XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGINEERING

A Dijkstra-Based Framework for Analyzing Urban Mobility from CDR-Derived OD Matrices: a case study of RMBH, Brazil

Henrique Costa Braga[1], Guilherme de Castro Leiva[1], Guilherme Mascarenhas Dias[2], João Vítor de Carvalho[3], Miguel Costa Ferreira[3], Thiago Magela Rodrigues Dias[3], Gray Farias Moita[4], Renato Guimarães Ribeiro[1]

[1]*Dept. of Transportation Engineering, Centro Federal de Educação Tecnológica de Minas Gerais – CEFET-MG*
Av. Amazonas, 5253 – Nova Suiça – CEP 30421-169, Belo Horizonte-MG, Brazil
henriquebraga@cefetmg.br, guilherme.leiva@cefetmg.br, renato.ribeiro@cefetmg.br
[2]*Instituto Federal de Minas Gerais -IFMG*
Rua Padre Alberico, 440, Bairro São Luiz; 35570-000 – Formiga – MG, Brazil
guilhermedias2501@gmail.com
[3]*Dept. of Computing, Centro Federal de Educação Tecnológica de Minas Gerais – CEFET-MG*
Rua Álvares de Azevedo, 400 – Bela Vista - CEP 35503-822, Divinópolis-MG, Brazil
vicjoao2345@gmail.com, miguelcfe.50@gmail.com, thiagomagela@cefetmg.br
[4]*Dept. of Mathematical and Computational Modeling and Dept. of Civil Engineering, Centro Federal de Educação Tecnológica de Minas Gerais – CEFET-MG*
Av. Amazonas, 7675 – Nova Gameleira – CEP 30510-000, Belo Horizonte-MG, Brazil
gray@cefetmg.br

Abstract. This study presents a computational framework for analyzing urban mobility patterns through interzonal crossing maps derived from origin–destination (OD) matrices inferred from mobile phone data. Using the Dijkstra algorithm applied to an interzonal graph representing the Metropolitan Region of Belo Horizonte (RMBH), Brazil, the method estimates the spatial distribution of travel flows by identifying the most frequently ideal traversed zones in interzonal trips. In addition to standard shortest-path allocation, the model incorporates a second-neighborhood approach that distributes flows across nearby alternative routes to better reflect real-world mobility dispersion. The results, based on OD data from November 2019, highlight critical zones subject to potential overuse, particularly in and around the city of Belo Horizonte. Heatmaps and binary crossing maps reveal structural bottlenecks and flow concentration patterns, offering valuable input for infrastructure planning, prioritization of interventions, and the development of more resilient and equitable mobility strategies. The proposed framework demonstrates strong potential for integration into broader decision-making processes in transportation planning.

Keywords: Dijkstra algorithm, Call Detail Record, Origin-Destination matrix, Interzonal crossing maps, Urban mobility.

1 Introduction

Understanding urban mobility patterns is essential for transport planning, public policy formulation, and the organization of urban space. In this context, so-called *big data* — such as those generated by mobile devices, sensors, and digital platforms — constitute a vast and valuable source of information on the daily movement of the population. However, processing and analyzing these data pose significant challenges, requiring advanced techniques in data handling and modeling [1, 2]. Overcoming these difficulties is crucial to transforming this massive volume of data into actionable knowledge capable of supporting more precise, efficient, and adaptive decision-making in urban mobility. In complex metropolitan regions, such as the Metropolitan Region of Belo Horizonte (RMBH), such analysis becomes even more strategic given the challenges related to equity, sustainability, and territorial pressures.

The origin–destination (OD) matrix is the classical tool for representing travel flows between geographic

zones, traditionally constructed through sample-based household surveys — a robust methodology, though costly and infrequently updated. In recent years, passive mobile phone data, such as Call Detail Records (CDRs), have enabled the inference of OD matrices with high spatial and temporal granularity, significantly enhancing the analytical potential for studying dynamic mobility phenomena. Despite these advantages, OD matrices inferred from CDRs that focus solely on origin–destination pairs present limitations, as they fail to capture the intermediate zones traversed during trips — regions that are often strategic for infrastructure and transport service operations.

