SUMO User Conference 2026

TIBconferencesession

TIB-OP_will_set_DOI

Published: 1970-01-01

Integrating Synthetic Populations and Activity Chains for Individual Emission Assessment in SUMO

Alix Ngari Lendoye[1,2], Corwin Fèvre[1], Tatiana Graindorge[1], and Alain Bouju[2]

[1]EIGSI La Rochelle, France

[2]La Rochelle University, L3i Laboratory, France

*Correspondence: Alix Ngari Lendoye, am.ngarilendoye@eigsi.fr

Abstract. Assessing the environmental impacts of mobility in peri-urban areas requires individual-level emission indicators that link travel practices to vehicle characteristics. While microscopic traffic simulators such as SUMO offer high representational accuracy, they lack realistic mobility demand as input. This paper proposes a reproducible workflow coupling synthetic demand generation with microscopic simulation to estimate individual emissions in peri-urban contexts. The approach integrates three components: Eqasim for synthetic population generation and activity-chain assignment; MATSim for plan consolidation and mobility-choice refinement; and SUMO for microscopic simulation, with activity locations anchored to the road network via points of interest and public transport integrated through GTFS feeds. Fleet composition is derived from local open data, translated into HBEFA4 emission classes, and distributed across the simulated population to improve emission realism. The workflow is demonstrated through a case study on eight peri-urban municipalities surrounding La Rochelle (Charente-Maritime, France). Its modular architecture and reliance on open data and open-source tools support reproducibility and transferability to other peri-urban territories.

Keywords: Activity-Based Demand Generation, Individual Emission Estimation, Synthetic Population

1. Introduction

Peri-urban areas concentrate a significant share of contemporary mobility challenges: extended home-to-work distances, limited public transport supply, and consequently high levels of car dependence [1]. In this context, assessing environmental impacts on a fine spatial and behavioral scale, linking emissions to individual travel practices and vehicle characteristics, is of particular relevance. This assessment can leverage agent-based modeling, microscopic traffic simulation, and emission models. However, producing credible individual-level indicators requires an accurate representation of the territory's population, its socio-demographic attributes, and the daily activity chains that structure mobility demand.

SUMO enables detailed microscopic representation of vehicle movements on a road network [2], but it is not designed to generate autonomously realistic descriptions of daily schedules and trips of residents [3]. As a result, even though traffic dynamics can be simulated with high fidelity, building a complete and consistent demand model that specifies who travels, when, for what purpose, and between which locations remains a methodological bottleneck, especially when individual-level interpretability is required.

This paper addresses precisely this gap by proposing an estimation framework for individual emissions in peri-urban contexts, enriching SUMO with an adapted demand-generation process. The pipeline combines three complementary components: Eqasim for synthetic population generation and activity-chain assignment from publicly available data; MATSim for plan consolidation and mobility-choice refinement within a multi-agent simulation framework; and SUMO for microscopic simulation, where the assignment of vehicle emission classes enables the computation of individual-level environmental indicators. The pipeline is applied to the La Rochelle Agglomeration Community (Charente-Maritime, France) and its peri-urban fringe, a context well suited to examining the relationship between car dependence, daily activity structures, and individual emissions.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 details the methodology, from demand generation to microscopic simulation and emission-class assignment. Section 4 presents the case study and implementation choices for the La Rochelle peri-urban area. Section 5 presents the resulting indicators for the target population. Section 6 discusses limitations and perspectives, and Section 7 concludes.

2. Related Work

Microscopic simulators such as SUMO offer fine-grained representations of traffic dynamics and are well suited to producing the detailed trajectory-level outputs required for emission estimation [2], [4]. However, constructing realistic mobility demand is a distinct methodological challenge. It involves specifying who travels, when, why, between which locations, and with which activity chains, and is typically addressed through dedicated demand models or synthetic population pipelines [5], [6].

SUMO provides several families of demand-generation tools, including random or heuristic generators (*randomTrips*, *OSMWebWizard*), count- or OD-matrix-based approaches (*od2trips*), and the activity-based generator *ActivityGen* [7]. From a scenario-engineering perspective, these tools lower the barrier to obtaining a first functional scenario, an explicit design objective of the SUMO ecosystem [4]. Nevertheless, ActivityGen supports only a limited set of activity types (work, school, leisure) and transport modes, and produces mobility wishes rather than fully specified activity chains in the sense of activity-based modeling (ABM) [8]. The SUMOPy demand generator recognizes that its virtual population can be rigid and that integrating OD patterns or survey data remains difficult. It positions itself as a step toward more flexible plans, while recognizing the difficulty of scaling to 24-hour multi-modal simulations [9]. More recently, SAGA (SUMO Activity GenerAtion) has further automated the construction of multi-modal scenarios from open data (notably OpenStreetMap), providing an end-to-end reproducible tool-chain [10]. Although SAGA represents a significant step toward accessible multi-modal scenario creation, its objective remains largely pragmatic: to produce plausible, heuristic-based scenarios rapidly rather than to implement the full behavioral

mechanisms of a state-of-the-practice ABM (hierarchical choices, intra-household coordination, endogenous location choice, tour formation, etc.). Thus, SAGA and related generators partially bridge the gap between demand based on OD/trips and activity, but do not substitute for ABM approaches when behavioral realism is required at the individual level [5].

Activity-Based Models (ABMs) were specifically designed to represent travel demand as the outcome of a temporally and spatially constrained activity program, with interdependencies across decision dimensions [5], [11]. Methodological reviews emphasize that these models provide an explicit representation of time constraints, activity-trip linkages, and intra-household interactions, elements that are difficult to reconstruct with conventional trip-generation approaches [5], [12]. ActivitySim is a prominent open-source implementation of this paradigm, adopted by multiple metropolitan planning organizations in the United States [13], [14]. Its hierarchical decision structure is designed to produce behaviorally consistent individual-level travel diaries from a synthetic population. It proceeds from long-term choices to daily scheduling, tour formation, and mode and time-of-day choice [15]. SimMobility represents another key reference, with a multi-scale architecture integrating short-term microscopic dynamics, mid-term activity scheduling, and long-term land-use interactions within a unified agent-based platform [16]. Together, these frameworks illustrate that microscopic traffic simulators (SUMO, Aimsun, Vissim) excel at representing network dynamics but, for individual-level analyzes, must be supplied with externally generated demand. This observation has motivated substantial research on demand–supply coupling: pairing an ABM or synthetic population pipeline with a traffic assignment or microscopic simulation model to combine behavioral realism with network fidelity. Recent work has formalized this approach through ActivitySim–MATSim co-simulation frameworks, treating demand scheduling and traffic dynamics as complementary but distinct components [17], [18]. Beyond travel demand realism, the environmental evaluation of such simulated scenarios raises additional methodological choices (e.g., fleet composition, emission factors, and occupancy assumptions), which can strongly affect results.

Our work follows this separation between demand construction and microscopic simulation. Eqasim and MATSim are used upstream to construct and refine the synthetic population, activity chains, schedules, and travel choices, while SUMO is used downstream for microscopic execution and emission estimation. Building on our previous review of environmental assessment methods for shared autonomous vehicles[19], this paper applies this coupling to a real peri-urban case study and provides a transparent workflow for producing individual-level emission indicators.

3. Methodology

To assess the environmental consequences of mobility behaviors in peri-urban areas, we implement a processing chain spanning from demand synthesis to microscopic simulation. The pipeline comprises three successive stages: synthetic population generation and activity-chain assignment using Eqasim [20]; multi-agent simulation and mode-choice refinement with MATSim [6]; and microscopic simulation in SUMO [4], including multimodal route conversion and vehicle fleet characterization. Figure 1 provides an overview of the complete workflow.

