sensors

MDPI

Article

Identification and Assessment of GPS-Derived Multimodal Urban Mobility Corridors: Spatial Stability, Network Connectivity, and Reproducibility in Quito, Ecuador

Juan José Molina-Campoverde *, Jhonny Pumacuro-Paneluisa and Jean Ramirez-Arias

Grupo de Ingeniería Automotriz, Movilidad y Transporte (GiAUTO), Carrera de Ingeniería Automotriz-Campus Sur, Universidad Politécnica Salesiana, Quito 170702, Ecuador; jpumacurop@est.ups.edu.ec (J.P.-P.); jramireza10@est.ups.edu.ec (J.R.-A.)
* Correspondence: jmolinac@ups.edu.ec

Abstract

Urban mobility analysis increasingly relies on GPS trajectories to identify spatial patterns and their relationships with transportation infrastructure. However, spatial concentrations do not necessarily represent connected, empirically supported, or reproducible mobility corridors. This study proposes a framework for identifying and assessing candidate multimodal mobility corridors in Quito, Ecuador, using 683 GPS-recorded trips and 1,342,876 observations across eight transportation modes. The methodology integrates grid-based trip incidence, bootstrap resampling, road-network reconstruction, empirical GPS support, and repeated spatial validation. Candidate concentrations were identified using a 250 m grid and a 3% trip-incidence threshold, with stability assessed through 300 bootstrap iterations. Empirical GPS support reached 66.7% for a public transport corridor, 46.1% for a private motorized corridor, and 29.9% for a multimodal corridor. Across five stratified 70/30 training–validation partitions, mean held-out GPS-trip coverage reached 57.96% overall, 73.16% for public transport, 82.86% for private motorized transport, and 49.35% for active mobility. Compared with a random spatial baseline matched to the buffered area of the detected patterns, coverage was higher overall (57.96% versus 16.16%), whereas private motorized transport showed slightly lower coverage than the baseline (82.86% versus 86.71%). These findings indicate that broad spatial correspondence does not necessarily imply localized pattern recurrence, emphasizing the importance of distinguishing reconstructed connectivity, empirical GPS support, and spatial reproducibility.

Keywords: urban mobility; GPS-derived mobility corridors; grid-based spatial incidence; road-network reconstruction; multimodal trajectory analysis

Received: 15 October 2026
Revised: 23 November 2026
Accepted: 24 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Urban mobility constitutes a fundamental dimension of sustainable urban development, influencing accessibility, transportation efficiency, spatial organization, and the interaction between populations and urban infrastructure. Recent advances in location-based technologies and large-scale mobility datasets have transformed the analysis of urban dynamics, enabling researchers to examine how travel behavior reflects changes in spatial structure and the distribution of urban activities [1]. The directionality of human movement also provides valuable information about monocentric and polycentric urban

configurations, emphasizing the importance of examining the spatial characteristics of recorded movements rather than relying exclusively on aggregated transportation indicators [2].

The increasing availability of georeferenced mobility observations has expanded the understanding of how urban populations interact with their environments. Large-scale mobility networks have revealed spatial inequalities and patterns of interaction that cannot be adequately characterized using residential locations alone, demonstrating the relevance of observed travel behavior for understanding urban segregation and accessibility [3]. Furthermore, mobility-based investigations have explored the relationship between urban spatial structures, human activity, and environmental exposure, highlighting the potential of movement data to support integrated assessments of urban sustainability [4].

From a methodological perspective, urban mobility analysis increasingly incorporates spatial networks, geographic representations, and data-driven approaches to characterize travel behavior. Recent research has proposed vector-field formulations to describe the direction and organization of mobility flows, providing alternative representations of spatial interactions across urban environments [5]. Similarly, network-based descriptions derived from GPS observations have demonstrated that the organization of human movement involves complex spatial relationships that extend beyond conventional origin–destination summaries. Consequently, identifying geographically concentrated mobility patterns can provide complementary evidence about the spatial organization of recorded trips and their relationship with transportation infrastructure [6].

GPS trajectory data are particularly relevant because they provide sequential observations of movement that can be examined in relation to geographic locations, travel conditions, and transportation networks. Recent applications have combined GPS trajectories with contextual geographic information to investigate the contribution of different activity-related trips to traffic congestion and the spatial distribution of transportation impacts [7]. Other studies have reconstructed daily mobility patterns from smartphone observations, demonstrating the potential of trajectory-processing techniques to identify meaningful travel structures and changes in mobility behavior. Nevertheless, the usefulness of these approaches depends on the quality, continuity, and spatial coverage of the underlying observations [8].

Despite these advances, mobility datasets remain subject to methodological uncertainty. Comparisons across large-scale mobility data sources have revealed substantial differences in the spatial networks and displacement distributions obtained from alternative collection and processing procedures. These discrepancies indicate that the geographic representation of mobility patterns may depend on the characteristics of the available data rather than exclusively on actual travel behavior [9]. Moreover, associating smartphone-based positioning observations with urban road infrastructure remains challenging because of positional uncertainty, irregular sampling, and complex road geometries. These considerations highlight the importance of transparent processing criteria and explicit spatial validation when identifying mobility structures from recorded trajectories [10].

The correspondence between observed GPS concentrations and existing road infrastructure represents an additional methodological challenge. Recent map-matching research has emphasized the importance of road topology, geometric consistency, and trajectory continuity when associating positioning observations with road segments, particularly in complex urban networks [11]. Similarly, uncertainty-aware trajectory modeling has demonstrated that spatial inference should account for the limitations and variability of observed movement data. For studies seeking to identify candidate mobility corridors, these findings support distinguishing spatial concentrations of GPS observations, their association with the road network, and the empirical evidence supporting the resulting structures [12].

These challenges are particularly relevant in Latin American metropolitan areas, where urban expansion, unequal accessibility, and heterogeneous transportation systems produce differentiated mobility conditions. Comparative research involving São Paulo, Santiago, and Bogotá has identified substantial differences in periurban travel behavior, modal choices, and dependence on central urban areas, reflecting the influence of metropolitan spatial organization on mobility opportunities [13]. In Santiago, the integration of smart-card and georeferenced transportation information has also supported detailed assessments of public transport crowding and its temporal variation. Such findings demonstrate the value of disaggregated mobility information for understanding transportation systems in cities characterized by spatial and socioeconomic heterogeneity [14].

Recent Latin American studies have further demonstrated the potential of digital mobility data to support transportation planning and infrastructure analysis. In Santiago, mobile phone observations and data-driven methods have been used to characterize urban freight movements and identify spatial patterns relevant to logistics planning [15]. In Bogotá, empirical investigations of road-space reallocation for cycling have examined changes in traffic speeds across affected and neighboring road segments, highlighting the importance of considering network-level interactions when evaluating transportation interventions. These applications illustrate the relevance of transportation-mode differentiation and spatial relationships in urban mobility analysis [16].

Within Ecuador, the Metropolitan District of Quito provides a relevant setting for examining spatial mobility patterns [17] because of its elongated urban configuration, geographic constraints, and diverse transportation services. Research on Quito's first metro line has demonstrated the importance of spatial accessibility analysis for understanding the distribution of urban opportunities and the differentiated benefits associated with public transport infrastructure [18]. At the national level, investigations into rapid transit implementation have identified institutional and planning barriers affecting the development of major transportation projects. These findings reinforce the need for reproducible, geographically explicit evidence capable of complementing conventional transportation planning approaches [19].

Previous research conducted in Ecuador has also demonstrated the feasibility of using smartphone-derived GPS measurements to investigate multimodal urban mobility. A study conducted in Cuenca developed a machine-learning-based approach for identifying transportation modes from smartphone sensor observations, including geographic coordinates, speed, altitude, and temporal information, demonstrating the potential of mobile sensing for transportation characterization in Ecuadorian cities [20]. At a broader spatial scale, international research has shown that natural geographic barriers and urban fragmentation can influence transportation connectivity, network detours, and urban development. These findings are relevant to geographically constrained metropolitan environments, where spatial concentrations of movement and road-network continuity should be evaluated as distinct characteristics [21].

Although existing studies have advanced mobility sensing, transportation-mode characterization, spatial accessibility, and road-network analysis, an important methodological challenge remains in determining whether GPS-derived spatial concentrations are stable, correspond to plausible road-network structures, and can be reproduced using different subsets of recorded trips. The heterogeneity of multimodal travel conditions further complicates the interpretation of spatial concentrations, particularly when mobility patterns are evaluated using aggregated transportation categories [22]. Furthermore, research on commuting-network structures has demonstrated the importance of examining spatial organization at different scales and distinguishing localized mobility relationships from broader regional structures. Accordingly, integrating trip-level spatial incidence, stability assessment, road-network association, and independent holdout evaluation can provide a more transparent basis for

interpreting candidate mobility corridors without treating sample-based concentrations as representative measures of metropolitan travel demand [23].

To address these challenges, the present study develops and applies a GPS-based analytical framework for identifying and assessing candidate multimodal mobility corridors in the Metropolitan District of Quito, Ecuador. The analysis uses 683 GPS-recorded trips comprising 1,342,876 georeferenced observations collected across eight transportation modes between July 2025 and August 2026. Recorded trajectories are aggregated into a regular spatial grid using distinct-trip incidence, and candidate spatial concentrations are evaluated through trip-level bootstrap resampling. Stable spatial components are subsequently associated with a reference road network and examined in terms of topological connectivity and empirical GPS support. Finally, spatial reproducibility is assessed through repeated transportation-mode-stratified training–validation partitions, Jaccard-based recurrence analysis, and a complementary chronological holdout assessment of subsequently recorded trips. By distinguishing spatial stability, road-network correspondence, and out-of-sample reproducibility, the study provides a systematic and transparent framework for identifying candidate mobility structures from multimodal GPS observations while explicitly acknowledging the limitations imposed by sample coverage, modal heterogeneity, and algorithmically reconstructed network connections.

2. Data and Methods

The methodological framework integrates GPS trajectory preprocessing, grid-based spatial pattern identification, road-network reconstruction, and complementary procedures for assessing the stability and reproducibility of candidate mobility corridors in the Metropolitan District of Quito, Ecuador. As illustrated in Figure 1, the analysis comprises 683 GPS-recorded trips across eight transportation modes. Spatial concentrations are identified through distinct-trip incidence and bootstrap resampling, subsequently associated with a reference road network, and evaluated using GPS proximity support. Methodological robustness is further examined through repeated stratified training–validation partitions, an area-matched spatial baseline, recurrence analysis, and sensitivity assessments of GPS data processing and spatial resolution. These complementary procedures distinguish spatial concentration, network connectivity, empirical trajectory correspondence, and reproducibility without assuming that the recorded sample represents metropolitan travel demand.

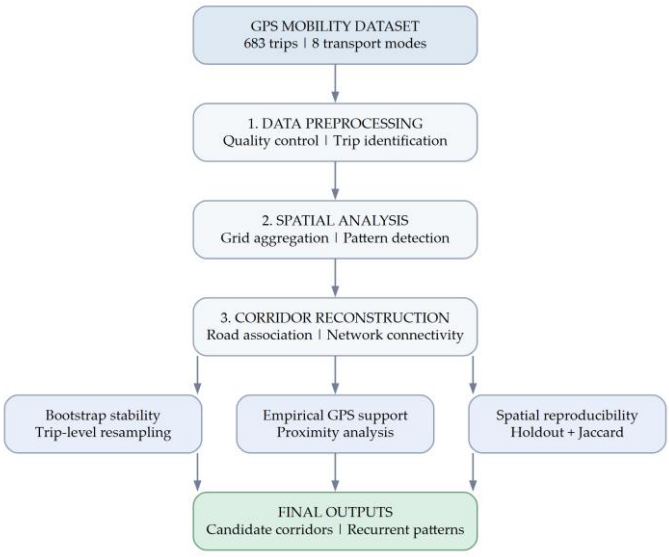
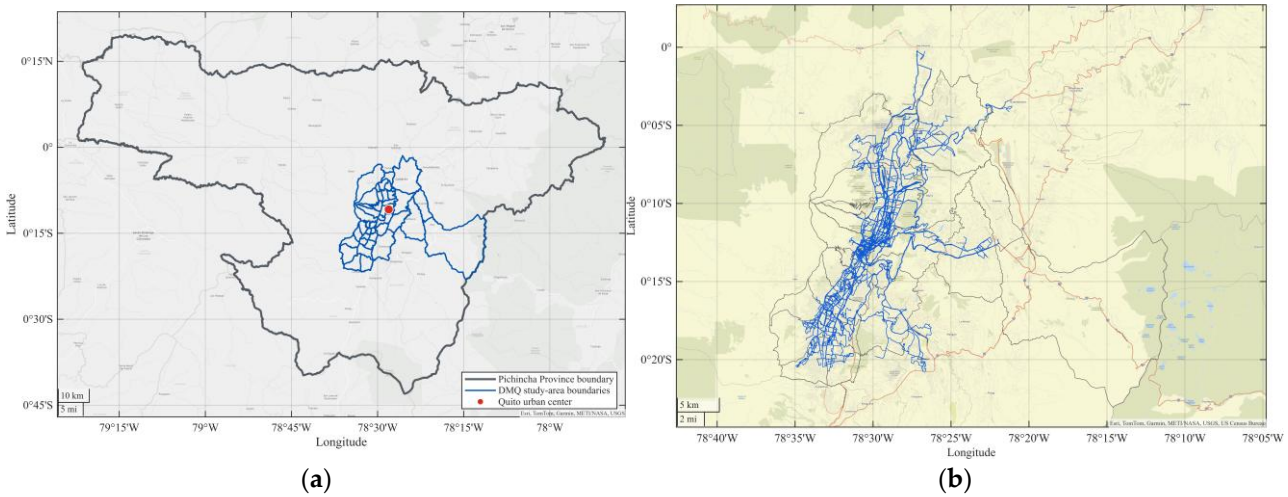


Figure 1. Overview of the methodological framework for identifying and assessing GPS-derived candidate mobility corridors.

