SUMO User Conference 2026

TIBconferencesession

TIB-OP_will_set_DOI

Framework for Generating Activity-Based Travel Demand Using Aggregated Mobility and Land-Use Big Data

Ngoc-An Nguyen[1], Joerg Schweizer[1], Cristian Poliziani[2], and Federico Rupi[1]

[1]Department of Civil, Chemical, Environmental and Materials Engineering, University of Bologna, 40126 Bologna, Italy

[2]Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA

*Correspondence: Joerg Schweizer, joerg.schweizer@unibo.it

Abstract. Activity-based models have emerged over the past four decades and have demonstrated a stronger capacity than traditional four-step models to capture the sequence of an individual's daily activities, associated trips, and travel behavior. However, the state of the art shows that demand generation in activity-based models still relies heavily on detailed travel diary data, which require extensive data-collection efforts. This paper introduces an activity-based modeling framework developed for the hybridPy/SUMO simulation environment. The aim is to generate individual plans based on generally available data such as aggregated population data, vehicle ownership and other mobility-related data, as well as building-scale land-use data from open sources. The model enables to generate a detailed daily travel plan for each agent, in which the land-use type of the visited locations must match the respective activity type. In addition, the plans must meet multiple constraints, such as building capacity, individual daily travel time budgets, and opening hours, where available. The overall framework comprises three key sub-modules: (i) household synthesis, which reconstructs a synthetic household population; (ii) individual disaggregation, which disaggregates households into individual-level attributes; and (iii) daily diary and schedule generation, which produces complete activity-travel diaries with primary activities and optional secondary activities. For each sub-module, the key input data, parameters, and outputs are presented to enable users to construct daily travel-demand scenarios. This framework offers several advantages in terms of reliable, behavior-based travel demand modeling and practical applicability across different urban contexts. By presenting a step-by-step model development workflow from the user's perspective, it enables scientists and transport planners to quickly and easily build large-scale microscopic travel-demand models without relying on disaggregated travel diary surveys or other location-specific statistics.

Keywords: daily activity-based model, activity-based model, travel demand generation, microscopic simulation

1. Introduction

Travel demand generation constitutes a fundamental component of transport modelling. Accurately estimating travel demand is a complex task that requires detailed individual-level data to ensure robust and reliable model outcomes. Since the 1950s, trip-based modelling – commonly referred to as the traditional four-step model (FSM) – has served as the dominant framework in transportation planning for several decades. FSMs represent travel demand at an aggregate level through a sequential process consisting of trip generation, trip distribution, mode choice, and route assignment [1]. Despite their widespread application, FSMs exhibit limitations, particularly in demand generation. Most notably, they oversimplify individual travel behaviour: (i) daily trips are treated as independent events; (ii) personal and infrastructure characteristics are averaged at the zonal level; and (iii) aggregate statistical techniques are applied without accounting for behavioural heterogeneity or individual variability. Such aggregation limits the ability to represent complex decision-making processes, thereby reducing the accuracy of travel demand forecasts under temporal interdependencies and changing socioeconomic conditions [2].

Activity-based models (ABMs) were introduced in the 1970s as an alternative to address the shortcomings of traditional four-step models (FSMs). They are grounded in the concept that travel demand is derived from individuals' daily activity patterns rather than isolated trips. ABMs represent travel behavior at the individual level and reproduce the full sequence of activities and associated trips over a day or even an entire week using behavioral-econometric frameworks [3]. A major strength of ABMs lies in their sensitivity to changes in transport supply conditions. Because the synthetic population is modeled at the individual level, agents can respond to modifications in networks, services, or policy interventions. This enables realistic simulation of what-if scenarios and supports more accurate predictions of travel demand. Consequently, ABMs provide a more behaviorally consistent representation of travel patterns and offer deeper insights into the potential impacts of transport policies [4].

