Exploiting Vehicular Data for Exposure-Aware Pedestrian Routing

Gurban Aliyev
University of Pisa, Italy
ISTI-CNR, Italy
g.aliyev@studenti.unipi.it

Mirco Nanni
KDD Lab, ISTI-CNR, Italy
mirco.nanni@isti.cnr.it

***Abstract*—Vehicular traffic is one of the major sources of air pollution in urban settings, making it essential to clearly understand how much and where vehicle emissions impact residents. Recent approaches manage to yield pollution maps at the microscopic level by processing GPS trajectories of vehicles. That is achieved by applying mathematical models to estimate instantaneous emissions from GPS data, extending estimates to areas without data through missing data imputation, and further considering air dispersion factors. In this work, we leverage such inferred knowledge to implement an emission-aware pedestrian routing strategy and to study its impact on the reduction of exposure to vehicular pollutants and walking time. The study is realized through simulations of large masses of pedestrians over a medium-sized city in Italy, analyzing the interplay between the two factors – exposure versus walking time – in terms of time efficiency of paths and changes over existing habits both at a global and at an individual level. Experiments suggest that exposure-aware routing can yield a significant margin of improvement in health over most paths with minor effects on mobility, making it feasible and effective.**

***Index Terms*—vehicular emissions, pedestrian routing, emission exposure.**

I. Introduction

Ambient air pollution poses a significant challenge to the sustainable development of rapidly expanding urban areas [1], in particular in cities. Indeed, major sources of anthropogenic emissions, such as energy production and transportation, are typically concentrated in urban regions. As a result, the high concentration of air pollutants leads to poor air quality in cities, with an increasing number of people being exposed to pollution each year. Given the negative impact on public health and the economy, the United Nations has called for efforts to reduce the adverse per capita environmental effects of urban areas [2]. Recent research [3]–[5] examines the spatial and temporal patterns of vehicular emissions using vehicle trajectories, providing means to assess the effectiveness of current green mobility policies [5], next-generation vehicle routing strategies [4], and the electrification of major polluters [3] in reducing emissions.

The general focus of this work is on studying the impact that vehicular emissions can have on citizens, especially in terms of emissions exposure of pedestrians, designing an exposure-aware pedestrian routing algorithm, and analyzing its effects through simulations. The fundamental observation is that although reducing overall emission levels is a key goal towards sustainability, the effects on human health of existing emissions strongly depend on where exactly they are generated and dispersed. Indeed, higher concentrations of emissions at the local level (for instance, nearby roads) are more likely to affect the safety of individual moving close to it, increasing their direct exposure to the pollutant and its effects on health. Thus, a first strategy to mitigate the effects of direct exposure can consist of deviating pedestrian routes to avoid more dangerous areas, clearly at the cost of following a path that is not necessarily the fastest one. Studying the effects of such routing deviations and the consequential trade-offs between exposure and time efficiency is central to understanding how much this strategy can be useful and realistically proposed to users. Indeed, on the one hand, if the decrease in exposure that can be achieved is very small the whole process is pointless; on the other hand, if significant decreases can be achieved only at the price of making routes much longer and wasting large amounts of time, the proposed path deviations become impractical and would never be adopted.

While the idea of separating pedestrian mobility from vehicle traffic is quite natural and various works considered it to some extent, existing studies are very limited and they do not explore in detail the trade-off options between exposure and time efficiency. Also, to the best of our knowledge no approaches have been proposed so far to automatically select an optimal trade-off at the level of the single (pedestrian) trip – more exactly, trade-offs are usually chosen manually at the global level, basically fixing once for all the thresholds and factors that define how much important is each of the two dimensions of the problem (travel time and exposure).

In this paper, we tackle the main open issues discussed above, providing the following contributions:

- we define an emission-aware path selection strategy based on a dual optimization problem involving both travel times and exposure as components of the cost function. The choice is trip-specific, meaning that different trips might choose a different importance weight for the exposure component of the cost;
- through simulations, we evaluate the trade-offs between time and exposure obtained in a medium-sized city;
- we study the spatial impact of exposure-aware routing to the road network structure, as well as the effects of using a trip-specific optimization strategy instead of a global

one.

In the rest of the paper, we discuss the background and relevant literature (Section II), introduce our exposure-aware pedestrian routing (Section III), and show experimental results assessing the impact of our routing strategy in terms of exposure-travel time trade-off and relations with the city structure (Section IV), and finally provide conclusive remarks (Section V).

II. Related Work

In this section, we discuss existing approaches related to the two main technical problems involved in the paper: estimating a map of emissions at the road level from mobility traces, and routing pedestrians (and similar mobility modalities) taking into consideration the amount of exposure to emissions.

A. Emissions Inference

Several works in the literature estimate vehicular emissions by exploiting the spatial precision and relative abundance of mobility data coming from vehicle fleets [5], typically applying microscopic emission models to GPS traces [1], [3] based on emission factors (EF) relating pollutants emitted vs kilometer traveled or liter of fuel consumed [6], [7]. While extremely useful and informative, such estimates might require some further processing to account for (at least) two factors: first, estimates might not cover some parts of the road network for a limitation of data; second, in reality emissions are affected by dispersion mechanisms that move them also far from the emission point.

Imputation Strategies Because of the limited amount of data collected, the problem of emission data sparsity on urban roads is common. However, according to recent surveys on traffic-related data imputation [8], there are currently no specific methods to impute missing emission values, while there are various attempts using state-of-the-art [9]–[12] approaches to estimate other road features. In [13] various imputation methods based on standard machine learning as well as road network embedding were experimented, showing that basic methods tend to perform better than more complex ones on this specific problem.

Dispersion Models. Studying individual and population exposure to air pollution is important to quantify the effect of changing travel behavior and traffic policies. Initial studies [14] focus only on the assessment of exposure in static locations (workplace, school, and home) using population census data. Later studies [15], [16] use more accurate mobile data to quantify average mobility exposure. More recent studies investigate the impact of travel behavior on traffic-induced emissions and subsequent exposure to traffic emissions [17]. However, these studies have some limitations. Firstly, these works disregard the dispersion of vehicular emissions both vertically and horizontally. There are various dispersion models to be considered, and the main approaches used in these models include Computational Fluid Dynamics (CFD), Gaussian, Lagrangian, and Box models [18], [19]. Depending on the dispersion levels, exposure to car emissions on roads and around roads can change. The concentration of on-road vehicular emissions in certain places can increase or decrease depending on different factors, such as weather conditions (temperature, precipitation, etc.) and building profile data.