2.1. Study Area and GPS Mobility Dataset

The study was conducted in the Metropolitan District of Quito (DMQ), Ecuador, an Andean metropolitan area characterized by complex topography, considerable elevation variations, and heterogeneous urban mobility conditions. The city's predominantly north–south urban configuration, together with its road network and multimodal transportation infrastructure, provides a relevant environment for investigating the spatial distribution of urban travel patterns. The transportation system includes private motorized vehicles, conventional buses, dedicated public transport corridors, cycling, and walking, enabling the analysis of mobility patterns across different transportation modes.

Figure 2 illustrates the geographic location of the study area within Pichincha Province, highlighting the selected study zones of the Metropolitan District of Quito considered in the GPS-based mobility analysis.



(**a**) (**b**)

Figure 2. Geographic location of the study area and spatial distribution of GPS-recorded mobility trajectories in Quito, Ecuador: (**a**) location of the study area within Pichincha Province and the Metropolitan District of Quito; (**b**) spatial coverage of the GPS-recorded trajectories across the urban road network and surrounding areas.

The analysis was based on a georeferenced dataset containing trips recorded between July 2025 and August 2026, comprising 683 GPS-recorded trips and 1,342,876 positioning records across eight transportation modes: cars, motorcycles, conventional buses, trolleybuses, Ecovía, other public transport corridors, bicycles, and walking. Each GPS trajectory contained geographic coordinates (latitude and longitude), altitude, elapsed time, and recorded speed. Individual trips were identified using a trip identifier (TripID) in combination with the transportation mode, allowing each trajectory to be processed independently while maintaining its modal classification. The GPS-recorded trip constituted the analytical unit; therefore, the 683 recorded trips should not be interpreted as 683 independent participants, as multiple trajectories may originate from the same individual.

GPS trajectories were collected through a university-based voluntary participation initiative using a non-probability recruitment approach. Contributors received verbal information about the study objectives, the recording procedure, and the voluntary nature of their participation. They used GPS Logger for Android (Mendhak; version not recorded; https://github.com/mendhak/gpslogger/? ; accessed on 23 November 2026) on their personal Android smartphones. Because the devices were participant-owned, smartphone manufacturers and models varied and were not recorded, configured with a nominal recording interval of 1 s (1 Hz), and manually started and stopped recordings during their journeys. Each trip was classified according to the transportation mode

reported by the contributor, without independent mode-label verification. GPS files were submitted to a restricted-access Microsoft OneDrive repository (Microsoft Corporation, Redmond, WA, USA; web-based service, version not applicable; https://www.microsoft.com/microsoft-365/onedrive/online-cloud-storage; accessed on 23 November 2026). No financial incentives or participant identifiers were collected. Consequently, repeated contributions from the same individual could not be identified, and independence between trips could not be established. All computational processing and spatial analyses were performed in MATLAB R2026a (The MathWorks, Inc., Natick, MA, USA).

For the multimodal analysis, the eight transportation modes were classified into three categories: public transport (conventional buses, trolleybuses, Ecovía, and other public transport corridors), private motorized transport (cars and motorcycles), and active mobility (cycling and walking). The resulting dataset comprised 261 public transport trips, 115 private motorized trips, and 307 active mobility trips. Additionally, all 683 trips were considered collectively to investigate spatial mobility patterns across transportation categories. The distribution of GPS-recorded trips by transportation mode is presented in Table 1.

In this study, the term "multimodal" refers to the joint analysis of GPS trajectories recorded across different transportation modes. The complete multimodal dataset combines all 683 trips, whereas category-specific analyses consider public transport, private motorized transport, and active mobility separately. Consequently, a spatial corridor identified from the combined dataset does not necessarily indicate that the same road segments were traversed by every transportation mode. Modal participation is represented by the recorded trips contributing to the dataset and should not be interpreted as a measure of metropolitan modal share or travel demand.

Table 1. Distribution of GPS-recorded trips by transportation mode.

Transport Category	Mode	Trips	Percentage
Public transport	Conventional bus	200	29.28
	Trolleybus	25	3.66
	Ecovía	23	3.37
	Other public transport corridors	13	1.90
Private motorized transport	Car	69	10.10
	Motorcycle	46	6.73
Active mobility	Bicycle	108	15.81
	Walking	199	29.14
Total		683	100.00

2.2. GPS Data Processing and Spatial Representation

The GPS dataset was processed to establish a consistent spatial representation of the recorded trips across transportation modes. Each trip was identified using its transportation mode and a unique trip identifier (TripID), which served as the analytical unit throughout the processing and spatial analysis. A supplementary GPS data-quality audit examined invalid or missing timestamps, repeated timestamps, duplicate consecutive observations, chronological inconsistencies, recording interruptions, spatial discontinuities, and unusually high implied speeds between consecutive GPS positions. Geographic coordinates were examined to exclude invalid or non-finite observations. The spatial analysis was based on observed GPS positions, without interpolating missing trajectory segments. The complete dataset of 683 trips was retained for the identification of spatial mobility patterns. The number of recorded trips should not be interpreted as the number of independent participants, as multiple trips could have been contributed by the same individual.

GPS trajectories were treated as trip-level observational records, with each trajectory identified by its transportation mode and TripID. The available analytical dataset did not establish a verified one-to-one relationship between trips and individual contributors; therefore, multiple trips could have originated from the same person. Although the subsequent training–validation procedure separated complete trips between subsets, it did not establish participant-level independence or temporal independence between the resulting partitions. These limitations were considered when interpreting the reproducibility results.

For trip-level characterization, travel duration was calculated from the difference between the first and last valid timestamps. Traveled distance was estimated using the Haversine formula between consecutive GPS observations, excluding intervals exceeding 20 s to avoid introducing artificial distances across recording gaps. Mean travel speed was calculated from the estimated distance and total trip duration. Stationary time was determined using a speed threshold of 0.5 m/s, with the proportion of stationary time weighted by the duration of the corresponding observation intervals.

The 20 s threshold was applied consistently to distance estimation and temporal coverage calculations, preventing discontinuous recording intervals from contributing artificial displacement to the trip-level indicators.

To ensure the reliability of trip-level descriptive indicators, a minimum temporal coverage threshold of 80% was applied. Temporal coverage was defined as the proportion of the total trip duration represented by consecutive GPS observations separated by no more than 20 s. This criterion resulted in 570 eligible trips (83.5% of the dataset), comprising 190 public transport trips, 101 private motorized trips, and 279 active mobility trips. These trips were used exclusively for descriptive mobility characterization, whereas all 683 recorded trips were retained for the subsequent spatial analyses. The sensitivity of the descriptive indicators to recording interruptions was examined using alternative temporal-gap thresholds of 10, 30, and 60 s, with 20 s retained as the reference criterion. In addition, the influence of potentially anomalous GPS displacements was assessed by recalculating distance and mean-speed indicators after excluding intervals with implied speeds exceeding 130 and 160 km/h. These speed thresholds were applied exclusively for sensitivity analysis and were not used to remove observations from the spatial corridor detection procedure.

Accordingly, the temporal coverage criterion was not used to exclude trajectories from spatial pattern identification, ensuring that the descriptive and spatial analyses were based on explicitly differentiated subsets.

For spatial aggregation, geographic coordinates were transformed into a local metric coordinate system using an equirectangular approximation referenced to latitude −0.20° and longitude −78.50°. GPS positions were assigned to regular square grid cells at spatial resolutions of 100, 250, and 500 m, with 250 m adopted as the principal resolution. Each grid cell was characterized by the number of distinct GPS-recorded trips containing at least one observed position within its boundaries. Repeated observations from the same trip within a given cell were counted only once, thereby limiting the influence of recording frequency and prolonged stationary periods on the spatial incidence indicator.

This trip-based counting approach reduced the influence of differences in GPS sampling density and trip duration, ensuring that cell incidence reflected the number of recorded journeys rather than the total number of positional observations.

The grid-based analysis was conducted for the complete multimodal dataset and separately for public transport, private motorized transport, and active mobility. This common spatial framework enabled the comparison of recorded trip concentrations across transportation categories. Figure 3 presents the spatial distribution of GPS-occupied grid cells for the four analytical groups, with cell intensity representing the number of distinct

recorded trips associated with each location. The same spatial representation was subsequently used for identifying candidate mobility corridors and evaluating their spatial stability and reproducibility.

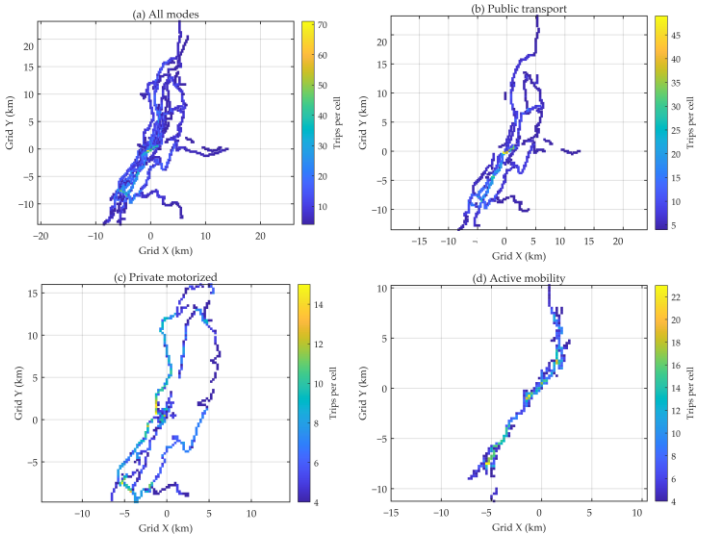


Figure 3. Spatial distribution of selected GPS-derived grid cells in the Metropolitan District of Quito for: (**a**) the complete multimodal dataset; (**b**) public transport trips; (**c**) private motorized trips; and (**d**) active-mobility trips. The values associated with the grid cells represent the number of distinct trips observed within each transportation category.

2.3. Spatial Pattern Identification and Road Network Association

Spatial mobility patterns were identified by analyzing the distribution of distinct GPS-recorded trips within the 250 m grid described in Section 2.2. For each transportation category, cells containing observations from at least 3% of the corresponding trips were selected as candidate spatial concentrations. The selection threshold was examined through a sensitivity analysis considering 2%, 3%, and 5% of recorded trips and grid resolutions of 100, 250, and 500 m. The analysis was performed independently for the complete multimodal dataset, public transport, private motorized transport, and active mobility. The 3% trip-incidence threshold was applied to the total number of trips in each transportation category, and spatial incidence was based on the number of distinct trips per grid cell, with repeated observations from the same trip counted only once.

To identify spatial concentrations that were less sensitive to the composition of the recorded sample, a trip-level bootstrap procedure was implemented using 300 resampling iterations with replacement. Each iteration preserved the original number of trips within the corresponding transportation category. For each grid cell, the bootstrap selection probability was calculated as the proportion of iterations in which the cell satisfied the original 3% trip-incidence threshold. Cells meeting this threshold in the original dataset and exhibiting a selection probability of at least 0.80 were retained as stable spatial concentrations. This procedure reduced the influence of random variations in the recorded sample and helped identify spatial concentrations consistently observed across multiple resampled datasets.

For the detailed road-network assessment, spatial components were ranked within each analytical group according to their spatial extent, measured by the number of occupied 250 × 250 m grid cells. Mean bootstrap selection probability was also retained as an indicator of spatial stability. The three highest-ranked components from each group were selected for subsequent road-network association and connectivity analysis. This selection

prioritized spatial extent rather than traffic demand, population representativeness, or road-network connectivity.

To characterize modal composition within the pooled spatial corridors, a complementary trip-level analysis was performed for all three selected pooled corridors (C17, C2, and C10). A trip was considered to intersect a corridor when at least one valid GPS observation fell within its selected 250 m grid cells. Each trip was counted only once per corridor, regardless of the number of recorded observations. For each transportation mode, the number of intersecting trips was divided by the total number of distinct trips intersecting that corridor. These percentages describe the modal composition of the sampled trajectories and are distinct from the road-network proximity support metrics reported in Section 3.3.