Table 1. *Comparison between Four-Step Models (FSMs) and Activity-Based Models (ABMs)*

Four-Step Models (FSMs)	**Activity-Based Models (ABMs)**
Zoning-level considerations	Individual-level behavior representation
Aggregated input data	Disaggregated input data
Trips modeled independently	Sequences of activities and trips modeled jointly
Aggregate statistical demand models	Behavioral econometric models
Limited responsiveness to supply changes	Individuals react to transport supply changes, enabling realistic what-if scenario predictions

ABMs have been increasingly applied to represent complex interactions between individuals and the transport system, capturing dynamic travel demand and emerging mobility technologies across diverse contexts [5]. Various approaches have been proposed for activity generation, including constraint-based, utility-based, rule-based, learning-based, and hybrid methods [6]. These frameworks model heterogeneous decision-making processes and are commonly used at regional and urban scales, primarily supporting mesoscopic or macroscopic simulations. However, large-scale activity-based models designed for microscopic simulation remain relatively scarce. This is mainly due to substantial challenges, such as the need for highly detailed data, extensive calibration, long runtimes for simulating individuals, advanced computational resources, and limited standardized software tools [4], [7]. Consequently, there is a growing demand for activity-based demand generation models that are computationally efficient and flexible enough to support large-scale microscopic scenarios and multi-model applications [4].

The current state of the art indicates that large-scale microscopic ABMs typically depend on detailed daily travel diary data. Existing mobility surveys, such as the German Mobility Panel (MOP) in Germany, Public-Use Microdata Samples (PUMS) in the United States, or French Mobility Data in France, provide highly useful and insightful information for understanding individual travel behavior and calibrating ABMs. However, such detailed travel diary data are not commonly available for many cities. Their collection requires extensive, costly, and administratively demanding survey efforts, and the resulting datasets may still suffer from underreporting, missing entries, or inconsistencies. Consequently, the applicability of conventional ABMs is often limited to cities where comprehensive travel diary surveys are available.

At the same time, the integration of land-use information and emerging open or big data sources remains limited in the literature. Moreover, most existing studies concentrate on peak-hour conditions, while only a few have successfully developed full-day activity-based models suitable for microscopic traffic simulations.

This paper proposes a framework for generating daily activity-based travel demand integrated within hybridPy. The proposed framework utilizes open data sources and publicly available statistical datasets that are commonly accessible in many cities. This significantly lowers data requirements and simplifies the process of constructing daily trip chains, making activity-based demand modeling more practical and transferable across different urban contexts. The paper focuses on presenting the overall methodological framework, data requirements, key input parameters, computational procedures, and model outputs. It aims to provide a step-by-step implementation guide that enables users to independently develop traffic demand scenarios in hybridPy.

The remainder of the paper is structured as follows. Section 2 reviews the state of the art. Section 3 presents the Bitjoin framework and implementation process for generating daily travel demand in hybridPy. Section 4 describes the case study and presents selected preliminary results. Lastly, Section 5 concludes the paper.

2. State of the art

This section presents software packages that produce activity based demand. The most relevant characteristics are listed and a comparison in table format is provided at the end. Each package is using different traffic simulator to evaluate trip experiences of single agents.

2.1 Eqasim

Eqasim (Equilibrium Activity Simulation) is an activity-based travel demand generation framework designed to produce behaviorally consistent daily travel plans for large-scale microscopic transport simulations [8], [9]. It generates individual activity schedules including activity types, locations, departure times, and transport modes. Eqasim is closely integrated with the mesoscopic (queue-based) traffic simulator MATSim [10] and is designed to operate within an iterative simulation framework that approximates stochastic user equilibrium.

Eqasim consists of four principal components: (i) generation of a synthetic population based on census microdata, (ii) assignment of daily activity patterns derived from travel survey data, (iii) spatial location choice for activities using accessibility-based utility models, and (iv) discrete mode choice modeling. The generated plans are subsequently simulated within MATSim, where iterative re-routing and time adjustment allow the system to converge toward equilibrium under network congestion.

The required data sources include:

- **Travel survey data:** Household travel diaries containing observed activity chains, trip purposes, timing, mode choice, and socio-demographic characteristics. These data are required to estimate behavioral choice models.
- **Census microdata:** Demographic distributions (age, employment, household size, income, car ownership) used to construct a synthetic population representative of the study region.
- **Land-use and facility data:** Spatial distribution of residences, workplaces, schools, retail, and leisure locations, typically represented in zones or activity facilities.
- **Transport network data:** A detailed multimodal network including road infrastructure and public transport schedules.
- **Cost and utility parameters:** Mode-specific travel times, monetary costs, and behavioral parameters required for discrete choice modeling.

