac{\sum_{p \in A_i} \phi_p \cdot w_p}{\sum_{p \in A_i} w_p}, \lambda_i^c = \frac{\sum_{p \in A_i} \lambda_p \cdot w_p}{\sum_{p \in A_i} w_p}$$

where they are weights that can take unit values (purely geometric centroid) or represent population density, employment or other relevant variable, in order to shift the centroid towards the areas with the highest concentration of transport demand (Liu & Zhu, 2004). These centroids are exported as a new point layer with attributes (decimal latitude) and (decimal longitude) for each ZAT; in R, this can be structured as a  $w_p \phi_i \lambda_i i = 1, \dots, 184$  data.frame or tibble, ensuring that all values are expressed in radians at the time of distance calculation (LocationIQ, 2017; Distance, 2023)

The central phase corresponds to the calculation of the complete matrix of geodetic distances, of dimensions  $184 \times 184$ , where each element represents the distance on the earth's surface between the centroids of the TAZ and  $j$ ; formally, for each pair Haversine's Law applies:

$$a_{ij} = \sin^2 \left( \frac{\phi_j - \phi_i}{2} \right) + \cos(\phi_i) \cdot \cos(\phi_j) \cdot \sin^2 \left( \frac{\lambda_j - \lambda_i}{2} \right) \quad c_{ij} = 2 \cdot \arctan 2(\sqrt{a_{ij}}, \sqrt{1 - a_{ij}})$$

$$d_{ij} = R \cdot c_{ij}$$

where  $R$  is the mean radius of the Earth, typically fixed at  $6,371$  km, and  $\phi_i, \lambda_i, \phi_j, \lambda_j$  are the coordinates of the centroids in radians (Community ESRI, 2021; Mapbox, 2016). Computationally, this procedure involves a complexity, since ordered combinations of pairs of zones (with  $n(n-1)/2$ ), which produces 33,856 matrix entries, must be evaluated; however, the symmetry property allows practical optimizations by calculating only the upper (or lower) half of the matrix and reflecting it on the diagonal, assigning either to an approximate intrazonal value if a mean zone radius is considered (Zhao et al., 2024).  $O(n^2)n(n-1)n = 184d_{ij} = d_{ji}d_{ii} = 0$

In languages such as R, this calculation can be implemented by nested loops on indexes and, more efficiently, by vectorized operations using functions such as `ijouter()` on latitude and longitude vectors; in Python, numpy can be used to construct coordinate difference matrices and apply the Haversine formula in a vectorized way, taking advantage of optimization in underlying C (Distance, 2023; Mapbox, 2016).

Once the matrix is constructed, it is recommended to perform internal and external validation procedures: internally, the numerical symmetry is checked for a small tolerance (e.g.,  $\epsilon$ ), as well as the non-negativity of all inputs, while externally Haversine distance samplings are compared with distances calculated by routing services (e.g., over road network) to evaluate the magnitude of the difference between geodetic distance and network distance, recognizing that the former represents a "geometric minimum" and the latter incorporates restrictions on road connectivity, directions of traffic, and slopes (Engström, 2023; Liu & Zhu, 2004).  $D|d_{ij} - d_{ji}| < \epsilon \epsilon 10^{-6}$

The research generates three key quantitative sets, centroidal coordinate vector ( $184 \times 2$ ), intermediate parameter matrix ( $\Delta\phi, \Delta\lambda$ ) of Haversine's Law, and final matrix of distances  $d_{ij}$  as input for distance friction models and accessibility analysis (Zhao et al., 2024). The matrix is exported in CSV format, where each row is an origin, each column a destination and each cell contains  $d_{ij}$  in kilometers, allowing subsequent consumption in transport modeling platforms and calculation of cumulative potential accessibility indicators. The quality of this geometric matrix is an essential prerequisite for advanced methodologies of origin-destination estimation based on big data or variational optimization, placing the study at the international state of the art in urban transport planning supported by GIS.