The spatial organization of stable cells was subsequently examined using four-neighbor adjacency, whereby cells sharing a common edge were assigned to the same connected component. Components containing at least three cells were retained as candidate mobility corridors. This procedure enabled the identification of spatially connected concentrations of recorded trips without relying on predefined transport routes. Additional geometric examinations considered spatial extent and connectivity under alternative neighborhood configurations. A minimum component size of three grid cells was used to avoid the inclusion of isolated or very small spatial clusters.

To establish correspondence with existing road infrastructure, the selected spatial components were associated with a georeferenced road network of the Metropolitan District of Quito. Road segments were identified according to their proximity to the centers of the selected grid cells, using a maximum association distance of 75 m. The associated road segments were then examined as a network to determine their topological connectivity. Where disconnected segments could be linked through the reference network, additional connecting segments were incorporated using a shortest-path procedure subject to geometric and distance constraints. Network nodes separated by less than 3 m were merged to reduce artificial topological breaks. Candidate connections were searched within a 750 m buffer, with a maximum anchor separation of 500 m, a maximum connector route length of 1200 m, and a maximum detour ratio of 2.5, considering up to three nearest anchor candidates. The parameters used for spatial identification and road network reconstruction are summarized in Table 2.

Table 2. Parameters used for spatial pattern identification and road network reconstruction.

Parameter	Value
Principal grid resolution	250 m
Spatial resolutions examined	100, 250, and 500 m
Trip-incidence thresholds examined	2%, 3%, and 5%
Principal trip-incidence threshold	3%
Bootstrap resampling iterations	300
Minimum bootstrap selection probability	0.80
Connectivity of stable grid cells	Four-neighbor adjacency
Minimum component size	3 grid cells
Road-segment association threshold	75 m
Node-merging tolerance	3 m
Road-network search margin	750 m
Maximum anchor separation	500 m
Maximum connector route length	1200 m
Maximum detour ratio	2.5
Nearest anchor candidates	3
Network connection method	Shortest-path algorithm

Figure 4 illustrates the road-network association and topological reconstruction of multimodal corridor C17, together with its segment-level GPS proximity support. Panel (a) distinguishes the reference road network, original GPS-associated road segments, and additional connectors introduced during reconstruction. Panel (b) identifies reconstructed segments supported by observations from at least two distinct GPS-recorded trips within a 75 m proximity threshold. This proximity-based assessment provides evidence of internal spatial consistency but does not constitute independent validation or confirm that the complete length of each segment was traversed.

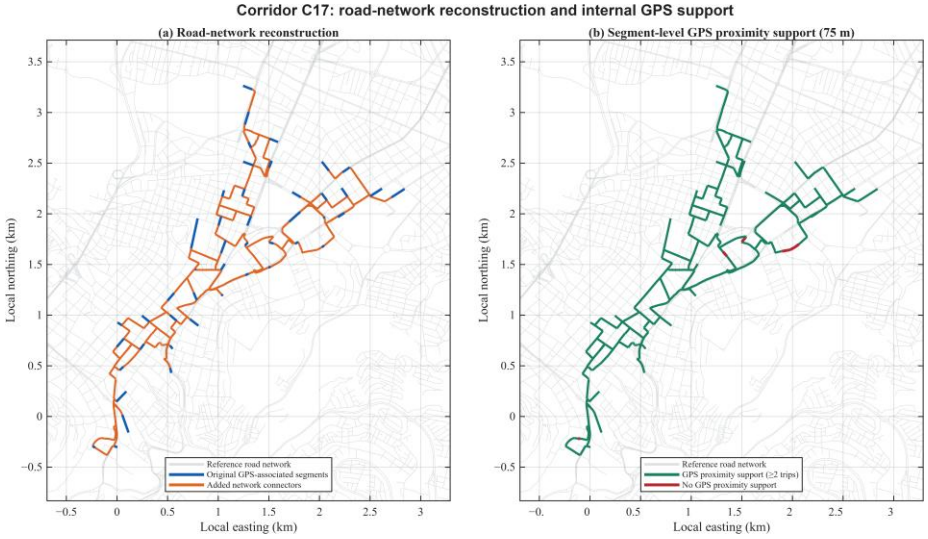


Figure 4. Road-network reconstruction and segment-level GPS proximity support for multimodal corridor C17 in Quito: (**a**) original GPS-associated road segments and network connectors introduced during topological reconstruction; (**b**) reconstructed segments supported by observations from at least two distinct GPS-recorded trips within a 75 m proximity threshold. Proximity-based support represents internal spatial consistency rather than independent validation or confirmed traversal of the entire segment.

2.4. Spatial Stability and Reproducibility Assessment

The spatial stability and reproducibility of the identified mobility patterns were assessed through a complementary framework comprising bootstrap resampling, GPS-based corridor support, and repeated training–validation partitions. These procedures were designed to examine the persistence of spatial concentrations under changes in the contributing trips and to quantify their correspondence with recorded mobility trajectories. All resampling and partitioning procedures were conducted at the trip level to prevent observations belonging to the same trip from being treated as independent sampling units.

Spatial stability was assessed using the trip-level bootstrap procedure described in Section 2.3. Stability was expressed as the empirical probability that each spatial cell satisfied the predefined 3% trip-incidence threshold across 300 resampling iterations. Cells meeting this threshold in the original dataset and exhibiting a bootstrap selection probability of at least 0.80 were classified as stable. Additionally, spatial agreement between the complete dataset and the subset of continuously recorded trips was examined using the Jaccard coefficient. Because this subset was contained within the original sample, the comparison was interpreted as an internal consistency assessment rather than an independent validation.

Empirical support for the reconstructed road-aligned corridors was subsequently evaluated by measuring the spatial proximity of GPS observations to their corresponding

road networks. Trajectories were sampled at 10-second intervals, and proximity thresholds of 25, 50, 75, and 100 m were considered. A principal distance threshold of 75 m was adopted for identifying supporting observations. A trip was classified as supporting a candidate corridor when at least five sampled GPS observations and 10% of its sampled positions fell within this threshold. This procedure quantified spatial correspondence without assuming that network-based connectors represented actual travel routes.

Spatial reproducibility was assessed using five repeated training–validation partitions, with 70% of GPS-recorded trips allocated to training and the remaining 30% reserved for validation. The partitions were stratified by transportation mode, and all observations belonging to the same trip were assigned exclusively to either the training or validation subset within each repetition. Candidate spatial patterns were identified using the same detection procedure applied to the complete dataset, including a 250 m grid, a minimum incidence threshold of 3% of the training trips, 300 trip-level bootstrap iterations, a minimum cell-selection probability of 0.80, four-neighbor connectivity, and a minimum connected-component size of three grid cells. This consistent configuration prevented differences between the training and validation analyses from being attributed to changes in the spatial detection algorithm.

To consolidate spatial components identified across the five training partitions, pairwise spatial correspondence was evaluated using the Jaccard similarity coefficient. Component pairs with a Jaccard coefficient of at least 0.25 were considered eligible for matching. For each candidate anchor component, the highest-similarity eligible component from each other partition was selected, allowing a maximum of one component per partition.

Candidate patterns were subsequently selected through a binary integer programming formulation that maximized the aggregate anchor-to-member Jaccard similarity, subject to the constraint that each component could belong to exactly one selected pattern. Singleton candidates were included to ensure complete component assignment. This optimization resolved competing correspondences within the generated candidate set without allowing components to be assigned to multiple patterns.

Potential splitting and merging configurations were handled through mutually exclusive candidate selection rather than explicitly classified as separate events. Pattern recurrence was calculated as the proportion of the five partitions represented in each consolidated pattern, with a recurrence threshold of 0.60 corresponding to presence in at least three partitions.

The reserved GPS trajectories were used exclusively to evaluate the spatial correspondence of the training-derived patterns, without contributing to their identification. GPS observations were sampled at 10-second intervals, and a proximity threshold of 150 m from the selected training-cell centers was applied. A validation trip was classified as supporting a spatial pattern when at least five sampled observations, representing at least 10% of its sampled positions, satisfied the proximity criterion.

To assess whether held-out trip coverage exceeded that expected from spatial extent alone, an area-matched random baseline was implemented across five transportation-mode-stratified 70/30 train–test partitions. For each partition and transportation category, the spatial detector was fitted using training trips only. The reference area was calculated as the geometric union of 150 m buffers centered on the grid cells retained by the detector, approximated using 48-sided polygons. Random baseline configurations were generated from training-occupied grid cells and matched to the detector's buffered union area within a relative tolerance of ±2%. Twenty random baseline replicates were evaluated per partition and category, yielding 400 baseline evaluations. Detector and baseline configurations were assessed using the same held-out GPS trips and proximity-support criteria. This procedure controlled for the effective buffered area while allowing spatial configuration and connectivity to differ between the detector and the random baseline.

Two complementary analyses were conducted using the 636 trips with assignable recording dates. First, five stratified random holdout partitions were evaluated within this subset. Second, a chronological holdout assigned 445 trips dated on or before 18 July 2026 to training and 191 trips dated from 19 July 2026 onward to validation. All trips sharing the same assigned recording date were kept within the same partition, with no overlap in assigned dates between training and validation. The remaining 47 trips were excluded because no usable recording date was available. However, dates for some trips were inferred from filenames and could not be independently confirmed against the original GPS records. Accordingly, the chronological analysis should be interpreted as a supplementary temporal sensitivity assessment rather than definitive evidence of temporal independence.

Finally, the recurrence of individual spatial patterns was examined by comparing the grid-cell sets obtained across the five training partitions. Spatial overlap was quantified using the Jaccard similarity coefficient (Equation (1)), defined as the ratio between the number of shared grid cells and the total number of distinct cells belonging to either pattern.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

where $J(A,B)$ is the Jaccard similarity coefficient between two spatial patterns; A and B represent their respective sets of grid cells; $|A \cap B|$ denotes the number of shared cells; and $|A \cup B|$ represents the total number of distinct cells belonging to either pattern. The coefficient ranges from 0 (no spatial overlap) to 1 (complete overlap).

$$R_i = \frac{n_i}{N} \quad (2)$$

where R_i represents the recurrence fraction of spatial pattern i, n_i is the number of repetitions in which the pattern was identified, and (N = 5) is the total number of training repetitions. Patterns with $R_i \geq 0.60$ were classified as recurrent.

Spatial-resolution sensitivity was assessed by repeating the corridor detection procedure using grid resolutions of 100, 250, and 500 m across the five training–validation partitions. The analysis considered the number of retained spatial components, their total area, and the proportion of held-out trips intersecting detected corridor cells. Cross-resolution spatial agreement was further evaluated using the intersection-over-union (IoU) coefficient and directional area-coverage measures. Corridor surfaces were represented on a common 50 m grid to enable consistent spatial comparison. These analyses were used to characterize the effects of grid resolution on spatial fragmentation, extent, and overlap, rather than to establish a statistically optimal resolution.

To facilitate methodological reproducibility, Algorithm 1 summarizes the complete analytical workflow, including GPS data-quality assessment, grid-based corridor identification, bootstrap stability, road-network reconstruction, empirical GPS support, repeated training–validation, and spatial sensitivity analyses.

Algorithm 1. GPS-based identification and reproducibility assessment of multimodal mobility corridors.

Input: GPS trajectories, transportation-mode labels, road-network dataset.
Output: Candidate mobility corridors, GPS support indicators, recurrent spatial patterns.
Procedure:
1. Load GPS trajectories and transportation-mode labels.
2. Perform GPS data-quality assessment:
- Check coordinates, timestamps, and duplicate records.

- Identify temporal gaps, spatial jumps, and speed anomalies.
- Assess temporal-coverage sensitivity.
3. For each transportation category:
   a. Project GPS coordinates to the local spatial system.
b. Construct a regular 250 m grid.
c. Count distinct GPS trips intersecting each grid cell.
d. Select cells meeting the 3% trip-incidence threshold.
e. Perform 300 trip-level bootstrap iterations.
f. Retain cells with selection probability ≥ 0.80.
g. Identify four-neighbor connected components.
h. Retain components containing at least three cells.
4. Associate selected spatial components with nearby road-network segments.
5. Reconstruct candidate corridor connectivity using constrained road-network connections.
6. Assess empirical GPS support:
- Sample GPS trajectories at 10-second intervals.
- Evaluate proximity to reconstructed corridors.
- Apply the principal 75 m proximity threshold.
- Require at least five supporting samples and at least 10% of sampled positions per trip.
7. Repeat the following procedure five times:
a. Stratify trips by transportation mode.
b. Assign 70% of trips to training and 30% to validation.
c. Apply the original spatial detector to training trips only, retaining the 3% trip-incidence threshold, 300 bootstrap iterations, minimum selection probability of 0.80, four-neighbor connectivity, and minimum component size of three cells.
d. Assess held-out trip coverage within 150 m of selected training-cell centers.
e. Calculate the geometric union area of 150 m buffers around the selected training-cell centers.
f. Generate 20 random baseline configurations per transportation category using training-occupied cells, matching the detector's buffered union area within ±2%.
g. Evaluate baseline coverage using the same held-out trips and GPS-support criteria.
8. Compare spatial patterns across training partitions:
a. Calculate pairwise Jaccard similarities.
b. Match components using a minimum Jaccard similarity of 0.25, ensuring that each component is assigned to no more than one pattern within a partition.
c. Calculate recurrence Ri across five partitions.
d. Classify patterns with Ri ≥ 0.60 as recurrent.
9. Evaluate sensitivity:
- Recurrence thresholds: 0.20 to 1.00.
- Grid resolutions: 100, 250, and 500 m.
- Cross-resolution spatial overlap using IoU.
10. Report corridor geometry, connectivity, empirical GPS support, held-out coverage, and spatial recurrence.