Eqasim outputs individual, time-consistent daily travel plans that are directly compatible with MATSim population files and suitable for equilibrium-based microscopic transport simulation.

2.2 ActivitySim

ActivitySim is an open-source, Python-based activity-based travel demand modeling platform developed by Los Alamos National Laboratory in collaboration with Metropolitan Planning Organizations (MPOs) [11]. It integrates PopulationSim [12] for synthetic population generation and follows the CT-RAMP [13] framework for activity generation and scheduling. Based on discrete choice theory and random utility maximization, ActivitySim models daily activity patterns and trip chains at the household and individual levels. ActivitSim is used in the simulation framework BRAM-Core traffic simulation framework [14], [15], where the generated plans are simulated with a mesoscopic simulator similar to MATSim.

ActivitySim operates through a sequence of interconnected components. It begins with the computation of accessibility measures derived from transport skims. Long-term choice models simulate workplace location, school location, and auto ownership. The Coordinated Daily Activity Pattern (CDAP) module determines household activity participation, followed by tour-level models (destination, time-of-day, and mode choice) and trip-level models that refine attributes for individual trip legs.

The main input data required for ActivitySim include:

- **Synthetic population data:** Disaggregated household and person attributes generated using PopulationSim from census microdata and control totals, travel diary data are typically needed for model estimation and calibration.
- **Land-use data (TAZ level):** Zonal residential and employment characteristics, activity opportunities, and planning data obtained from census, government records, or GIS databases.
- **Transport network skims (OMX format):** Pre-computed O–D matrices of travel time, cost, and distance for different modes, usually generated by an external assignment model.
- **Accessibility indicators:** Logsum-based or cumulative opportunity measures derived from skims and land-use data.

ActivitySim produces detailed trip outputs that can be aggregated into O–D matrices by time period and mode for subsequent traffic assignment

2.3 TAPAS

TAPAS (Travel and Activity Pattern Simulation) is an activity-based travel demand generator that produces daily activity schedules and corresponding trip chains for a synthetic population [16], [17]. Instead of directly generating origin-destination matrices, TAPAS models individual behavior: it determines which activities persons perform during a day (e.g., work, shopping, leisure), where these activities take place, when they start and end, and which transport mode is used. The resulting person-based travel demand can be exported and used as input for microscopic traffic simulators such as SUMO.

The TAPAS workflow consists of four main components: (i) generation of a statistically representative synthetic population, (ii) assignment of daily activity patterns based on travel survey data, (iii) spatial location choice for each activity using accessibility measures, and (iv) mode choice and temporal scheduling. Behavioral models (typically discrete choice models) estimated from empirical travel diary data to simulate how each individual chooses activities, destinations, departure times, and travel modes.

The most important input data categories are:

- **Travel survey data:** Household travel diaries providing observed activity chains, departure times, trip purposes, mode choices, and socio-demographic characteristics. These data are essential for estimating activity generation and mode choice models.

- **Census or population statistics:** Demographic distributions (age, household size, employment status, car ownership, etc.) used to create a synthetic population representative of the study area.
- **Land-use and spatial data:** Geographic information on residential areas, workplaces, schools, retail facilities, and other activity locations, represented as individual, discrete destination places; traffic analysis zones (TAZs) are used mainly for aggregation and for travel time and distance matrices.
- **Accessibility and travel time data:** Skim matrices (travel times, distances, generalized costs) for different transport modes, derived from the transport network.
- **Cost and behavioral parameters:** Mode-specific attributes such as travel costs, value of time, fuel prices, parking costs, and public transport fares.

Based on these inputs, TAPAS produces time-consistent, individual daily travel schedules that can be converted into trip tables or directly into person-based route files for SUMO.

2.4 SAGA

SAGA (SUMO Activity Generation Application) is a structured framework for generating activity-based travel demand for the microscopic traffic simulator SUMO. It provides a configurable pipeline for creating daily activity schedules and corresponding trips, typically relying on SUMO's built-in `activitygen` tool for plan generation. SAGA is designed for scenario-oriented applications and enables reproducible demand generation with moderate data requirements.