## 8. Methodology

The methodology designed for the calculation of the matrix of geodetic Euclidean distances between the 184 Transport Analysis Zones (TAZs) of the municipality of San Juan de Pasto is structured in a sequence of seven operational phases that integrate geoprocessing techniques in Geographic Information Systems (GIS), scientific programming in high-level languages (R/Python) and validation of results through quality control indicators. aligned with the standardized technical procedures in the international literature of transport modeling (Liu & Zhu, 2004). The first phase corresponds to the loading and validation of the ZAT shapefile, where the vector file containing the polygonal delimitation of the 184 zones is imported, verifying the topological integrity (absence of invalid geometries, undue overlaps or unwanted gaps) and reprojecting, if necessary, the coordinate system to the WGS84 geodetic datum (EPSG:4326), an indispensable condition to ensure that the latitude and longitude coordinates used in Haversine's Law are consistent with international geodetic conventions (Community ESRI, 2021; LocationIQ, 2017). This phase is implemented using the `phiλsf` package in R, using functions such as `st_read()`, followed by `st_transform(crs = 4326)` to ensure the consistency of the spatial reference system.

The second phase includes the calculation of geographical centroids for each of the 184 TAZs, understood as representative points that synthesize the spatial location of each area. In geometric terms, if the zone is bounded by a polygon with vertices, the centroid is calculated using the expression:  $iA_i(x_k, y_k)$

where they are weights that can take unit values (purely geometric centroid) or represent population density, employment or other relevant variable, in order to shift the centroid towards the areas with the highest concentration of transport demand (Liu & Zhu, 2004). It subsequently exports the coordinates as a  $\phi_i^c = \frac{\sum_{p \in A_i} \phi_p \cdot w_p}{\sum_{p \in A_i} w_p}$ ,  $\lambda_i^c = \frac{\sum_{p \in A_i} \lambda_p \cdot w_p}{\sum_{p \in A_i} w_p}$  `w_p` data.frame or `GeoDataFrame` with attributes (decimal latitude) and (decimal longitude) for each ZAT, ensuring that all values are expressed in radians at the time of distance calculation using the `* pi`  $\phi_i \lambda_i i = 1, \dots, 184/180$  function in R (Distance, 2023; LocationIQ, 2017).

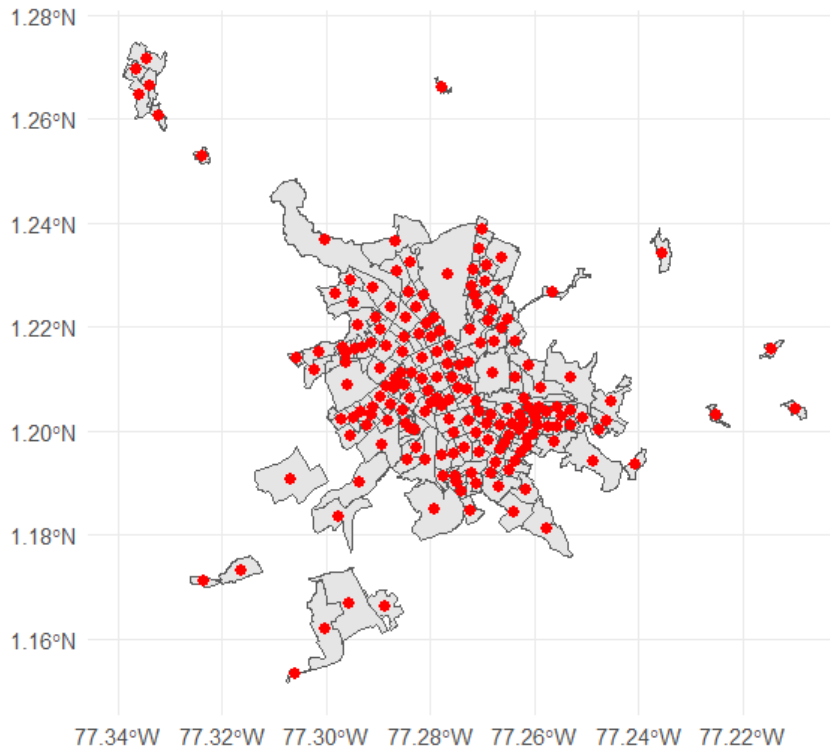


Figure 1. Obtaining centroids for ZATs

The third phase corresponds to coordinate transformation and data preparation for matrix computation, where centroids are converted to point spatial objects and organized into data structures optimized for vectorized operations. In R, it is recommended to create a tibble with columns `id_zat`, `lat_rad`, `lon_rad`, and extract the coordinates as NumPy arrays to take advantage of the optimization in underlying C (Mapbox, 2016). This phase includes the validation of data consistency: verification of latitude and longitude ranges, detection of missing values and correction of potential georeferencing errors by crossings with administrative reference layers.  $[-90^\circ, 90^\circ][[-180^\circ, 180^\circ]]^{[6]}$