The principal parameters and decision criteria used in the analytical workflow are summarized in Table 3.

Table 3. Parameters and criteria applied to evaluate the spatial stability, GPS-based support, and reproducibility of candidate mobility patterns.

Assessment Procedure	Parameter	Value
Bootstrap stability	Resampling unit	GPS-recorded trip
	Number of iterations	300
	Trip-incidence threshold	3%
	Minimum selection probability	0.80
GPS-based corridor support	Temporal sampling interval	10 s
	Spatial proximity thresholds	25, 50, 75, and 100 m
	Principal proximity threshold	75 m
	Minimum supporting observations	5 GPS samples
	Minimum within-buffer proportion	10%
Training–validation	Training proportion	70%
	Validation proportion	30%
	Number of repetitions	5
	Grid resolution	250 m
	Minimum trip incidence per cell	3% of training trips
	Bootstrap resampling iterations	300
	Minimum bootstrap selection probability	0.80
	Grid-cell connectivity	Four-neighbor
	Minimum component size	3 grid cells
	Validation proximity threshold	150 m
	Minimum supporting GPS observations	5 samples
	Minimum supporting observation proportion	10%
	Spatial baseline	Area-matched random baseline
	Baseline buffer radius	150 m
	Baseline replicates	20 per partition and category
	Maximum relative area mismatch	±2%
	Chronological validation subset	636 dated trips
	Chronological training trips	445
	Chronological validation trips	191
	Excluded trips without assignable dates	47
	Random holdout sensitivity analysis	5 splits using 636 dated trips
	Chronological holdout repetitions	1
Spatial recurrence	Similarity metric	Jaccard coefficient
	Minimum Jaccard similarity	0.25
	Pattern matching method	Binary integer programming over anchor-based candidate patterns.
	Minimum recurrence	3 of 5 repetitions (60%)
	Recurrence thresholds examined	0.20, 0.40, 0.60, 0.80, and 1.00
Spatial-resolution sensitivity	Grid resolutions examined	100, 250, and 500 m
	Spatial overlap metric	Intersection-over-union (IoU)
	Common comparison grid	50 m
	Number of train–test splits	5

Note: Bootstrap resampling and training–validation partitioning were performed at the GPS-recorded trip level. The same spatial detection criteria were applied to the complete dataset and training subsets. Spatial recurrence was evaluated across five partially overlapping training partitions; therefore, recurrence frequencies represent descriptive reproducibility rather than independent statistical evidence. The area-matched baseline and spatial-resolution sensitivity analyses were used as complementary assessments of the detected mobility patterns.

Figure 5 illustrates the spatial overlap assessment using a recurrent public transport pattern (P1) identified across two training partitions. Panels (a) and (b) show the spatial distribution of grid cells associated with the pattern in partitions 1 and 5, respectively, revealing a similar spatial configuration despite differences in the training samples. Panel (c) distinguishes the grid cells shared by both partitions (green) from the cell identified exclusively in partition 5 (orange). All five cells identified in partition 1 were also present in partition 5, which contained one additional cell. The resulting Jaccard similarity coefficient of 0.833 indicates that approximately 83.3% of the distinct grid cells in the union of both components were shared. This example illustrates how spatial correspondence was quantified to assess pattern recurrence across training partitions, without implying that the corresponding road segments were necessarily traversed by the same trips.

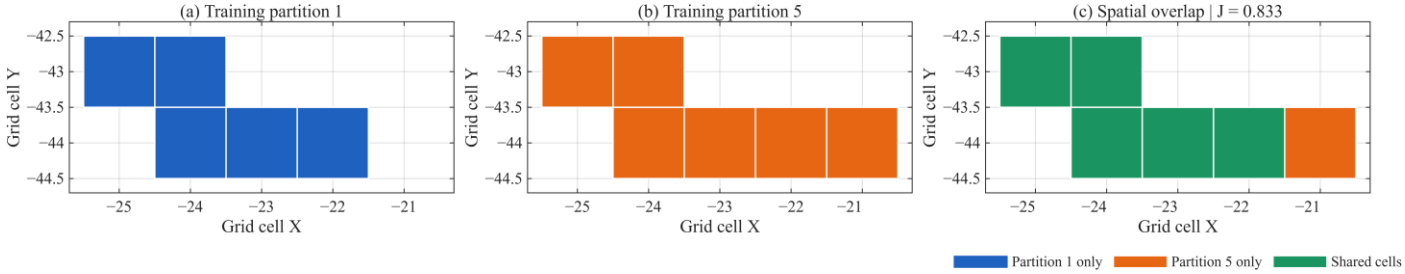


Figure 5. Spatial overlap of recurrent public transport mobility pattern P1 across two training partitions: (**a**) grid cells identified in partition 1; (**b**) corresponding grid cells identified in partition 5; and (**c**) spatial comparison highlighting shared cells (green) and the additional cell identified exclusively in partition 5 (orange). The Jaccard similarity coefficient was 0.833.

3. Results

3.1. Multimodal GPS Mobility Characteristics

The analysis was based on 683 GPS-recorded trips comprising 1,342,876 observations collected across eight transportation modes in the Metropolitan District of Quito between July 2025 and August 2026. These trips were grouped into public transport (261 trips), private motorized transport (115 trips), and active mobility (307 trips), as described in Table 1. For the descriptive characterization of trip-level mobility indicators, a minimum temporal coverage threshold of 80% was applied, resulting in 570 eligible trips (83.5% of the original dataset). This criterion reduced the influence of discontinuous GPS recordings on traveled distance and speed calculations. The complete dataset of 683 trips was retained for the subsequent spatial analyses.

Table 4 summarizes the median trip-level characteristics of the 570 eligible trips, while Figure 6 illustrates their distributions across the three transportation categories. Trip durations were relatively similar, with median values of 30.80 min for public transport, 35.90 min for private motorized transport, and 31.47 min for active mobility (Figure 6a). Nevertheless, the boxplots indicate differences in the dispersion of recorded travel durations, particularly between private motorized transport and active mobility.

The descriptive analysis revealed differences in travel distance and speed across transportation categories (Figure 6b,c). Private motorized transport exhibited the highest median trip distance (13.61 km) and median of trip-level mean speeds (22.88 km/h), followed by public transport (9.06 km and 16.31 km/h, respectively). Active mobility presented substantially shorter median trip distances (2.66 km) and median trip-level mean speeds (4.80 km/h). These differences highlight the distinct operational characteristics of the transportation categories and provide context for interpreting the spatial distributions and extent of the mobility patterns identified in subsequent analyses.

Stationary time also varied across transportation categories (Figure 6d). Public transport exhibited the highest median stationary proportion (31.50%), followed by private motorized transport (24.11%) and active mobility (11.16%). These differences may reflect variations in traffic interruptions, scheduled stops, and other temporary stationary conditions; however, their individual causes cannot be distinguished from the available GPS observations.

The GPS data-quality assessment identified no invalid coordinates, timestamps, or trip identifiers, and no duplicate consecutive observations. Sensitivity analysis of the temporal-gap threshold showed that the number of eligible trips varied from 570 at 10 and 20 s to 572 and 577 at 30 and 60 s, respectively. The corresponding mean temporal coverage increased from 90.55% to 91.41%, indicating limited sensitivity of trip eligibility to the selected interruption threshold. Additionally, 52 GPS observation intervals exhibited implied speeds exceeding 130 km/h, affecting 23 recorded trips. Excluding these intervals reduced the total estimated travel distance across the dataset by only 0.0398%, without changing the number of eligible trips. These findings indicate that the aggregate descriptive mobility indicators were relatively insensitive to the examined GPS recording interruptions and high-speed displacement anomalies, although they do not constitute independent validation of the recorded trajectories.

Table 4. Descriptive statistics of GPS-recorded trips by transportation category.

Transportation Category	Trips (*n*)	Duration (min)	Distance (km)	Mean Speed (km/h)	Stationary Time (%)
Public transport	190	30.80	9.06	16.31	31.50
Private motorized transport	101	35.90	13.61	22.88	24.11
Active mobility	279	31.47	2.66	4.80	11.16
All modes	570	31.72	5.57	11.14	21.41

Note: Values represent trip-level medians, except for the number of trips. Only trips meeting the minimum temporal coverage threshold of 80% were included (n = 570). Stationary time was calculated using a speed threshold of 0.5 m/s.

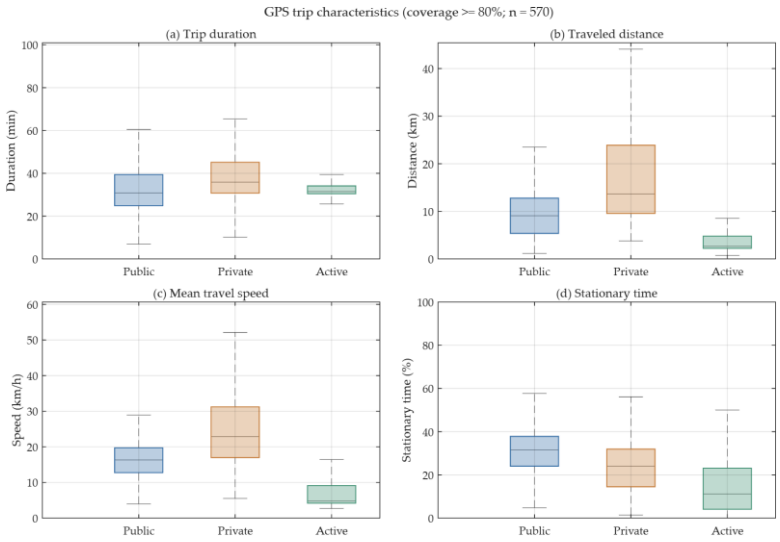


Figure 6. Distribution of GPS-derived trip characteristics across public transport, private motorized transport, and active mobility: (**a**) trip duration, (**b**) traveled distance, (**c**) mean travel speed, and (**d**) stationary time proportion. Only trips meeting the minimum temporal coverage threshold of 80% were included (n = 570).

3.2. Spatial Distribution and Road Network Reconstruction of Mobility Corridors

The spatial analysis of the 683 GPS-recorded trips revealed candidate mobility corridors with distinct geographic configurations across the Metropolitan District of Quito.

The identification procedure was applied to the complete multimodal dataset and separately to public transport, private motorized transport, and active mobility, allowing spatial concentrations to be examined at different levels of modal aggregation. Figure 3 illustrates the distribution of selected grid cells across these categories, highlighting differences in the geographic extent and concentration of recorded mobility. These patterns reflect the spatial coverage of the sampled trajectories rather than a comprehensive representation of travel demand throughout the city.

Following the identification of spatially connected grid-cell patterns, selected candidate corridors were associated with the reference road network. This procedure identified road segments located within the specified spatial association tolerance and subsequently evaluated their topological connectivity. Where disconnected segments could be linked through the reference network, additional road segments were incorporated subject to the maximum connection distance and detour constraints described in Section 2.3. Figure 4 illustrates the road-network reconstruction of multimodal corridor C17 and the corresponding segment-level GPS proximity support. Panel (a) distinguishes the originally associated road segments from the additional network connectors, whereas panel (b) identifies segments supported by observations from at least two distinct GPS-recorded trips within a 75 m proximity threshold.

A complementary GPS proximity assessment of multimodal corridor C17 showed that 45 of its 46 originally associated road segments (97.8%) were supported by observations from at least two distinct recorded trips within 75 m. These supported segments represented 3.839 km of the 3.866 km originally associated road length (99.29%). Among the 567 additional network connectors, 557 segments (98.2%) met the same support criterion, corresponding to 16.811 km of the 17.128 km added road length (98.15%). Overall, 204 of the 683 recorded trips contributed GPS proximity support to at least one reconstructed C17 segment. These findings indicate substantial spatial correspondence between the reconstructed road network and the recorded GPS trajectories, although proximity-based support does not establish complete segment traversal or independent validation.