SAGA follows a probabilistic, rule-based activity generation approach. A synthetic population is defined using statistical distributions for household and person attributes. Predefined activity chain templates (e.g., home–work–home) are assigned according to configurable probabilities. Activities are spatially allocated using zones or network edges derived from land-use classifications. Departure times and activity durations are sampled from specified time windows. Generated trips are subsequently routed using SUMO's routing tools to ensure network consistency.

The essential input requirements are:

- **SUMO network data:** A road network file (`.net.xml`) and optionally traffic analysis zones (`.taz.xml`).
- **Population and activity configuration:** Statistical definitions of household types, participation rates, activity templates, and time windows.
- **Spatial / land-use information:** Optional but recommended zone definitions or edge classifications for residential, work, and other activity locations.
- **Optional calibration targets:** Aggregate traffic counts or OD matrices for scenario tuning.

SAGA produces time-consistent daily activity schedules and corresponding SUMO route files that can be directly used in microscopic traffic simulations. It emphasizes structured scenario generation rather than survey-calibrated behavioral estimation or equilibrium modeling.

2.5 ABIT

ABIT (Activity-Based Incremental Transport Model) is an activity-based travel demand model that generates weekly, 24/7 activity and travel schedules for a synthetic population [18]. It is behaviorally estimated using data from the German Mobility Panel (MOP), which provides multi-year household travel diaries including trip purposes, modes, distances, timing, and socio-demographic attributes. ABIT is tightly integrated with the SILO land-use model [19], which supplies and updates the synthetic population over time, and with MATSim [20], which provides transport supply indicators (e.g., travel times and accessibility).

ABIT consists of several interconnected modules. First, MOP travel diary data are processed into weekly activity schedules and are used to estimate behavioral and transition models. Second, based on the synthetic population provided and annually updated by SILO, ABIT generates and updates individual weekly activity patterns, distinguishing mandatory and discretionary activities. Third, habitual mode choice is modeled using MOP-based parameters and transport supply indicators (e.g., travel times and accessibility) provided by MATSim, household vehicles are allocated to members with the highest expected benefit from car use. Fourth, life-event modules rely on demographic changes simulated by SILO, and mobility changes derived from MATSim network conditions, to trigger adjustments in travel behavior.

The main input data categories include:

- **Travel survey data (MOP):** Multi-year household travel diaries with trip purposes, times, distances, modes, socio-demographics, and vehicle availability.
- **Synthetic population data:** Census-based demographic and household attributes generated and updated by SILO.
- **Transport network data:** Road and public transport networks used to generate skim matrices and accessibility indicators.
- **Transport supply indicators:** Travel times and service attributes derived from MATSim.
- **Life-event and demographic transition data:** Information on household changes (e.g., marriage, birth, relocation, car ownership changes) used to update travel behavior over time.

As output, ABIT produces detailed individual daily and weekly schedules and tour-based travel plans that can be directly converted into MATSim plans files for dynamic traffic assignment and simulation.

2.6 Software comparison

Table 2 summarizes the main characteristics of the above-mentioned software packages. It becomes clear that models using more sophisticated choice models require a higher more data to calibrate.

Table 2. *Compact comparison of activity-based demand generation frameworks.*

Characteristic	**TAPAS** [16]	**Eqasim** [8]	**SAGA** [21]	**ActivitySim** [11]	**ABIT** [18]
Primary ecosystem	SUMO contrib	MATSim	SUMO contrib	Python MPOs	SILO / MATsim
Activity-based (individual plans)	✓	✓	✓	✓	✓
Survey-estimated choice models	✓	✓	✗	✓	✓
Household coordination	~	~	✗	✓	~
Equilibrium integration	~[a]	✓	~[a]	~[a]	✗[a]
Data intensity	High	High	Low–Med	Very High	High
Traffic assignment	SUMO	MATSim	SUMO	BEAM-Core	MATSim

✓ supported; ✗ not supported; ~ partially supported / depends on configuration and coupling.

[a] Typically coupled to external assignment / traffic simulation for congestion feedback rather than providing an internal equilibrium loop by itself.

3. Daily travel demand generation using the Bitjoin method

Before diving into the demand generation procedure itself, the overall scenario-building process with hybridPy is presented. In this way it becomes clearer where and how the activity generator is integrated. Thereafter, the single steps are detailed in consecutive subsection.