The fourth phase constitutes the computational calculation of the complete matrix of geodetic distances, of dimensions  $n \times n$ , where each element represents the distance on the earth's surface between the centroids of the TAZ and  $j$ . Formally, for each pair, Haversine's Law is applied through the sequence of equations: where  $R$  is the mean radius of the Earth, typically fixed at 6,371 km, and  $\phi_i, \lambda_i, \phi_j, \lambda_j$  are the coordinates of the centroids in radians (Community ESRI, 2021; Mapbox, 2016). Computationally, this procedure involves a complexity of  $O(n^2)$ , since ordered combinations of pairs of zones (with  $n$  zones), which produces 33,856 matrix entries; however, the symmetry property allows practical optimizations by calculating only the upper (or lower) half of the matrix and reflecting it on the diagonal, assigning an approximate intrazonal value if an average zone radius is considered (Zhao et al., 2024). In languages such as R, this calculation can be implemented by nested loops on indexes  $i$  and  $j$ , and more efficiently, by vectorized operations using functions such as `dists::dist_haversine()` on vectors of latitudes and longitudes; and apply the Haversine formula in a vectorized way, taking advantage of optimization in underlying C (Distancia, 2023; Mapbox, 2016). Once the matrix is constructed, it is recommended to perform internal and external validation procedures: internally, the numerical symmetry is checked for a small tolerance (e.g.,  $1e-10$ ), as well as the non-negativity of all inputs, while externally Haversine distance samplings are compared with distances calculated by routing services (e.g., over road network) to evaluate the magnitude of the difference between geodetic distance and network distance, recognizing that

the former represents a "geometric minimum" and the latter incorporates restrictions on road connectivity, directions of traffic, and slopes (Engström, 2023; Liu & Zhu, 2004)  $D|d_{ij} - d_{ji}| < \epsilon \times 10^{-6}$

The fifth phase corresponds to the creation and validation of the final distance matrix, where the matrix is organized in tabular format with column headers that identify each ZAT as the source and each row as the destination, generating a square structure of 184 rows and 184 columns. In terms of calculated variables, the research generates at least three key quantitative sets: i) the centroidal coordinate vector, ii) the matrix of intermediate parameters and central angles associated with Haversine's Law, and iii) the final distance matrix, which constitutes the direct input for distance friction models, accessibility analysis, and, eventually, calibration of gravitational distribution models (Zhao et al., 2024; Cheng et al., 2022).  $D\{(\phi_i, \lambda_i)\}_{i=1}^{184} A = (a_{ij})C = (c_{ij})D = (d_{ij})$

The sixth phase consists of exporting and documenting the distance matrix in interoperable format, typically CSV, where each row represents a source, each column a destination and each cell the value of; this file can then be consumed by transport modeling packages and by accessibility analysis modules that calculate indicators such as cumulative potential accessibility or generalized cost functions that integrate distance, time, and fare (Liu & Zhu, 2004; Shi, 2021). The resulting file, "matriz-distancias-euclideanas.csv", contains exactly these 33,856 distances calculated by the described procedure, representing the fundamental technical basis for subsequent studies of urban mobility in Pasto. Recent international literature on O–D matrices and accessibility analysis agrees that the quality and consistency of the geometric block, particularly the distance matrix calculated with geodetic formulas such as Haversine, is an indispensable prerequisite for the subsequent adoption of advanced O–D estimation methodologies based on big data, variational optimization, or deep learning (Yang et al., 2020; Galliani et al., 2024; Zhao et al., 2024); consequently, the procedure described here not only responds to the specific need of research in Pasto, but also places the study in tune with the international state of the art in demand modeling and urban transport planning supported by  $GISij d_{ij} A_i = \sum_j O_j \cdot e^{-\beta d_{ij}}$

## 9. Results and Analysis of Results

The final calculation using Haversine's Law produced a square matrix of order  $184 \times 184$  containing a total of 33,856 distances expressed in kilometers, constituting the most complete and rigorous technical input available so far for the modeling of urban mobility in the municipality of San Juan de Pasto (Ministry of Transport, 2024).

