What spatial resolution and regional scope are required to quantify accessibility? A multi-scale approach and evaluation

Marco Miotti[1]

[1]Department of Civil, Environmental, and Geomatic Engineering, ETH Zurich

Abstract

Accurate transport and land use modeling requires balancing high spatial resolution with broad geographical coverage. While regional analysis demands extensive scope, capturing active mobility requires micro-scale granularity, often leading to computationally prohibitive Origin-Destination matrices. This study addresses this constraint by introducing a multi-scale accessibility framework utilizing asymmetric matrices. By nesting a high-resolution hexagonal grid within standard Traffic Analysis Zones (TAZs), the trade-off between computational cost and model accuracy is optimized. A sensitivity analysis, calibrated on data from the Zurich metropolitan area in Switzerland, reveals that a cell resolution of 100 m is critical for quantifying micro-accessibility. A high-resolution buffer of 1 km is identified as the optimal 'walkability horizon' to capture local behavior. Furthermore, results demonstrate that regional cutoffs must be function-dependent: a 40 km radius is necessary to represent commuting labor markets, whereas $5-10$ km suffices for discretionary service trips. This hybrid approach enables the integration of precise pedestrian metrics into regional Land Use Transport Interaction (LUTI) models without sacrificing operational efficiency.

Background

Many fundamental applications in transport and land use modeling, such as accessibility analysis, require the estimation of travel distances or other costs between a vast number of spatial points [1]. As the spatial resolution of input data increases, the computational demand grows exponentially. For a study area comprising 10^6 (one million) individual locations, a full origin-destination (OD) matrix requires the evaluation of 10^{12} pairs. This combinatorial explosion poses significant challenges when quantifying high-resolution accessibility metrics, such as cumulative opportunities (e.g., the number of destinations reachable within 500 meters or 15 minutes).

To mitigate these computational scaling issues at the micro scale, a standard approach is to aggregate individual locations into larger spatial units, such as raster grids or Traffic Analysis Zones (TAZs) (e.g., [1, 2]). By mapping millions of points to a manageable number of zones, the dimensionality of the

OD matrix is drastically reduced, making the calculation of regional travel patterns computationally feasible. In Switzerland, the National Transport Model (NPVM) utilizes a system of approximately 8,000 TAZs. In the Swiss context, these TAZs are relatively small, with an average residential population of roughly 1,000 inhabitants per zone [3].

However, such zonal aggregation often fails to provide sufficient spatial resolution to accurately capture micro-scale accessibility, particularly for active modes of transport like walking [4, 5]. These scaling issues, defined by the conflict between high spatial resolution and broad geographical coverage, are a recurring constraint in transportation modeling [1, 6] and are part of the Modifiable Areal Unit Problem (MAUP) [7, 1, 6]. If zones are approximately $1\ \mathrm{km}^2$ in dense urban areas—and often exceeding $10\ \mathrm{km}^2$ in suburban or rural contexts, the centroid-to-centroid approximation masks critical local variability. For instance, determining whether a specific destination (e.g., a supermarket or transit stop) is available within a 300-meter or 5-minute walk is mathematically impossible if the spatial unit of analysis itself exceeds the walkability threshold.

At the macro scale, questions remain regarding the spatial extent required to capture meaningful accessibility without incurring unnecessary computational debt. While it is intuitive that a location 500 km away is irrelevant for daily commuting or shopping trips, determining the precise cutoff threshold is complex. The challenge lies in defining a boundary that is sufficiently large to capture typical daily mobility behavior as accurately as possible, yet sufficiently compact to maintain matrix sparsity.

To address these conflicting scaling requirements–identifying ideal spatial resolution at the local level while ensuring sufficient coverage at the regional level–a multi-scale approach utilizing two distinct geographical layers is introduced. Various parameter configurations are tested, identifying the Pareto frontier that minimizes computational cost while maximizing the accuracy of accessibility indicators relative to a disaggregated baseline.

This framework serves as a foundation for designing computationally efficient and precise accessibility evaluations, transport analyses, and Land Use Transport Interaction (LUTI) models. Crucially, the proposed method enables the accurate representation of all land-based modes within a single integrated model, capturing the granularity required for local walking trips alongside the regional scope necessary for daily car or public transit commutes.