In this paper, we adopt the simple solution provided in [13], which allows us to obtain a full map of dispersed emissions, which form the basis for further simulations. The approach consists of the following components:

- Raw emissions estimate: following [3], we make use of GPS traces of vehicles moving in a city to estimate their contribution to the emissions over the road segments traversed during their trips. The mathematical model for instantaneous emissions comes from [1], which computes them for each trajectory point p as $E_p^j = f_1^j + f_2^j v_p + f_3^j v_p^2 + f_4^j a_p + f_5^j a_p^2 + f_6^j v_p a_p$, where v_p and a_p are the instantaneous speed and acceleration of the vehicle. Emission factors f_i are determined by the type of pollutant and the engine type (petrol, diesel, etc.).
- Imputation: emissions for the road edges without data were estimated by adopting an XGBoost-based regression model in cascade with a classifier that separates high-emission roads from low-emission ones, and using a wide range of descriptive variables for nodes and edges.
- Dispersion of emissions: we adopted a simplified version of the standard Gaussian plume model proposed in [20] which computes the pollutant concentration in a given position (x, y) relative to the emission location and at height z, and wind direction along the x axis:

$$C(x,y,z) = \frac{Q}{2\pi u \sigma_y \sigma_z} \cdot e^{-\frac{y^2}{2\sigma_y^2}} \cdot \left[e^{-\frac{(z-H_e)^2}{2\sigma_z^2}} + e^{-\frac{(z+H_e)^2}{2\sigma_z^2}} \right]$$

  where Q is the emission rate of the point source and σ_y and σ_z are the dispersion coefficients (lateral and vertical). The formula is simplified by assuming a fixed wind (u) in all directions.

The result of this process is a map that associates each location of the area of interest – discretized into grid cells – to its corresponding emission concentration level. Experiments in this work will revolve around a use case in the city of Pisa, Italy, whose emission concentration map is depicted in Figure 2.

B. Emission-aware routing

Eco-routing has been explored in multiple domains, with prior studies focusing on vehicle routing for emissions mitigation and pedestrian routing for pollution exposure reduction.

Vehicle Routing. Several studies have proposed eco-routing algorithms to reduce the total emissions from private and commercial vehicles. Demir et al. [21] introduced the bi-objective pollution-routing problem (PRP), which optimizes vehicle routes based on fuel consumption and CO_2 emissions, balancing cost and environmental impact. However, this approach focuses on reducing total emissions rather than minimizing exposure. Luo et al. [22] extended this concept by developing pollutant exposure-aware vehicle routing, where

road segments with high human exposure were penalized in route selection.

Pedestrian Routing for Exposure Reduction. A more relevant line of research investigates pedestrian and cyclist routing to reduce personal exposure to pollutants. Davies and Whyatt [23] introduced a least-cost path approach to minimize pedestrian exposure by integrating GIS-based routing with pollutant dispersion modeling. Yoon et al. [24] further refined this by using a network centrality-based approach to identify low-exposure pedestrian routes in Seoul. Willberg et al. [25] applied multi-exposure optimization, integrating air pollution, noise, and greenery exposure to optimize cyclist routes.

All the solutions and studies mentioned above are rather limited in scope and do not explore in detail what are the effects of associating different importance weights to exposure vs. travel time. Moreover, such weight is assumed to be the same for all trips, and there is no clear way to automatically choose it. Our work tackles exactly these points.

III. Methodology

A. Road-level potential exposure

The amount of emissions that walking pedestrians absorb while moving can be estimated as the integral of the emissions the individual is exposed to in each of the points they visit along their movement trajectory: $exposure(T) = \int_{\bar{x}\in T} e(\bar{x}) d\bar{x}$. Making use of the emission map mentioned in Section II-A, which assigns emission concentrations to each cell of a regular grid, the computation can be approximated by translating the trajectory T to the sequence S of cells visited and, for each cell $s \in S$, its traversal time $time(s)$, obtaining $exposure(T) \sim \sum_{s\in S} e(s) \cdot time(s)$. We then consider a simplified scenario where the movement speed v_p of pedestrians and their breath intensity are constant, thus making $time(s) = length(T \cap s)/v_p$ (i.e. it is proportional to the length of the trajectory segment contained in the cell) and the overall absorption dependent only on the concentration.

This view can be expanded at the level of roads: any pedestrian walking a specific road segment will absorb an amount equal to the sum of contributions $e(s)$ for all subsegments s it contains. This allows us to precompute an *exposure cost* of each road segment.

B. Time vs. Exposure trade-off

Given a road segment r, we denote its precomputed exposure cost as $C_{exp}(r)$ and its associated travel time as $C_{time}(r)$. Then, given a parameter $w \in \mathcal{R}, w \in [0, 1]$, we compute the trade-off cost $C_{tradeoff}(r, w)$ as:

$$C_{tradeoff}(r, w) = w \cdot C_{time}(r) + (1 - w) \cdot C_{exp}(r)$$

To make the two measures easier to compare[1], we assume their values have been normalized with a standard min-max operator, mapping them into the interval $[0, 1]$, namely:

$$C_{exp}(r) = \frac{C_{exp}(r)^{orig} - \min_r C_{exp}(r)^{orig}}{\max_r C_{exp}(r)^{orig} - \min_r C_{exp}(r)^{orig}}$$

and similarly for $C_{time}(r)$. By definition, $C_{tradeoff}(r, 0) = C_{exp}(r)$ and $C_{tradeoff}(r, 1) = C_{time}(r)$, while values $w \in (0, 1)$ represent different blendings of the two components.

C. Global trade-off selection

Given a set of pedestrian OD pairs, we simulate their optimal routes based on $C_{time}(r)$ and on $C_{tradeoff}(r, w)$ for various values of w, where least-cost routes are computed using Dijkstra's algorithm. We then measure the percentage differences of the latter compared to the former for each r in terms of travel time and CO_2 exposure. The aggregation of such values will produce a pair $(time\_deviation_w, exp\_deviation_w)$ associated with w. By repeating the process for several values of w, we will obtain several deviation pairs. By construction, deviations for $w = 1$ are zero; any smaller w will weigh more the exposure impact and less the travel time, thus will tend to reduce the former and increase the latter; by iteratively decreasing w, then, we obtain a sequence of deviation pairs where each is worse than the previous ones in terms of time and better in terms of exposure. Since none of them is better than any other on both dimensions, they form a Pareto front of locally optimal choices. An example can be seen in Figure 9, which will be discussed in detail in the next section.