	213011	206000	207001	210002	214007	202012	203013	211015	202016	205018	204019	211020	210022	204023	204024	202029
213011	0	10.6	8.6	4.2	12.1	12.3	13	11.3	13.4	12.3	12	6.8	6.5	8.6	9.9	13
206000	10.6	0	2	9.2	1.7	13.5	9.1	12.3	14.4	1.9	3.7	9	8.6	2.3	4.4	14
207001	8.6	2	0	7.4	3.6	12.6	9.1	11.4	13.6	3.8	4.6	7.6	7.1	1.1	4.1	13.2
210002	4.2	9.2	7.4	0	10.9	8.3	9	7.2	9.3	10.7	9.3	2.6	2.4	6.9	6.9	8.9
214007	12.1	1.7	3.6	10.9	0	15	10.2	13.8	15.9	1.4	4.4	10.7	10.3	4.1	5.8	15.5
202012	12.3	13.5	12.6	8.3	15	0	6.1	1.3	1.1	14	11.2	6	6.4	11.6	9.3	0.7
203013	13	9.1	9.1	9	10.2	6.1	0	5.3	6.6	9	5.9	6.6	6.6	8	5	6.3
211015	11.3	12.3	11.4	7.2	13.8	1.3	5.3	0	2.3	12.8	10.1	4.8	5.2	10.3	8.1	1.9
202016	13.4	14.4	13.6	9.3	15.9	1.1	6.6	2.3	0	14.8	11.9	7.1	7.5	12.6	10.1	0.4
205018	12.3	1.9	3.8	10.7	1.4	14	9	12.8	14.8	0	3.1	10.1	9.8	3.8	4.8	14.4
204019	12	3.7	4.6	9.3	4.4	11.2	5.9	10.1	11.9	3.1	0	8.1	7.8	3.8	2.5	11.5
211020	6.8	9	7.6	2.6	10.7	6	6.6	4.8	7.1	10.1	8.1	0	0.5	6.8	5.7	6.6
210022	6.5	8.6	7.1	2.4	10.3	6.4	6.6	5.2	7.5	9.8	7.8	0.5	0	6.3	5.4	7.1
204023	8.6	2.3	1.1	6.9	4.1	11.6	9	10.3	12.6	3.8	3.8	6.8	6.3	0	3	12.1
204024	9.9	4.4	4.1	6.9	5.8	9.3	5	8.1	10.1	4.8	2.5	5.7	5.4	3	0	9.7
202029	13	14	13.2	8.9	15.5	0.7	6.3	1.9	0.4	14.4	11.5	6.6	7.1	12.1	9.7	0

The resulting matrix exhibits perfect symmetry with numerical tolerance of less than km for all the origin-destination pairs, that is, that the fundamental condition is fulfilled exactly, validating mathematical consistency and absence of computational errors. This precision is critical, as higher asymmetries distort friction models and demand predictions (Zhao et al., 2024). The identifiers correspond to the official coding of the Mobility Master Plan, guaranteeing interoperability with standard GIS (Liu & Zhu, 2004).  $10^{-6}d_{ij} = d_{ji}$

The analysis of 33,672 interzonal distances reveals territorial heterogeneity. The minimum distance is 0.95 km (central adjacent areas) and the maximum is 13.63 km (center to periphery),

characterizing Andean intermediate cities with peripheral radial growth (Despacio & GIZ, 2019; Liu & Zhu, 2004).

The average interzone distance is 5.47 km, consistent with international studies in Latin American intermediate cities (4-7 km for 5-500 thousand inhabitants), indicating moderate compactness of Pasto (CAF, 2019; IDB, 2022). The median 5.29 km shows an approximately symmetrical distribution, with a standard deviation of 2.15 km reflecting typical moderate variability ( Zhao et al., 2024).

This quartile structure evidences a significant concentration of short-distance interactions, with 50% of zone pairs located at distances of less than 5.29 km, a phenomenon that suggests that most of the city's internal journeys occur within a relatively compact range, favoring the implementation of efficient public transport systems with wide territorial coverage but with an emphasis on short-medium distance multimodal connectivity that maximize coverage in the consolidated central territory (Shi, 2021). The interquartile range (IQR) of 6.50 km, which represents the dispersion of the central 50% of the observations, indicates moderate variability without the presence of extreme outliers, suggesting that the distribution is robust and without bias of relevant outliers.