The spatial detection procedure identified 54 components across the four analytical groups: 14 for all transport modes, 13 for public transport, 23 for private transport, and 4 for active mobility. The three largest components from each group were retained for detailed road-network assessment, resulting in 12 selected components comprising 726 of the 1012 identified grid cells (71.7%).

The complementary modal-composition analysis showed that corridors C17, C2, and C10 were intersected by 247, 184, and 68 distinct GPS-recorded trips, respectively (Table 5). All eight transportation modes were represented in each corridor, although their contributions varied. Bus trips accounted for the largest proportion in C17 (38.06%) and C2 (26.09%), followed by walking trips (19.03% and 23.91%, respectively). In contrast, C10 exhibited greater contributions from trolleybus (26.47%) and Ecovía (20.59%) trips. These results indicate that the pooled spatial concentrations include trajectories from multiple transportation modes, while revealing differences in modal composition among corridors.

Table 5. Modal composition of distinct GPS-recorded trips intersecting selected pooled mobility corridors.

Transportation Mode	C17 (n = 247)	C2 (n = 184)	C10 (n = 68)
Car	6.88%	12.50%	7.35%
Bicycle	11.74%	11.96%	4.41%
Bus	38.06%	26.09%	17.65%
Walking	19.03%	23.91%	13.24%
BRT corridor	2.83%	2.17%	7.35%
Ecovía	4.86%	5.98%	20.59%

Motorcycle	9.31%	9.78%	2.94%
Trolleybus	7.29%	7.61%	26.47%

Note: Percentages represent the distribution of distinct GPS-recorded trips intersecting each corridor's selected 250 m grid cells. Each trip was counted once per corridor, although a trip could intersect multiple corridors. These percentages describe the available sample and do not represent population-level modal shares or verified traversal of reconstructed road segments.

Road-network reconstruction revealed substantial differences in topological connectivity across the selected corridors (Table 6). Multimodal corridors C17 and C10 formed single connected components, whereas C2 retained five components. Public transport corridor C4 exhibited 13 connected components, with 52.5% of its reconstructed length contained in the largest component. Private motorized corridors showed greater fragmentation, particularly C3 (27 components) and C6 (20 components), while C22 had the lowest largest-component proportion (23.0%). Active mobility corridors generally exhibited fewer connected components. These differences demonstrate that reconstructed network connectivity varies considerably across transportation categories and does not necessarily indicate observed travel continuity

The number of additional road segments reflects the extent of algorithmic network reconstruction rather than observed travel demand. Although the GPS proximity assessment provides evidence of spatial correspondence, it does not confirm continuous traversal of individual segments or establish independent validation of the reconstructed network.

Table 6. Road network reconstruction and topological connectivity characteristics of selected GPS-derived candidate mobility corridors.

Category	Corridor	Original Segments	Original Length (km)	Original GPS-Supported (km)	Added Segments	Added Length (km)	Added GPS-Supported (km)	Connected Components	Largest Component (%)
All modes	C17	46	3.866	3.538	567	17.128	15.643	1	100.0
	C2	38	3.232	2.009	333	10.805	8.887	5	91.2
	C10	17	1.155	0.827	206	6.315	5.606	1	100.0
Public transport	C4	163	11.309	5.573	2049	58.006	38.799	13	52.5
	C21	28	2.371	0.909	210	9.205	5.441	4	51.5
	C12	17	0.960	0.587	149	4.988	4.093	1	100.0
Private motorized	C3	152	10.532	4.815	1426	43.451	29.152	27	61.5
	C6	138	11.393	6.766	1231	40.834	28.338	20	64.9
	C22	31	1.648	1.285	239	7.018	5.937	12	23.0
Active mobility	C2	29	2.847	1.566	317	9.482	7.681	2	98.6
	C6	10	0.799	0.675	79	2.886	2.825	2	88.0
	C9	9	0.606	0.472	63	2.739	2.644	1	100.0

Note: Original segments represent road segments associated with selected GPS-derived spatial cells. Added segments correspond to network-based connections introduced during reconstruction. GPS-supported length represents the portion of each segment located within 30 m of GPS observations from the corresponding transportation category, estimated using longitudinal sampling at intervals of up to 10 m. GPS proximity provides indirect spatial support and does not confirm actual traversal or temporal continuity. The largest-component percentage represents the proportion of total reconstructed road length contained in the largest connected component.

For corridor C17, the reconstructed network comprised 3.866 km of originally associated road segments and 17.128 km of added connectors. At a 30 m GPS proximity threshold, 3.538 km (91.52%) of the original road length and 15.643 km (91.33%) of the added connector length showed spatial correspondence with GPS observations. These results

indicate substantial proximity-based support for both segment groups, although they do not establish that the reconstructed road segments were actually traversed.

3.3. Spatial Stability and Empirical GPS Support

The spatial stability of the identified mobility patterns was examined through trip-level bootstrap resampling, while their empirical correspondence with recorded trajectories was assessed using GPS-based proximity criteria. These complementary analyses addressed two distinct dimensions of robustness: the persistence of spatial concentrations under changes in the sampled trips and the proportion of recorded journeys spatially associated with the reconstructed candidate corridors. The bootstrap analysis identified cells that consistently satisfied the selection criterion across repeated resampling iterations, providing an internally stable basis for candidate corridor identification. However, this stability does not establish that the corresponding reconstructed road segments were traversed by the recorded trips.

Empirical GPS support varied considerably across the 12 reconstructed candidate corridors, as summarized in Table 6 and illustrated in Figure 7. Using a reference proximity threshold of 75 m, public transport corridor C4 exhibited the highest trip-level support, with 174 of 261 trips (66.7%) satisfying the minimum spatial correspondence criteria. Private motorized corridors C3 and C6 were supported by 53 of 115 trips (46.1%) and 44 of 115 trips (38.3%), respectively. Within the complete multimodal dataset, corridor C17 was supported by 204 of 683 trips (29.9%), followed by C2 with 132 trips (19.3%). These differences indicate that the reconstructed corridors captured varying proportions of the mobility trajectories recorded within each transportation category.

The trip-level support reported in Table 7 differs from the segment-level proximity assessment presented for corridor C17 in Section 3.2. Whereas trip-level support required a minimum number and proportion of sampled GPS observations near the reconstructed corridor, segment-level support assessed whether individual road segments were located within 75 m of observations from at least two distinct recorded trips. Consequently, a high percentage of supported road segments does not necessarily imply that a similarly high proportion of recorded trips satisfied the corridor-level support criteria.

Table 7. Empirical GPS support for selected reconstructed mobility corridors at a 75 m proximity threshold.

Category	Corridor	Supporting Trips (n)	Trip Support (%)	GPS Observations Near Corridor (%)
All modes	C17	204	29.9	12.0
	C2	132	19.3	6.5
	C10	45	6.6	2.1
Public transport	C4	174	66.7	33.5
	C21	26	10.0	6.2
	C12	12	4.6	1.5
Private motorized	C3	53	46.1	15.5
	C6	44	38.3	17.5
	C22	8	7.0	1.4
Active mobility	C2	52	16.9	8.7
	C6	30	9.8	2.8
	C9	30	9.8	3.6

Note: Trip support represents the percentage of trips satisfying the minimum correspondence criteria of five sampled GPS observations and at least 10% of sampled points within 75 m of a reconstructed corridor. GPS observations were sampled at 10-second intervals. Percentages are calculated relative to the corresponding transportation category.

The proportion of sampled GPS observations located near the reconstructed corridors also differed across transportation categories. At the 75 m threshold, public transport corridor C4 accounted for 33.5% of the sampled GPS observations within its corresponding category, whereas private motorized corridors C3 and C6 accounted for 15.5% and 17.5%, respectively. For active mobility, corridor C2 was supported by 52 of 307 trips (16.9%), with 8.7% of sampled observations located within the proximity threshold. These findings demonstrate that trip-level support and point-level spatial coverage capture different characteristics of corridor correspondence and should therefore be interpreted jointly.

Figure 7 compares the trip-level support percentages of the 12 candidate corridors. Public transport corridor C4 exhibited the highest support, followed by private motorized corridors C3 and C6. In contrast, several corridors showed support below 10%, including public transport corridor C12, private motorized corridor C22, and active mobility corridors C6 and C9. This variation indicates that spatial concentrations identified through grid-based analysis may exhibit substantially different levels of correspondence with the recorded GPS trajectories after road-network reconstruction.

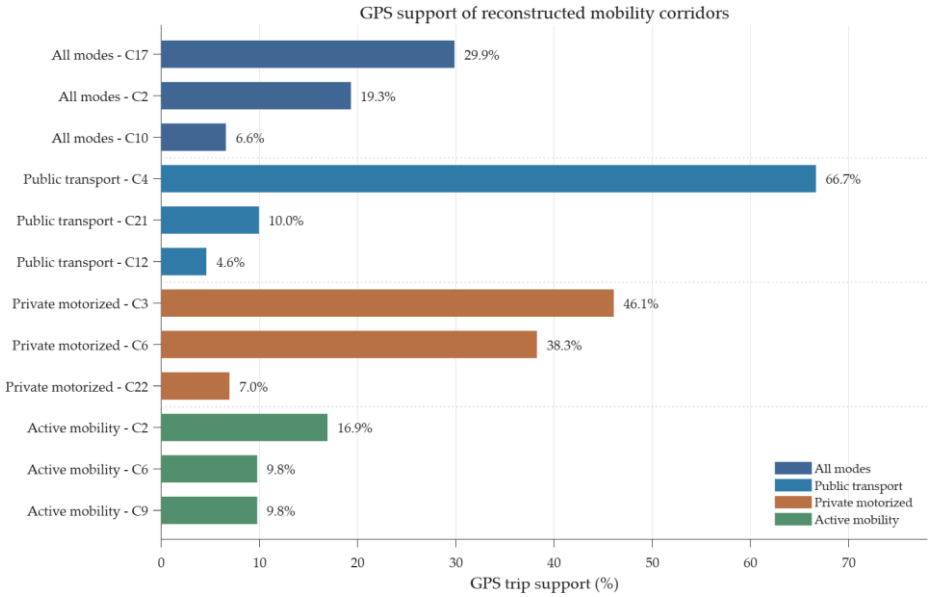


Figure 7. Trip-level GPS support for 12 selected candidate mobility corridors across transportation categories at a 75 m proximity threshold. Supporting trips were required to have at least five sampled GPS observations and at least 10% of sampled points within the specified proximity distance. Percentages were calculated relative to the total number of trips in each transportation category.

Overall, the results indicate that high topological connectivity does not necessarily imply extensive empirical support from GPS trajectories. For example, multimodal corridor C17 formed a single connected component following road-network reconstruction (Section 3.2), but only 29.9% of all recorded trips satisfied the specified corridor-level support criteria. The segment-level assessment presented in Table 6 identified GPS proximity support for 91.52% of the originally associated road length and 91.33% of the added connector length at a 30 m threshold. Furthermore, because spatial proximity was evaluated against the reconstructed corridors, including network-based connectors, the resulting support percentages cannot independently confirm that the additional road segments were traversed. These findings emphasize the distinction between geometric continuity and empirical spatial correspondence, motivating the out-of-sample reproducibility assessment presented in Section 3.4.

3.4. Spatial Reproducibility and Mode-Specific Differences

Spatial reproducibility was evaluated using five repeated, transportation-mode-stratified 70/30 train–test partitions, with 477 trips assigned to training and 206 reserved for validation in each partition. Using a 150 m validation proximity threshold, mean held-out trip coverage reached 57.96% for the complete multimodal dataset, 73.16% for public transport, 82.86% for private motorized transport, and 49.35% for active mobility. The corresponding area-matched random baseline yielded 16.16%, 61.22%, 86.71%, and 19.73%, respectively. The detector exceeded the baseline by 41.80, 11.94, and 29.62 percentage points for the complete dataset, public transport, and active mobility, respectively, whereas private motorized transport showed 3.85 percentage points lower coverage than the baseline. The relative difference in buffered union area remained below the predefined 2% tolerance across all baseline configurations. Figure 8 compares detector and baseline coverage across transportation categories and illustrates the variation among the five partitions.

A complementary analysis was performed using the 636 trips with assignable recording dates. Five stratified random holdout partitions yielded mean coverage values of 64.00%, 74.59%, 81.94%, and 56.47% for all modes, public transport, private motorized transport, and active mobility, respectively. A single chronological partition, using 18 July 2026 as the training cutoff, yielded corresponding values of 70.68%, 64.29%, 92.86%, and 31.29%. The chronological validation subset comprised 14 public transport trips, 14 private motorized trips, and 163 active-mobility trips. These results are descriptive because the modal composition differed substantially between partitions and some assigned recording dates could not be independently verified against the original GPS files. Consequently, the chronological analysis provides a preliminary temporal sensitivity assessment rather than confirmation of independent temporal generalization. The corresponding comparison is presented in Table 8.

Table 8. Comparison of held-out GPS-trip coverage across the original random, date-restricted random, and chronological validation procedures.