3.1 Procedure for creating a traffic scenario in hybridPy

The sequential procedures for creating a traffic scenario in hybridPy consists of four main stages corresponding to 14 steps:

- **Stage 1 – Launching hybridPy and creating a scenario**
  - **Step 1 – Launch hybridPy:**
    Open a command prompt, move to the hybridPy folder, and run:
    python hybridPy_gui.py.
  - **Step 2 – Set up a new scenario:**
    Navigate to *Scenario/Create/New.../Run*. Provide a scenario name and description, then save it in the working directory as an *.obj* file (avoid using spaces in the filename).
- **Stage 2 – Generating the supply model**
  - **Step 3 – Build the transport network:**
    Create a grid network using *Network/Generate/Grid Network.../Run*, or load a network from OSM data via *Network/Import/From Osm.xml.../Run*.
  - **Step 4 – Edit the transport network:**
    Select *Network/Edit/Network.../Run* to open SUMO netedit and modify the network settings as needed.
  - **Step 5 – Define land use:**
    Run *Land Use/Facilities/Generate Facilities.../Run* to create random facilities, or import OSM land-use data through *Network/Import/From Poly File.../Run*.
  - **Step 6 – Create traffic analysis zones:**
    In the Object Editor, choose “Add a zone” to manually draw zones, or import

  them using *Landuse/Zones/Import From Shape File.../Run*. Zone attributes can be modified under *Landuse/Browse/Zones*.
  - **Step 7 – Generate parking facilities:**
    Access *Landuse/Parking/Generate Parking.../Run* and configure the main parameters for on-street parking.
  - **Step 8 – Develop public transport infrastructure and services:**
    Under *Network/Edit/Network*, use the bus-stop tool to insert stops along road segments and save the updated network. Then, in the Object Editor, select "Add Public Transport Lines" to define routes and stop sequences, or import GTFS data via *Demand/Public Transport/Import From GTFS*.
- **Stage 3 – Generate the activity-based travel demand using the Bitjoin method**
  - **Step 9 – Generate Facility-Based Households:**
    Run *Demand/Bitjoin/Household Generation.../Run* and input the necessary aggregated household statistics.
  - **Step 10 – Create the virtual population:**
    Execute *Demand/Bitjoin/Virtual Population Generation.../Run* using the complete set of aggregated demographic data.
  - **Step 11 – Generate daily activities and schedules:**
    Run *Demand/Bitjoin/Daily Activity Generation.../Run* and specify all required activity-related parameters.
- **Generating travel plans and executing microscopic simulations in SUMO**
  - **Step 12 – Create daily travel plans:**
    Execute *Demand/Bitjoin/Travel Plan Generation.../Run* and choose the desired transport modes.
  - **Step 13 – Select Travel plan:**
    Run *Demand/Bitjoin/Travel Plan Selection.../Run* with the selected mode choice model applied.
  - **Step 14 – Run the microscopic simulation:**
    Start the simulation via *Simulation/Microscopic Simulation.../Run* and configure all necessary simulation parameters.

The detailed procedure for Stage 3: the framework of the Bitjoin method, which is the key focus of this paper is presented in Section 3.2.

3.2 The framework of the Bitjoin method

3.2.1 Primary data requirements and sources

The primary input datasets for these modules include:

- **Road network data:** Provide a road network imported from OpenStreetMap [22] and converted to a SUMO-compatible XML network using `netconvert` [23]. If needed, refine the network in `netedit` [24] using satellite imagery, Google Maps street views, and on-site surveys. The network should include edges/lanes, speed limits, mode access rights, connectivity, junctions/nodes, traffic lights, and parking.
- **Traffic light data:** Provide signalized junctions with signal phases, timings, and control logic. Signal plans can be generated from OSM nodes via `netconvert` [23] and refined through field inspections to match observed operations.