While all solutions on the Pareto front are locally optimal, there exist a few heuristics to identify a solution that provides the best trade-off between the two dimensions. Among them, we adopt the so-called *hyper-rectangle volume* method [21]. This consists of associating each point on the Pareto front to a measure of the potential set of other solutions that it could dominate, simply estimated as the area of the (hyper-)rectangle having opposing vertices at the point itself and an ideal point representing the theoretical worst possible solution. In our specific case, that can be expressed as:

$$\begin{aligned} HV(w) = \quad & (time\_deviation_{max} - time\_deviation_w) \times \\ & (exp\_deviation_{max} - exp\_deviation_w) \end{aligned} \tag{1}$$

Then, the HV-based optimum value is chosen as $w^* = \arg\max HV(w)$. We define the ideal worst case in Equation 1 as $time\_deviation_{max} = 1.1 \cdot \max_w time\_deviation_w$ for time deviations, and similarly for $exp\_deviation_{max}$. Notice that the 1.1 factor is introduced to avoid the case $HV(w) = 0$, which would otherwise systematically happen on the extreme cases ($w = 0$ and $w = 1$) creating issues when the Pareto front contains very few points.

D. Local (trip level) trade-off selection

The selection mechanism described above seeks a single value of w that overall fits best all the trips considered. While the aggregate trade-off is optimized, it does not prevent single trips from adopting a choice that results in being unbalanced, meaning that the corresponding path is either much longer yet only slightly healthier, or (on the opposite)

[1] We remark that normalization is not strictly needed. However, if omitted, the weight w would play both the role of *blending degree* of the two components (time vs exposure) and *translation coefficient* between two different dimensionalities (seconds vs milligrams of CO_2), making the interpretation of results in later sections more difficult.

just slightly shorter yet much less healthy. A natural evolution of the heuristics consists thus in applying the same selection approach separately to each trip, building its corresponding Pareto front, and selecting the w that maximizes $HV(w)$.

Figure 1 depicts three examples of Pareto front obtained on different trips. We can immediately observe that the front is less smooth than the global front example shown in Figure 9, which could be expected, since the variability of routes for a single trip is clearly more limited than the sum of effects of hundreds of different trips, and slightly changing the value of w often does not affect the routing of a trip. A second observation is that in some cases the variations of w produce variations of routes (and thus exposure and time measures) only close to one edge of the front, corresponding to a strong emphasis on one of the two dimensions, e.g. the blue line (where several w produce quasi time-optimal solutions) or the green one (where many w values yield similar, quasi exposure-optimal results). Finally, there are cases where the Pareto front contains just one or, as in the case of the red line in the plot, two points. Here, a straightforward definition of HV would produce the same zero volume for both solutions, choosing w randomly, which motivated our definition of the worst-case reference point in Section III-C.

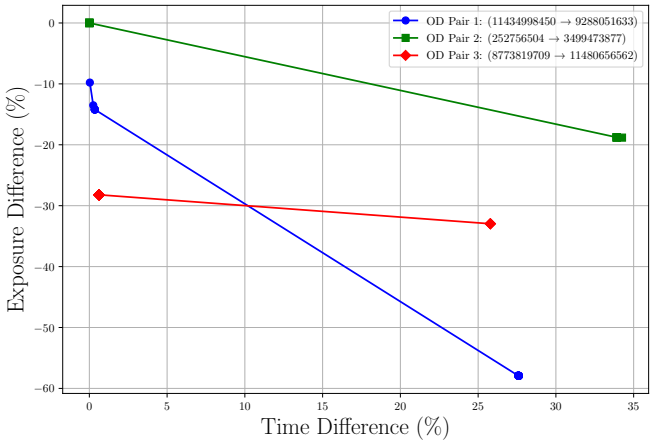

Fig. 1: Pareto fronts for three selected OD pairs, showing the trade-off between time and exposure differences. Each line represents a different OD pair, where the x-axis indicates the percentage increase in travel time and the y-axis represents the percentage reduction in exposure.

IV. Experiments

In this section, we will describe the experimental process adopted to study the impact of routing on exposure and travel times over a use case centered in a medium-sized city in Italy[2]. We compare along several dimensions the solutions obtained with different strategies: the standard fastest-path routing, the global trade-off strategy described in Section III-C, and the local (trip-specific) trade-off strategy shown in Section III-D.

[2]The code to reproduce our experiments are available at https://github.com/baygaliyev/exawaro

A. Experiments setup

Data and use case. Experiments are centered in the city of Pisa, in Italy. The city has several interesting features that are relevant to our study: it has a historical center with several pedestrian roads, as well as many roads with resident-only access and others with a significant traffic load, all within a compact area. Also, the city hosts many schools and a university, with significant flows of people walking/cycling between work/study locations and public transport facilities. Finally, Pisa is also a major iconic tourist attraction, which adds flows of visitors simply wandering around the city without clear origin-destination schedules.

Leveraging the work done in paper [13], we make use of an emission concentration map derived from a GPS dataset of private vehicles in the city, also used in other studies for similar purposes (e.g. [3]), covering ca. 11000 vehicles over one year. The map is shown in Figure 2. The low-emission areas are mainly concentrated in the historical center and pedestrian streets (center of the map), as well as around parks (lower-right and upper-left) and some peripherical areas (e.g. the lower part, where only airport-related services are located).

