How Routing Strategies Impact Urban Emissions

Giuliano Cornacchia  
University of Pisa, ISTI-CNR  
Pisa, Italy  
giuliano.cornacchia@phd.unipi.it

Matteo Böhm  
Sapienza University of Rome  
Rome, Italy  
bohm@diag.uniroma1.it

Giovanni Mauro  
University of Pisa, IMT, ISTI-CNR  
Pisa, Italy  
giovanni.mauro@phd.unipi.it

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

Dino Pedreschi  
University of Pisa  
Pisa, Italy  
dino.pedreschi@unipi.it

Luca Pappalardo  
ISTI-CNR  
Pisa, Italy  
luca.pappalardo@isti.cnr.it

ABSTRACT

Navigation apps use routing algorithms to suggest the best path to reach a user’s desired destination. Although undoubtedly useful, navigation apps’ impact on the urban environment (e.g., carbon dioxide emissions and population exposure to pollution) is still largely unclear. In this work, we design a simulation framework to assess the impact of routing algorithms on carbon dioxide emissions within an urban environment. Using APIs from TomTom and OpenStreetMap, we find that settings in which either all vehicles or none of them follow a navigation app’s suggestion lead to the worst impact in terms of CO₂ emissions. In contrast, when just a portion (around half) of vehicles follow these suggestions, and some degree of randomness is added to the remaining vehicles’ paths, we observe a reduction in the overall CO₂ emissions over the road network. Our work is a first step towards designing next-generation routing principles that may increase urban well-being while satisfying individual needs.

CCS CONCEPTS

• Human-centered computing → Empirical studies in ubiquitous and mobile computing; • Computing methodologies → Multi-agent planning.

KEYWORDS

human mobility, social AI, routing, navigation systems, urban sustainability, vehicular traffic, traffic simulation

1 INTRODUCTION

To tackle the pressing challenge of climate change [7] and the urgent call by the United Nations to reduce the adverse environmental impact of cities [56], there is an increasing effort to study the impact of human mobility on urban well-being dimensions, such as traffic congestion, air pollution, and carbon dioxide emissions. An element that enriches the complexity of this challenge is the widespread diffusion of GPS navigation apps such as TomTom, Google Maps, and Waze, which use routing algorithms, heuristics and AI to suggest the best path to reach a user’s desired destination. Although undoubtedly useful, particularly when exploring an unfamiliar city, GPS navigation apps may also cause several issues in the urban environment: since they are typically optimised to keep an individual’s trip as short as possible, they do not care about collective effects on the city, such as whether the traffic can be absorbed by the streets, compromises safety or creates more pollution [20, 41, 54]. Many documented cases show that GPS navigation apps may create chaos: for example, they sometimes divert heavy traffic through side roads in nearby towns, with such an impact on the residents’ life that experts stated that these apps are a new agent claiming a “right to the city” [24, 25].

Beyond the anecdotal, preliminary research show that the impact of navigation apps on the urban environment is mixed [22, 52]. On the one hand, navigation apps may provide benefits in mitigating carbon dioxide emissions [5]; on the other hand, they may increase the population exposure to pollution in densely-populated areas [49]. Overall, existing studies are sporadic and yield contradictory results, leading to a picture of the navigation apps’ impact on the urban environment that is mainly unclear and incomplete. In particular, the literature still lacks a rigorous framework to assess and compare the impact of navigation apps on urban well-being.

In this paper, we design a simulation framework – TraffiCO – to assess the impact of GPS navigation apps on urban well-being in terms of carbon dioxide (CO₂) emissions. The framework uses GPS data to reconstruct a city’s mobility demand, and navigation apps’ APIs to obtain routing suggestions (paths on the road network) for a set of vehicles’ origin and destination. Then, TraffiCO relies on a traffic simulator that considers all aspects of vehicular mobility (e.g., jams, queues at traffic lights, roads capacity), to generate trajectories that describe when each vehicle visits each road in its path. Finally, the framework use an emission model to estimate CO₂ emissions for each segment in the road network.
We apply TraffiCO to the city of Milan (Italy) to study how emissions changes as we increase the fraction of vehicles that follow navigation apps’ routing. We find that settings in which either all vehicles or none of them follow a navigation apps’ routing suggestion lead to the highest amount of CO\(_2\) emissions. These settings also correspond to the most uneven distribution of the CO\(_2\) emissions across the roads, with a few roads suffering the greater quantity of emissions. In contrast, a scenario where just a fraction of the vehicles (around 50\%) follows navigation apps’ suggestion and the remaining part follow a randomly perturbed fastest path, reduces the vehicles’ overall emissions and distributes them more evenly on the road network. We also find that the navigation apps’ routing affects the spatial distribution of emissions, making Milan’s external ring road more polluted while unloading the internal roads from the emissions. Notably, adding perturbation to the vehicles’ path is beneficial, as it reduces the overall emissions, distributes them more evenly, and reduces the vehicles’ average travel time.

Our simulation framework is a useful tool to assess and compare routing strategies, helping drivers, institutions, and policymakers understand the impact of navigation apps on the urban environment. Moreover, TraffiCO is a first step towards designing and testing next-generation routing principles that may increase urban well-being while satisfying individual needs.

Open Source
We provide the implementation of TraffiCO, the code and the link to the data to reproduce our study at https://bit.ly/traffico2_gh.