Transportation Category	Original Random Holdout (683 Trips)	Date-Restricted Random Holdout (636 Trips)	Chronological Holdout (636 Trips)
All modes	57.96 ± 4.10	64.00 ± 1.73	70.68
Public transport	73.16 ± 1.65	74.59 ± 5.27	64.29
Private motorized	82.86 ± 7.28	81.94 ± 10.35	92.86
Active mobility	49.35 ± 4.90	56.47 ± 5.26	31.29

Note: Values are percentages. Random holdout results are expressed as mean ± standard deviation across five stratified 70/30 partitions. The original random holdout used an area-matched spatial baseline, whereas the complementary date-restricted analyses were conducted separately. The chronological holdout comprised 445 training trips recorded on or before 18 July 2026 and 191 validation trips assigned to subsequent dates, with no overlap in assigned dates. Forty-seven trips without usable dates were excluded. Some dates were inferred from filenames and could not be independently verified. Chronological results correspond to a single exploratory partition and should not be interpreted as evidence of temporal independence.

To evaluate the reproducibility of localized spatial concentrations, patterns identified across the five training partitions were matched using the Jaccard similarity coefficient and an anchor-based candidate generation followed by binary integer optimization, with a minimum similarity threshold of 0.25. At the reference recurrence threshold of 0.60, the complete multimodal dataset contained 13 recurrent patterns among 27 identified patterns (48.1%). Public transport presented 8 recurrent patterns among 30 (26.7%), private motorized transport exhibited 18 among 41 (43.9%), and active mobility contained 4 among 9 (44.4%). Overall, 43 of the 107 patterns identified across the four analysis

categories (40.2%) satisfied the recurrence criterion. These findings indicate differences in the persistence of spatial concentrations across transportation categories, without implying that the repeated partitions constituted independent samples.

The sensitivity analysis showed that the number of recurrent spatial patterns decreased as the minimum recurrence threshold increased (Figure 9). For the complete multimodal dataset, 27 patterns satisfied a threshold of 0.20, compared with 16 at 0.40, 13 at 0.60, and 8 at both 0.80 and 1.00. Public transport exhibited a reduction from 30 to 3 patterns between the lowest and highest thresholds, while private motorized transport decreased from 41 to 6 and active mobility from 9 to 3. These findings indicate that the identification of recurrent patterns depends on the selected persistence criterion. The reference threshold of 0.60 required a pattern to be identified in at least three of the five partitions, providing an intermediate criterion between occasional detection and recurrence across all partitions.

Spatial-resolution sensitivity revealed substantial differences in corridor fragmentation, geographic extent, and held-out trip intersection (Figure 10). For the complete multimodal dataset, the mean number of retained components decreased from 32.8 at a 100 m grid resolution to 14.4 at 250 m and 4.6 at 500 m. Conversely, the mean corridor area increased from 2.36 to 10.24 and 32.70 km^2, respectively. The proportion of held-out trips intersecting detected corridor cells increased from 54.83% at 100 m to 72.78% at 250 m and 84.10% at 500 m. These results demonstrate a trade-off between spatial detail, corridor fragmentation, and geographic coverage.

Cross-resolution spatial agreement was relatively limited, with intersection-over-union (IoU) values of 0.214 between 100 and 250 m, 0.312 between 250 and 500 m, and 0.072 between 100 and 500 m for the complete multimodal dataset. Although coarser grids covered most of the areas identified at finer resolutions, their substantially larger spatial extent resulted in relatively low intersection-over-union values when corridor surfaces were compared on a common 50 m grid. The 250 m reference grid therefore represents a practical compromise between spatial fragmentation and geographic generalization rather than a statistically optimal resolution. Overall, the combined validation, recurrence, and spatial-resolution analyses demonstrate that the identified mobility patterns are sensitive to both the sampling configuration and the spatial representation adopted, see Table 9.

Table 9. Recurrence of local GPS-derived mobility patterns across five training partitions.

Transportation Category	Identified Patterns (*n*)	Recurrent Patterns (*n*)	Recurrence (%)
All modes	27	13	48.1
Public transport	30	8	26.7
Private motorized transport	41	18	43.9
Active mobility	9	4	44.4
Total	107	43	40.2

Note: Spatial patterns were matched across five partially overlapping training partitions using a minimum Jaccard similarity threshold of 0.25. Anchor-based candidate patterns were generated and subsequently consolidated through binary integer optimization. Patterns identified in at least three partitions ($Ri \geq 0.60$) were classified as recurrent. Percentages represent the proportion of recurrent patterns within each transportation category. The total aggregates the four analysis categories and does not represent a count of geographically unique patterns. Recurrence frequencies provide descriptive evidence of spatial reproducibility rather than independent statistical inference.

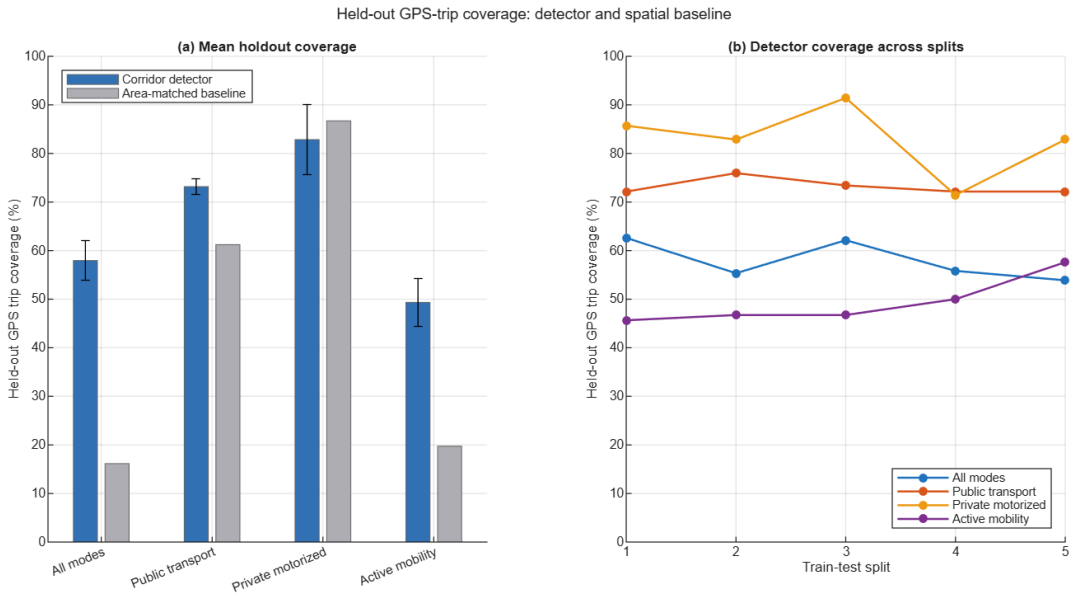


Figure 8. Held-out GPS-trip coverage across five stratified 70/30 training–validation partitions. (**a**) Mean coverage of the corridor detector compared with a random spatial baseline matched to the union area of the 150 m buffers (±2% tolerance), using 20 baseline repetitions per split and transportation category. Error bars indicate the standard deviation of detector coverage across the five partitions. (**b**) Detector coverage for each transportation category across the five training–validation partitions.

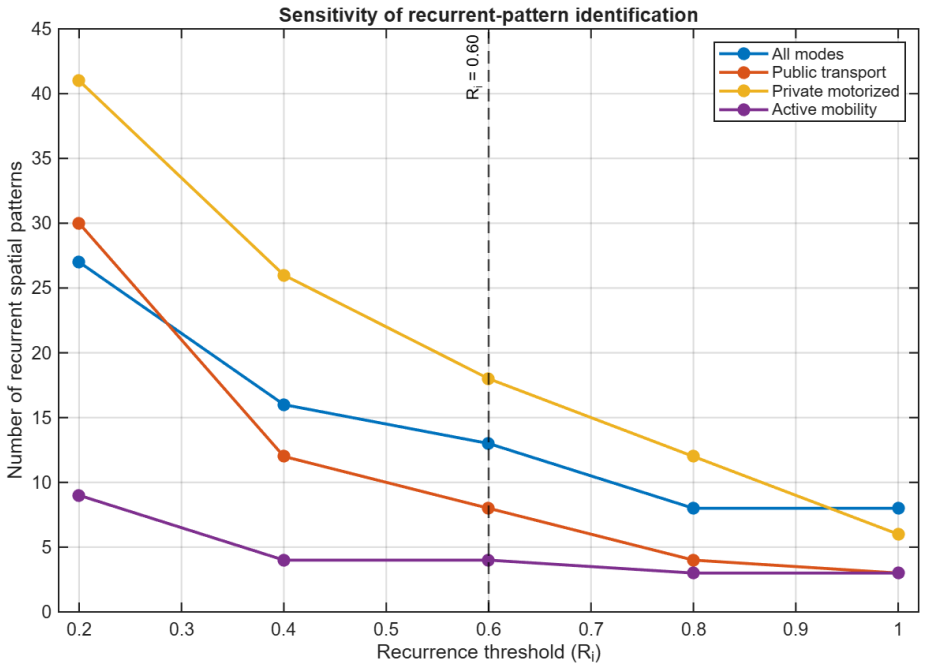


Figure 9. Sensitivity of recurrent spatial pattern identification to the recurrence threshold (Ri) across transportation categories. Pattern recurrence was calculated from five partially overlapping training partitions using anchor-based candidate generation and binary integer optimization. The dashed vertical line indicates the adopted recurrence threshold (Ri = 0.60), corresponding to identification in at least three partitions.

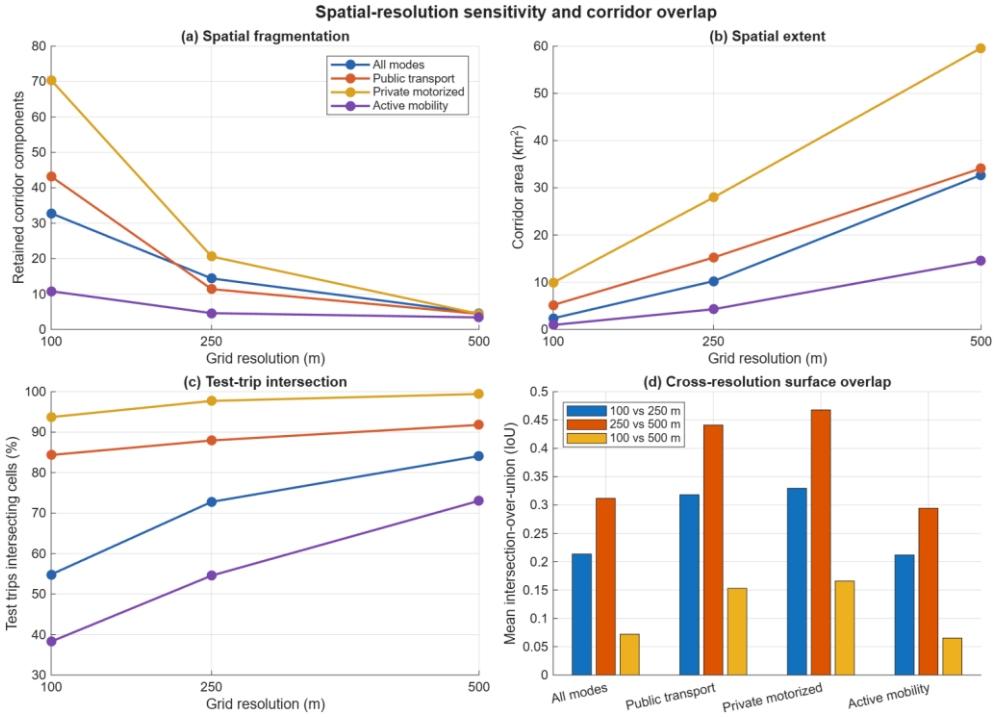


Figure 10. Sensitivity of GPS-derived corridor detection to spatial grid resolution: (**a**) mean number of retained corridor components; (**b**) mean corridor area; (**c**) percentage of held-out trips intersecting detected corridor cells; and (**d**) mean intersection-over-union (IoU) between corridor surfaces obtained at different spatial resolutions. Values represent averages across five stratified train–test splits.

4. Discussion

The results demonstrate that multimodal GPS trajectories can be used to identify spatially concentrated mobility patterns in the Metropolitan District of Quito, although their interpretation depends on the distinction between spatial incidence, road-network reconstruction, empirical GPS support, and reproducibility. The analysis of 683 trips and 1,342,876 georeferenced observations revealed differences in travel characteristics and corridor configurations across public transport, private motorized transport, and active mobility. More importantly, the results indicate that the identification of spatial concentrations alone is insufficient to establish whether the resulting structures correspond to frequently observed travel corridors or remain reproducible under changes in the sampled trips.