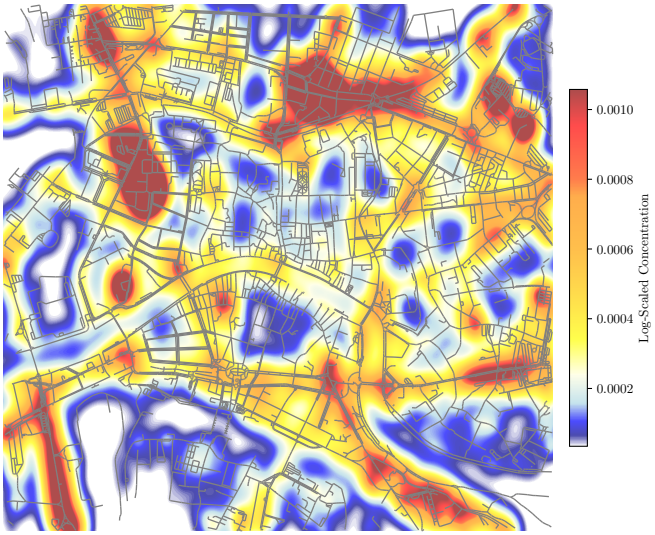

Fig. 2: Emission concentration levels of CO_2 estimated from [13] (Area: 3 km × 3 km square on the center of Pisa).

Combining the emission concentration map with the road network (as provided by OpenStreetMap), we derived two key attributes for each road segment: the walking time, simply inferred as the ratio of segment length over a fixed walking speed, set to 5 km/h; and the total exposure resulting from traversing the whole segment, estimated as described in Section III. Figure 3 shows the log-scaled distributions of walking time (left) and personal exposure (right) across the road network in the center of Pisa. Darker red segments on the left map indicate roads with higher travel times, typically corresponding to longer or less direct paths, while lighter yellow segments represent shorter travel durations. The right map illustrates personal exposure levels, where red segments

indicate high-exposure roads, often in high-traffic areas, while blue segments correspond to low-exposure paths, likely in pedestrianized zones or low-emission areas.

Generation of OD Pairs. To analyze the trade-off between travel time and pollution exposure, we generated a set of 50,000 random origin-destination (OD) pairs over the city and computed the corresponding routes. We selected OD pairs based on the spatial structure of the road network to ensure a realistic range of travel distances. The OD pairs were sampled from the set of all nodes in the network while enforcing the following constraints:

- The distance between the origin and destination had to be within 200 m to 1500 m.
- Nodes were randomly sampled from the road network graph to ensure diversity in selected routes.
- The Haversine function was used to compute the great-circle distance between sampled nodes.

A depiction of the sampled OD pairs is given in Figure 4. Here we see that both origins (4-left) and destination locations (4-center) are well spread across the whole city, with a slightly higher density around some key areas, like the train station (lower-center), the city center and parks/university (lower-right). Finally, Figure 4-right shows a few origin points together with all their associated destination, providing an idea of the spatial range of the simulated trips.

Computation of Fastest and Trade-Off Cost Routes. For each OD pair, we computed two types of routes:

1) **Fastest Route**: the shortest path was determined by minimizing the walking time, using the $C_{time}(r)$ cost for each road segment r.
2) **Trade-Off Cost Route**: the cost function $C_{tradeoff}(w, r)$ combining walking time and pollution exposure for each road segment r was used, allowing the path to deviate from the shortest path to reduce exposure. We selected representative weights w using an exponential scaling approach, ranging from 0.8^1 to 0.8^{19}, to ensure a well-distributed set of weights and achieve a smoother Pareto front.

B. Impact of standard routing

Standard routing strategies usually minimize travel time, aiming for efficiency, and disregarding other factors. As a reference for the discussion, we show in Figure 5 the distributions of the fastest travel durations (left) and exposure levels (right) across all OD pairs obtained with the quickest path criterion. The travel time histogram exhibits a left-skewed distribution, with most trips concentrated between 500 and 1500 seconds, while longer trips are progressively less frequent. The exposure distribution is also left-skewed, with the majority of trips experiencing lower exposure values. Still, a secondary peak appears around the value of 2.0, indicating a subset of trips with significantly higher exposure levels. This suggests that in many cases following the fastest path indirectly reduces exposure, since it reduces the exposure time, and that is enough in situations where the emission concentrations are not strongly skewed. However, in a significant portion of cases, the exposure-blind heuristics probably make pedestrians cross high-concentration areas, causing an increase in exposure.

C. Impact of global trade-off

As discussed before, adopting a trade-off cost in the routing of pedestrians can result in different routes depending on the importance we give to exposure vs. travel time. Figure 6 shows a (real) example of the effect of different weights in path selection. The visualization overlays three routes on the CO_2 concentration heatmap, demonstrating how varying the weight parameter affects the trade-off between travel time and exposure:

a) Exposure-Minimizing Route ($w = 0.01$): The red route represents the most exposure-aware path, which prioritizes avoiding high-pollution areas at the cost of increased travel distance. This route bypasses regions with high concentration levels, favoring lower-exposure alternatives even if they result in a longer journey.

b) Balanced Route ($w = 0.50$): The blue route seeks a middle ground, optimizing both travel time and exposure reduction. It follows a more direct path while making slight deviations to avoid highly polluted areas.

c) Time-Minimizing Route ($w = 0.99$): The green route, optimized for minimum travel time, follows a nearly direct trajectory regardless of pollution levels.

Figure 7 presents the distribution of travel time (left) and exposure (right) obtained by the trade-off heuristics by globally fixing three different values to the weight parameter w – this is the counter-part of Figure 5 for a trade-off schema. As the time weight increases (favoring travel time minimization), the distribution of travel durations is more concentrated towards lower values, with the highest weight ($w = 0.8$) resulting in the shortest routes. Correspondingly, higher exposure weights yield a distribution of exposure levels closer to low values, also greatly reducing the significance of the second peak, i.e. not only does the overall exposure decrease, but the cases of critically high exposure are strongly reduced.

An alternative perspective is given in Figure 8. Here we try to understand how much different are the suggested routes in terms of path similarity. The latter is computed using a variant of the Jaccard similarity, namely we measure the path length shared by two trips, normalized by their overall length:

$$J(A, B) = \frac{\sum_{r \in A \cap B} length(r)}{\sum_{r \in A \cup B} length(r)}$$

where A and B are two routes, represented as the set of road segments traversed.