2 RELATED WORK
Environmental impact of vehicular traffic
The environmental impact generated by vehicles (e.g., air pollution) is becoming increasingly evident in urban environments. Existing methods to quantify vehicles’ emissions range between two extremes. On the one hand, some approaches rely on measurements performed on small samples of vehicles with high spatio-temporal resolutions, such as those coming from particulate sensors [21] or portable emissions measurement systems (PEMS) [16, 40]. These sensors measure emissions in real-world driving conditions, producing accurate estimates but hardly generalisable patterns due to the limited sample size. For example, two studies [16, 40] analyse emissions from PEMS of one and three vehicles, finding that the highest emissions are associated with the urban part of the route, flat roads, and low speed.

On the other hand, some studies cover a region’s entire fleet, such as those using odometer readings from annual safety inspections. These data describe each vehicle’s age, fuel type, engine volume, and mileage, used in macroscopic models to estimate annual emissions. Two studies [14, 29] use odometer readings to compute mean annual emissions for UK postcode areas and explore the built-environment effects (e.g., work accessibility) on the vehicles’ annual miles travelled in Boston. Unfortunately, odometer readings miss critical information such as instantaneous speed and acceleration [17, 23, 32, 60], making it challenging to track emissions over time and map them to suburban areas.

Somewhere between these two extremes lie works that use GPS traces, which describe human mobility in great detail [8, 39] and can cover a representative fraction of the vehicle fleet [11, 47]. These data allow computing instantaneous speed and acceleration, which can be used within microscopic models to obtain emissions estimates in high spatio-temporal resolution. Several studies use GPS traces to analyse the vehicles’ emissions at different spatio-temporal scales [11, 37, 46], investigate the relationship between emissions and the urban environment [51], vehicle miles travelled and fuel consumption [57], or trip rates and travel mode choice [13]. Other studies concentrate on congestion-related emissions [27] or braking [15], emissions associated with ride-hailing [55] and bus stops’ positioning [59], the impact of urban policies [50], methods for emission modelling [6, 61], and air quality monitoring [21].

Microscopic traffic simulation
Microscopic traffic simulators are crucial to simulate vehicular traffic given a mobility demand. A notable example is SUMO (Simulation of Urban MOBility), an open-source, portable, and multi-modal tool designed to handle traffic simulations on road networks [2, 35, 38]. SUMO allows controlling several aspects of traffic, from fuel consumption to vehicle emissions and routing strategies. Several works use SUMO to study the impact of vehicular traffic on the urban environment. Alazzawi et al. [2] simulate the introduction of a fleet of automated shared vehicles into Milan, finding that a fleet of 9500 of these vehicles helps mitigate traffic congestion and emissions. Malik et al. [42] propose a traffic system that re-routes an emergency vehicle in case of traffic jams to reduce travel time and pollution. Kraijewicz et al. [34] use optical information systems to optimise traffic lights over junctions better than traditional approaches. Zubillaga et al. [63] compare the traditional traffic-light coordination (green-wave method) with a self-organising method that adapts to traffic demands, showing how the latter is way better.

Urban routing and impact of navigation systems
People’s natural routing choices may significantly deviate from the optimal route. The origin of these sub-optimal human routes may lie in several factors, e.g., the environment in which one grew up [19], the subjective perception of space [45], the presence of landmarks [26], and even the usage of an electric vehicle [31]. Zhu et al. [62] find that only 34\% of trips, mostly very short and very long ones, follow the shortest time path. Lima et al. [36] find that 53\% of the drivers’ routes are not optimal and that most drivers use a few preferred routes for their journeys. Xu et al. [58] find similar results using location-based services data, which are less accurate but more pervasive than vehicular trajectory data. Bongiorno et al. [12] find that people increasingly deviate from the shortest path as the distance between origin and destination increases. Similarly, Manley et al. [43] show that people’s route choice results from multiple decisions made at each anchor of the route. Given the evident uncertainty in people’s routing strategies, the impact of navigation apps’ routing suggestions on the urban environment is not negligible [41]. For this reason, some effort has been put into developing “environmentally-friendly” navigation apps that minimise fuel consumption instead of travel time [9]. Studies on the effects of this eco-routing, either as fuel savings [22] or system-wide impact [1, 52], find that the urban impact of navigation apps is mixed: green navigation apps can reduce CO\(_2\) emissions but...
increase the population exposure to nitrogen oxides [49]. Following the same research line, Mehrvarz et al. [44] find that the fastest route suggested by navigation apps may not optimise fuel consumption. In contrast, Arora et al. [5] show that following Google Maps’ routing strategy saves 1.7% of CO₂ emissions and 6.5% travel time.

Position of our work
Existing studies are sporadic and yield contradictory results, and the impact of navigation apps on urban well-being is mainly unclear and incomplete. Moreover, the existing literature lacks a rigorous framework to assess and compare the various navigation apps and routing strategies. We fill this gap by proposing TraffiCO₂, a simulation framework to estimate the adverse impact of different routing strategies on total CO₂ emissions and their distribution of a city’s road network. Thus, our work extends the branch of literature on sustainable urban mobility by providing a rigorous experimental framework to disentangle the impact of routing criteria on the urban environment.