One widely adopted technique for visualizing intermediate zones is the use of desire lines [3, 4]. These contribute to understanding aggregated mobility flows between zones by offering a direct, intuitive, and visually expressive representation of trips. By connecting OD pairs with straight lines whose thickness is proportional to the flow intensity, desire lines facilitate the identification of key trip generators and attractors, revealing structural mobility axes and relevant territorial asymmetries. They are particularly useful in the early stages of diagnosis and planning, as they synthesize large datasets in a format accessible even to non-specialist audiences. Moreover, their construction is relatively simple, especially when based on aggregated OD matrices, making them an attractive tool for technical teams working in urban and transport planning.

Nonetheless, despite their visual appeal, desire lines present significant limitations for more detailed mobility analyses. In areas with dense connectivity or a high number of zones, these lines tend to overlap excessively, compromising the readability of the graph. As such, they should be considered a preliminary analytical tool, typically requiring complementary methods for deeper analysis. As an alternative and complementary visualization technique, this study proposes the use of **interzonal crossing maps**.

These maps record how frequently each territorial zone is traversed during simulated interzonal trips. To construct them, Dijkstra's algorithm [5, 6] is applied to an interzonal graph in which edge costs are defined by the geodetic distances between zone centroids, allowing for the estimation of shortest paths between OD pairs. In this model, the entire flow indicated by the OD matrix is assigned to the lowest-cost path between each origin–destination pair, and the sum of flows passing through each zone along the minimum path provides an estimate of the total traversal load on that unit. This metric provides strategic insight, revealing zones with potential overuse and indicating where road infrastructure and transport services may face greater demand.

Additionally, this study proposes an extension to the model by incorporating variability in route choice. Instead of assuming that all trips strictly follow the shortest path, it is assumed that a portion of travelers opt for adjacent or nearby alternative routes, resulting in weighted crossing maps. This refinement accounts for individual route decisions, operational detours, and inherent uncertainties in real-world mobility behavior. Although the optimal path remains a theoretical reference — given that modal, operational, or local infrastructure aspects are not explicitly modeled at this stage — the weighted crossing maps provide a more realistic representation of the spatial distribution of theoretical interzonal flows calibrated by the OD matrix.

To illustrate its applicability, the proposed methodology was applied to the OD matrix of the RMBH for the years 2019 and 2021, inferred from anonymized mobile phone data [3]. The resulting interzonal crossing maps reveal patterns of concentration, redundancy, and potential overload in travel flows, offering valuable technical inputs for infrastructure planning, the identification of priority corridors, and the development of regional mobility strategies more attuned to the territorial complexity of the metropolitan area.

2 Methodology

The methodology adopted in this study aims to explore the interzonal structure of the RMBH. For this purpose, the OD (origin–destination) matrix for the RMBH, covering the period from 2019 to 2021, was used. This matrix was inferred from anonymized mobile phone data, following the methodology developed for the RMBH Development Agency [3], and comprises 393 geographic zones with aggregated flows between zone pairs.

The zonal mesh was represented as a directed and connected graph, where each zone corresponds to a node, and the connections between adjacent zones are modeled as weighted edges. The edge weights were defined based on the shortest geodetic distance between zone centroids, although the proposed structure allows for the future incorporation of more sophisticated criteria, such as estimated travel time or generalized multimodal cost.

For each OD pair, Dijkstra's algorithm is applied to identify the lowest-cost path between zones. However, the methodological approach proposed in this work also incorporates a logic for distributing flows between main routes and nearby alternative routes, representing the concept of a "second neighborhood" — that is, paths with

only slightly higher cost than the minimum but spatially contiguous to the optimal route. This approach aims to more directly capture, albeit in a simplified way, the spatial dispersion observed in real-world mobility networks, reflecting both operational variations and individual preferences.