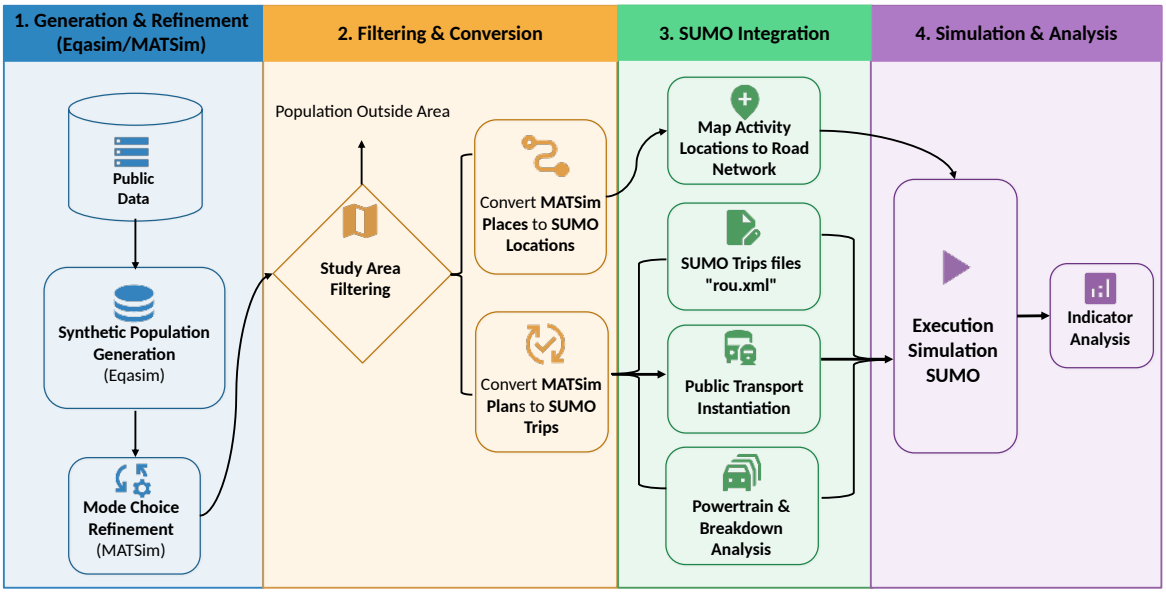

Figure 1. *Overview of the proposed end-to-end workflow for individual-level emission estimation.*

3.1 Mobility Demand Generation with Eqasim

The first stage consists of generating realistic individual-level mobility demand from publicly available data. We use Eqasim [21], a reproducible open-source framework designed to generate synthetic populations and associated travel demand for agent-based transport simulation. The pipeline integrates four processing blocks: socio-demographic synthesis, activity-chain assignment, activity location, and trip generation. Internal mechanisms are described in detail in [20]; the present section focuses on adaptation choices relevant to a generic peri-urban deployment.

3.1.1 Input Data

Eqasim combines five families of public datasets to build the synthetic travel demand:

- **Population census**: provides the micro-sample of households and individuals, with socio-demographic attributes (age, sex, employment status, car ownership, etc.) used to seed the synthetic population.
- **Origin–destination flow data**: enables the assignment of workplace and study municipalities for each synthetic individual.
- **Income data**: used to attach a household income level consistent with the municipal distribution.
- **Household Travel Survey (HTS)**: supplies activity sequences, schedules, and modal practices, matched to synthetic individuals by socio-demographic profile.
- **Spatial opportunity datasets**: geo-referenced databases of addresses, establishments, and facilities used to locate activities explicitly in space.

The availability and format of these datasets vary across national contexts. The pipeline accommodates this variability through a configurable data-loading layer.

3.1.2 Synthetic Population

The synthetic population determines who travels, where individuals reside, and their socio-demographic attributes. Following [20], the census provides a micro-sample of households associated with a statistical weight w_i indicating how many real households each observation represents. The population is generated by direct sampling: each household is replicated proportionally to w_i. Since w_i is not generally an integer, Eqasim applies stochastic rounding to convert weights into integer multipliers m_i, preserving the correct statistical expectation while producing a discrete agent population. The procedure is deterministic once a random seed is fixed.

Running a full-scale simulation is computationally expensive. A sampling rate is therefore applied to reduce the number of agents while maintaining acceptable statistical representativeness.

The census is structured at the household level; Eqasim preserves household–person links throughout, avoiding the artificial disaggregation of family structures. Synthetic households are associated with a residential spatial unit (sub-municipal zone or municipality) in accordance with the geographic precision available in the census data, which varies according to the statistical disclosure rules of the national statistical office.

3.1.3 Activity Chains

Once the synthetic population is constructed, each agent is then assigned a daily activity chain describing activity types (e.g., home–work–home, home–shopping–leisure–home), schedules, and transport modes. This assignment relies on the Household Travel Survey (HTS), which provides behavioural information on day structures, temporal profiles, and modal practices rather than fine spatial precision.

Survey days are transferred to synthetic agents based on socio-demographic proximity (age, sex, socio-professional category, etc.), producing an activity chain that constitutes a temporal and behavioral skeleton of the day. Following [20], five modal classes are retained (driver, passenger, public transport, bicycle, walking) and six activity categories (home, work, education, leisure, shopping, other). The transfer also enriches the synthetic population with variables not systematically present in the census, such as driving license ownership or public transport subscription. The precise spatial location of activities is addressed in a subsequent step, distinguishing primary activities (home, work, education) from secondary ones using spatial opportunity datasets.

3.1.4 Eqasim Outputs

At the end of its execution, Eqasim produces a mobility demand in a standardized tabular structure, usable directly for analysis or as input to MATSim. The main output files are described in Table 1.

3.2 MATSim Simulation and Mode-Choice Refinement

The Eqasim outputs describe an initial demand. Simulation is then performed with MATSim [6], used here as a multi-agent simulation engine providing network assignment and

Table 1. *Main Eqasim output files and their role in the workflow.*

File	Content	Role in the workflow
`meta.json`	Generation metadata (seed, sampling rate, date)	Traceability and reproducibility
`households.csv`	Household attributes (car/bicycle availability, income, etc.)	Socio-demographic baseline
`persons.csv`	Individual attributes (age, sex, employment, driving licence, etc.)	Agent reference
`activities.csv`	Activities per person (type, start/end time, location)	Activity schedules and locations
`trips.csv`	Trips between activities (times, mode, purposes)	Demand description
`activities.gpkg`	Georeferenced activities (points)	GIS checks and spatial diagnostics
`trips.gpkg`	Georeferenced trips (lines)	GIS checks and spatial diagnostics
`homes.gpkg`	Home locations (linked to households)	Spatial distribution of residences

agent interaction. Within the Eqasim framework, MATSim handles mobility execution and network interactions while Eqasim provides integration tooling and modules [21].

The demand is converted into a MATSim scenario and simulated iteratively. At convergence, travel times and traffic conditions are consistent with network assignment, and trips are made compatible with capacity constraints and agent interactions, a prerequisite for conversion to SUMO. Mode choice is handled through the Discrete Mode Choice (DMC) module [22], called during MATSim's replanning phase. We rely on the DMC implementation distributed with the Eqasim framework, which allows the mode of a subset of agents to be re-evaluated at each iteration based on trip attributes produced by MATSim routing (in-vehicle time, waiting time, transfers, access/egress, costs). At the end of this stage, stabilized plans contain, for each agent, an activity chain with modes and schedules consistent with the simulation.

3.3 Microscopic Simulation with SUMO

The final stage converts the stabilized MATSim plans into a microscopic SUMO simulation [4], yielding fine-grained trajectory outputs and an operational framework for computing individual-level and aggregate environmental indicators. In the current implementation, the coupling is one-way: SUMO is used for microscopic traffic execution and emission estimation, but the resulting travel times are not fed back into the upstream demand-generation or MATSim mode-choice stages.

3.3.1 Network Construction

The SUMO road network is built from OpenStreetMap data using `netconvert`, covering the study area defined by the analyst. The network extraction should encompass the full spatial extent of the target population's activity space, including a buffer to limit boundary effects.