The boxplot distributions of measure J for lower weights (e.g., $w = 0.014$) exhibit greater variability, while higher weights (e.g., $w = 0.8$) result in greater path similarity with the fastest route, indicated by the flat boxplots. Mean (green triangles) and median (orange horizontal lines) values with weight, while the median values rise above the mean

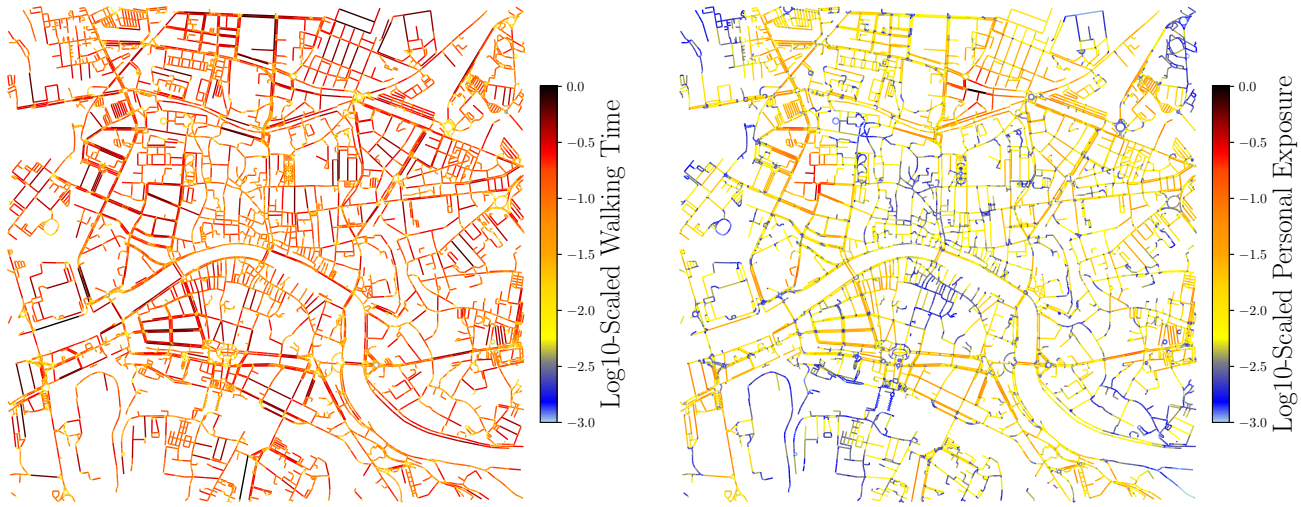

Fig. 3: Visualization on the road network of (log-scaled) walking time (left) and personal exposure (right). The color intensity represents the magnitude, with darker roads indicating longer travel times (left) or higher exposure levels (right).

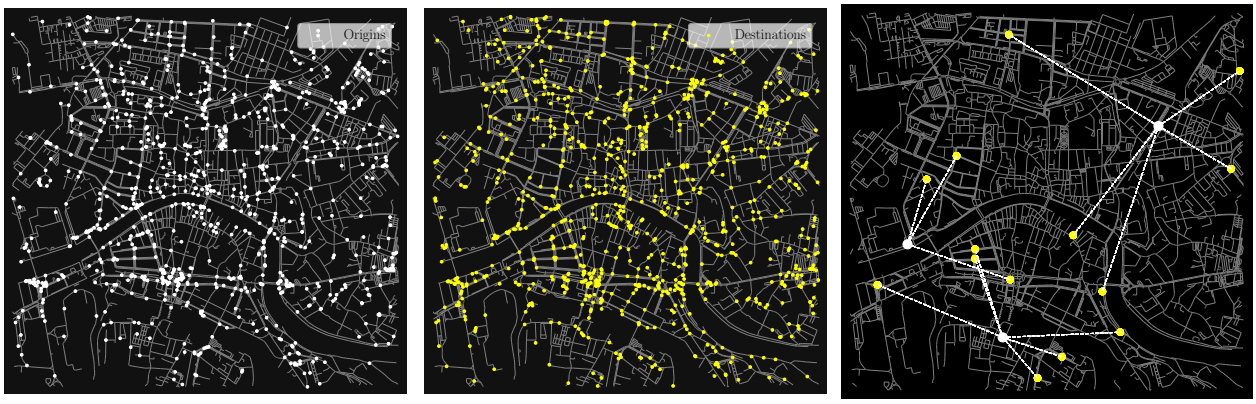

Fig. 4: Spatial distribution of origin (left) and destination (middle) points from 50,000 OD pairs plotted on the road network of Pisa. Origins are marked in white, and destinations are marked in yellow. On the right, three randomly selected origins with their respective destinations are plotted.

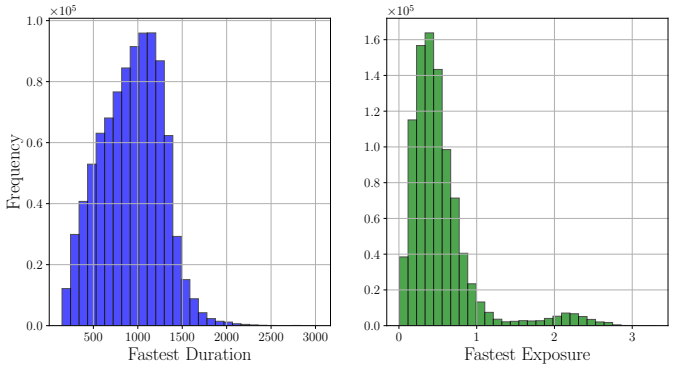

Fig. 5: Distribution of durations (left) and exposure values (right) obtained from fastest-time routes for the analyzed OD pairs.

values, indicating that more than half of the OD routes become identical to the fastest routes.

The Pareto front in Figure 9 depicts the trade-off between average time difference and average exposure reduction across different route optimization weight choices (w). As expected, lower exposure values are achieved at the cost of increased travel time, forming a convex frontier. The highlighted best point (red) represents the optimal trade-off w^* found by maximizing $HV(w)$, balancing a 15% reduction in exposure with a moderate increase in travel time (ca. 5%) at $w = 0.107$. This point is close to the steepest section of the curve, indicating that further reductions in exposure would require disproportionately higher increases in travel time. The overall shape confirms the diminishing returns of exposure reduction, making the optimal weighting choice dependent on user preferences.