3 SIMULATION FRAMEWORK
We design a simulation framework – TraffiCO₂ – to generate a realistic urban traffic considering different routing strategies (Figure 1). First, we use real data to generate a mobility demand describing trips (origin-destination pairs) in an urban environment. Second, we transpose each trip into a path on the road network using some routing algorithm, obtaining a multiset of routed paths. Third, we use an agent-based model (SUMO) that, considering realistic aspects of vehicular mobility (e.g., jams, traffic lights, slowdowns), simulates an urban traffic based on the multiset of routed paths. Finally, we compute the vehicles’ travel time and the CO₂ emissions on each road from the trajectories generated by SUMO.

3.1 Road Network
It describes the road infrastructure where the vehicles move during the simulation. The road network is a directed graph \( G = (V, E) \), where \( V \) is the set of nodes representing road intersections and \( E \) is the set of edges representing roads. Both nodes and edges may have attributes, such as: traffic lights, number of lanes, road speed limit and type (e.g., motorway, secondary road). These attributes are used by SUMO to simulate realistic aspects of vehicular mobility.

3.2 Mobility Demand
The mobility demand \( D = \{T_1, \ldots, T_N\} \) is a multiset of \( N \) trips (one per each vehicle) within an urban environment. A single trip \( T_v = (o, d) \) for a vehicle \( v \) is defined by its origin location \( o \) and destination location \( d \). To compute \( D \), we first divide the urban environment into squared tiles of a given side (Figure 1a). Second, we use real mobility data to estimate the flows between the tiles, thus obtaining an origin-destination matrix \( M \) where an element \( m_{o,d} \in M \) describes the number of vehicles’ trips that start in tile \( o \) and end in tile \( d \) (Figure 1b.1).

Then, we iterate \( N \) times the following procedure. A vehicle’s \( v \) trip is a pair \( T_v = (e_o, e_d) \) generated by selecting at random a matrix element \( m_{o,d} \in M \) with a probability \( p_{o,d} \propto m_{o,d} \) and uniformly at random two edges \( e_o, e_d \in E \) within tiles \( o \) and \( d \), respectively (Figure 1b.2, b.3).

3.3 Paths generation
We translate the \( N \) trips in \( D \) into \( N \) paths obtaining a multiset \( \mathcal{D} \) of routed paths within the urban environment. Each path \( P_v(e_o, e_d, R) = (e_o, \ldots, e_d) \) of a vehicle \( v \) is a sequence of edges on the road network connecting \( e_o \) and \( e_d \) (Figure 1b.5), obtained by some routing algorithm \( R \). When a vehicle’s path is generated by a routing algorithm \( R \), we say that the vehicle is \( R \)-routed. In \( \mathcal{D} = \{P_1, \ldots, P_N\} \), the routed paths (Figure 1c) are generated independently by (different) routing algorithms.

3.4 Traffic Simulation
We simulate the vehicular traffic generated by the routed paths in \( \mathcal{D} \) using SUMO (Simulation of Urban Mobility) [2, 35, 38] (Figure 1d). SUMO explicitly models each vehicle’s physics and dynamics, including their routes through the road network, allowing us to simulate vehicular traffic realistically, including traffic jams, queues at traffic lights, and slowdowns due to heavy traffic. SUMO outputs an urban traffic (Figure 1e), i.e., a multiset \( S(\mathcal{D}, \mathcal{T}) = \{\overline{P_1}, \ldots, \overline{P_N}\} \) where:

- \( \mathcal{T} = \{t_1^{(1)}, \ldots, t_N^{(N)}\} \) is the sequence of departure times of the \( N \) paths in \( \mathcal{D} \), where \( t_1^{(1)} \in \mathcal{T} \) is a timestamp chosen uniformly at random from the simulation interval (e.g., 1 hour);
- a vehicle \( v \)'s trajectory \( \overline{P_v} \in S(\mathcal{D}, \mathcal{T}) \) is defined as:
  \[
  \overline{P_v} = ((e_o, t_v^{(1)}), \ldots, (e_d, t_v^{(m)}))
  \]
  where \( m \) is the length of path \( \overline{P_v} \).

3.5 Vehicle Emissions
We assess the impact of the urban traffic \( S(\mathcal{D}, \mathcal{T}) \) using a model that estimates the vehicles’ emissions based on their trajectories \( \overline{P_1}, \ldots, \overline{P_N} \). Specifically, we use the HBEFA3 emission model based on the Handbook of Emission Factors for Road Transport (HBEFA) database [30]. The HBEFA3-based model estimates the vehicle’s instantaneous CO₂ emissions relying on the following function, which is linked to the power the vehicle’s engine produces in each trajectory point \( j \) to overcome the driving resistance force [33]:

\[
E(j) = c_0 + c_1 s + c_2 s^2 + c_3 s^3 + c_4 s^4 + c_5 s^5
\]

where \( s \) and \( a \) are the vehicle’s speed and acceleration in point \( j \), respectively, and \( c_0, \ldots, c_5 \) are parameters changing per emission type and vehicle taken from the HBEFA database.

We compute the amount of CO₂ emissions on each edge \( e \in E \) by summing all the emissions corresponding to any vehicle \( v \)'s trajectory point that fall on \( e \), i.e., \( E(e) = \sum_{j} \sum_{\overline{P_v} \in \mathcal{D}} E(j) \). Finally, we construct a weighted road network \( \overline{G} = (V, \overline{E}) \) where each edge \( \overline{e} \in \overline{E} \) is associated with the attribute \( E(e) \) describing the amount of CO₂ emissions on it.