3.3.2 Post-Processing and Spatial Filtering of MATSim Outputs

After MATSim simulation, outputs contain all agents and trips generated at the full simulation extent. Spatial filtering is applied to restrict the dataset to the study area while preserving a coherent mobility context. Two complementary populations are retained:

- **Target population**: residents of the peri-urban municipalities of interest.
- **System population**: individuals whose complete set of activities falls within the agglomeration perimeter, to maintain realism in network interactions.

The final retained population is the union of these two sets. Filtering is applied consistently to all tabular outputs (CSV) and geographic files (GeoPackage), as well as to the MATSim population XML, to ensure downstream consistency.

3.3.3 Extraction and Anchoring of Activity Locations

Converting MATSim plans to SUMO requires resolving activity locations into network positions. In MATSim, activities reference facilities defined in a dedicated file; SUMO does not natively handle this concept. A set of Points of Interest (POIs) is therefore constructed from the facilities used by the filtered population, and each POI is anchored to the road network via mode-compatible edges.

Facility coordinates, expressed in the source coordinate reference system, are transformed to WGS84 (EPSG:4326). For a facility i with Lambert-93 coordinates (x_i, y_i):

$$(\lambda_i, \phi_i)_{\text{WGS84}} = \mathcal{T}_{2154 \to 4326}(x_i, y_i) \tag{1}$$

where $\mathcal{T}$ denotes the geodetic transformation (implemented using `pyproj`).

For each POI, a multimodal network anchoring is computed: one edge compatible with pedestrian traffic (`vClass=pedestrian`) and one compatible with car traffic (`vClass=passenger`). The nearest compatible edge is found by iterative radius search using geometric progression:

$$r_0 = 20\,\text{m}, \quad r_{k+1} = 2\,r_k, \quad r_k \leq 4000\,\text{m} \tag{2}$$

Candidate edges are filtered to exclude internal edges, edges shorter than 5 m, and edges that do not allow the required vehicle class. A local connectivity check further ensures that selected edges have both incoming and outgoing connections within a search depth of two hops, avoiding assignment to isolated or disconnected network fragments. The POI coordinate is then projected onto the selected lane geometry to obtain a curvilinear position, slightly offset from the lane endpoints to prevent initialization artifacts. Results are exported as a tabular file (`facilities2sumo_multimode.csv`) providing, for each POI, the edge, lane, position, and anchoring distance by mode.

3.3.4 Integration of Public Transport

Public transport is integrated using GTFS data, processed with the standard SUMO tool-chain:

- `netconvert` generates the road network (`*.net.xml`) and a candidate stop file (`busstop.add.xml`) via the `--ptstop-output` option.
- `gtfs2pt.py` imports the GTFS feed, reconstructing line geometries by shortest-path routing between stops when no explicit shape is available, and producing a PT vehicle file (`*.rou.xml`) and an additional file describing stops and routes.

Because GTFS encodes calendar-dependent service, a simulation date must be specified (`--date YYYYMMDD`) to instantiate the services active on the chosen day. A weekday is recommended to represent a standard service level (regular peak frequencies), avoiding weekend or holiday configurations.

3.3.5 Conversion of MATSim Plans into Multimodal SUMO Trips

The core conversion step transforms MATSim activity chains (activities and legs) into a `*.rou.xml` file directly executable in SUMO. Each agent is described as a `<person>` element containing `<stop>` elements (activities) and inter-activity segments (`<walk>`, `<personTrip>`, `<ride>`). Technical MATSim activities (e.g., `interaction`) are filtered out; consecutive activities referencing the same facility are merged.

Trips are summarized by dominant mode following a priority rule: car > public transport > bicycle > walking. Before writing each segment, a network feasibility check is performed via shortest-path computation on the SUMO network; agents for whom no valid path exists are excluded and logged rather than producing an inconsistent route file. Mode-specific segment generation proceeds as follows:

- **Walking** (`<walk>`): anchored on pedestrian edges; a short adjustment walk is prepended if the activity anchor differs from the departure edge.
- **Car** (`<personTrip modes="car">`): anchored on passenger edges; driver/passenger roles are preserved when available in the MATSim plan.
- **Bicycle** (`<personTrip modes="bicycle">`): generated from the pedestrian anchors of the origin and destination facilities, while route feasibility is checked on the SUMO network using (`vClass=pedestrian`). If no bicycle-compatible path exists between these anchors, the corresponding agent is excluded and logged.
- **Public transport** (`<ride>`): reconstructed as an access walk + in-vehicle ride + egress walk sequence, consistent with the GTFS-based supply imported into SUMO. Walking time is estimated as:

$$t_{\text{walk}} = \frac{d_{\text{eucl}}}{v_{\text{walk}}} \times \alpha \tag{3}$$

with $v_{\text{walk}} = 1.3$ m/s and $\alpha = 1.30$ (detour correction factor). Vehicle departures are selected by binary search to satisfy:

$$d_0 \geq t_a - \Delta_b \tag{4}$$

where t_a is the agent's arrival time at the boarding stop and Δ_b is the GTFS time offset of that stop relative to the trip's first stop. When transfers are required, two `<ride>` segments are chained, with the transfer point determined by spatial proximity between stops. A lightweight pedestrian connectivity check (breadth-first search) is applied to reduce inconsistencies in discontinuous pedestrian graphs.

Conversion failures (unmapped POIs, missing routes, unknown edges) are recorded in a log file with status codes (`ok`, `skipped_nan`, `skipped_no_route`, etc.), enabling transparent quantification of agent losses.

3.3.6 Vehicle Fleet Characterization and Emission-Class Assignment (HBEFA4)

To ground emission estimates in local conditions, the passenger car fleet composition is derived from publicly available municipal vehicle registration data. Fleet counts are aggregated by fuel type and environmental classification category (or equivalent national proxy), and each combination is mapped to an HBEFA4 [23] emission class using an explicit correspondence table. When the mapping is ambiguous, a deterministic assumption (e.g., most recent applicable standard) is applied to ensure reproducibility.

The share of each emission class k in the fleet is computed as:

$$p_k = \frac{N_k}{\sum_j N_j} \times 100 \tag{5}$$

The expected number of simulated individuals assigned to class k is:

$$q_k = N \times \frac{p_k}{100} \tag{6}$$

where N is the number of individuals performing at least one car trip in the route file. Allocation is performed using the largest-remainder method to minimize deviation from target proportions, with deterministic ordering (fixed random seed and sorted person identifiers) to ensure reproducibility. The selected emission class is applied to all car trips of the individual, and corresponding `<vType emissionClass="HBEFA4/..."/>` definitions are injected into the route file.

4. Application to the La Rochelle Peri-Urban Area

The methodology is applied to a case study centered on the La Rochelle Agglomeration Community (Communauté d'Agglomération, CdA), an inter-municipal authority grouping 28 municipalities around the city of La Rochelle, on the Atlantic coast of southwestern France, in Charente-Maritime (a French department, i.e., a county-level administrative unit). This section documents the datasets, parameters, and implementation choices used for this context.

4.1 Study Area and Network

The SUMO network is extracted from OpenStreetMap using the following bounding box:

```
bbox = (-1.571, 45.800, -0.533, 46.353)
```

This extent covers approximately 80×60 km (west–east $\times$ south–north), encompassing the La Rochelle CdA and its main surroundings. The CdA comprises 28 municipalities over 327 km^2, with 181,057 inhabitants in 2022. It combines a dense urban core with a dispersed peri-urban fringe, making it a relevant context for analyzing car-dependent mobility. The environmental analysis focuses specifically on residents of eight peri-urban municipalities in this fringe: Bourgneuf, Montroy, Clavette, La Jarrie, Saint-Médard-d'Aunis, Saint-Christophe, Croix-Chapeau, and Salles-sur-Mer.

4.2 Input Data for Charente-Maritime

The Eqasim pipeline is applied at the scale of the Charente-Maritime department (Department 17). In the French context, most socio-demographic and spatial statistical data used in the pipeline are provided by INSEE, the French National Institute of Statistics and Economic Studies. Table 2 details the data sources used.