For a better reading, the Pareto front of *median* time and exposure differences are shown in Figure 10, which incor-

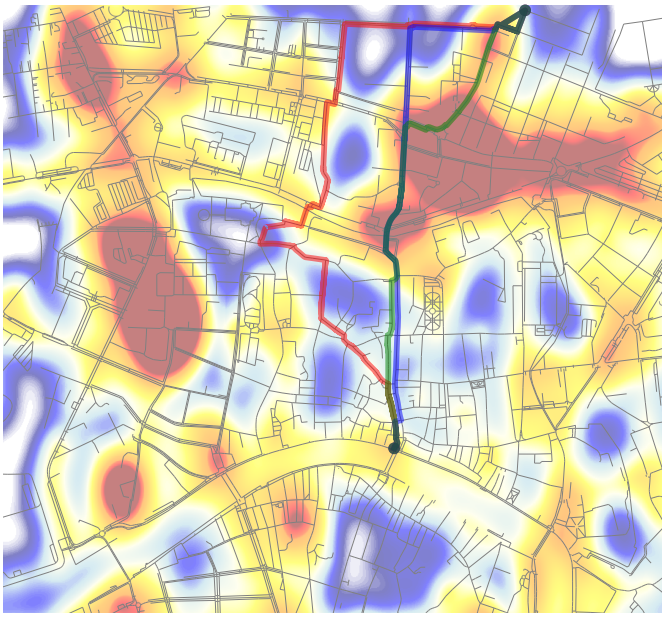

Fig. 6: Visualization of three routes from the northern outskirts of Pisa to Ponte di Mezzo, overlaid on the CO_2 concentration map. Each route corresponds to a different weight (0.01, 0.50, 0.99), represented in distinct colors (red, green, and blue, respectively). The visualization highlights the impact of weight variation on path selection, showcasing the trade-off between minimizing travel time and reducing exposure.

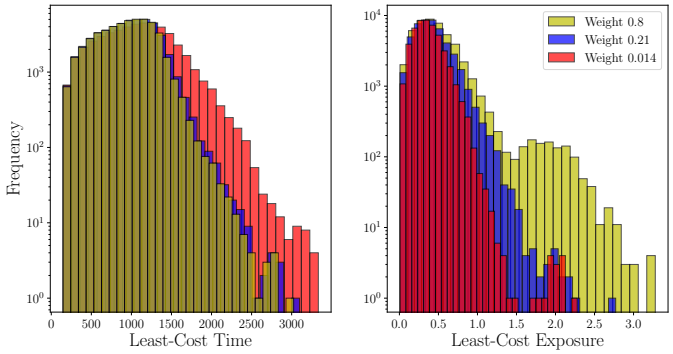

Fig. 7: Distribution of least-cost travel time (left) and least-cost exposure (right) for three weight values – 0.8 (yellow), 0.21 (blue), and 0.014 (red) – in the routing cost function.

porates variability in exposure reduction through the 25th-75th percentile range (shaded area) and boxplots. The trend confirms that higher exposure reductions require increasing travel times, with diminishing returns at the lower end. The red-highlighted best point ($w = 0.055$), selected using the hyper-volume criterion, represents the optimal balance between time increase and exposure reduction. The spread of the percentiles indicates greater variability in exposure reductions for intermediate weightings, while extreme values tend to be more stable.

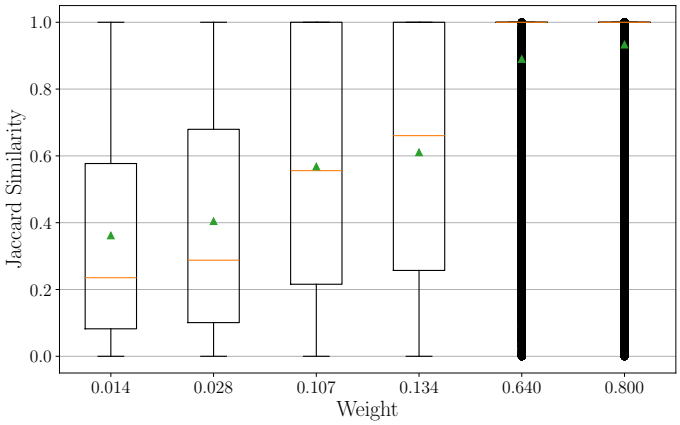

Fig. 8: Boxplots of weighted Jaccard similarity values across selected weights. The medians are represented by orange horizontal lines, while the means are indicated by green triangles.

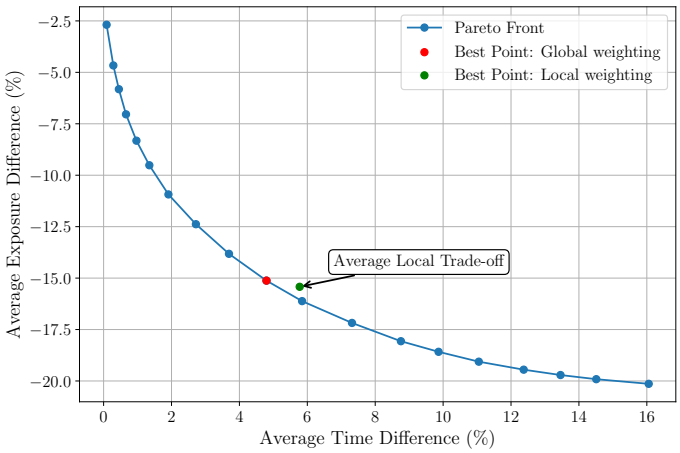

Fig. 9: Pareto front based on average differences in travel time and CO_2 exposure across weights, computed over 50,000 OD routes in Pisa.

D. Local trade-off

Adopting the local, trip-specific selection of the best w generally yields different results from choosing the same w for all the trips. The first evidence of this can be seen in the already discussed Figure 9, the average impact of the local trade-off is plotted together with the Pareto front of the global approach. As we can see (and expect) the selected point is slightly off the global Pareto front, yet it is not dominated by any solution of the former, meaning that it is another locally optimal solution. Also, we can see that the local trade-off optimum gives a slightly higher emphasis to the exposure component as compared to the global one.