Methods

To reconcile the computational efficiency of zonal models with the precision required for active transport analysis, a hybrid, multi-scale approach is employed. This method differentiates spatial resolution based on proximity. For the high-resolution layer, a grid of hexagonal cells with diameter d is superimposed over the study area. To maintain matrix sparsity and computational efficiency, cells containing no relevant attributes (i.e., zero residential population, employment, or points of interest) are pruned

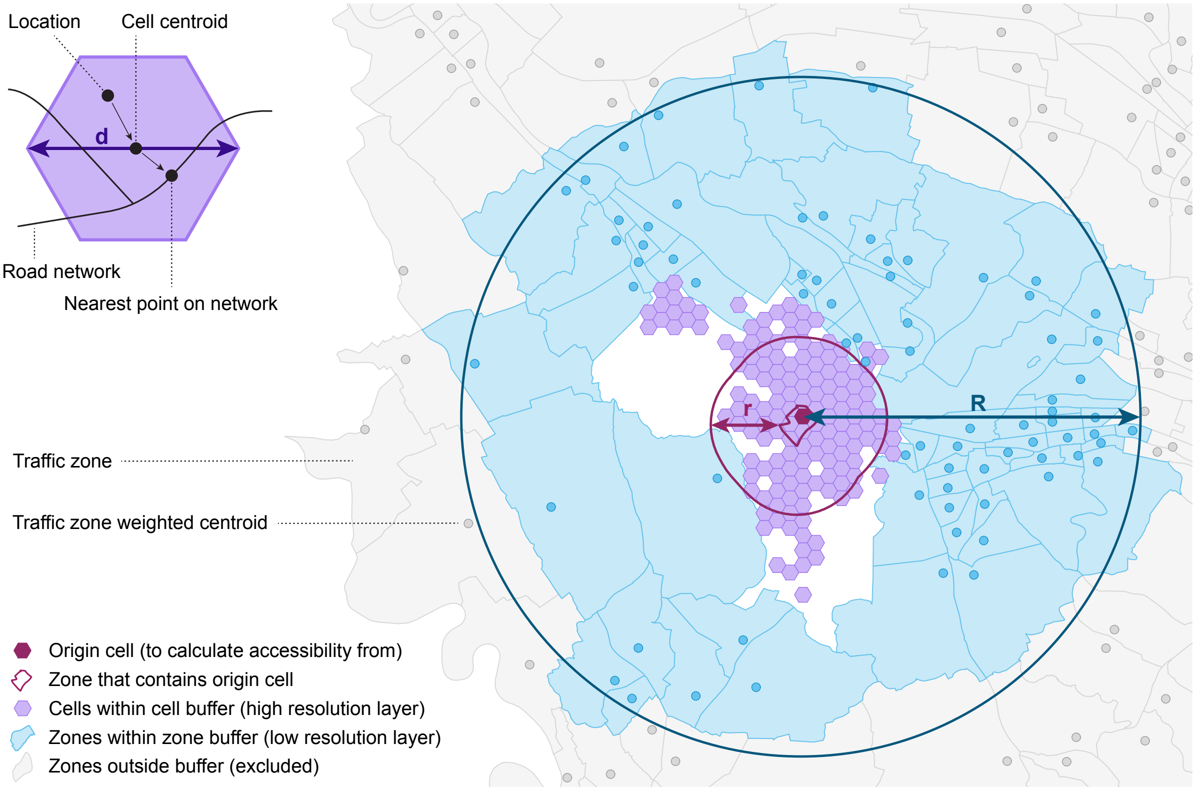

Figure 1: Schematic of the multi-scale accessibility analysis framework, highlighting cell diameter d, spatial buffer r, and cut-off radius R. Note that traffic zone relationships are symmetrical: if the centroid of zone b does not fall within the buffer r of zone a, but the centroid of zone a falls within the buffer r of zone b, cells within b are dissolved with respect to cells within zone a *and* vice-versa. This is so that return trip distances or travel times can be calculated accurately.

from the grid and excluded from evaluation. A spatial buffer r is defined around the TAZ containing the trip origin. Any destination located within a TAZ whose centroid falls within this buffer is resolved at the fine-grained hexagonal cell level. Beyond this buffered area, destinations are aggregated to the standard TAZ level, extending up to a maximum regional analysis radius R (see Figure 1). Specifically, three different cell diameters d (100 m, 250 m, and 500 m) and three different buffers r (0, 1000 m, and 2000 m) are evaluated (Table 1). A buffer of 0 meters means that only the TAZ that the location is in is dissolved into cells; no neighboring zones are considered at the cell level. Finally, a baseline version that does not make use of cells, only TAZs, is included.