Table 2. *Public datasets used in the Eqasim pipeline for Charente-Maritime (Department 17).*

Category	Source	Role in pipeline	Application (Dept. 17)
Socio-demographic data	Population Census (INSEE)	Synthetic household/person generation	Residents of Dept. 17
OD flows	INSEE (commuting matrix)	Workplace/school municipality assignment	Flows involving Dept. 17
Income data	Filosofi (INSEE)	Household income assignment	Municipal distributions
Mobility survey	EMP 2019	Activity chains and modal practices	National profiles applied locally
Spatial opportunities	BPE, SIRENE, IRIS, BD TOPO, BAN	Activity location	Objects in Dept. 17

4.3 Synthetic Population: Sampling and Spatial Resolution

Charente-Maritime had 668,160 inhabitants in 2022. To limit computational cost while retaining statistical representativeness, a sampling rate is applied. Since our analysis focuses on the eight peri-urban municipalities (about $N \approx 15{,}000$ inhabitants), we estimated a target sampling rate for this sub-area using the classical formula for proportion estimation with finite-population correction [24], setting $p = 0.5$ (maximum variance):

$$n_0 = \frac{Z^2 \cdot p(1-p)}{e^2} \tag{7}$$

$$n = \frac{n_0}{1 + \dfrac{n_0 - 1}{N}} \tag{8}$$

where N is the total population size, Z is the critical value associated with the chosen confidence level, e is the desired margin of error (expressed as a proportion), and p is set to 0.5. Table 3 reports required sample sizes for various combinations of confidence level and margin of error, computed using Equations (7)–(8).

Table 3. *Required sample sizes by confidence level and margin of error ($N = 15{,}000,\ p = 0.5$).*

Confidence level	Margin of error (e)	Sample size (n)	Sampling rate (n/N) %
95%	5%	374	2.49%
95%	3%	996	6.64%
99%	5%	637	4.25%
99%	3%	1,646	10.97%

For a 99% confidence level and a 3% margin of error, the required sample size is $n \approx 1{,}646$, i.e., about 11% of the target population. Accordingly, we generate the synthetic population by sampling 11% of individuals and then report the effective share retained in the eight municipalities after filtering.

Charente-Maritime presents a heterogeneous spatial structure: 463 municipalities, of which only 15 are subdivided into IRIS units (Îlots Regroupés pour l'Information Statistique), i.e., French sub-municipal statistical zones defined by INSEE for the dissemination of local socio-demographic data. In total, the department contains 94 IRIS units. For the majority of the territory, low-density rural municipalities, the geographic precision available in the census is at the municipal level. We follow the operational rule of the reference Eqasim pipeline [20]: IRIS when available, municipality otherwise. When only departmental information is present (typically for very small municipalities), a municipality is sampled with probability proportional to population totals, and an IRIS is then sampled within that municipality.

4.4 Activity Chains: EMP 2019

Activity chains are derived from the *Enquête Mobilité des Personnes* (EMP 2019), a national household travel survey conducted between May 2018 and April 2019 among approximately 20,000 households in metropolitan France. EMP 2019 was selected because it predates the COVID-19 pandemic, providing behavioral patterns representative of pre-disruption travel practices. Applying a national survey to a specific department rests on the principle that the HTS is not expected to provide fine spatial representativeness, but rather to supply robust behavioral distributions (day structures, temporal profiles, modal practices) applicable to synthetic agents matched by socio-demographic profile.

4.5 MATSim Configuration

The demand generated for Charente-Maritime is converted into a MATSim scenario and simulated for several iterations, at which point travel times and modal shares are considered stabilized. The DMC module [22] is used without additional re-estimation of utility parameters, relying on the default Eqasim calibration.

4.6 Spatial Filtering

After MATSim, filtering retains: (i) residents of the eight target peri-urban municipalities (target population, $n = 1{,}647$); and (ii) individuals whose activities all fall within the 28 CdA municipalities (system population, $n = 17{,}096$). In total, the filtered dataset contains $n = 17{,}478$ individuals. The filtering is applied to `17households.csv`, `17persons.csv`, `17activities.csv`, `17trips.csv`, `17pt_legs.csv`, their GeoPackage equivalents, and `17population.xml.gz`.

4.7 Public Transport: Yélo Network

Public transport is represented using GTFS data from the Yélo network (La Rochelle agglomeration), retrieved from the national open data portal [25]. The network comprises regular urban routes (including high-frequency bus lines) and a night service (routes N1–N3, operating on Fridays and Saturdays from approximately 22:00 to 01:00, complemented by on-demand night service). A weekday (Tuesday, March 17, 2026) is selected as the simulation date to instantiate a standard service level and exclude night and weekend-specific configurations.

4.8 Conversion of the Retained Population to SUMO

The retained population, obtained after spatial filtering, is converted into SUMO person plans using the procedure described in Section 3.3.5. SUMO simulations are conducted on the filtered synthetic sample only, not on the full CdA population. Therefore, the results describe the simulated sample rather than a full-scale reconstruction of traffic conditions in the agglomeration. During conversion, 16,952 out of 17,478 individuals were successfully written to the SUMO route file, corresponding to a conversion rate of 97%. The remaining 526 agents were excluded because no valid path could be found or because of unresolved network/PT mapping issues. These exclusions are logged by cause to quantify conversion losses and ensure traceability. Importantly, all 1,647 individuals belonging to the target population were successfully converted.

4.9 Vehicle Fleet and Emission Classes

The local fleet composition is derived from the *Données sur le parc de véhicules au niveau communal* dataset [26], filtered to retain passenger cars (*Véhicule Particulier*) registered in the 28 CdA municipalities as of January 1, 2022. The French Crit'Air classification [27] serves as the operational proxy for inferring Euro standard, as it encodes both powertrain type and emission standard in a publicly available sticker category (Crit'Air 0 to Crit'Air 5). The mapping from (fuel type, Crit'Air) to HBEFA4 [23] emission class is detailed in Table 4.

Table 4. *Mapping from fuel type and Crit'Air category to SUMO HBEFA4 emission classes.*

Fuel	Crit'Air	SUMO emission class	Model
Diesel	Crit'Air 2	`HBEFA4/PC_diesel_Euro-6ab`	HBEFA4
Diesel	Crit'Air 3	`HBEFA4/PC_diesel_Euro-4`	HBEFA4
Diesel	Crit'Air 4	`HBEFA4/PC_diesel_Euro-3`	HBEFA4
Diesel	Crit'Air 5	`HBEFA4/PC_diesel_Euro-2`	HBEFA4
Diesel	Unclassified	`HBEFA4/PC_diesel_Euro-1`	HBEFA4
Diesel HNR	Crit'Air 2	`HBEFA4/PC_diesel_Euro-6ab`	HBEFA4
Diesel HR	Crit'Air 1	`HBEFA4/PC_PHEV_diesel_Euro-6ab`	HBEFA4
Electric	Crit'Air E	`HBEFA4/PC_BEV`	HBEFA4
Petrol	Crit'Air 1	`HBEFA4/PC_petrol_Euro-6ab`	HBEFA4
Petrol	Crit'Air 2	`HBEFA4/PC_petrol_Euro-4`	HBEFA4
Petrol	Crit'Air 3	`HBEFA4/PC_petrol_Euro-3`	HBEFA4
Petrol	Unclassified	`HBEFA4/PC_petrol_Euro-1`	HBEFA4
Petrol HNR	Crit'Air 1	`HBEFA4/PC_petrol_Euro-6ab`	HBEFA4
Petrol HNR	Crit'Air 2	`HBEFA4/PC_petrol_Euro-4`	HBEFA4
Petrol HR	Crit'Air 1	`HBEFA4/PC_PHEV_petrol_Euro-6ab`	HBEFA4
Gas	Crit'Air 1	`HBEFA4/PC_LPG_petrol_Euro-6_(LPG)`	HBEFA4
Gas	Crit'Air 3	`HBEFA4/PC_LPG_petrol_Euro-3_(LPG)`	HBEFA4

When the Crit'Air–Euro correspondence is ambiguous, an explicit and documented assumption is applied (most recent applicable standard) to maintain determinism. For example, diesel plug-in hybrids (diesel HR) are assigned to Crit'Air 1, following the Crit'Air rule for all plug-in hybrid vehicles, whereas non-rechargeable diesel hybrids (diesel HNR) are mapped according to the diesel Euro-standard rule. Fleet proportions p_k are computed using Equation (5) and assigned to simulated car users using the largest-remainder method (Equation (6)), with a fixed random seed to ensure reproducibility.