The identification of GPS-derived spatial concentrations is consistent with the findings of Shi et al., who developed a framework for detecting recreational hotspots from volunteered GPS trajectories while accounting for spatial heterogeneity. Their research demonstrated the importance of considering uneven geographic distributions when identifying meaningful concentrations of urban activity. In the present study, spatial concentrations were identified using the number of distinct trips recorded within regular grid cells, followed by trip-level bootstrap resampling. Unlike approaches focused primarily on localized activity hotspots, the proposed framework extends the analysis toward connected spatial components and their association with road infrastructure. Nevertheless, the observed concentrations should be interpreted as patterns within the sampled trajectories rather than as direct measurements of metropolitan travel demand [24].

The sensitivity of spatial identification to analytical parameters is another relevant consideration. Yan et al. demonstrated that conventional hotspot detection methods may be affected by the use of global parameters when analyzing geographically heterogeneous

data, motivating the development of adaptive detection strategies. In the present study, this concern was addressed through sensitivity analyses involving grid resolutions of 100, 250, and 500 m, with 250 m adopted as the principal spatial configuration and a minimum incidence threshold of 3% of distinct training trips applied to identify candidate grid cells. Furthermore, 300 trip-level bootstrap iterations were used to retain cells with a selection probability of at least 0.80. The resolution sensitivity analysis showed that finer grids produced more fragmented spatial patterns, whereas coarser grids increased corridor coverage at the expense of spatial specificity. For all transportation modes, the mean number of retained components decreased from 32.8 at 100 m to 14.4 at 250 m and 4.6 at 500 m, while the corresponding detected area increased from 2.36 to 10.24 and 32.70 km², respectively. Moreover, the relatively low intersection-over-union values between resolutions (0.214 for 100–250 m and 0.312 for 250–500 m) demonstrated that the spatial boundaries of the identified patterns remained dependent on grid resolution. Accordingly, the 250 m configuration represents a practical compromise between spatial fragmentation and excessive aggregation rather than a universally optimal resolution. These procedures reduce dependence on a single sample realization, although they do not eliminate the influence of grid geometry, spatial resolution, or uneven modal coverage. Consequently, the identified corridors represent stable concentrations under the specified analytical configuration rather than spatial structures that are independent of parameter selection [25].

Road-network reconstruction revealed substantial differences in topological connectivity among the selected candidate corridors. Multimodal corridor C17 formed a single connected component after the incorporation of 567 additional road segments, whereas public transport corridor C4 retained 13 connected components and required 2049 additional segments. These findings illustrate that spatially concentrated GPS observations do not necessarily translate into continuous road-aligned structures without algorithmic intervention. Wöltche's research on map matching emphasized the challenges introduced by positioning errors, incomplete trajectory information, and discrepancies between recorded movements and digital road networks. In contrast to trajectory-level map matching, the present study associated aggregated spatial concentrations with nearby road segments and subsequently introduced network-based connectors. Therefore, the reconstructed connectivity should be understood as a topological property of the candidate corridor rather than evidence that all connecting segments were traversed by the recorded trips [26].

The empirical GPS support analysis provided an additional distinction between network continuity and observed spatial correspondence. At a 75 m proximity threshold, public transport corridor C4 exhibited the highest support, with 174 of 261 trips (66.7%) satisfying the correspondence criteria. Private motorized corridors C3 and C6 were supported by 46.1% and 38.3% of their respective trips, whereas multimodal corridor C17, despite forming a single connected component, was supported by 29.9% of all recorded trips. A complementary segment-level assessment showed that 99.29% of the originally associated road length and 98.15% of the additional connector length in C17 were supported by GPS observations from at least two distinct trips within 75 m. Although these findings indicate substantial internal spatial correspondence, they do not demonstrate that every reconstructed segment was fully traversed by the recorded vehicles. Furthermore, the segment-level proximity percentages and the proportion of trips supporting C17 describe different properties and should not be interpreted as equivalent measures of corridor utilization. These differences indicate that a geometrically continuous reconstructed corridor may capture only a limited proportion of the observed journeys. They also demonstrate that the number of added road segments and the resulting network connectivity should not be used as substitutes for empirical evidence of corridor utilization.

The complementary modal-composition analysis further clarifies the interpretation of pooled mobility corridors. Although all eight transportation modes contributed GPS

observations to C17, C2, and C10, their relative contributions were heterogeneous. Bus trips predominated in C17 and C2, whereas trolleybus and Ecovía trips were more prominent in C10. These findings support the interpretation of multimodal corridors as spatial concentrations derived from mixed-mode trajectories, rather than evidence of balanced modal participation or shared traversal of identical road segments. The observed percentages reflect the composition of the available sample and cannot be extrapolated to metropolitan modal shares or transportation demand.

Differences between transportation categories further emphasize the importance of mode-specific spatial interpretation. Public transport exhibited the highest median stationary proportion (31.50%), while private motorized transport presented the greatest median traveled distance (13.61 km) and highest median of trip-level mean speeds (22.88 km/h). Active mobility showed shorter median distances (2.66 km) and lower speeds (4.80 km/h), reflecting the different travel characteristics represented in the dataset. Gao et al., in their investigation of map matching for urban cycling trajectories, highlighted the importance of accounting for cycling-specific network accessibility, positional uncertainty, and route-selection characteristics. These considerations are relevant to the present findings because the road infrastructure and spatial behavior associated with walking and cycling may differ substantially from those of motorized transport. Accordingly, the comparatively limited support of selected active mobility corridors should not be interpreted as evidence of low active mobility demand, but rather in relation to the spatial coverage of the recorded sample, the selected candidate patterns, and the road-network representation used in the analysis [27].

The repeated training–validation assessment revealed a distinction between held-out GPS-trip coverage and the reproducibility of localized spatial patterns. Across five transportation-mode-stratified partitions, mean held-out trip coverage reached 57.96% for all modes, 73.16% for public transport, 82.86% for private motorized transport, and 49.35% for active mobility. Compared with a spatial baseline matched to the union area of the 150 m buffers, the proposed detector achieved higher coverage for all modes (16.16% for the baseline), public transport (61.22%), and active mobility (19.73%). However, private motorized transport exhibited slightly lower coverage than its baseline (86.71%), indicating no coverage advantage in this category. These findings show that spatial coverage alone is insufficient to establish the reproducibility of localized mobility concentrations and should be interpreted relative to the geographic extent of the detected patterns.

The complementary validation analyses revealed heterogeneous spatial correspondence across transportation categories. Within the subset of 636 trips with assignable dates, overall coverage was 64.00% under repeated random partitions and 70.68% under a single chronological partition. Active mobility exhibited lower chronological coverage (31.29%) than under date-restricted random partitions (56.47%), whereas private motorized transport showed higher chronological coverage (92.86%). These differences cannot be attributed solely to temporal changes because the validation subsets differed substantially in modal composition. Moreover, some recording dates were inferred from filenames and could not be independently verified. Consequently, the chronological results should be interpreted as exploratory evidence of sensitivity to the assigned temporal partition, rather than confirmation of temporal generalization or independence between contributors.

The spatial recurrence analysis provided a more selective assessment of pattern reproducibility. Using a minimum Jaccard similarity of 0.25 and a recurrence threshold of $Ri \geq 0.60$, 43 of the 107 identified patterns (40.2%) were recurrent across the four analysis categories. The proportion of recurrent patterns was 48.1% for all modes (13 of 27), 26.7% for public transport (8 of 30), 43.9% for private motorized transport (18 of 41), and 44.4% for active mobility (4 of 9). The recurrent public transport pattern P1 illustrated in Figure 5 was identified in all five training partitions, with a Jaccard similarity of 0.833 between

partitions 1 and 5. Nevertheless, the relatively limited recurrence observed in several categories indicates that localized spatial concentrations remain sensitive to the composition of the training sample. Because the partitions partially overlap, these recurrence frequencies represent descriptive evidence of spatial consistency rather than independent statistical validation.

Several limitations should be considered when interpreting these findings. First, the GPS dataset represents recorded trips rather than a statistically representative sample of all journeys in Quito. Consequently, differences in trip counts, geographic coverage, and transportation-mode composition may influence the identification and relative support of spatial concentrations. Second, the fixed spatial grid introduces sensitivity to cell size and alignment, even when alternative resolutions are examined. Third, the road-network reconstruction incorporates algorithmically selected connections that have not been independently verified as traveled routes. Fourth, the empirical GPS support and holdout analyses use different spatial representations and proximity thresholds: 75 m from reconstructed road networks and 150 m from training-derived grid-cell centers, respectively. Their percentages therefore cannot be directly compared as equivalent validation metrics. Finally, the repeated 70/30 partitions provide internal out-of-sample assessments, while the complementary chronological holdout evaluates spatial correspondence with subsequently recorded trips. However, this assessment was based on a single temporal partition, with substantial differences in transportation-mode composition and limited validation-trip counts for public and private motorized transport. Therefore, the results do not establish long-term temporal stability, participant-level independence, or generalization to independently collected mobility datasets. Furthermore, the area-matched baseline approximately controls for spatial coverage but does not reproduce the connectivity, geometry, or geographical distribution of the detected corridors. Although binary integer optimization resolved competing component assignments within the generated candidate set, the resulting spatial patterns remained dependent on the anchor-based candidate generation procedure and the selected Jaccard similarity threshold (0.25). Therefore, the optimization should not be interpreted as an exhaustive search over all possible cross-partition component groupings. Accordingly, the reported recurrence rates should not be interpreted as statistically independent estimates of underlying metropolitan mobility structures.

The proposed framework should also be distinguished from approaches that explicitly examine the relationship between mobility demand and transport supply. While demand–supply assessments investigate the correspondence between observed travel requirements and available transportation services, the present study focuses on the spatial stability, network association, and reproducibility of GPS-derived mobility concentrations. Accordingly, the identified candidate corridors provide complementary spatial evidence but do not directly quantify transport demand, service adequacy, or unmet mobility needs. Integrating these dimensions would represent a relevant extension of the proposed framework [28].

The selection of the three spatially largest components per analytical group was intended to support a detailed assessment of road-network association and connectivity. Consequently, the resulting 12 corridors should not be interpreted as a comprehensive representation of all detected spatial components or as the most heavily traveled routes. Furthermore, repeated random partitions may share observations, and the absence of participant identifiers prevented participant-level separation between training and validation subsets. Therefore, the reported recurrence measures describe spatial consistency under trip-level resampling rather than independent participant-level reproducibility.

From an applied perspective, the proposed framework can support the exploratory identification of candidate mobility corridors and recurrent spatial concentrations that

warrant more detailed investigation. Its principal contribution lies in combining distinct-trip spatial incidence, bootstrap-based stability assessment, road-network association, empirical GPS correspondence, and repeated spatial reproducibility analysis within a common multimodal workflow. Future research should incorporate additional GPS observations across different periods and geographic sectors, evaluate the sensitivity of results to alternative grid alignments and network representations, and independently verify selected road connections using observed trajectories or reference routes. Future validation should also consider geographically and temporally independent GPS datasets, alternative spatial matching strategies, and baseline configurations that account for both spatial extent and network connectivity. Such developments would help distinguish persistent urban mobility structures from patterns influenced by sampling characteristics and analytical assumptions.

5. Conclusions

This study developed and applied a GPS-based analytical framework to identify and assess candidate multimodal mobility corridors in the Metropolitan District of Quito, Ecuador. The analysis integrated 683 GPS-recorded trips and 1,342,876 georeferenced observations across eight transportation modes, combining grid-based spatial incidence, trip-level bootstrap resampling, road-network reconstruction, empirical GPS support, and repeated spatial reproducibility assessment. The proposed approach enabled the identification of spatially concentrated mobility patterns while distinguishing their bootstrap-based spatial stability, topological connectivity, and correspondence with observed trajectories.

The results revealed differences in mobility characteristics and spatial corridor configurations across transportation categories. Private motorized transport exhibited the greatest median trip distance (13.61 km) and the highest median of trip-level mean speeds (22.88 km/h), whereas public transport presented the highest median stationary proportion (31.50%). Active mobility was characterized by shorter median distances (2.66 km) and lower speeds (4.80 km/h). These differences highlight the importance of considering transportation-mode heterogeneity when interpreting spatial concentrations derived from multimodal GPS observations. The complementary modal-composition assessment showed that the selected pooled corridors C17, C2, and C10 contained GPS-recorded trips from all eight transportation modes, although their relative contributions varied substantially across corridors.

Road-network reconstruction and empirical GPS support provided complementary evidence regarding the identified candidate corridors. Although several reconstructed corridors exhibited substantial topological connectivity, their correspondence with recorded trips varied considerably. Public transport corridor C4 achieved the highest empirical support among the selected candidates, with 66.7% of public transport trips satisfying the established GPS proximity criteria. In comparison, private motorized corridors C3 and C6 reached 46.1% and 38.3%, respectively, while multimodal corridor C17 was supported by 29.9% of all trips. For C17, the segment-level assessment showed GPS proximity support for 91.52% of the originally associated road length and 91.33% of the added connector length at a 30 m threshold. Nevertheless, this spatial correspondence does not confirm actual traversal of the reconstructed segments. These findings demonstrate that reconstructed network continuity does not necessarily imply extensive observed travel along the resulting corridor.