This dual-layer spatial representation informs the construction of two asymmetrical cost matrices: A cell-to-cell matrix (M_{local}) that contains precise travel costs for all pairs where the destination lies within the high-resolution buffer area, and a zone-to-zone matrix ($M_{regional}$) that contains aggregated travel costs for pairs extending beyond the local buffer up to the regional limit R. For any given origin-destination query, the algorithm performs a hierarchical lookup. It first queries M_{local} to capture micro-accessibility; if the pair falls outside the high-resolution scope, it queries $M_{regional}$. Pairs exceeding the

Table 1: Number of cells as a function of cell diameter d and number of cell-to-cell origin-destination (OD) pairs as a function of cell diameter d and high-resolution radius r.

		OD pair count		
d	Cell count	r=0	r=1000	r=2000
100	138,004	8,642,152	59,692,588	132,830,040
250	44,903	1,393,133	6,806,269	14,920,663
500	19,955	359,927	1,397,161	3,044,593

regional radius R are discarded. This hierarchical structure allows for the precise evaluation of local walkability while maintaining the capacity to model regional interaction potential.

Using these matrices, accessibility to three different destination types is quantified: grocery stores, gastronomy-related leisure amenities, and work places. These different types of destinations are designed to capture different types of travel behavior: grocery and gastronomy trips tend to be rather short; though the former can be more restricted in terms of modal choice because groceries need to be carried home. Work trips tend to be longer. Grocery and gastronomy locations are gathered from OpenStreetMap [8] using OSMNx [9]. Employment opportunities are based on the number of employees in each location and are obtained from the Swiss Federal Statistical Office [10], which available at a hectare (100 x 100 m) grid level.

To define the employment opportunity count for each cell, the 100 m grid employment data is mapped to our custom hexagonal grid, which employees distributed proportionally according to area overlap. To convert employees to employment opportunities, the number of employees in that cell is divided by 50, rounded up to the nearest integer. An employer with 290 employees according to the employment data therefore counts as 6 employment destinations; an employer with 10 employees counts as one destination. This approach allows us to reduce the scale of employment destinations to a level comparable to other destination types while considering the fact that large employers will attract more trips than smaller ones.

Using the origin-destination distance matrices M_{local} and $M_{regional}$ as well as the destination counts in each cell and zone, accessibility to each destination type from each cell is calculated. To do so, the 'cumulative opportunities' approach is used, where the accessibility for location i and a given destination type is equal to the count of destinations that fall within a given distance bracket $q : [q_l, q_h)$:

$$A_{i,q} = \sum_{j:\ q_l \leq d_{i \to j} < q_h} 1 \tag{1}$$

These accessibility indicators are then used to evaluate which spatial resolution of the multi-scale grid, composed of b and r, is sufficient. To do so, accessibility to each of the three destination types (gro-

Table 2: Cumulative count bins used to count the number of destinations of each type accessible within a given distance from each origin location.

	Network type	Lower boundary	Upper boundary
Bin 1	Walk	0	250 m
Bin 2	Walk	250 m	500 m
Bin 3	Walk	500 m	750 m
Bin 4	Walk	750 m	1 km
Bin 5	Car	1 km	4 km
Bin 6	Car	4 km	10 km
Bin 7	Car	10 km	40 km

cery stores, gastronomy locations, and work places) is correlated with two indicators of travel behavior. These indicators include travel distance (in meters, continuous) and whether the trip was made with an active mode (walking, cycling) or not. Travel behavior data is obtained from the Swiss Mobility and Transport Microcensus (MTMC, 2015 and 2021 survey years) [11]. Specifically, trips made by walking or by car for any of the three modeled purposes (grocery shopping, gastronomy-related leisure trips, and commutes to the workplace) are used.

To correlate the first indicator, travel distance, with accessibility, a linear regression using log-transformed accessibility features and dependent variable is used:

$$\ln\left(d_{k:i\to j}\right) = c + \sum_{q} \ln\left(A_{i,q,\text{type}(k)}\right) + \epsilon_k \tag{2}$$

where $d_{k:i\to j}$ is the travel distance for trip k from location i to location j; $A_{i,q,\text{type}(k)}$ is accessibility metric q, measured at location i, for the destination type (trip purpose) of k; and ϵ_k is the error term.