5. Results

To obtain emissions at the individual level, an allocation rule must be defined, since SUMO computes emissions at the vehicle level. For each car trip, emissions reported directly by SUMO in the `tripinfo.xml` file (CO_2, PM_x, etc.) are extracted and linked to a person using the identifier scheme adopted in the simulation, whereby private vehicles carry the person's identifier. In the analysis, we restrict emission accounting to individuals **from the target population** who perform at least one car trip, yielding $n = 1{,}293$ persons.

Allocation is based on the role associated with each car segment: when an individual is a driver (`role="car"`), 100% of the trip emissions are attributed to them; when they are a passenger (`role="car_passenger"`), 50% are attributed to them, under the simplifying assumption of a two-occupant vehicle with equal sharing. Formally, for a trip j with total emissions E_j and an allocation factor $f_j \in \{1, 0.5\}$ depending on the role, the emissions allocated to the person i over a day are:

$$E_i^{\text{alloc}} = \sum_{j \in \mathcal{T}(i)} f_j E_j, \tag{9}$$

where $\mathcal{T}(i)$ denotes the set of car trips associated with person i. This approach yields per-person, per-day emission estimates suitable for descriptive statistics and distributional analyzes.

The descriptive statistics in Table 5 summarize the individually allocated emissions for car trips only (driver allocation factor $= 1$, passenger allocation factor $= 0.5$). Two findings stand out immediately. First, distributions are highly dispersed, as evidenced by large standard deviations and wide interquantile ranges (P10–P90). Second, the mean consistently exceeds the median for all pollutants, indicating right-skewed distributions: most individuals exhibit moderate emission levels, while a smaller subgroup concentrates substantially higher values.

For CO_2, the mean emission allocated is 5.58 kg per person per day (median: 4.174 kg), with a P10 of 0.943 kg and a P90 of 11.638 kg. The gap between the mean and the median, combined with the wide P10–P90 spread, highlights strong heterogeneity in daily car-related emissions: a fraction of individuals experiences markedly more emissive days than the typical profile. This is confirmed by Figure 2, which displays a pronounced right tail indicative of a high-emitter subpopulation. To assess emission inequality, we use the Lorenz curve and the Gini coefficient: the Lorenz curve compares the cumulative share of emissions with the cumulative share of individuals, while the Gini coefficient summarizes the degree of concentration from 0 (perfect equality) to 1 (maximum inequality). The curve reveals a non-negligible concentration of emissions,

Table 5. *Descriptive statistics of individually allocated car emissions.*

Indicator	Mean	Median	Std. Dev.	P10	P90
CO_2 allocated (kg)	5.580	4.174	4.887	0.943	11.638
CO allocated (g)	19.330	2.911	40.162	0.414	56.544
HC allocated (mg)	494.12	281.68	800.29	52.78	1,066.72
NO_x allocated (g)	13.535	4.375	20.507	0.435	37.124
PM_x allocated (g)	0.917	0.568	1.087	0.113	2.164

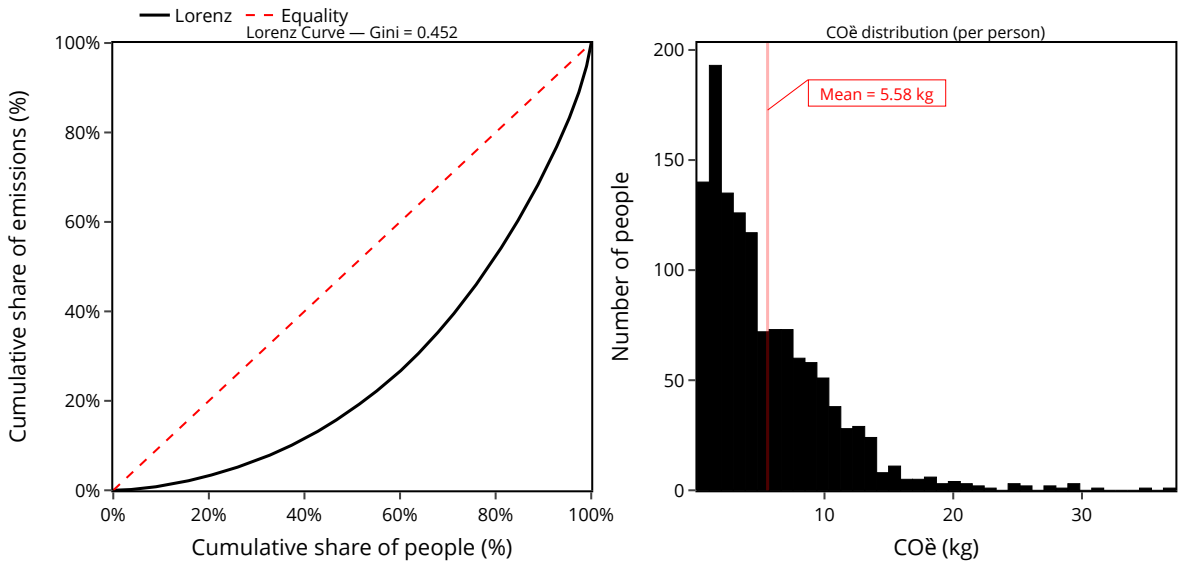

Figure 2. *Distribution of individually allocated daily car CO_2 emissions and associated inequality.*

with a Gini coefficient of 0.452 reflecting an unequal, though not extreme, distribution consistent with a high-emitter effect driven by differences in distances traveled, trip frequencies, and vehicle characteristics.

For particulate matter (PM_x), the mean allocated emission is 0.917 g per person per day (median: 0.568 g), with P10 at 0.113 g and P90 at 2.164 g. Figure 3 illustrates the corresponding distribution and its concentration, through the histogram and the Lorenz curve. As observed for CO_2, the mean-to-median gap confirms a right-skewed distribution; however, the degree of concentration is stronger, with a Gini coefficient of 0.531. PM_x emissions are thus more strongly dominated by a minority of individuals, consistent with the sensitivity of aggregate particulate emissions to vehicle profiles and traffic conditions such as cold starts, frequent stops, and congestion.

Table 6. *Summary statistics of daily distance traveled (km): full sample vs. top-50 highest CO_2 emitters.*

Group	n	Mean	Median	P10	P90
All individuals	1,293	31.07	25.83	5.00	59.37
Top-50 emitters	50	104.49	95.93	60.53	158.03

To interpret this heterogeneity, we examine the 50 highest CO_2 emitters. This group accounts for 14.8% of total CO_2 emissions and is used to identify the drivers of the upper tail of the distribution. The most direct determinant of road-transport emissions is total distance traveled per day. The results reveal a clear mobility gap between the full sample and the top-50 group, as summarized in Table 6.

Across the full population analyzed ($n = 1{,}293$), the mean daily distance traveled is 31.07 km. By contrast, the 50 highest emitters travel on average 104.49 km per day, approximately 3.4 times the population mean. This over-mobility is also visible in Figure 4: the distance distribution for the top-50 is markedly shifted toward longer values, and the distance-vs.-CO_2 scatter plot reveals an overall increasing relationship. A substantial share of the heterogeneity in CO_2 emissions can therefore be attributed to a small group of individuals undertaking considerably longer car-based journeys on a daily basis.