The spatial reproducibility assessment yielded mean held-out GPS-trip coverage of 57.96% for all transportation modes, 73.16% for public transport, 82.86% for private motorized transport, and 49.35% for active mobility. Relative to the area-matched spatial baseline, the detector showed higher coverage for all modes (16.16%), public transport (61.22%), and active mobility (19.73%), but slightly lower coverage for private motorized

transport (86.71%). Furthermore, 43 of the 107 identified spatial patterns (40.2%) were recurrent in at least three of the five training partitions. These findings indicate that held-out trip coverage and localized spatial recurrence capture complementary aspects of spatial reproducibility, while reconstructed network connectivity must be distinguished from empirically observed travel.

Complementary validation using trips with assignable recording dates yielded 64.00% overall coverage under repeated random partitions and 70.68% under a single chronological holdout. However, unequal modal representation, uncertainty in some assigned recording dates, and the absence of participant identifiers limit the interpretation of temporal independence and generalizability. The chronological assessment should therefore be regarded as exploratory.

The principal contribution of this research is a reproducible analytical framework that integrates complementary spatial assessment procedures to identify and interpret candidate mobility corridors from multimodal GPS data. Sensitivity analyses further demonstrated that spatial resolution influences corridor fragmentation, geographical extent, and overlap, with the 250 m grid providing a practical compromise between spatial detail and aggregation under the examined conditions. However, the findings are limited by the geographic and modal coverage of the recorded sample, sensitivity to spatial aggregation parameters, and the inclusion of algorithmically reconstructed road connections that were not independently verified as traveled routes. Additionally, the partially overlapping training partitions provide descriptive evidence of spatial recurrence rather than independent statistical validation. Consequently, the identified patterns should not be interpreted as representative measures of metropolitan travel demand. Future research should incorporate larger and temporally independent GPS datasets, evaluate alternative spatial resolutions and network representations, and verify candidate corridors against observed routes and transportation infrastructure. These developments would strengthen the applicability of the framework to urban mobility monitoring and transportation planning in Quito and other metropolitan environments.

Author Contributions: J.J.M.-C.; methodology, J.J.M.-C., J.P.-P. and J.R.-A.; software, J.P.-P. and J.R.-A.; validation, J.J.M.-C., J.P.-P. and J.R.-A.; formal analysis, J.J.M.-C., J.P.-P. and J.R.-A.; investigation, J.J.M.-C., J.P.-P. and J.R.-A.; resources, J.J.M.-C.; data curation, J.P.-P. and J.R.-A.; writing—original draft preparation, J.J.M.-C.; writing—review and editing, J.J.M.-C., J.P.-P. and J.R.-A.; visualization, J.J.M.-C., J.P.-P. and J.R.-A.; supervision, J.J.M.-C.; project administration, J.J.M.-C.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universidad Politécnica Salesiana, Ecuador, through research project No. 64-04-2026-05-29, entitled "Spatiotemporal Characterization of Dynamic Urban Mobility Patterns Using GPS Data and Machine Learning Models".

Institutional Review Board Statement: This study involved a non-interventional secondary analysis of de-identified GPS trip files. No names, identification numbers, email addresses, demographic variables, participant codes, or correspondence keys were collected. The analytical dataset could not be linked by the research team to specific individuals.

Informed Consent Statement: Not applicable. The dataset consisted of voluntarily contributed GPS trip files with no direct participant identifiers or correspondence key, and the analysis was conducted at the trip and aggregated spatial levels.

Data Availability Statement: The GPS trajectory data supporting the findings of this study are not publicly available due to privacy and confidentiality considerations. Although direct personal identifiers were not collected, the spatial and temporal characteristics of individual trajectories may present a risk of indirect participant identification. Requests for access to non-identifiable aggregated

or processed data may be directed to the corresponding author and will be considered subject to applicable privacy and institutional requirements.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GPS	Global Positioning System
GIS	Geographic Information System
DMQ	Metropolitan District of Quito
OSM	OpenStreetMap
OD	Origin–Destination
CI	Confidence Interval
J	Jaccard Similarity Coefficient

References

1. Xu, Y.; Olmos, L.E.; Mateo, D.; Hernando, A.; Yang, X.; González, M.C. Urban Dynamics through the Lens of Human Mobility. *Nat. Comput. Sci.* **2023**, *3*, 611–620. https://doi.org/10.1038/s43588-023-00484-5.
2. Zhao, P.; Wang, H.; Liu, Q.; Yan, X.Y.; Li, J. Unravelling the Spatial Directionality of Urban Mobility. *Nat. Commun.* **2024**, *15*, 4507. https://doi.org/10.1038/s41467-024-48909-7.
3. Nilforoshan, H.; Looi, W.; Pierson, E.; Villanueva, B.; Fishman, N.; Chen, Y.; Sholar, J.; Redbird, B.; Grusky, D.; Leskovec, J. Human Mobility Networks Reveal Increased Segregation in Large Cities. *Nature* **2023**, *624*, 586–592. https://doi.org/10.1038/s41586-023-06757-3.
4. Huang, X.; Jiang, Y.; Mostafavi, A. The Emergence of Urban Heat Traps and Human Mobility in 20 US Cities. *npj Urban Sustain.* **2024**, *4*, 6. https://doi.org/10.1038/s42949-024-00142-3.
5. Liu, E.; Mazzoli, M.; Yan, X.Y.; Ramasco, J.J. A Generalized Vector-Field Framework for Mobility. *Commun. Phys.* **2024**, *7*, 190. https://doi.org/10.1038/s42005-024-01672-z.
6. Zhong, Z.; Takayasu, H.; Takayasu, M. Human Mobility Description by Physical Analogy of Electric Circuit Network Based on GPS Data. *Sci. Rep.* **2024**, *14*, 13380. https://doi.org/10.1038/s41598-024-63719-z.
7. Kan, Z.; Liu, D.; Yang, X.; Lee, J. Measuring Exposure and Contribution of Different Types of Activity Travels to Traffic Congestion Using GPS Trajectory Data. *J. Transp. Geogr.* **2024**, *117*, 103896. https://doi.org/10.1016/J.JTRANGEO.2024.103896.
8. Mourtakos, V.; Mantouka, E.G.; Fafoutellis, P.; Vlahogianni, E.I.; Kepaptsoglou, K. Reconstructing Mobility from Smartphone Data: Empirical Evidence of the Effects of COVID-19 Pandemic Crisis on Working and Leisure. *Transp. Policy* **2024**, *146*, 241–254. https://doi.org/10.1016/J.TRANPOL.2023.11.018.
9. Gallotti, R.; Maniscalco, D.; Barthelemy, M.; De Domenico, M. Distorted Insights from Human Mobility Data. *Commun. Phys.* **2024**, *7*, 421. https://doi.org/10.1038/s42005-024-01909-x.
10. Stipancic, J.; Saunier, N.; Navidi, N.; Racine, E.B.; Miranda-Moreno, L.; Labbe, A. Evaluating Map Matching Algorithms for Smartphone GNSS Data: Matching Vehicle Trajectories to an Urban Road Network. *Transp. Res. Procedia* **2025**, *82*, 303–322. https://doi.org/10.1016/J.TRPRO.2024.12.045.
11. Alkhazraji, H.; Kanakis, T.; Al-Sherbaz, A.; Bhaya, W.S.; Agyeman, M.O. Enhancing Map Matching Accuracy Using Backtracking Technique. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *132*, 103988. https://doi.org/10.1016/J.JAG.2024.103988.
12. Li, G.; Li, Z.; Knoop, V.L.; van Lint, H. Unravelling Uncertainty in Trajectory Prediction Using a Non-Parametric Approach. *Transp. Res. Part C Emerg. Technol.* **2024**, *163*, 104659. https://doi.org/10.1016/J.TRC.2024.104659.
13. Tiznado-Aitken, I.; Vecchio, G.; Guzman, L.A.; Arellana, J.; Humberto, M.; Vasconcellos, E.; Muñoz, J.C. Unequal Periurban Mobility: Travel Patterns, Modal Choices and Urban Core Dependence in Latin America. *Habitat Int.* **2023**, *133*, 102752. https://doi.org/10.1016/J.HABITATINT.2023.102752.
14. Basso, F.; Frez, J.; Hernández, H.; Leiva, V.; Pezoa, R.; Varas, M. Crowding on Public Transport Using Smart Card Data during the COVID-19 Pandemic: New Methodology and Case Study in Chile. *Sustain. Cities Soc.* **2023**, *96*, 104712. https://doi.org/10.1016/J.SCS.2023.104712.
15. Basso, F.; Cox, T.; Pezoa, R.; Maldonado, T.; Varas, M. Characterizing Last-Mile Freight Transportation Using Mobile Phone Data: The Case of Santiago, Chile. *Transp. Res. Part A Policy Pract.* **2024**, *186*, 104149. https://doi.org/10.1016/J.TRA.2024.104149.

16. Pena-Bastidas, J.; Liu, J.; Jones, S. Measuring Traffic Speed Change After the Reallocation of Road Space for Cycling: A Data-Driven Analysis for Bogotá. *Cities* **2024**, *153*, 105296. https://doi.org/10.1016/J.CITIES.2024.105296.
17. Molina-Campoverde, J.J.; García, E.S.G.; Pilamunga, A.A.G. Identification of Recurrent Urban Vehicle Emission Hotspots Using Repeated GPS-Based Real-Driving Measurements and Vehicle-Specific Power: A Case Study in Quito, Ecuador. *Urban Sci.* **2026**, *10*, 614. https://doi.org/10.3390/URBANSCI10110614.
18. Quezada Larriva, A.; Orellana, D.; Guerrero Balarezo, M.L.; García, J.A.; Cárdenas Villenas, G.; Osorio Guerrero, P. Impact of Quito's First Metro Line on the Accessibility to Urban Opportunities. *J. Transp. Geogr.* **2023**, *108*, 103548. https://doi.org/10.1016/J.JTRANGEO.2023.103548.
19. Arias, J.F.; Bachmann, C. Quantifying the Relative Importance of Rapid Transit Implementation Barriers: Evidence from Ecuador. *J. Transp. Geogr.* **2022**, *100*, 103311. https://doi.org/10.1016/J.JTRANGEO.2022.103311.
20. Molina-Campoverde, J.J.; Rivera-Campoverde, N.; Molina Campoverde, P.A.; Bermeo Naula, A.K. Urban Mobility Pattern Detection: Development of a Classification Algorithm Based on Machine Learning and GPS. *Sensors* **2024**, *24*, 3884. https://doi.org/10.3390/S24123884.
21. Wang, L.; Saiz, A.; Li, W. Natural Fragmentation Increases Urban Density but Impedes Transportation and City Growth Worldwide. *Nat. Cities* **2024**, *1*, 642–653. https://doi.org/10.1038/s44284-024-00118-5.
22. Fu, X.; van Lierop, D.; Ettema, D. Is Multimodality Advantageous? Assessing the Relationship between Multimodality and Perceived Transport Adequacy and Accessibility in Different Travel Contexts. *Transp. Res. Part A Policy Pract.* **2024**, *179*, 103893. https://doi.org/10.1016/J.TRA.2023.103893.
23. Marin, V.; Molinero, C.; Arcaute, E. Uncovering Structural Diversity in Commuting Networks: Global and Local Entropy. *Sci. Rep.* **2022**, *12*, 1684. https://doi.org/10.1038/s41598-022-05556-6.
24. Shi, H.; Huang, H.; Ma, D.; Chen, L.; Zhao, M. Capturing Urban Recreational Hotspots from GPS Data: A New Framework in the Lens of Spatial Heterogeneity. *Comput. Environ. Urban Syst.* **2023**, *103*, 101972. https://doi.org/10.1016/J.COMPENVURBSYS.2023.101972.
25. Yan, Y.; Quan, W.; Wang, H. A Data-Driven Adaptive Geospatial Hotspot Detection Approach in Smart Cities. *Trans. GIS* **2024**, *28*, 303–325. https://doi.org/10.1111/TGIS.13137.
26. Wöltche, A. Open Source Map Matching with Markov Decision Processes: A New Method and a Detailed Benchmark with Existing Approaches. *Trans. GIS* **2023**, *27*, 1959–1991. https://doi.org/10.1111/TGIS.13107. Correction in *Trans. GIS* **2024**, *28*, 697–698. https://doi.org/10.1111/tgis.13162.
27. Gao, T.; Daamen, W.; Krishnakumari, P.; Hoogendoorn, S. Map-Matching for Cycling Travel Data in Urban Area. *IET Intell. Transp. Syst.* **2024**, *18*, 2178–2203. https://doi.org/10.1049/ITR2.12567.
28. Arman, A.; Badii, C.; Bellini, P.; Bilotta, S.; Nesi, P.; Paolucci, M. Analyzing Demand with Respect to Offer of Mobility. *Appl. Sci.* **2022**, *12*, 8982. https://doi.org/10.3390/APP12188982.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.