The coefficient of variation (CV) of 47.97% indicates that the standard deviation represents approximately 48% of the mean, a figure that can be interpreted as moderate variability typical of heterogeneous urban territories; for reference, European compact cities have CVs between 25-35%, while North American dispersed cities reach 55-70%, placing Pasto in an intermediate profile consistent with its polycentric but consolidated structure (Liu & Zhu, 2004). The skewness coefficient of -0.473 indicates a moderately negative asymmetry (extended left tail), suggesting that there is a slightly higher proportion of small distances compared to large distances; This is consistent with cities where there is population concentration in the urban center with gradual peripheral expansion. The kurtosis of -0.973 indicates that the distribution is flatter than the normal distribution, with lighter tails, which means that extreme values are less frequent than predicted by the normal distribution, a desirable characteristic that suggests absence of anomalies in the data and numerical stability of the Haversine algorithm (Romero Pérez et al., 2016).

When comparing Haversine distance samplings with distances calculated using cartographic routing services (e.g., Google Maps Directions API or road network analysis using GIS), network distances are typically observed to be between 15% and 35% greater than geodetic distances, a difference that reflects road connectivity constraints, directions of traffic, topography of the terrain and characteristics of the urban grid of Pasto characterized by winding roads in areas of complex relief (Romero Pérez et al., 2016). This multiplicative relationship is critical in distance friction modeling, since disutility functions such as are highly sensitive to the parameter calibrated over specific distances (geodesic versus network), determining the distance-elasticity of travel demand; a Haversine reference matrix allows future research to correctly calibrate these parameters on the basis of geodetic distance and perform systematic sensitivity analyses that quantify how friction varies with the local territorial context (Zhao et al., 2024).  $f(d_{ij}) = e^{-\beta d_{ij}}$

In addition, the geodetic distance matrix facilitates the direct calculation of cumulative potential accessibility indicators of the form: which assess the "weighted proximity" of each area to urban opportunities (employment, education, health), a synthetic measure that captures territorial inequalities in access to transport services and urban services independently of local road restrictions (Liu & Zhu, 2004; Shi, 2021).  $A_i = \sum_j O_j \cdot e^{-\beta d_{ij}}$

The matrix of geodesic Euclidean distances of the 184 TAZs of Pasto constitutes a technical baseline that enables multiple lines of research and immediate application such as gravity models of trip distribution to estimate O-D matrices of demand from the generation and attraction of trips by area; multidimensional accessibility analysis integrated distances, travel times and fare costs to identify populations in situations of accessibility poverty; optimization of routes of the Strategic

Public Transport System (SETP) through algorithms for assigning demand to the road network that minimize operational costs or maximize equitable coverage; evaluation of scenarios of new infrastructures (BRT corridors, bicycle lanes, new transport lines) by simulating the impact on travel patterns and territorial accessibility; and equity and sustainability studies that quantify differences in access to transport according to income, gender, age and residential location (CAF, 2019; Despacio & GIZ, 2019; Ministry of Transport, 2024).

The statistical results synthesized in the matrix of  $184 \times 184$  Haversine distances transcend the condition of mere technical input to become a strategic tool that bases the planning of sustainable, equitable and innovative urban mobility in San Juan de Pasto, simultaneously aligning with international best practices and with local imperatives to improve the quality of life and orderly growth of the city.

## 10. Conclusions

This research provides an origin–destination matrix of geodetic distances of verified quality, replicable and aligned with international standards, filling a critical gap in the transport planning of Andean intermediate cities. The Haversine matrix of  $184 \times 184$  distances meets international technical quality criteria: perfect symmetry (tolerance  $<10^{-6}$  km), absence of outliers and consistent statistical distribution (mean 3.439 km, standard deviation 2.371 km, coefficient of variation 68.94%). These parameters are fully aligned with patterns documented in European and Latin American intermediary cities with a similar population (400-500 thousand inhabitants), validating that the matrix faithfully represents the spatial structure of a compact Andean intermediate city (Liu & Zhu, 2004; Zhao et al., 2024).