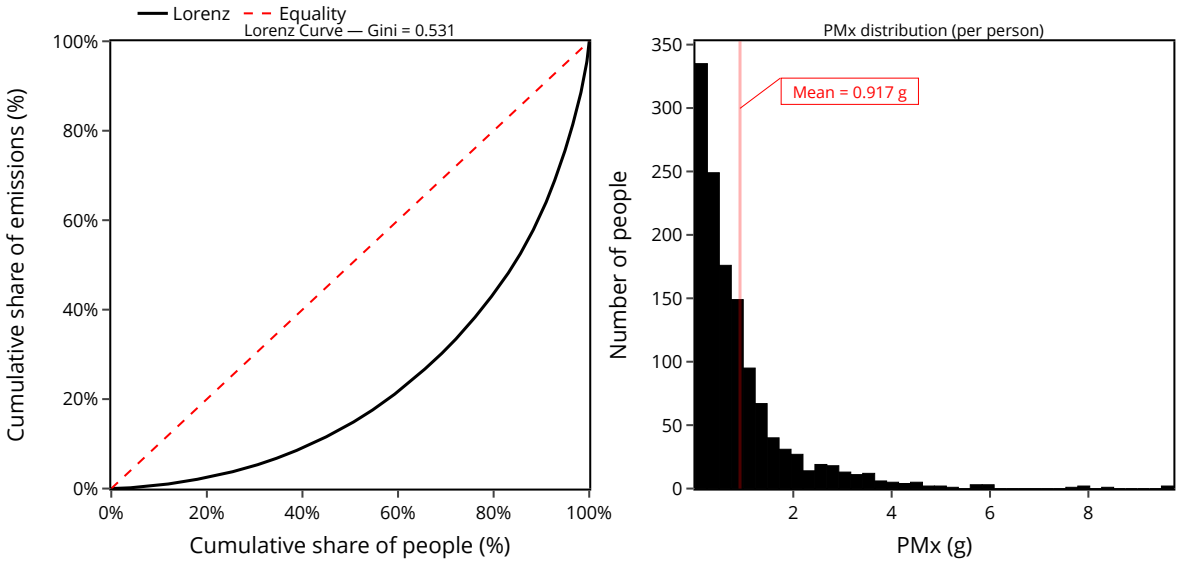

Figure 3. *Distribution of individually allocated daily car PM_X emissions and associated inequality.*

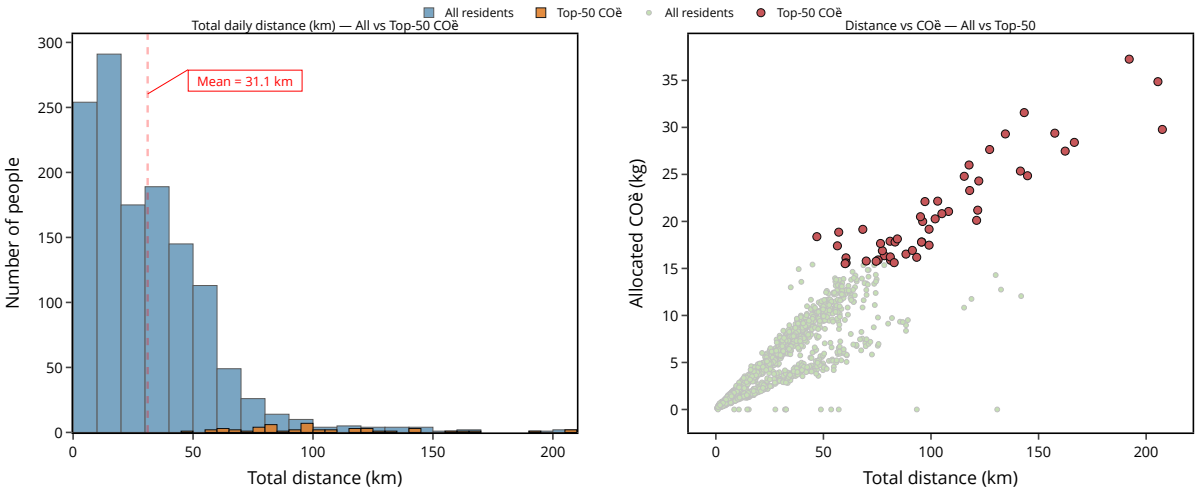

Figure 4. *Comparison of total daily distance for the full sample and the top-50 CO_2 emitters (left) and relationship between distance traveled and allocated CO_2 emissions(right).*

We now examine a second potential driver of heterogeneity: powertrain type. The objective is to assess whether the highest emitters also exhibit a vehicle-fleet profile that differs from the rest of the population.

The powertrain distribution (Figure 5) reveals a clear contrast: diesel vehicles are over-represented among the 50 highest emitters, while petrol vehicles are under-represented. This pattern is summarized by the top-50-to-fleet ratio: diesel yields a ratio of 1.08 (approximately 8% more prevalent in the top-50 than in the overall fleet), while petrol yields a ratio of 0.84 (approximately 16% less prevalent). Minority powertrains—Battery Electric Vehicles (BEV), Liquefied Petroleum Gas (LPG), and Plug-in Hybrid Electric Vehicles (PHEV)—remain marginal in the overall fleet and are almost absent from the top-50, limiting their statistical explanatory power for the highest individual emissions in this sample.

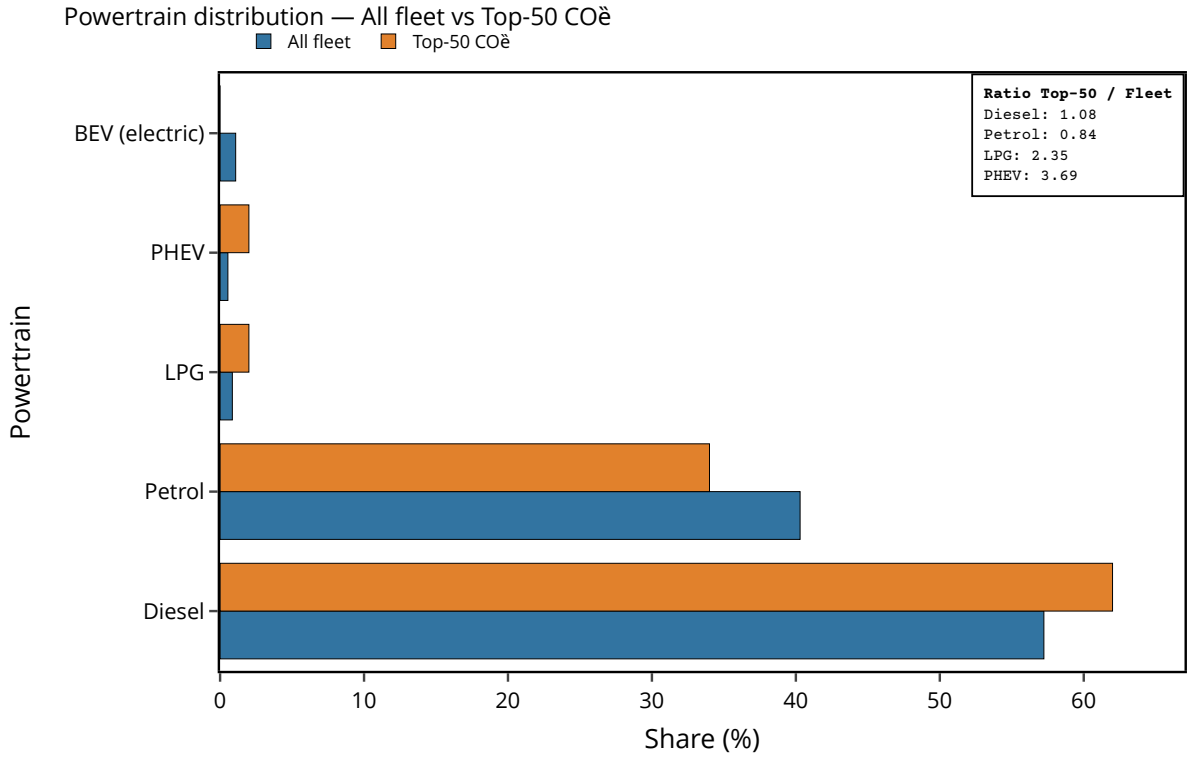

Figure 5. *The powertrain distribution (overall fleet vs. top-50 CO_2 emitters).*

Taken together, these results suggest that the observed heterogeneity arises from a combination of factors: a dominant behavioral effect, driven by substantially higher distances traveled among the highest emitters, and a secondary structural effect linked to the over-representation of diesel vehicles within this subgroup. In other words, the highest individual emissions appear to be driven primarily by more intensive daily mobility and are further reinforced by a more diesel-oriented fleet composition.