With this structure, different flow allocation scenarios can be simulated. For example, one may assume that 50% of the interzonal demand follows the shortest route, while the remaining 50% is distributed among neighboring paths, proportionally to their topological similarity or relative cost. This redistribution enhances the ability to represent the pressure exerted on zones that, although not origins or destinations, are strategically crossed during trips.

For each considered route — main or alternative — all intermediate zones traversed are recorded. From these records, the **interzonal crossing map** is constructed: a spatial representation that quantifies how frequently each zone is crossed in the simulated trips. These maps can be generated based on either the absolute number of travelers in each zone or the travel density per unit area. They may be visualized as heatmaps using either a decimal or logarithmic scale, depending on the value distribution. Alternatively, binary maps can be produced using a selected percentage threshold of the maximum recorded value.

To identify adjacent geographic zones, an algorithm was developed within the QGIS 3.42.3–Münster environment, capable of detecting neighborhood relationships between polygons based on their spatial boundaries. Initially, direct contiguity is tested using the *touches()* function, which checks whether two geometries share edges. If this criterion is not met, the algorithm performs a second, tolerance-based test by applying a small buffer to the geometries to account for potential spatial inaccuracies and then assesses whether the modified geometries intersect. Comparisons are optimized to avoid self-comparisons and duplicates, using nested loops that evaluate only unique polygon pair combinations. At the end of the process, identified neighboring pairs are recorded bidirectionally and exported in tabular format with the respective zone identifiers.

The spatial modeling was conducted in QGIS using the Open Street Map-compliant shapefile of zones provided by ADRMBH. Computational routines were developed in Python 3.12.10, within the isolated QGIS Processing Toolbox environment, using the following libraries: *pandas* (for tabular data manipulation), *geopandas* (a geospatial extension of pandas), *folium* and *streamlit_folium* (for interactive map generation), and *streamlit* (for building an exploratory web interface).

This framework, illustrated in Figure 1, enables both the automated execution of the modeling process and the visual communication of results, supporting exploratory analyses and potential future applications in various urban mobility contexts.

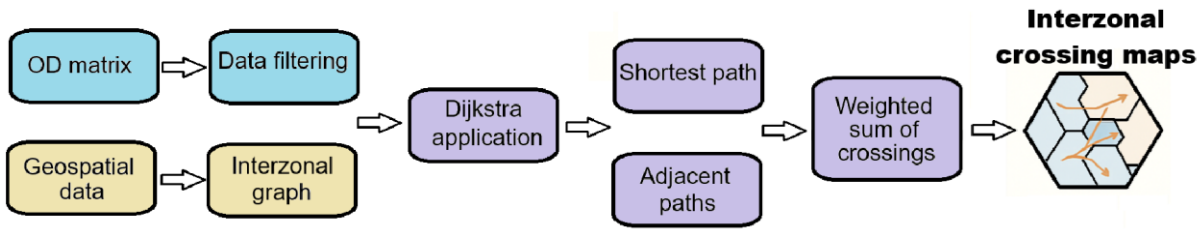


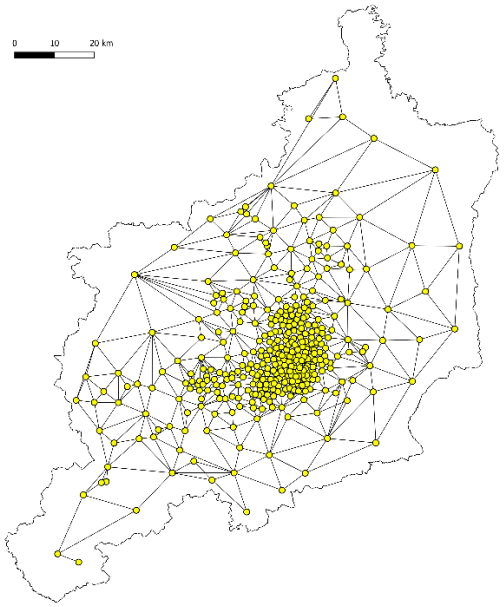
Figure 1. Flowchart of the methodological framework used to generate the weighted interzonal crossing maps.