4 EXPERIMENTAL SETUP
We apply TraffiCO₂ to a 45 km² area in the city centre of Milan, Italy, for which we have GPS data describing 17k private vehicles traveling between April 2nd and 8th, 2007 (114k GPS points). We split Milan into squared tiles with 1 km side, detect the tiles where each
vehicle starts and stops [28, 48], and compute the origin-destination matrix $M$ of vehicles’ flows.

We download Milan’s road network $G = (V, E)$ from OpenStreetMap ($|V| = 5551$ nodes and $|E| = 36,945$ edges) and preprocess it to fix incorrect information regarding turns, intersections, road interruptions, the number of lanes per road, and other inaccuracies that characterise these data [4].

Based on the matrix $M$ and the road network $G$, we compute the mobility demand $D$ with $N = 15,000$ vehicles using the procedure described in Section 3.2. We choose $N = 15,000$ because it minimises the difference between the distribution of travel time of real trajectories and simulated ones, a common way to assess the realism of a simulated urban traffic [4] (see Appendix B for detail).

To translate the mobility demand $D$ into the multiset of routed paths $\mathcal{D}$, we consider two routing algorithms: OpenStreetMap (OSM) and TomTom (TT). OSM is a public voluntary geographic information system, which provides APIs\(^1\) to generate paths between locations. TT is a commercial navigation system service that provides APIs\(^2\) for routes generation. For OSM and TT, we use a routing principle that suggests a path between two locations as a trade-off between travel time and distance. Note that OSM and TT implement this principle differently, i.e., they may provide different paths for the same origin-destination pair.

We obtain the path of vehicles that do not follow navigation apps’ suggestions using Duarouter (DR), a routing algorithm provided by SUMO\(^3\) that suggests the fastest path (i.e., shortest travel time) between two edges on the road network. The fastest path may be perturbed using a randomisation parameter $w \in [1, +\infty)$, where $w = 1$ means no randomisation (i.e., the fastest path). The higher $w$, the more randomly perturbed the fastest path is. Therefore, DR with $w > 1$ allows us to model the driving behaviour of vehicles that do not strictly follow navigation apps’ suggestions and, simultaneously, to model the imperfection and non-rationality of human drivers [53]. Indeed, individuals get distracted when driving (e.g., take wrong turns), and they lack complete knowledge of the city’s traffic, the road network, and the best path to reach a destination. In our experiments, we use $w = 5$ because it is the value that minimises the difference between the distribution of travel time of real trajectories and simulated ones (see Appendix B). Figure 2 shows four examples of paths generated by OSM, TT, DR with $w = 1$ and DR with $w = 5$ between the same origin-destination pair (trip).

---

\(^1\)https://openrouteservice.org/dev/#/api-docs
\(^2\)https://developer.tomtom.com/routing-api
\(^3\)https://sumo.dlr.de/docs/duarouter.html

---

Figure 1: Schema of the TraffiCO simulation framework. (a) The city is split into squares using scikit-mobility [48]. (b.1) A trip is created by selecting at random an origin-destination pair from the OD matrix and two edges on the road network. (b.5) Some routing algorithm is used to convert each trip into a path on the road network. (c) Steps b.1-b.5 are repeated $N$ times ($N =$ number of vehicles) to obtain a multiset of routed paths. (d) A traffic simulator (SUMO) is used to simulate the urban traffic generated by the routed paths (e).

Figure 2: Examples of routed paths between an origin and destination pair according to OSM (blue), TT (red), DR with $w = 5$ (orange), and DR with $w = 1$ (black).
Given a routing algorithm $R \in \{\text{OSM, TT}\}$, we study its impact on the urban environment generating 11 multisets of routed paths $\mathcal{D}_i^R$, $i = 0, \ldots, 10$. In each multiset $\mathcal{D}_i^R$, the paths are chosen uniformly at random among the $N$ paths that are $R$-routed ($R \in \{\text{OSM, TT}\}$) and the remaining paths are routed by DR with $w = 5$. For example, for $i = 5$ and $R = \text{TT}$, $\mathcal{D}_5^T$ contains 50% of the paths routed by TT and the remaining vehicles routed by DR with $w = 5$. Similarly, $i = 7$ means that 70% of the vehicles are TT-routed and 30% are DR-routed ($w = 5$).

To make experiments more robust, for each $i = 0, \ldots, 10$, we generate ten times, each one with a different choice of $R$-routed vehicles that are chosen uniformly at random. Finally, we generate the urban traffic $S(\mathcal{D}_i^R, T)$ for each multiset of routed paths $\mathcal{D}_i^R$, where departure times in $T$ are chosen uniformly at random between 0 and 3600 seconds, i.e., we simulate one hour of traffic in Milan. Then, for each urban traffic $S(\mathcal{D}_i^R, T)$, we compute the CO2 emissions for each trajectory and aggregate them at the edge level, obtaining the weighted road network $G_i^R$. From $G_i^R$, we obtain the overall CO2 emissions in Milan $E_i^R = \sum_{e \in E} E(e)$, where $E(e)$ is the amount of emissions on edge $e$.