To correlate the second indicator, the choice to walk, with accessibility, a logistic regression is employed, with accessibility for each distance bracket again being log-transformed:

$$p(\text{walk}_{k:i\to j}) = \frac{1}{1 + e^{c + \sum_q \ln\left(A_{i,q,\text{type}(k)}\right)}} \tag{3}$$

Model accuracy is then evaluated using R^2, indicating the degree to which the predictors explain the variance in the dependent variable:

$$R^2 = 1 - \frac{\sum\left(\hat{y_k} - y_k\right)}{\sum\left(y_k - \overline{y_k}\right)} \tag{4}$$

where y either corresponds to $ln(d)$ (distance model) or $p(\text{walk})$ (walking choice model).

The multi-scale accessibility framework presented here is tested for the metropolitan area of Zurich, Switzerland with an additional buffer of 50 km. Results only evaluated within the metropolitan area

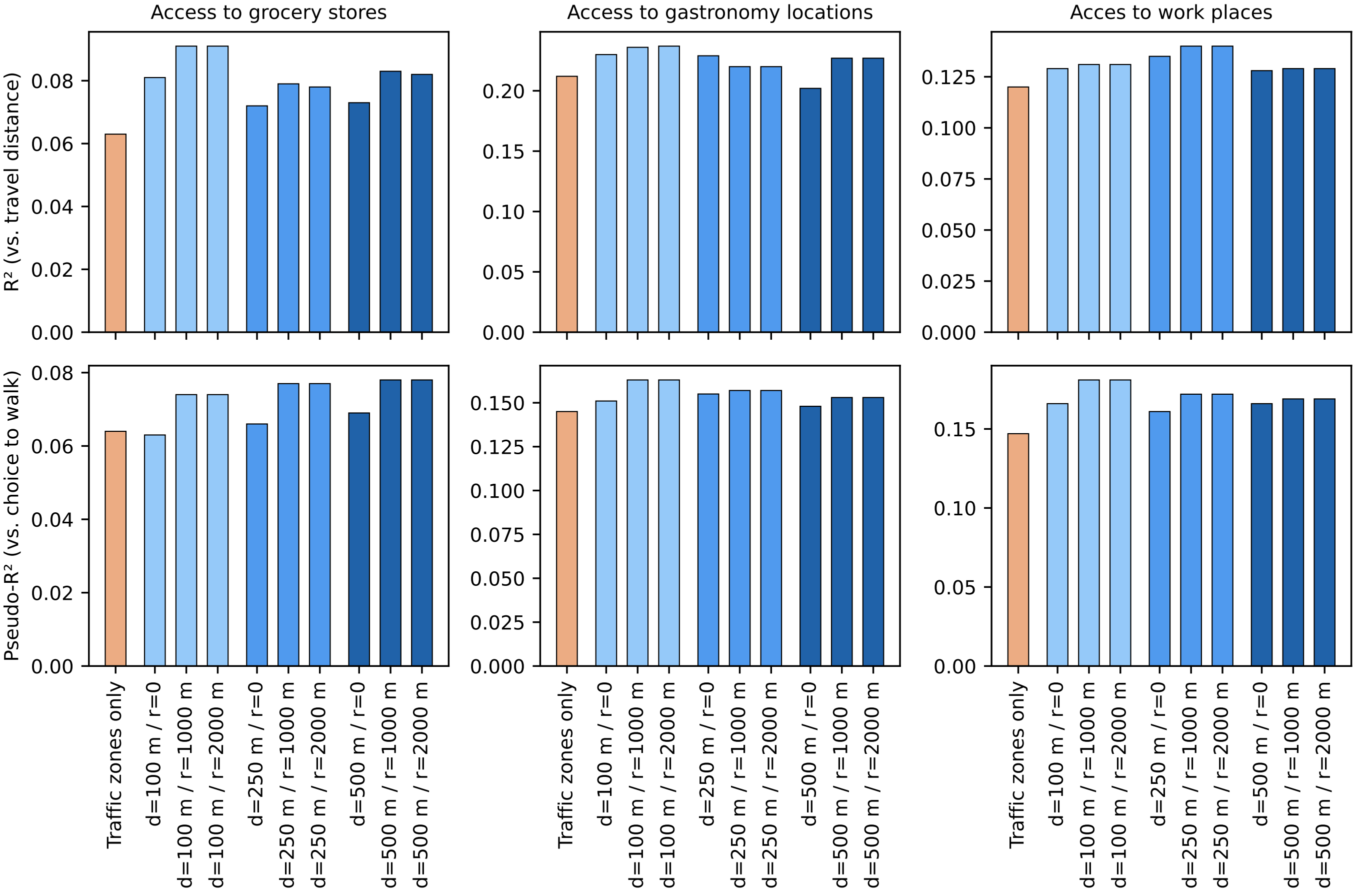

Figure 2: Evaluation of spatial resolution (cell diameter d) and radius of high-resolution layer (r) in terms of how well accessibility metrics calculated based on those layers explain the variance in trip distance (top row) and the choice to walk (bottom row). In both cases, accessibility is quantified at the trip origin of each trip.

and buffer of 10 km around to avoid border effects.

Findings

The introduction of a high-resolution hexagonal layer yields a consistent accuracy gain over the traffic-zone baseline (Figure 2). Crucially, the magnitude of this improvement is inversely related to trip length: the model excels at capturing the nuances of short-range grocery trips, while the gains for longer commute trips are present but less pronounced.

Reducing the cell diameter d to 100 m sharpens the accuracy for grocery and gastronomy trips, confirming that micro-scale accessibility requires micro-scale resolution. However, workplace trips do not benefit from this finer granularity. Likely, they are hitting a ‘data floor’ imposed by the 100 m grid resolution of the input employment data.