3 Results and Discussion

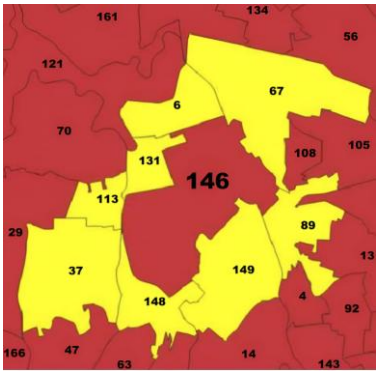
The RMBH, composed of 34 municipalities, was subdivided into 393 study zones (ZEs) during the development of the 2019–2021 OD matrix by ADRMBH. These zones were defined based on census data reflecting urban homogeneity and territorial articulation [3]. As a result, the distribution of ZEs is not uniform across municipalities. For instance, the city of Belo Horizonte is divided into 165 zones, whereas municipalities such as Rio Acima and Confins each consist of only one zone.

For the application of Dijkstra's algorithm, this network of study zones is modeled as a graph, in which each node represents a ZE and is positioned at its respective geographic centroid. The edges correspond to the shortest geodetic distances between the centroids (nodes) of adjacent zones, with their lengths calculated using the Haversine formula [7, 8]. Although simplified, this representation provides sufficient robustness to simulate interzonal routes compatible with the spatial resolution of the available data. Figure 2 illustrates several of these

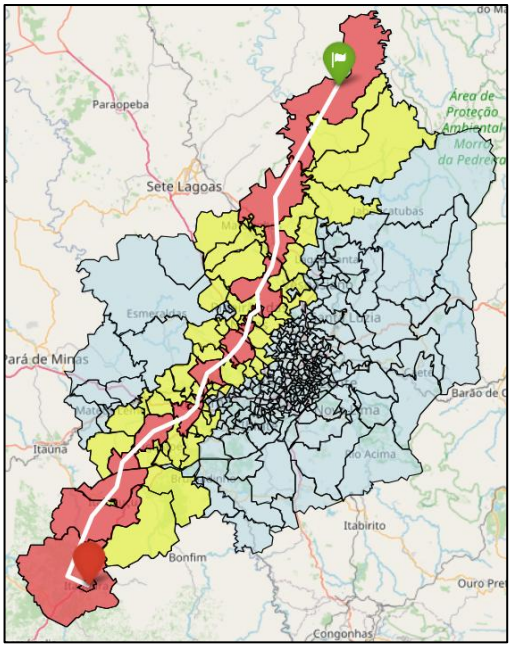
steps.



(a)



(b)



(c)

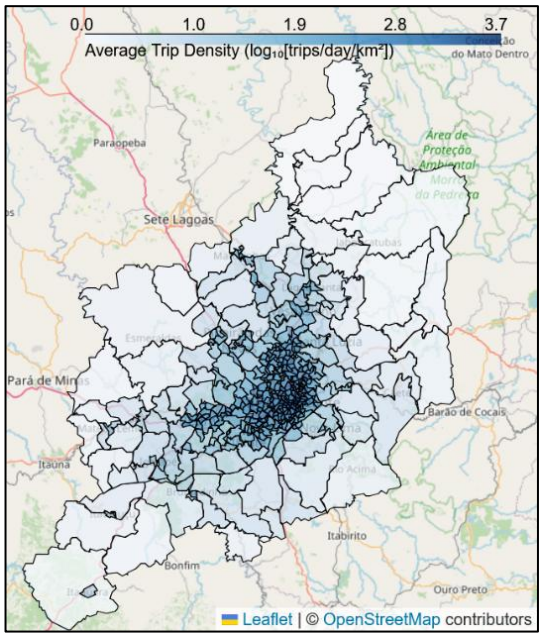
Origin ZE: 1
Destination ZE: 276

🔢 Number of ZEs in the shortest path: 21
🚩 ZEs in the shortest path (in red): 1 (origin) → 312 → 335 → 332 → 329 → 347 → 345 → 355 → 351 → 222 → 249 → 178 → 197 → 194 → 169 → 271 → 304 → 280 → 279 → 275 → 276 (destination)
Number of ZE in the adjacent path (in yellow): 67