6. Discussion

In this section, we assess the credibility of the emission estimates produced by our pipeline. We first assess the plausibility of the simulated emissions by using institutional reports at the La Rochelle scale as external order-of-magnitude references, then describe the assumptions and simplifications that constitute the main sources of uncertainty.

6.1 Order-of-magnitude comparison

To assess the plausibility of the simulated emissions, we compare our results with two official documents covering the La Rochelle CdA: the *Bilan Carbone Territoire 2019* [28] and the *Plan Climat Air Energie Territorial 2024* [29]. *Bilan Carbone* 2019 reports approximately 283,400 tCO_2/year attributed to road passenger travel (private cars, buses, and motorized two-wheelers) at the CdA scale. Compared to the reference population of approximately 162,400 inhabitants, this corresponds to roughly 4.78 $kgCO_2e$ per inhabitant per day. Over one simulated day, our model yields approximately 83 tCO_2/day for 16,952 agents (target population and system population), i.e., approximately 4.89 $kgCO_2$/agent/day, which is close to the per-capita average reported in the territorial carbon assessment.

The simulated per-capita estimate is close to the territorial reference value, suggesting a plausible order of magnitude rather than a strict quantitative validation. The simulated population does not represent the full CdA population, but rather residents of the eight peri-urban municipalities and individuals whose daily activity chains are fully captured within the CdA perimeter. Moreover, the territorial assessment and the simulation rely on different accounting frameworks and population scopes.

Table 7. *Comparison of daily road-transport CO_2 emissions: simulation vs. territorial reference.*

Source	Population	tCO_2/day	$kgCO_2$/person/day
Simulation	16,952	82.93	4.89
Territorial reference [28], [29]	162,381	776.49	4.78

6.2 Limitations

Our approach relies on a multi-step processing chain, and several methodological choices may influence the results, particularly with respect to simulated distances and individual emission estimates.

POI-to-network anchoring

Activity locations are projected onto the SUMO network through a heuristic nearest-edge procedure. Although effective at scale, this may occasionally assign activities to unsuitable or poorly connected edges, affecting distances, speed profiles, and emissions.

Agent exclusion and selection bias

Agents for whom conversion fails (no valid route, unknown edge, or disconnected network components) are excluded to ensure simulation stability. This introduces a selection bias: excluded agents may be more peripheral, involve more complex multimodal chains, or correspond to atypical activity locations, potentially distorting aggregate statistics and distributional analyzes.

Activity simplifications

Removing interaction activities and merging consecutive activities at the same location improves SUMO plan consistency, but may slightly alter daily schedules, peak-hour patterns, and emission estimates.

Fleet vintage mismatch

The vehicle fleet is derived from a 2022 snapshot, whereas the order-of-magnitude reference is based on a 2019 territorial report. This temporal inconsistency may introduce biases related to changes in diesel/petrol market shares and fleet renewal rates, contributing to discrepancies in emission levels, particularly for pollutants sensitive to vehicle emission standards.

Traffic dynamics calibration

The SUMO car-following parameters were not specifically calibrated against local traffic counts or speed observations. Since emission estimates are sensitive to speed and acceleration profiles, this may affect the representativeness of absolute emission values. Future work will therefore include calibration of traffic dynamics parameters when suitable local validation data are available.

Emission allocation assumption

Individual emission allocation relies on a simplified occupancy model: driver $= 1$ and passenger $= 0.5$ (equal sharing between two occupants). Actual vehicle occupancy varies substantially in practice. This assumption may underestimate the driver's attributed impact in predominantly single-occupant vehicles, or overestimate the passenger's share when average occupancy differs from two. A dedicated sensitivity analysis using alternative occupancy scenarios (e.g., $f \in \{1/1,\ 1/0.33\}$ based on empirically plausible occupancy rates) would strengthen the robustness of the results. This allocation should therefore be interpreted as an accounting convention rather than as an explicit matching between passengers and drivers within the simulated population.

7. Conclusion

This paper presented a reproducible workflow for generating individual-level emission indicators from a microscopic SUMO simulation in peri-urban contexts, addressing a key gap in SUMO-based environmental studies: the construction of realistic mobility demand encompassing synthetic population, activity chains, daily schedules, and activity locations. The proposed pipeline combines Eqasim for synthetic population and activity-chain generation, MATSim for plan consolidation and mobility-choice refinement, and a dedicated conversion layer to SUMO including POI-to-network spatial anchoring, GTFS-based public transport integration, and emission-class assignment calibrated to reflect the local passenger-car fleet.

Applied to the La Rochelle peri-urban area, the pipeline produces per-person, per-day emission estimates and reveals strong inter-individual heterogeneity. For car-related CO_2, the distribution is right-skewed and moderately unequal (Gini ≈ 0.46), while

PM_x exhibits stronger concentration (Gini ≈ 0.55). The top-50 CO_2 emitters alone account for 14.8% of total CO_2 emissions within the study population, with their profile pointing to a dominant behavioral driver, substantially higher daily car distances, compounded by a secondary structural effect: over-representation of diesel vehicles relative to the fleet average. An external comparison with territorial references at the CdA scale confirms that the simulated daily CO_2 emissions are plausible at the order-of-magnitude level, while underscoring the importance of population definitions and accounting conventions when comparing model outputs to aggregated institutional reports.

Future work will focus on the eight peri-urban municipalities involved in the Yélo DETA project, a pilot initiative funded by Bpifrance under the national program France 2030 [30]. Building on the current baseline scenario, we will simulate the deployment of this on-demand automated transport service and assess its effects on travel behavior.

Data availability statement

The code, configuration files, and selected outputs used in this study are publicly available at https://github.com/NGALENAL1004/SUMO-demand2traffic. The upstream Eqasim France pipeline is available at https://github.com/eqasim-org/eqasim-france.

Author contributions

All authors have contributed to the article.

Alix Ngari Lendoye: conceptualization, investigation, methodology, software, visualization, and writing. Corwin Fèvre, Tatiana Graindorge, and Alain Bouju contributed to conceptualization, supervision, validation, and manuscript review.

Competing interests

The authors declare no competing interests.

Funding

This research was supported by the **Yélo DETA** project, funded by **Bpifrance** through the **France 2030** program.

References

[1] E. Flores-Juca, J. García-Navarro, E. Mora-Arias, and J. Chica, "La segregación espacial desde la perspectiva de la movilidad cotidiana y la densidad de las zonas periurbanas de cuenca en ecuador," *EURE (Santiago)*, vol. 49, no. 147, pp. 1–22, 2023, ISSN: 0250-7161. DOI: 10.7764/eure.49.147.04. [Online]. Available: https://doi.org/10.7764/eure.49.147.04.

[2] D. Krajzewicz, M. Behrisch, P. Wagner, R. Luz, and M. Krumnow, "Second generation of pollutant emission models for SUMO," in *Simulation of Urban Mobility*, Cham: Springer, 2015, pp. 231–257. DOI: 10.1007/978-3-319-15024-6_12.

[3] F. B. Ø. Carlsen, J. Jenner Rasmussen, M. M. Sørensen, N. Ø. Jensen, and M. Albano, "Generation of realistic activity scenarios for sumo," in *MobiQuitous '20: MobiQuitous 2020 - 17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, New York, NY, USA: Association for Computing Machinery, 2020, pp. 357–365, ISBN: 978-1-4503-8840-5. DOI: 10.1145/3448891.3448916. [Online]. Available: https://doi.org/10.1145/3448891.3448916.