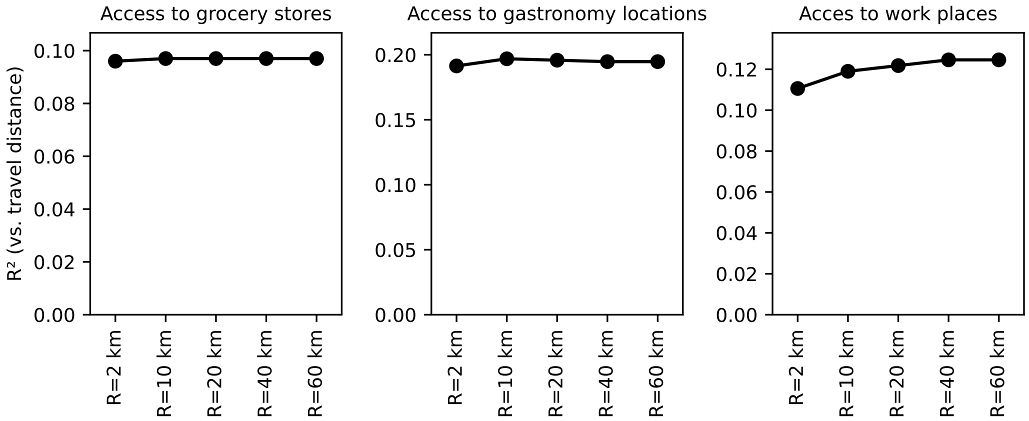

Figure 3: Evaluation of spatial coverage R (maximum euclidean distance between origin and destination considered when calculating accessibilities) for the model in Figure 2 with d = 1000 m and r = 1000 m.

Furthermore, a distinct range of influence for the high-resolution buffer zone is identified. Expanding the buffer radius r to 1 km significantly improves model performance by smoothing the transition between cells and zones. However, extending this boundary to 2 km offers diminishing returns. This plateau aligns with the physiological limits of active transport: since most walking trips occur within a Euclidean distance of 1 km, a larger high-resolution buffer adds computational cost without capturing relevant behaviors.

Finally, our analysis of the total system extent R reveals that the optimal cutoff is strictly function-dependent. Commuting models require a regional scope, peaking in accuracy at a 40 km radius. In contrast, non-work activities (grocery, gastronomy)—which are generally secondary to home and work locations—are adequately captured within a compact 5 km radius, allowing for aggressive matrix pruning without data loss.

Discussion

This study demonstrates that the trade-off between computational tractability and spatial precision is not a zero-sum game. It presents evidence that accurately quantifying micro-accessibility—particularly for active modes—requires a spatial resolution as fine as 100 m. In traditional zonal models, this granularity would induce a combinatorial explosion; however, this work shows that this can be circumvented through a hybrid-resolution approach. By treating space as a continuum of relevance, map destinations near the origin are effectively mapped at high resolution while aggregating distant locations into Traffic Zones (TAZs).