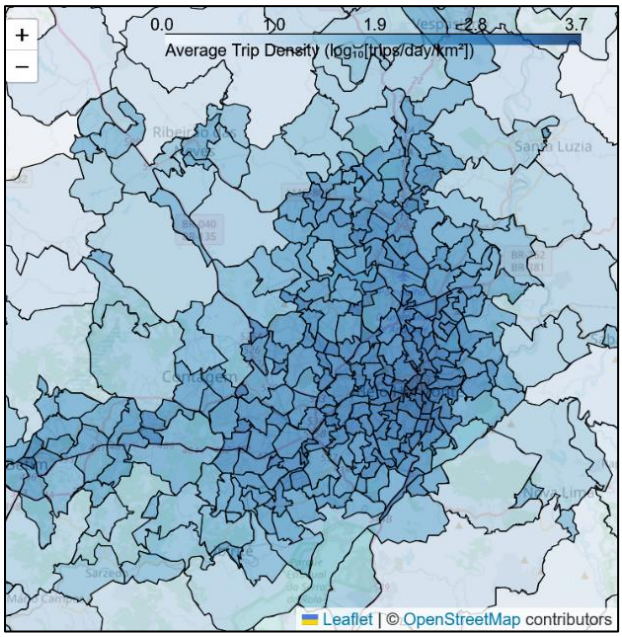
Figure 2. (a) The RMBH graph composed of 393 study zones (ZEs); (b) a zoomed-in view of a central area of the RMBH with selected ZEs identified; and (c) the optimal route obtained via Dijkstra's algorithm between ZEs 1 and 276, with traversed ZEs highlighted in red and their respective adjacent zones in yellow.

Figure 2a presents the graph generated from the 393 ZEs, where each node corresponds to the centroid of a zone. Figure 2b illustrates the application of the adjacency detection algorithm, using ZE 146 at the center (in red) and its adjacent zones (in yellow) in the surrounding area. Figure 2c shows the shortest path computed between ZE 1 (the municipality of Baldim) and ZE 276 (the municipality of Itaguara), with traversed zones marked in red and their adjacent zones in yellow. This subfigure also includes the numeric identification of both the zones along the optimal path and their adjacent ZEs.

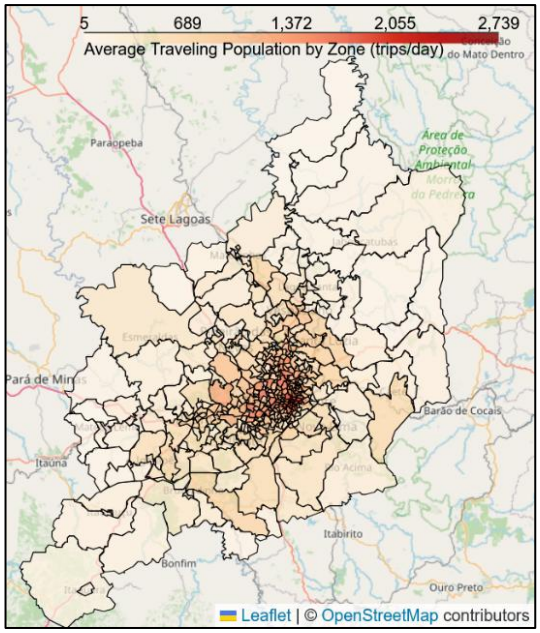
As a continuation and for illustrative purposes, Figure 3 presents an example application of the interzonal crossing maps in the form of a heatmap for the entire RMBH, considering all 393 study zones.