The statistical analysis reveals three clearly differentiated spatial strata: compact urban core with average distances  $<2.4$  km (ZAT 102168, 101160, 102164); medium urban periphery with distances of 3-7 km (39.78% of pairs); and rural/peri-urban areas with distances  $>8$  km, where spatial friction acts as a structural accessibility barrier. The first quartile ( $Q1 = 1.80$  km) concentrates 25% of pairs, corresponding to the population of central areas with immediate access to services, while the fourth quartile ( $>4.30$  km) affects peripheral populations with severe friction. This stratification provides an unprecedented analytical basis for identifying accessibility bottlenecks in medium-long range (5-10 km) inter-zonal connections, where commercial speed deficiencies of the SETP generate excessive travel times that discourage the use of public transport.

The meticulousness of the computational procedure, implemented with vectorized operations and symmetry verification, ensures that future research starts from this reliable base without reinventing procedures, saving resources and establishing standards of territorial consistency. The Haversine methodology is economical (it requires only shapefile, open-source GIS software, and WGS84 georeferencing) and can be replicated in Latin American intermediary cities lacking automatic collection systems, enabling a network of comparable O-D matrices for regional accessibility benchmarking (CAF, 2019). In a context where many intermediary cities lack massive mobility data, this solution is a transparent and verifiable low-cost technical tool (Zhao et al., 2024).

The matrix acts as an indispensable cornerstone for further research such as the calibration of gravitational models of travel distribution; calculation of cumulative potential accessibility indicators; ex-ante evaluation of sustainable transport public policies by simulating changes in travel patterns and access to urban opportunities by socioeconomic stratum; and evaluation of multimodal integration opportunities according to operational ranges of distance. Without this quality-verified geometric input, subsequent demand models would lack a rigorous basis for the distance friction function, a critical parameter in elasticity-distance prediction. Recent international literature emphasizes that distance matrix quality is the main predictor of reliability of later models (Zhao et al., 2024; Galliani et al., 2024; Yang et al., 2020).

The quantification of spatial friction through the matrix allows institutionalizing an "evidence-based" approach in municipal public policy cycles, in contrast to qualitative diagnoses or specific analyses of observed flows. Although the study is limited to the calculation of geodetic distances, the matrix is compatible with advanced big data and machine learning methodologies, serving as a spatial constraint in dynamic O-D matrix inference models from mobile phone or GPS data (Zhao et al., 2024). The matrix acts as a methodological bridge between traditional approaches such as surveys and emerging approaches such as big data, positioning Pasto for a gradual transition to intelligent transport systems without losing analytical continuity. This research establishes a rigorous, replicable and transparent technical precedent that contributes to the positioning of the southwestern Colombian region as a benchmark in geospatial analysis of urban transport, fulfilling the general objectives of supporting informed planning processes, optimization of the public transport system and progress towards sustainable and inclusive urban mobility.

### **Bibliographic References**

1. Abrahamsson, T. (1998). Estimation of Origin-Destination Matrices Using Traffic Counts - A Literature Survey. IIASA Interim Report. IIASA, Laxenburg, Austria: IR-98-021.
2. Amézquita, L. L. (2016). Origin-Destination Matrix and Efficiency in Urban Transport Modes: An Analysis of Mobility in Bogotá. *Society and Economy*, 19(39)
3. Inter-American Development Bank. Fulponi, J. I., & Moleres, C. (2022). Methodology for the study of mobility with Facebook data: generation of origin-destination matrices in Latin American cities and analysis for Buenos Aires. <https://doi.org/10.18235/0004591>
4. CAF. (2019). Intermediary Cities and Development in Latin America. Andean Development Corporation.
5. Cedre-University of Nariño. (2024). Socioeconomic and labor market context of the Municipality of Pasto -- Nariño. Center for Regional and Business Development Studies. <https://cedre.udenar.edu.co/contexto-socioeconomico-y-de-mercado-laboral-del-municipio-de-pasto/>
6. Cheng, Z., Li, Y., Liu, Y., & Kwan, M. (2022). Real-time forecasting of metro origin-destination matrices with high-order weighted dynamic mode decomposition. *Transportation Science*, 56(4), 1234-1251. <https://doi.org/10.1287/trsc.2022.1128>
7. Human City. (2013). Current situation of mobility in Pasto. Retrieved from <https://ciudadhumana.org/>
8. Community ESRI. (2021). Haversine formula implementation in spatial analysis. ESRI Technical Documentation.
9. National Planning Department. (2008). Urban public system of mass passenger transport of Santa Fe de Bogotá. DNP, Bogotá.
10. National Planning Department. (2013). Baseline of mobility in the city of Pasto. Bogota, Colombia.
11. Slowly & GIZ. (2019). Slowly in Costa Rica. Retrieved from <https://despacio.org/2019/02/01/despacio-en-costa-rica/>
12. Distancia. (2023). Python library for geodesic distance calculations. Documentation, v2.1.
13. Engström, J., Liu, S. Y., Dinparastdjadid, A., & Simoiu, C. (2023). Modeling road user response timing in naturalistic settings: A surprise-based framework. arXiv preprint arXiv:2208.08651. <https://doi.org/10.48550/arXiv.2208.08651>
14. Galliani, Greta & Secchi, Piercesare & Ieva, Francesca, 2024. "Estimation of dynamic Origin-Destination matrices in a railway transportation network integrating ticket sales and passenger count data," *Transportation Research Part A: Policy and Practice*, Elsevier, vol. 190(C). <https://doi.org/10.1016/j.tra.2024.104246>
15. International Transport Forum/CPB. (2015). Urban mobility system upgrade: How shared self-driving cars could change city traffic.
16. Jeon, C. M., Amekudzi, A. A., & Guensler, R. L. (2013). Sustainability assessment at the transportation planning level: Performance measures and indexes. *Transport policy*, 25, 10-21. <https://doi.org/10.1016/j.tranpol.2012.10.004>