5 Results

We study how the distribution of the CO2 emissions across Milan’s roads changes varying the percentage of $R$-routed vehicles, i.e., varying $i = 0, \ldots, 10$ of $\mathcal{D}_i^R$, $R \in \{\text{OSM, TT}\}$. In general, emissions distribute across roads in a heterogeneous way: a few grossly polluted roads coexist with roads with significantly fewer emissions (Figure 3). Indeed, these distributions are associated with a Gini index $g_\text{osm} \in [0.864, 0.876]$ and $g_\text{tt} \in [0.860, 0.868]$ (Figure 4a) and are well approximated by a truncated power-law with the exponent $\alpha_\text{osm} \in [1.76, 1.89]$ and $\alpha_\text{tt} \in [1.76, 2.00]$ (Figure A.1). See Appendix A for details on the curve fitting. In particular, the distributions are the least uneven when 50% and 70% of the vehicles are OSM-routed and TT-routed, respectively (Figure 3).

The analysis of how the total CO2 emissions ($E$) varies with the percentage of $R$-routed vehicles also reveals a clear pattern: when either all vehicles are $R$-routed or none of them, the overall emissions are maximised (Figure 4b). In contrast, the overall emissions are minimised when just half of the vehicles are $R$-routed. In other words, $E_5^R < E_0^R$ and $E_7^R < E_{10}^R$. Note that the total emissions when all vehicles are TT-routed are much heavier than when all are OSM-routed (Figure 4b). This result suggests that TT’s routing algorithm recommends paths that generate a smaller adverse impact on urban well-being than OSM.

The fraction of $R$-routed vehicles also influence the spatial distribution of emissions in the city. In particular, comparing the two scenarios that maximise the total emissions (0% and 100% of $R$-routed vehicles) with the scenario that minimises them (50% of $R$-routed vehicles) helps understand where vehicles are being routed, consequently revealing emissions hot spots. In Figure 5a, we show the difference between the per-road emissions (normalized by the road length) when none of the vehicle is OSM-routed and 50% of them are (i.e., $E_0^\text{osm}(e) - E_5^\text{osm}(e)$, $\forall e \in E$). Similarly, Figure 5b shows the normalized emissions difference when all vehicles are OSM-routed and 50% of them are (i.e., $E_1^\text{osm}(e) - E_{10}^\text{osm}(e)$, $\forall e \in E$). We find that when 100% of vehicles are OSM-routed, the emissions are more concentrated towards Milan’s ring road (Figure 5b). In contrast, when none of them is OSM-routed, the emissions are more concentrated towards the city centre (Figure 5a). We find similar results for TT (Figure A.4).

Impact of randomization. We investigate the impact of DR’s randomisation parameter $w$ on the total emissions, their distribution over the road network, and the vehicles’ average travel time. For this purpose, we repeat the simulation of the urban traffic varying $w = 1, \ldots, 15$. We remind that the higher $w$, the more randomly perturbed DR’s fastest path is (see Appendix C).
Figure 4: Gini index of the CO\textsubscript{2} distribution (a) and total CO\textsubscript{2} emissions (b) varying the percentage of R-routed vehicles, for OSM (blue) and TT (red). In the error bars, points indicate the average Gini index (a) and the total CO\textsubscript{2} (b) over ten simulations with different choices of R-routed vehicles chosen uniformly at random. Vertical bars indicate the standard deviation.

Figure 5: The difference in the total CO\textsubscript{2} emitted on each road (in mg per meter of road) when: (a) none of the vehicles is OSM-routed and 50% of them are \((E^{\text{OSM}}_0(e) - E^{\text{OSM}}_5(e), \forall e \in E)\); (b) all vehicles are OSM-routed and 50% of them are \((E^{\text{OSM}}_{10}(e) - E^{\text{OSM}}_5(e), \forall e \in E)\). Red roads indicate a positive difference; blue ones indicate a negative difference.

First, changing \(w\) does not affect considerably the shape of the emission distributions: the Gini index still ranges in \(g_{\text{OSM}} \in [0.860, 0.879]\) and \(g_{\text{TT}} \in [0.858, 0.879]\) (Figure 6a, d) and the truncated power-law exponent in \(a_{\text{OSM}} \in [1.71, 1.94]\) and \(a_{\text{TT}} \in [1.71, 2.08]\). Second, the higher \(w\), the more even the distributions (Figure 6a, d), the lower the emissions (Figure 6b, e), and the shorter the average travel time (Figure 6c, f). Third, the total emissions are minimised when 40-60% of the vehicles are R-routed, and so are the average travel time and the distributions’ Gini index (Figure 6).

6 DISCUSSION

Our study provide several interesting results. We discuss and interpret them in the following paragraphs.
Navigation apps do impact on emissions. In general, CO₂ emissions are unevenly distributed across the city roads, and this unevenness is exacerbated when all vehicles or none of them are \( R \)-routed. In contrast, the total CO₂ emissions as well as their heterogeneity are minimised when around 50-70% of the vehicles are \( R \)-routed. These results clearly show that navigation apps do have a non-negligible adverse impact on the urban environment.

**Path perturbation is beneficial.** We also find that path perturbation (through DR’s \( w \) parameter) is beneficial: the more we randomise the vehicles’ paths, the shorter their travel time and the lower the amount of emissions in the city. This may be because the vehicles’ perturbed paths are more “diverse” (i.e., they spread over more roads), thus reducing congestion and consequently the amount of emissions and travel time. The role of path randomisation is interesting and deserves further investigation.

**Trips’ spatial distribution matters.** Urban planners and policymakers may be interested in investigating the impact of navigation apps on the spatial distribution of vehicles’ emissions. Our study shows that, in Milan, the more vehicles are \( R \)-routed, the less emissions concentrate in the city center and the more in the external ring road. Presumably, this may be because the navigation apps route the vehicles preferably on the city’s arterial roads (such as the ring road) to keep the individual path as fast as possible.