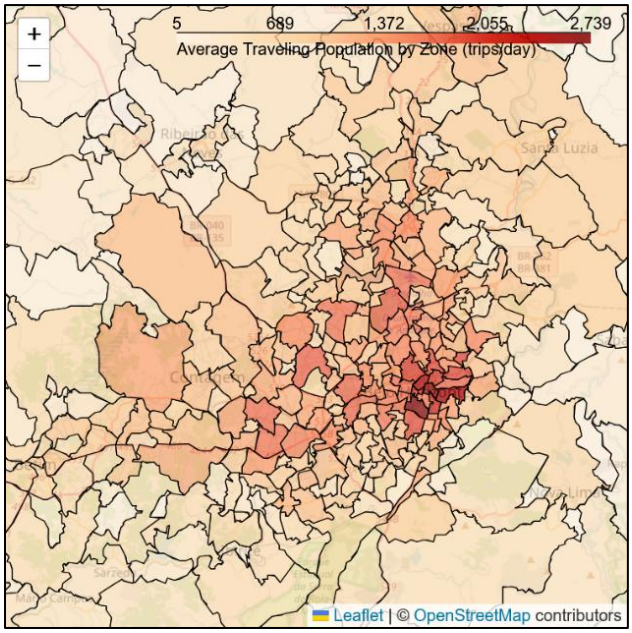
(a) Entire RMBH.



(b) Highlight of the central region.



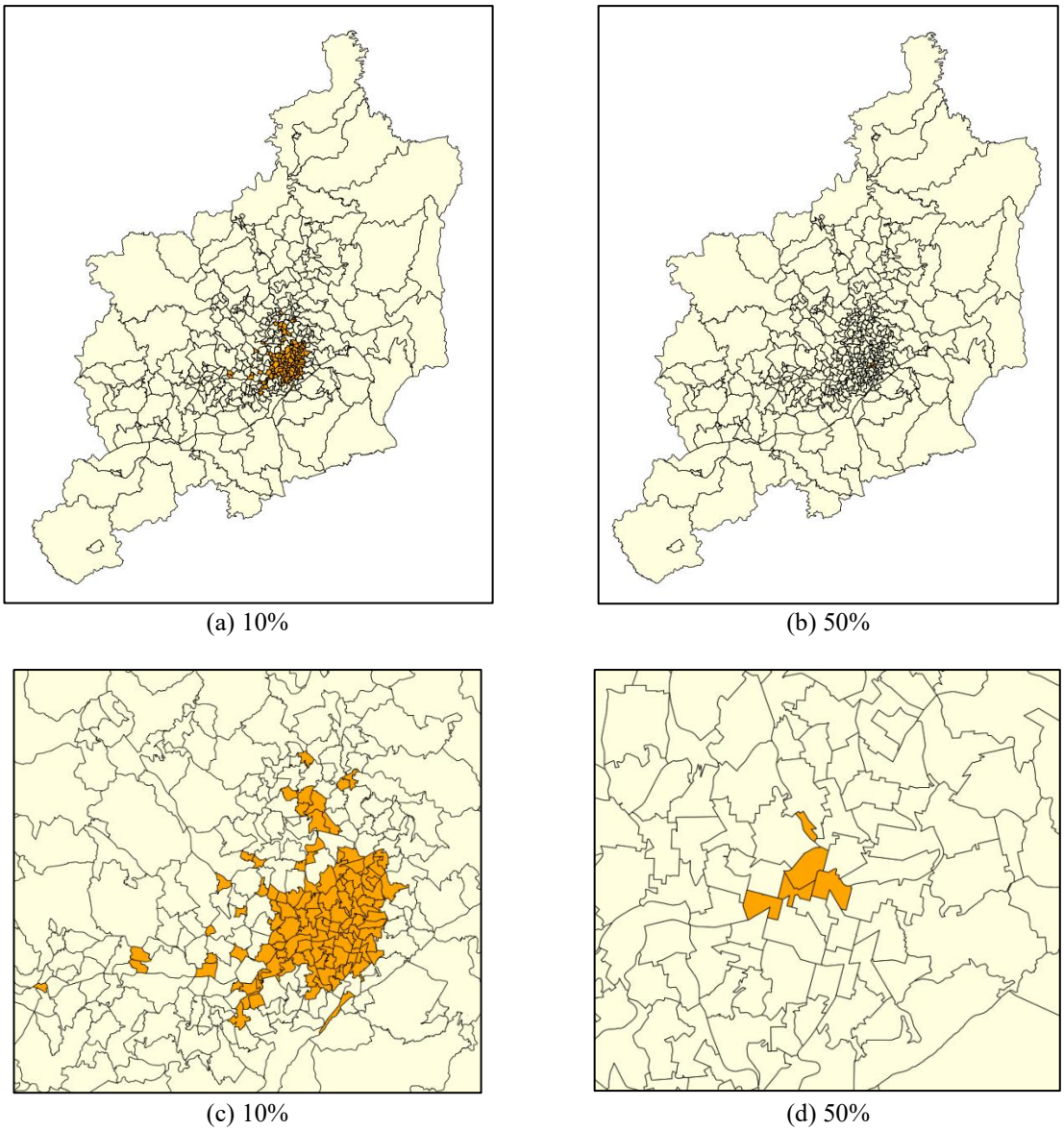
(c) Entire RMBH.