[4] P. A. López, M. Behrisch, L. Bieker-Walz, *et al.*, "Microscopic traffic simulation using SUMO," in *Proceedings of the 21st IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 2575–2582. DOI: 10.1109/ITSC.2018.8569938.

[5] S. Rasouli and H. J. P. Timmermans, "Activity-based models of travel demand: Promises, progress and prospects," *International Journal of Urban Sciences*, vol. 18, no. 1, pp. 31–60, 2014. DOI: 10.1080/12265934.2013.835118.

[6] A. Horni, K. Nagel, and K. W. Axhausen, Eds., *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press, 2016. DOI: 10.5334/baw.

[7] Eclipse SUMO Development Team, *Introduction to demand modelling in SUMO*, SUMO Documentation, Accessed: 2024, 2024. [Online]. Available: https://sumo.dlr.de/docs/Demand/Introduction_to_demand_modelling_in_SUMO.html.

[8] Eclipse SUMO Development Team, *Activity-based demand generation (ActivityGen)*, SUMO Documentation, Accessed: 2024, 2024. [Online]. Available: https://sumo.dlr.de/docs/Demand/Activity-based_Demand_Generation.html.

[9] J. Schweizer, F. Rupi, F. Filippi, and C. Poliziani, "Generating activity based, multi-modal travel demand for SUMO," in *EPiC Series in Engineering*, vol. 2, 2018, pp. 118–133. [Online]. Available: https://www.semanticscholar.org/paper/Generating-activity-based%2C-multi-modal-travel-for-Schweizer-Rupi/df9e5421a43907d43b2ff32cc225506d99d22fdf.

[10] L. Codecà, J. Erdmann, V. Cahill, and J. Haerri, "SAGA: An activity-based multi-modal mobility scenario generator for SUMO," in *SUMO User Conference 2020 – From Traffic Flow to Mobility Modeling*, 2020. [Online]. Available: https://github.com/lcodeca/SUMOActivityGen.

[11] J. L. Bowman and M. E. Ben-Akiva, "Activity-based disaggregate travel demand model system with activity schedules," *Transportation Research Part A: Policy and Practice*, vol. 35, no. 1, pp. 1–28, 2001. DOI: 10.1016/S0965-8564(99)00043-9.

[12] N. Rezvany, M. Kukic, and M. Bierlaire, "A Review of Activity-based Disaggregate Travel Demand Models," *Findings*, Dec. 2024. DOI: 10.32866/001c.125431.

[13] ActivitySim Consortium, *Activitysim: An open platform for activity-based travel modeling*, https://activitysim.github.io, Accessed: 2024, 2020.

[14] B. Stabler and J. Doyle, "Development of a common open platform for activity-based travel demand modeling: Activitysim," in *TRB Planning Applications Conference*, Raleigh, NC, USA, May 2017.

[15] G. S. Macfarlane and N. J. Lant, "Estimation and simulation of daily activity patterns for individuals using wheelchairs," Utah Department of Transportation (UDOT), Research & Innovation Division, Final Report UT-21.10, Jun. 2021. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/56982.

[16] M. Adnan, F. C. Pereira, C. M. Lima Azevedo, *et al.*, "SimMobility: A multi-scale integrated agent-based simulation platform," in *Transportation Research Board 95th Annual Meeting*, Washington, D.C., 2016. [Online]. Available: https://eprints.soton.ac.uk/390938/.

[17] J. Li, H. Zhou, M. Snelder, B. van Arem, and J. Gao, "An activity- and agent-based co-simulation framework for the metropolitan rotterdam the hague region," *Procedia Computer Science*, vol. 257, pp. 959–965, 2025, Presented at ABMTrans 2025 (Patras, Greece, April 22–24, 2025), ISSN: 1877-0509. DOI: 10.1016/j.procs.2025.03.123. [Online]. Available: https://doi.org/10.1016/j.procs.2025.03.123.

[18] D. Ziemke, K. Nagel, and C. R. Bhat, "Expanding the analysis scope of a MATSim transport simulation by integrating the FEATHERS activity-based demand model," *Procedia Computer Science*, vol. 184, pp. 753–760, 2021. DOI: 10.1016/j.procs.2021.03.094.

[19] A. N. Lendoye, T. Graindorge, C. Fèvre, and A. Bouju, "Modeling methods for the environmental impact assessment of shared autonomous vehicles," *Procedia Computer Science*, vol. 274, pp. 686–695, 2025, 37th European Modeling & Simulation Symposium (EMSS 2025), within the 22nd International Multidisciplinary Modeling & Simulation Multiconference (I3M 2025). DOI: 10.1016/j.procs.2025.12.067. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050925037883.

[20] S. Hörl and M. Balac, "Synthetic population and travel demand for Paris and île-de-France based on open and publicly available data," *Transportation Research Part C: Emerging Technologies*, vol. 130, p. 103 291, 2021. DOI: 10.1016/j.trc.2021.103291.

[21] S. Hörl and M. Balac, "Introducing the eqasim pipeline: From raw data to agent-based transport simulation," in *Procedia Computer Science*, vol. 184, 2021, pp. 712–719. DOI: 10.1016/j.procs.2021.03.089.

[22] S. Hörl, M. Balac, and K. W. Axhausen, "A first look at bridging discrete choice modeling and agent-based microsimulation in MATSim," in *Procedia Computer Science*, vol. 130, 2018, pp. 900–907. DOI: 10.1016/j.procs.2018.04.087.

[23] INFRAS, "Handbook emission factors for road transport (HBEFA), version 4.1," INFRAS, Bern, Tech. Rep., 2019. [Online]. Available: https://www.hbefa.net.

[24] W. G. Cochran, *Sampling Techniques*, 3rd. New York: John Wiley & Sons, 1977.

[25] Nouvelle-Aquitaine Mobilités, *Réseau urbain yélo: Open datasets (gtfs, netex, siri) for ca de la rochelle*, Dataset (produced by Yélo). Licence: ODbL. Last update: 27 Feb 2026, 2026. [Online]. Available: https://transport.data.gouv.fr/datasets/arrets-horaires-et-parcours-theoriques-des-reseaux-naq-lro-nva-m-1 (visited on 03/01/2026).

[26] Service des données et études statistiques (SDES), *Données sur le parc de véhicules au niveau communal*, Jeu de données. Parc de véhicules au 1er janvier de chaque année depuis 2011. Dernière mise à jour : 27 Feb 2026. Licence Ouverte / Open Licence, 2026. [Online]. Available: https://www.data.gouv.fr/datasets/donnees-sur-le-parc-de-vehicules-au-niveau-communal (visited on 03/01/2026).

[27] Ministère de la Transition écologique, *La vignette Crit'Air: Classification des véhicules*, https://www.ecologie.gouv.fr/vignette-critair, Accessed: 2024, 2016.

[28] Communauté d'Agglomération de La Rochelle, "Bilan Carbone® Territoire de la CdA La Rochelle – Année 2019 : Note méthodologique," Communauté d'Agglomération de La Rochelle, La Rochelle, France, Tech. Rep., Mar. 2019.

[29] Communauté d'Agglomération de La Rochelle, *Plan Climat Air Energie Territorial (PCAET) – Version finale 2024*, Online publication (Calaméo), Adopté le 14 mars 2024, 2024. [Online]. Available: https://www.calameo.com/books/001297424a456871fa51b (visited on 02/28/2026).

[30] T. Graindorge, E. Talhi, R. Campbell, and T. Raimbault, "Automated mobility-on-demand services in the less dense and rural area," in *Transport Transitions: Advancing Sustainable and Inclusive Mobility*, ser. Lecture Notes in Mobility, C. McNally, P. Carroll, B. Martinez-Pastor, B. Ghosh, M. Efthymiou, and N. Valantasis-Kanellos, Eds., Cham: Springer, 2026, pp. 516–522. DOI: 10.1007/978-3-032-06763-0_74.