**Different navigation apps, different impact.** Our TraffiCO₂ simulation framework is a useful tool to compare the impact of different navigation apps on urban well-being. In our study, we find that TomTom is better than OpenStreetMap in minimising the emissions distribution’s inequality, and the city’s total emissions and average travel time. Although TomTom’s algorithm is not public, we may suppose that TomTom’s heuristics are more sophisticated, e.g., they use more or higher-quality information than OpenStreetMap, such as traffic data, detailed information about typical road speed, capacity, and length.

**A novel what-if analysis tool.** TraffiCO₂ is a simulation framework to compare the impact of different routing strategies on urban well-being in terms of amount of emissions and their spatial and...
7 CONCLUSION

Scenarios where either all vehicles or none of them follow a navigation app’s suggestion lead to the highest CO₂ emissions and the most uneven distribution of emissions per road. We show that when just a fraction of the vehicles (around half of them) follows the navigation app’s suggestions, such an adverse impact is minimised. This minimisation also holds when introducing more randomness in non-R-routed paths, which leads to a reduction of the vehicles’ average travel time, overall CO₂ emissions, and inequality in the distribution of CO₂ emissions across the road network.

We plan to improve and extend this study in several directions. First and foremost, we plan to extend the set of navigation apps considered (e.g., to Google Maps) and to study the environmental impact of a fleet of vehicles that use various navigation apps. In this work, we focus on light-duty vehicles only and assume that all vehicles carry the same engine type. Including heavy-duty vehicles (e.g., buses, trucks) and considering different vehicle engine ages and types (e.g., diesel, LPG, petrol) would provide a more complete mosaic of the emissions on the road network. We also look forward to apply our framework to various cities, to investigate how the impact of navigation apps varies with city size, shape, road network, and other characteristics. Moreover, we could study the impact of navigation apps in terms of other pollutants, such as nitrogen oxides, ozone, particulate matter, and volatile organic compounds, using emissions models similar to those used in this paper.

In the meantime, our work is a first step towards designing next-generation routing algorithms that, as our results suggest, should consider some degree of path randomisation to increase urban well-being while still satisfying individual needs.

ACKNOWLEDGMENTS

This work has been partially supported by EU project H2020 HumaneAI-net G.A. 952026, EU project H2020 SoBigData++ G.A. 871042, and EU project H2020 Humane-CHIST-ERA-19-XAI-010, by MUR (grant No. not yet available), FWF and Validating a Large-Scale Digital Twin of Barcelona’s Traffic with Empirical Data. ISPRS International Journal of Geo-Information 11, 1 (2022). https://doi.org/10.3390/ijgi11010024

REFERENCES

[1] Kyoungho Ahn and Hesham A. Rakha. 2013. Network-wide impacts of eco-routing strategies: A large-scale case study. Transportation Research Part D: Transport and Environment 25 (2013), 119–130. https://doi.org/10.1016/j.trd.2013.09.006

[2] Sabina Alazzawi, Mathias Hummel, Pascal Kordt, Thorsten Sickenberger, Christian Wieswietz, and Oliver Wohla. 2018. Simulating the Impact of Shared, Autonomous Vehicles on Urban Mobility – a Case Study of Milan. In 32MO2018: Simulating Autonomous and Intermodal Transport Systems (EPIC Series in Engineer- Vol. 2), Evamarie Weijsner, Leonhard Lückken, Robert Hillebrich, Yun-Pang Pf/ottenried, Jakob Ermann, Laura Birkert-Walzl, and Michael Behrisch (Eds.). EasyChair. 94–116. https://doi.org/10.29007/2w4h

[3] Jeff Alstott, Ed Bullmore, and Dietmar Plenz. 2014. poweRlaw: A Python Package for Analysis of Heavy-Tailed Distributions. PLOS ONE 9, 1 (2014), 1–11. https://doi.org/10.1371/journal.pone.0085877

[4] Javier Argota Sánchez-Vaquero. 2022. Getting Real: The Challenge of Building and Validating a Large-Scale Digital Twin of Barcelona’s Traffic with Empirical Data. ISPRS International Journal of Geo-Information 11, 1 (2022). https://doi.org/10.3390/ijgi11010024

[5] Neha Arora, Theophile Cabannes, Sanjay Ganapathy Subramaniam, Yechin Li, Preston McAfee, Marc Nunkesser, Carolina Osorio, Andrew Tomkins, and Ivel Tagoerus. 2021. Quantifying the sustainability impact of Google Maps: A case study of Salt Lake City.

[6] H. M. Abdul Aziz and Satish V. Ukkusuri. 2018. A novel approach to estimate emissions from large transportation networks: Hierarchical clustering-based link-driving schedules for EPA-MOVES using dynamic time warping measures. International Journal of Sustainable Transportation 12, 3 (2018), 192–204. https://doi.org/10.1080/15568318.2017.1346732

[7] Xuebei Bai, Richard Dawson, Diana Urge Vorsatz, Gian Delgado, Alyiu Salisu Barau, Shibhaikan Dhakal, David Dodman, Lykke Leonardsen, Valerie Masson-Delmotte, Debra Roberts, and Seth Schipper. 2018. Six Research Priorities for Cities and Climate Change. Nature 555 (02 2018). https://doi.org/10.1038/s41586-018-02409-z

[8] Hugo Barbosa, Marc Barthelemy, Gourab Ghoshal, Charlotte B James, Maxime Lencormand, Thomas Louail, Ronaldo Meneses, José J Ramasco, Filippo Simini, and Marcello Tomassini. 2018. Human mobility. Models and applications. Physics Reports 734 (2018), 1–74.