Our sensitivity analysis provides concrete heuristics for calibrating this multi-scale framework. A 'walkability horizon' of approximately 1 km (r) is identified as the critical boundary for the high-resolution layer. Within this radius, aggregating points into zones masks the local availability of amenities; beyond it, the marginal gain of high resolution diminishes rapidly as mode choice shifts toward motorized transport. Simultaneously, this work provides evidence that the macro-scale cutoff (R) must be function-dependent. While a 5 km radius suffices for capturing discretionary service trips (e.g., grocery), a regional scope of at least 40 km is strictly necessary to represent daily commuting labor markets. This dual-layer optimization places the model on the Pareto frontier, minimizing matrix sparsity without sacrificing the behavioral realism of the accessibility indicators.

While this analysis utilized Euclidean distance as the primary measure of separation, the implications extend to more complex impedance functions. In operational transport models relying on network travel times or generalized costs, benefits of this hybrid approach could be even more pronounced. Network discontinuities and 'last-mile' barriers—often smoothed over by zonal centroids—are captured explicitly in a high-resolution grid. Therefore, these findings likely represent a conservative estimate of the accuracy gains; shifting to time-based metrics would likely exacerbate the errors found in the zonal baseline, further justifying the need for the granular buffering proposed here.

Ultimately, this work suggests that the standard spatial units used in regional modeling are insufficient for the questions facing modern planners. As policy focus shifts toward sustainable, active mobility and the '15-minute city,' our modeling tools must evolve to measure these granular dynamics. The proposed multi-scale framework allows for the integration of pedestrian-scale metrics into regional Land Use Transport Interaction (LUTI) models without prohibitive computational costs. By ensuring that active transport is measured with the same rigor as regional commuting, this method provides a foundation for more equitable transport analysis, ensuring that the walkability of a neighborhood is defined by its street network, not by the arbitrary boundaries of a traffic zone.

Acknowledgments

This work was supported by InnoSuisse as part of 'The Blue City Project' [grant number PFFS-21-03] and by a Mobility Initiative grant funded through the ETH Zurich Foundation [grant number 2024-HS-252].

References

[1] M.-P. Kwan and J. Weber, "Scale and accessibility: Implications for the analysis of land use–travel interaction," *Applied Geography*, vol. 28, pp. 110–123, Apr. 2008.

[2] L. M. Martínez, J. M. Viegas, and E. A. Silva, "A traffic analysis zone definition: A new methodology and algorithm," *Transportation*, vol. 36, pp. 581–599, Sept. 2009.

[3] Federal Office for Spatial Development (ARE), Switzerland, "NPVM 2016: Zonenstruktur und Verkehrsnetze," tech. rep., June 2017.

[4] J. Hewko, K. E. Smoyer-Tomic, and M. J. Hodgson, "Measuring Neighbourhood Spatial Accessibility to Urban Amenities: Does Aggregation Error Matter?," *Environment and Planning A: Economy and Space*, vol. 34, pp. 1185–1206, July 2002.

[5] S. Liu and X. Zhu, "Accessibility Analyst: An Integrated GIS Tool for Accessibility Analysis in Urban Transportation Planning," *Environment and Planning B: Planning and Design*, vol. 31, pp. 105–124, Feb. 2004.

[6] N. Kuehnel, D. Ziemke, R. Moeckel, and K. Nagel, "The end of travel time matrices: Individual travel times in integrated land use/transport models," *Journal of Transport Geography*, vol. 88, p. 102862, Oct. 2020.

[7] D. W. S. Wong, "The Modifiable Areal Unit Problem (MAUP)," in *WorldMinds: Geographical Perspectives on 100 Problems: Commemorating the 100th Anniversary of the Association of American Geographers 1904–2004* (D. G. Janelle, B. Warf, and K. Hansen, eds.), pp. 571–575, Dordrecht: Springer Netherlands, 2004.

[8] OpenStreetMap Contributors, "OpenStreetMap." https://www.openstreetmap.org/, 2024.

[9] G. Boeing, "Modeling and Analyzing Urban Networks and Amenities with OSMnx," tech. rep., 2025.

[10] "Swiss Federal Statistical Office." https://www.bfs.admin.ch/content/bfs/en/home.html, 2025.

[11] Federal Office for Spatial Development (ARE), Switzerland, "Mobility and Transport Microcensus 2021." https://www.are.admin.ch/are/en/home/verkehr-und-infrastruktur/grundlagen-und-daten/verkehrsverhalten.html.