For a more detailed comparison, in Figure 11 each point represents an OD pair, where the x-axis shows the percentage difference in travel time between the local and the global trade-offs, and the y-axis shows the same percentage difference for exposure. The color represents the local (trip-specific) best

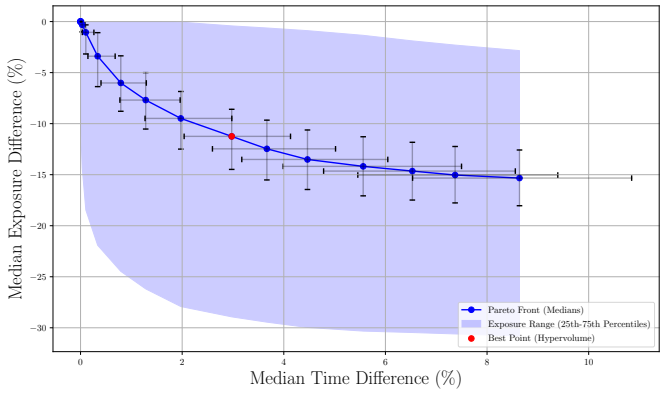

Fig. 10: Pareto front based on median differences in travel time and CO_2 exposure across weights, with 25th–75th percentile ranges shown for both metrics.

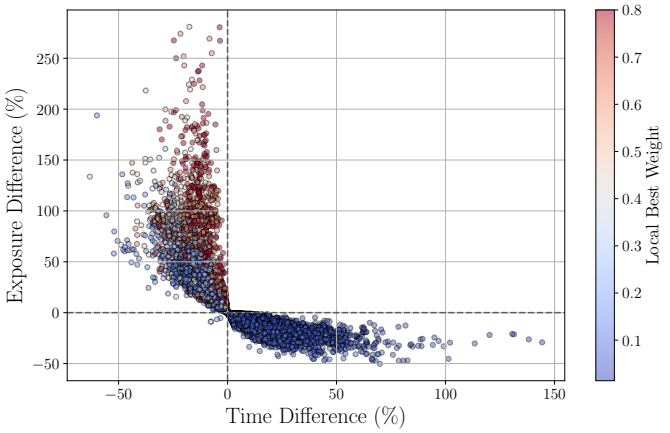

Fig. 11: Comparison of time and CO_2 exposure differences between the local best weighting and global weighting approach ($w = 0.107$).

weight. The dashed lines at $x = 0$ and $y = 0$ indicate where the local best weighting performs identically to the global weighting. Points in the bottom-left quadrant indicate OD pairs where the local best weighting reduces both time and exposure, while the empty top-right quadrant indicates that there is no case where the local best weighting increases both.

The strong concentration of points in the bottom-right region of the plot suggests that, in most cases, the local best weighting leads to a reduction in exposure, often at the cost of increased travel time. Meanwhile, a smaller number of points in the top-left quadrant indicates cases where local best weighting decreases travel time while increasing exposure. The presence of several OD pairs with high exposure differences, particularly for larger w_{local} values (red-colored points), suggests that time-minimizing weightings tend to result in substantial deviations from the global weighting, particularly in highly polluted areas. Exposure reductions mostly range between 0% and 50%, while increases in exposure are more pronounced, spanning from 0% to 200%, with some outliers exceeding 200%. In addition, the cluster of points near the $x = 0$ axis but with high positive exposure differences implies that, in some cases, the local best weighting does not significantly change travel time but may still lead to higher exposure levels. The resulting trade-off is reflected in Figure 9, where the green point has shifted toward longer travel time and higher exposure

Figure 12 depicts the difference in road segment usage between a global weighting approach and an adaptive weighting approach. The color coding indicates the relative preference for each road segment: green segments are more frequently used under adaptive weighting, while red segments are more frequently used under the global strategy. The width of each road segment in the figure represents how frequently that road is used across multiple OD pairs. Thicker segments correspond to roads that are used more frequently in either routing strategy. Notably, major roads are often thick red segments, suggesting that global weighting leads to repeated use of key corridors without optimizing for exposure. Conversely, adaptive weighting distributes route choices across the network, leading to broader use of alternative paths. The heatmap highlights areas with high CO_2 concentration, and a notable correlation exists between red-dominated roads and high-emission zones. This suggests that least-cost routing strategies under the global weighting do not actively avoid pollution hotspots. Conversely, the adaptive approach, which tailors weights per OD pair, is observed to reroute away from these areas, thereby reducing exposure. This results in a preference for pedestrian-friendly streets, such as Corso Italia, Borgo Stretto, and Viale Ciccone, where the concentration of vehicular emissions tends to be lower.

V. Conclusion

In this paper, we investigated the impact of vehicular emissions on pedestrian exposure and introduced an exposure-aware pedestrian routing strategy that balances travel time and pollution avoidance. By integrating emission concentration data – obtained through our 4-step pipeline – with a bi-objective optimization approach, we demonstrated that rerouting pedestrians away from high-exposure areas can significantly reduce their pollution intake, at the smaller cost of increased travel time. Through extensive simulations in an urban environment, we analyzed the aggregated trade-offs between exposure reduction and travel efficiency, providing insights into how different route optimization strategies affect pedestrian trips in terms of time and health impact. Additionally, we compared the global and local weighting for route optimization. Findings highlight that adaptive, trip-specific optimization yields better results than static, globally defined trade-offs in many cases. However, the average trade-off for locally weighted routes is slightly less favorable than the globally optimized weighting, highlighting the complexity of adaptive weighting. Furthermore, the spatial analysis of exposure-aware routing revealed that certain areas of the city pose higher pollution risks, reinforcing the need for targeted

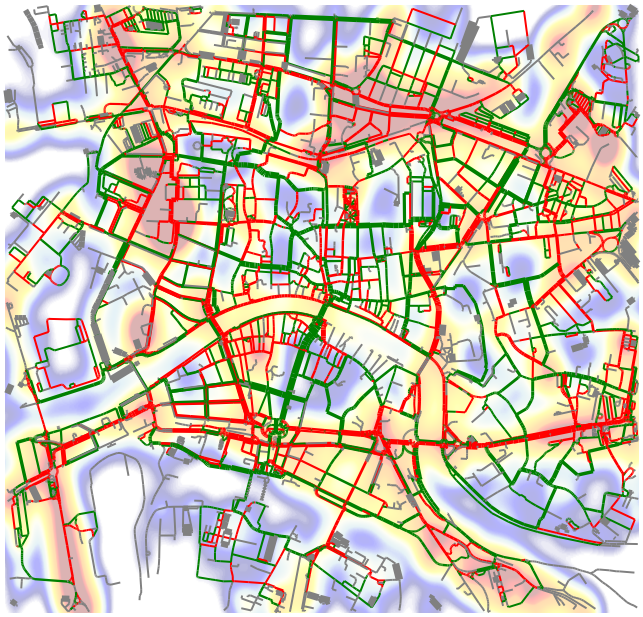

*Fig. 12: Visualization of road segment usage differences between the global route weighting ($w = 0.107$) and the optimal local weighting. **Green** segments indicate roads used more frequently in the optimal local weighting, **red** segments indicate roads used more frequently in the global weighting strategy, and **grey** segments indicate roads used equally in both.*

pedestrian infrastructure planning. Future work could expand this study by incorporating pedestrian mobility demand and integrating car routing. The latter presents a greater challenge, as optimizing vehicular routes also depends on driving speed.