- **Public transport data:** Provide a GTFS dataset ([25]) describing routes, stops, stop-level timetables, and daily trips. In `hybridPy`, stops are mapped to the network by stop ID/name and associated edges, and routes are map-matched to sequences of links with service frequency. Bus stop locations and service frequency are then used to enable public transport availability for the virtual population.
- **Aggregate household statistics data:** Provide aggregated household-level statistics for the study area, including household size, income, household member age/gender composition, and vehicle ownership. These statistics are typically sourced from censuses, household socioeconomic surveys, and statistical yearbooks and are used to reconstruct household distributions and generate a virtual population.
- **Aggregate personal travel statistics:** Provide aggregated travel-behavior statistics describing activity participation (e.g., activity shares by time or trips), personal vehicle ownership, and daily travel time budget (TTB). These inputs are typically derived from travel surveys and may also be sourced from mobile phone, smart card, or GPS-based datasets. They are used to parameterize activity scheduling, mode choice, and travel time allocation.
- **Land-use big data:** Provide land-use attributes extracted from OSM, including building facilities (IDs, land-use categories, names, footprint polygons, and coordinates) and associated POIs (type, name, and operating times). These data are used to disaggregate person locations (e.g., homes and workplaces) and to link activities to specific places with realistic time constraints. Building areas and operating hours support timing availability and service capacity inputs for the activity location and time choice models (see Table 3).

Table 3. *Description of land-use types extracted from the OSM database*

Land-use type	**Description**
Industrial	Employment areas related to manufacturing activities, storage facilities, and logistics operations.
Commercial	Areas dedicated to retail activities, service provision, and office-based business operations.
Education	Educational institutions including schools, universities, and vocational or training centers.
Leisure	Facilities for cultural events, sports, recreation, and entertainment purposes.
Residential	Housing areas and residential buildings where inhabitants reside.
Mixed	Buildings containing several points of interest (POIs) with combined functions (e.g., residential and commercial use).

3.2.2 Processing steps

To execute the Bitjoin framework, navigate to *Demand/Bitjoin* in the main menu. A set of sequential sub-modules becomes available for generating activity-based travel demand, see Figure 1. These sub-modules guide the user through the complete demand generation process, including: (i) household generation (Step 9), (ii) virtual population synthesis (Step 10), and (iii) daily activity scheduling (Step 11). Travel plan generation and travel plan selection in SUMO are also implemented to create daily travel plans, which are then fed into dynamic traffic assignment models to simulate agent-based transportation systems. The travel plans can be re-selected due to changes in travel times resulting from the traffic assignment after traffic simulation iterations