(d) Highlight of the central region.

Figure 3. Crossing maps represented as heatmaps, generated for the time window between 12:00 and 13:00 on weekdays during November 2019, based on the RMBH OD matrix (average over 13 days). Top: heatmaps in logarithmic scale of the traveling population; bottom: heatmaps in decimal scale of trip density per area.

Additionally, Figure 4 presents examples of binary interzonal crossing maps, generated for the time window between 12:00 and 13:00 on weekdays during November 2019, based on the RMBH OD matrix (averaged over 13 days), using trip density per area. Finally, Figure 5 illustrates the evolution of the number of ZEs exceeding specific thresholds of flow density relative to the maximum value — which, in this case, is 5,535.6 trips/(hour·km²) — offering a quantitative perspective on the spatial concentration of interzonal traffic observed in the simulation from Figure 4. This maximum density value serves as a reference point for threshold-based filtering, enabling comparative assessments of spatial distribution patterns across the metropolitan area.



(a) 10% (b) 50%

(c) 10% (d) 50%

Figure 4. Crossing maps represented as binary maps, generated for the time window between 12:00 and 13:00 on weekdays during November 2019, based on the RMBH OD matrix (average over 13 days), using trip density per area. Top: maps of the entire RMBH; bottom: zoomed-in view of the activated region. On the left, zones with less than 10% of the maximum value are excluded; on the right, zones below 50% of the maximum value are excluded.

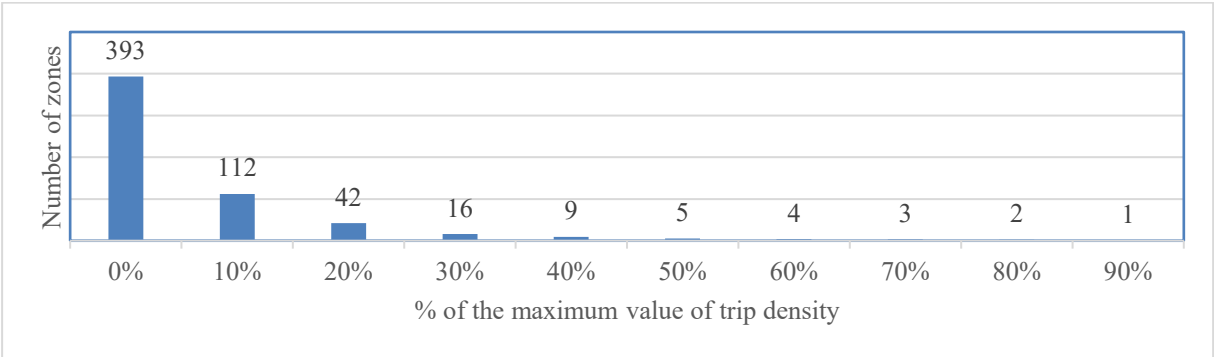


Figure 5. Evolution of the number of study zones (ZE) exceeding a given percentage of flow density relative to the maximum flow, based on the simulation presented in Figure 4.