References

[1] M. Nyhan and et al., "Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model," *Atmospheric Environment*, vol. 140, pp. 352–363, 2016.

[2] *Transforming Our World: The 2030 Agenda for Sustainable Development*. Springer Publishing Company, 6 2018.

[3] M. Böhm, M. Nanni, and L. Pappalardo, "Gross polluters and vehicle emissions reduction," *Nature Sustainability*, vol. 5, no. 8, pp. 699–707, 6 2022.

[4] G. Cornacchia, M. Böhm, G. Mauro, M. Nanni, D. Pedreschi, and L. Pappalardo, "How routing strategies impact urban emissions," in *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, ser. SIGSPATIAL '22. New York, NY, USA: Association for Computing Machinery, 2022.

[5] M. N. Rahman and A. O. Idris, "TRIBUTE: Trip-based urban transportation emissions model for municipalities," *International Journal of Sustainable Transportation*, vol. 11, no. 7, pp. 540–552, 2017.

[6] Y. Huang and et al., "Remote sensing of on-road vehicle emissions: Mechanism, applications and a case study from Hong Kong," *Atmospheric Environment*, vol. 182, pp. 58–74, 2018.

[7] C. K. Gately, L. R. Hutyra, S. Peterson, and I. S. Wing, "Urban emissions hotspots: Quantifying vehicle congestion and air pollution using mobile phone GPS data," *Environmental Pollution*, vol. 229, pp. 496–504, 2017.

[8] R. K. C. Chan, J. M.-Y. Lim, and R. Parthiban, "Missing traffic data imputation for artificial intelligence in intelligent transportation systems: Review of methods, limitations, and challenges," *IEEE Access*, vol. 11, pp. 34 080–34 093, 2023.

[9] X. Yao, Y. Gao, D. Zhu, E. Manley, J. Wang, and Y. Liu, "Spatial origin-destination flow imputation using graph convolutional networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 12, pp. 7474–7484, 2021.

[10] Z. Liu and et al., "Learning geo-contextual embeddings for commuting flow prediction," in *AAAI Conference on Artificial Intelligence*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:211037977

[11] T. S. Jepsen, C. S. Jensen, and T. D. Nielsen, "Relational fusion networks: Graph convolutional networks for road networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 418–429, 2022.

[12] Y. Chang, E. Tanin, X. Cao, and J. Qi, "Spatial structure-aware road network embedding via graph contrastive learning," in *Proceedings 26th International Conference on Extending Database Technology, EDBT 2023, Ioannina, Greece, March 28-31, 2023*. OpenProceedings.org, 2023, pp. 144–156.

[13] G. Aliyev and M. Nanni, "From gps traces to individual emission exposure: A data-driven four-step process," in *Intelligent Transport Systems*, A. Kocian, P. Milazzo, A. L. Henriques Martins, M. Nanni, and L. Pappalardo, Eds. Cham: Springer Nature Switzerland, 2025, pp. 64–82.

[14] P. de Souza and et al., "Quantifying disparities in air pollution exposures across the united states using home and work addresses." *Environmental science & technology*, vol. 58, no. 1, pp. 280–290, 2024.

[15] M. Nyhan and et al., "Exposure track - the impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution," *Environmental science & technology*, vol. 50, pp. 9671–81, 9 2016.

[16] B. Dewuf and et al., "Dynamic assesment of exposure to air pollution using mobile data," *International Journal of Health Geographics*, vol. 15, 2016.

[17] Q. Li, S. Liang, Y. Xu, L. Liu, and S. Zhou, "Assessing personal travel exposure to on-road PM2.5 using cellphone positioning data and mobile sensors," *Health & Place*, vol. 75, p. 102803, 5 2022.

[18] A. Leelőssy and et al., "Dispersion modeling of air pollutants in the atmosphere: a review," *Open Geosciences*, vol. 6, no. 3, pp. 257–278, 2014.

[19] M. Liang, Y. Chao, Y. Tu, and T. Xu, "Vehicle pollutant dispersion in the urban atmospheric environment: A review of mechanism, modeling, and application," *Atmosphere*, vol. 14, no. 2, 2023.

[20] G. A. Briggs, "Diffusion estimation for small emissions. preliminary report," 5 1973.

[21] E. Demir, T. Bektaş, and G. Laporte, "The bi-objective pollution-routing problem," *European Journal of Operational Research*, vol. 232, pp. 464–478, 2 2014. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0377221713006486

[22] J. Luo, M. J. Barth, and K. Boriboonsomsin, "Vehicle routing to mitigate human exposure to traffic-related air pollutants," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 11 2018, pp. 2765–2770. [Online]. Available: https://ieeexplore.ieee.org/document/8569501/

[23] G. Davies and D. Whyatt, "A least-cost approach to personal exposure reduction," *Transactions in GIS*, vol. 13, pp. 229–246, 4 2009. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2009.01150.x

[24] S. Yoon, Y. Moon, J. Jeong, C.-R. Park, and W. Kang, "A Network-Based Approach for Reducing Pedestrian Exposure to PM2.5 Induced by Road Traffic in Seoul," *Land*, vol. 10, p. 1045, 10 2021.

[25] E. Willberg, A. Poom, J. Helle, and T. Toivonen, "Cyclists' exposure to air pollution, noise, and greenery: a population-level spatial analysis approach," *International Journal of Health Geographics*, vol. 22, p. 5, 2 2023. [Online]. Available: https://ij-healthgeographics.biomedcentral.com/articles/10.1186/s12942-023-00326-7