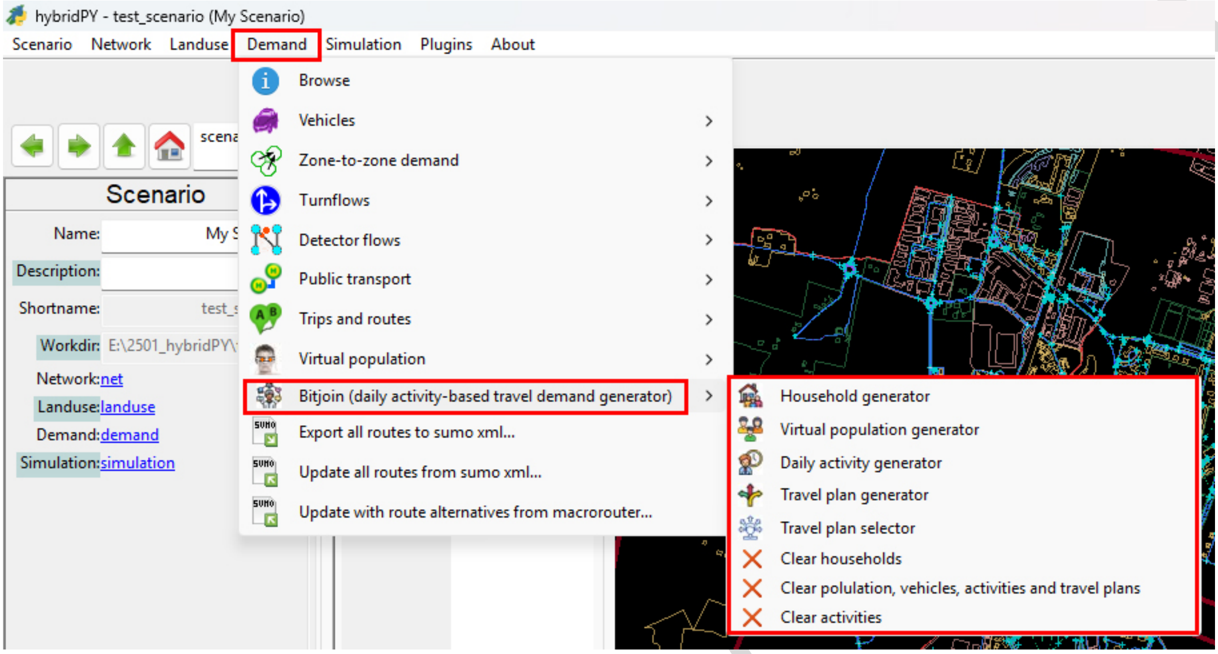


Figure 1. *Navigation to the Bitjoin module menu*

The processing framework, including the Bitjoin sub-modules and their corresponding inputs and outputs, is presented in Figure 2.

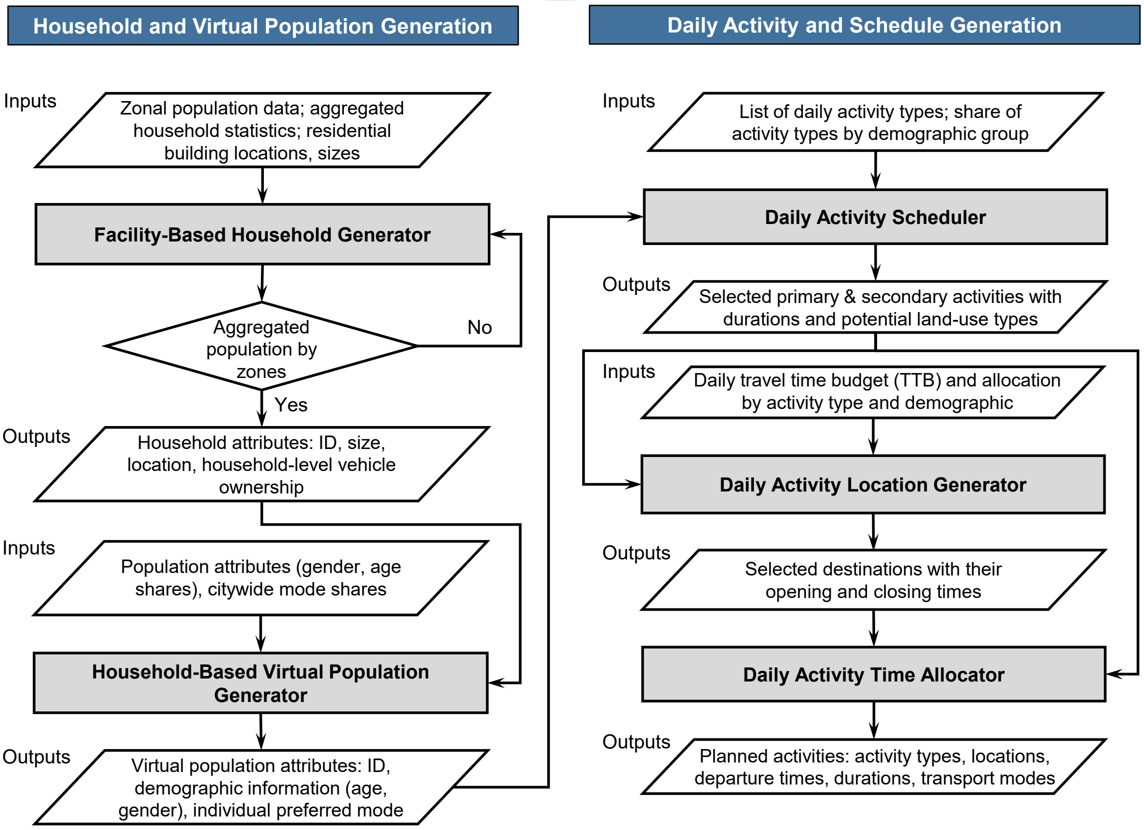


Figure 2. *Framework for synthetic population and daily travel plan generation modules*