[9] Matthew Barth, Kanbor Boonoomsomsin, and Alex Vu. 2007. Environmentally-Friendly Navigation. In 2007 IEEE Intelligent Transportation Systems Conference, 684–689. https://doi.org/10.1109/ITSC.2007.4357672

[10] Philippe Besse, Brendan Guillouset, Jean-Michel Loubes, and Francois Royer. 2015. Review & Perspective for Distance Based Trajectory Clustering. (08 2015).

[11] Matteo Böhm, Mirco Nanni, and Luca Pappalardo. 2022. Gross polluters and vehicle emissions reduction. Nature Sustainability (2022). https://doi.org/10.1038/s41893-022-00930-y

[12] Christian Bonigjorno, Yulan Zhou, Marta Kryven, David Theurel, Alessandro Rizzo, Paolo Santi, Joshua Tenenbaum, and Carlo Ratti. 2021. Vector-based pedestrian navigation in cities. Nature Computational Science 1 (2021), 678–685. https://doi.org/10.1038/s43588-021-00130-y

[13] Robert Cervero and Kara Kockelman. 1997. Travel demand and the 3Ds: Density, diversity, and design. Transportation Research Part D: Transport and Environment 2, 3 (1997), 199–219. https://doi.org/10.1016/S1361-9200(97)00009-6

[14] Tim Chatterton, Jo Barnes, R. Eddie Wilson, Jillian Anable, and Sally Cairns. 2015. Use of a novel dataset to explore spatial and social variations in car type, size, usage and emissions. Transportation Research Part D: Transport and Environment 39 (2015), 151–164. https://doi.org/10.1016/j.trd.2015.06.003

[15] Jinyu Chen, Wenjing Li, Haoran Zhang, Wenxiao Jiang, Yi Sui, Xuan Song, and Ryouzuke Shibasaki. 2020. Mining urban sustainable performance: GPS data-based spatio-temporal analysis on on-road braking emission. Journal of Cleaner Production 270 (2020), 122489. https://doi.org/10.1016/j.jclepro.2020.122489

[16] Huan S. Chong, Sangji Kwon, Yunsung Lim, and Jongtae Lee. 2020. Real-world fuel consumption, gaseous pollutants, and CO₂ emissions of light-duty diesel vehicles. Sustainable Cities and Society 53 (2020), 101925. https://doi.org/10.1016/j.scs.2019.101925

[17] Arti Choudhary and Sharad Ghokhale. 2016. Urban real-world driving traffic emissions during interruption and congestion. Transportation Research Part D: Transport and Environment 43 (2016), 59–70. https://doi.org/10.1016/j.trd.2015.12.006

[18] Aaron Clauset, Cosmo Rohilla Shalizi, and M. E. J. Newman. 2009. Power-Law Distributions in Empirical Data. SIAM Rev. 51, 4 (2009), 661–703. https://doi.org/10.1137/070711011

[19] Antoine Coutrot, Ed Manley, S. Goodroe, C. Gahnström, Gabriele Filomena, Demet Yesiltepe, Ruth Dalton, Jan Wiener, C. Hölscher, Michael Hornberger, and Hugo Spiers. 2022. Entropy of city street networks linked to future spatial navigation ability. Nature 604 (04 2022). https://doi.org/10.1038/s41586-022-04486-7

[20] Koers De Baets, Sven Vlassenkroet, Kobe Boussauw, Dirk Lauwers, Georges Allaert, and Philippe De Maeyer. 2014. Route choice and residential environment: introducing liveability requirements in navigation systems in Flanders. Journal of Transport Geography 37 (2014), 19–27. https://doi.org/10.1016/j.jtrangeo.2014.04.005

[21] Priyanka deSouza, Amin Anjumshoah, Fabio Duarte, Ralph Kahn, Prashant Kumar, and Carlo Ratti. 2020. Air quality monitoring using mobile low-cost sensors mounted on trash-trucks: Methods development and lessons learned. Sustainable Cities and Society 60 (2020), 102239. https://doi.org/10.1016/j.scs.2020.102239

[22] Eva Ericsson, Hannah Larsson, and Karin Brunell-Freij. 2006. Optimizing route choice for lowest fuel consumption – Potential effects of a new driver support tool. Transportation Research Part C: Emerging Technologies 14, 6 (2006), 369–383. https://doi.org/10.1016/j.trc.2006.10.001
APPENDIX

A CURVE FITTING

We fit the same distributions using a maximum-likelihood \[3, 18\]. In particular, we fit five models to the data – power-law, truncated power-law, lognormal, exponential, and stretched exponential – and compare pairwise their goodness-of-fit with a log-likelihood ratio test. We choose the best fitting model as the one that wins the highest number of comparisons.

The best fit for the distributions is always a truncated power-law (essentially, a power-law with an exponential cutoff), with probability density function \( p(x) \propto x^{-\alpha} e^{-\lambda x} \). Figure A.1 shows how the truncated power-law’s exponent \( \alpha \) varies with \( D_i^{(R)} \), \( i = 0, \ldots, 10 \) and \( R \in \{ \text{OSM}, \text{TT} \} \).