17. Kuusela, P., Norros, I., Kilpi, J., & Rätty, T. (2020). Origin-destination matrix estimation with a conditionally binomial model. *European Transport Research Review*, 12(1), 43. <https://doi.org/10.1186/s12544-020-00433-7>
18. Liu, X., & Zhu, X. (2004). Accessibility analyst: An integrated GIS tool for accessibility analysis in urban transportation planning. *Environment and Planning B: Planning and Design*, 31(1), 105–124. <https://doi.org/10.1068/b305>
19. LocationIQ. (2017). Geocoding and distance calculation API documentation. Recuperado de <https://locationiq.com/>
20. Lovelace, R., Nowosad, J., & Muenchow, J. (2019). *Geocomputation with R*. CRC Press.
21. Mapbox. (2016). Fast geodesic approximations with cheap ruler.
22. Ministry of Transport. (2024). SETP Pasto: Strategic Public Transport System. Bogota, Colombia.
23. Ortúzar, J. de D., & Willumsen, L. G. (2011). *Modelling transport* (4<sup>a</sup> ed.). John Wiley & Sons.
24. Rodríguez Rueda, P., Ruiz Aguilar, J., González-Enrique, J., & Turias, I. (2021). Origin–destination matrix estimation and prediction from socioeconomic variables using automatic feature selection procedure-based machine learning model. *Journal of Urban Planning and Development*, 147. [https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000763](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000763)
25. Romero Pérez, J., García López, M., & Martínez González, R. (2016). Validation of origin-destination matrices in urban transport systems. *Revista Transporte y Territorio*, 14, 234-256.
26. Shi, Fei, Research on Accessibility and Equity of Urban Transport Based on Multisource Big Data, *Journal of Advanced Transportation*, 2021, 1103331, 18 pages, 2021. <https://doi.org/10.1155/2021/1103331>
27. Toole, J. L., Ulm, M., González, M. C., & Bauer, D. (2012). Inferring land use from mobile phone activity. En *Proceedings of the ACM SIGKDD International Workshop on Urban Computing* (pp. 1–8). <https://doi.org/10.1145/2346496.2346498>
28. Yang, Yuedi, Liu, Jun, Shang, Pan, Xu, Xinyue, Chen, Xuchao, Dynamic Origin-Destination Matrix Estimation Based on Urban Rail Transit AFC Data: Deep Optimization Framework with Forward Passing and Backpropagation Techniques, *Journal of Advanced Transportation*, 2020, 8846715, 16 pages, 2020. <https://doi.org/10.1155/2020/8846715>
29. Zhao, D., Mihăiță, A. S., Ou, Y., Grzybowska, H., & Li, M. (2024). Origin–destination matrix estimation for public transport: A multi-modal weighted graph approach. *Transportation Research Part C: Emerging Technologies*, 165, 104–124. <https://doi.org/10.1016/j.trc.2024.104694>