Preliminary results indicate, both qualitatively and quantitatively, that part of the city of Belo Horizonte and part of its immediate surroundings constitute critical areas in terms of trip density, thus requiring a differentiated approach compared to other parts of the RMBH. In this context, interzonal crossing maps emerge as a promising tool for advancing urban mobility studies. By enabling the identification of zones with high recurrence of traversal, interzonal bottlenecks, flow concentration patterns, and structural redundancies, these maps offer a complementary analytical perspective to traditional origin–destination matrices.

Moreover, their application significantly contributes to diagnosing overload in transit areas, regardless of whether they serve as origins or destinations, allowing for a more refined understanding of the spatial logic of metropolitan flows. As such, interzonal crossing maps provide valuable technical input for integrated mobility planning, supporting the definition of investment priorities in infrastructure, the restructuring of public transport networks, and the improvement of territorial management strategies. These capabilities underscore their potential utility in future studies, particularly those aimed at developing more equitable, efficient, and sustainable urban mobility systems.

4 Conclusions

The results obtained through the proposed framework demonstrate the feasibility and analytical value of interzonal crossing maps in enhancing the understanding of urban mobility flows based on CDR-derived OD matrices. By capturing not only origin and destination pairs but also the frequency with which zones are traversed, the approach reveals hidden spatial patterns, including redundant routes, interzonal bottlenecks, and zones prone to overload. The integration of flow redistribution through second-neighborhood routing further increases the realism of the model, acknowledging variability in individual travel decisions and network dynamics.

Applied to the case of the RMBH, the method successfully identifies areas—especially Belo Horizonte and its surroundings—that require differentiated planning efforts due to their high interzonal traffic density. These findings reinforce the potential of the proposed methodology to support more informed and equitable transportation planning, enabling the design of infrastructure interventions and policies that are sensitive to the complexity and heterogeneity of metropolitan mobility.

Acknowledgements. The authors thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG for the financial support provided under Call No. 006/2024 – Process APQ-04889-24.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] D. Tzika-Kostopoulou; E. Nathanail; K. Konstantinos. “Big data in transportation: a systematic literature analysis and topic classification*”. Knowledge and Information Systems*, vol. 66, n. 8, pp. 5021-5046, 2024.
[2] V. M. Urbano; M. Arena; G. Azzone. “Big data for decision-making in public transport management: a comparision of diferente data sources”. *Research in Transportation Business & Management*, vol. 59, A 101298, pp. 1-13, 2025.
[3] ADRMBH - Agência de Desenvolvimento da Região Metropolitana de Belo Horizonte, “Relatório de Metodologia e Resultados OD RMBH: desenvolvimento de uma matriz origem/destino com dados de telefonia móvel para a Região Metropolitana de Belo Horizonte”. Belo Horizonte: ADRMBH. 2022. Disponível em: https://www.agenciarmbh.mg.gov.br/pesquisa-od/. Acesso em: 26/junho/2025.
[4] S.-I. Kinoshita, Y. Bando and H. Sayama, “Spatio-temporal differences in bike sharing usage: a tale of six cities”, 2024. Pre-print. DOI: 10.48550/arXiv.2412.19294
[5] E. W. Dijkstra, “A note on two problems in connexion with graphs”. *Numerische Mathematik*, vol. 1, p. 269–271, 1959. DOI: 10.1007/BF01386390.
[6] P.M. Casas, L. Kraus, A. Sedeño-Noda and R. Borndörfer, “Targeted Multiobjective Dijkstra Algorithm”. *Networks*, vol. 82, pp. 277-298, 2023.
[7] R. W. Sinnott, “Virtues of the Haversine”. *Sky and Telescope*, vol. 68, n. 2, p. 159, 1984.
[8] E. Maria, E. Budiman; Haviluddin; M. Taruk. “Measure distance locating nearest public facilities using Haversine and Euclidean Methods”. *Journal of Physics: Conference Series*, vol. 1450, A 012080, 2020. DOI: 10.1088/1742-6596/1450/1/012080.