![Figure A.1: Exponent \( \alpha \) of the truncated power-law fit to the CO\(_2\) emissions distribution across the roads varying the percentage of R-routed vehicles, for OSM (blue) and TT (red). In the error bars, points indicate the average \( \alpha \) over ten simulations with different R-routed vehicles chosen uniformly at random. Vertical bars indicate the standard deviation.](image)

B TRAFFIC CALIBRATION

To generate a plausible mobility demand reflecting real traffic patterns, we tune the number \( N \) of vehicles in the simulation, the value of the parameter \( w \in [1, +\infty) \) for DR, and the quantity \( c \) of extra vehicles to insert during the simulation, where \( c = \text{none} \) means no extra vehicles, \( x_{\text{start}} \) means to add \( N \times x_{\text{start}} \) extra vehicles at the start of the simulation; similarly, \( x_{\text{end}} \) means to add \( N \times x_{\text{end}} \) extra vehicles after the last trip’s departure time (see Table AT.1 for details).

To assess the realism of the mobility demand \( D \), we generate a multiset of routed paths \( D_i^{(R)} \) (no R-routed vehicles) and simulate the vehicular traffic using SUMO. Then, we use the Jensen–Shannon (JS) divergence \[39\] to compute the distance between the travel time distribution in real data and that of the simulated vehicles (see Figure A.2). The JS divergence ranges in \([0, 1]\) (the higher, the more similar the two distributions are) and is defined as \[39\]:

\[
JS(P||Q) = \frac{1}{2}KL(P||M) + \frac{1}{2}KL(Q||M)
\]

where \( P \) and \( Q \) are two density distributions, \( M = \frac{1}{2}(P + Q) \), and \( KL \) is the Kullback–Leibler divergence (KL), defined as:

\[
KL(P||Q) = \sum_{x \in X} P(x) \log \left( \frac{P(x)}{Q(x)} \right)
\]

We also assess the realism of the simulation using the absolute difference in seconds between the real and simulated average travel time \( |\Delta t| \) and the number of teleports during the simulation. SUMO makes a teleport whenever a vehicle waits too long stuck in gridlock: the vehicle is teleported onto the next free edge on its path. We simulate each \( D_0^{(R)} \) five times and consider the average values of the above three metrics.

We run several configurations \( C_{(N, w, c)} \) composing all combinations of \( N, w, \) and \( c \) in Table AT.1. We select configuration \( C_{(15k, 5, \text{none})} \), as it is the best on two out three of the evaluation criteria: the JS difference wrt the minimum is only of \( 0.002 \), \( |\Delta t| \approx 25s, 3298 \) teleports, and \( w = 5 \) (see Table AT.2).

C PERTURBATION OF THE FASTEST PATH

To model the driving behavior of vehicles that are not R-routed (i.e., they do not follow any navigation apps’ suggestion) and the imperfection and non-rationality of human drivers \[53\], we perturb the fastest path between an origin and destination pair. DR allows perturbing the fastest path using a randomisation parameter \( w \in [1, +\infty) \), where \( w = 1 \) means no randomisation (i.e., the fastest path), and the higher \( w \), the more randomly perturbed the fastest path is. To confirm that the perturbation of a path from an origin to a destination grows with \( w \), we take randomly 15,000 origin–destination pairs computing the path suggested by DR for different
values of $w \in \{x \in \mathbb{N}|1 \leq x \leq 20\}$. Next, we measure the perturbation of a path as the Symmetrized Segment-Path Distance (SSPD) [10] between the perturbated paths ($w > 1$) and the fastest path ($w = 1$). Figure A.3 shows that increasing the value of $w$ results in paths that have a greater average SSPD with respect to the fastest path, and hence that the perturbation grows with increasing $w$.

Figure A.3: The average Symmetrized Segment-Path Distance (SSPD) computed between 15k perturbated paths (for different values of $w \in \{x \in \mathbb{N}|2 \leq x \leq 20\}$) and the fastest path ($w = 1$). The higher $w$, the higher the a path’s perturbation.

**D SPATIAL DISTRIBUTION OF EMISSIONS**

In Figure A.4a, we show the difference between the per-road emissions (normalized by the road length) when none of the vehicle is TT-routed and when 50% of them are (i.e., $E^{(TT)}_0(e) - E^{(TT)}_5(e)$, $\forall e \in E$). Similarly, Figure A.4b shows the normalized emissions difference when all vehicles are TT-routed and 50% of them are (i.e., $E^{(TT)}_10(e) - E^{(TT)}_5(e)$, $\forall e \in E$). The results are the same as for OSM: when 100% of vehicles are tt-routed, the emissions are more concentrated towards Milan’s ring road (Figure A.4b). In contrast, when none of them is TT-routed, the emissions are more concentrated towards the city centre (Figure A.4a).

Figure A.4: The difference in the total CO$_2$ emitted on each road (in mg per meter of road) when: (a) none of the vehicles is TT-routed and 50% of them are $(E^{(TT)}_0(e) - E^{(TT)}_5(e), \forall e \in E)$; (b) all vehicles are TT-routed and 50% of them are $(E^{(TT)}_10(e) - E^{(TT)}_5(e), \forall e \in E)$. Red roads indicate a positive difference; blue ones indicate a negative one.