3.2.3 Generating Facility-Based Households (Step 9)

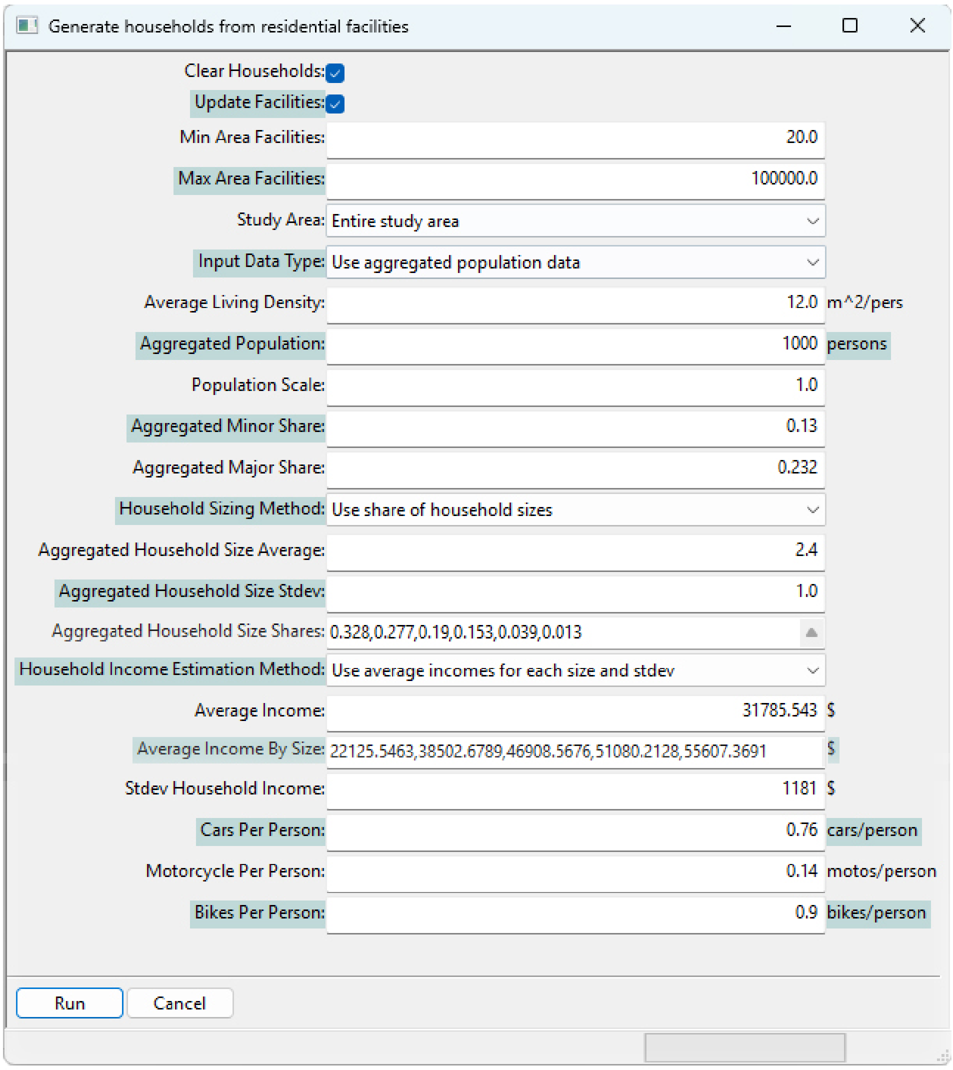
A capacity-based probabilistic allocation framework is used to convert an input population into synthetic households that are spatially distributed across *”residential”* buildings. The population input can be either aggregated for the entire study area or disaggregated by zone. Household attributes including household size, age composition, vehicle ownership, and income are generated to match aggregate census statistics.

Key Inputs

- **Statistical population and household demographic information**
  - Total population: either global for the entire study area or zonal totals by zone.
  - Household size statistics: either (i) Normal distribution parameters (mean and standard deviation), or (ii) an empirical discrete distribution (size categories with associated shares/probabilities).
  - Age-group shares: share of minors and share of elders (adults may be derived as the remainder, if required).
  - Income statistics: income distribution parameters (e.g., mean and standard deviation). The mean income may be specified globally or conditional on household size.
  - Ownership rates for cars, motorcycles, and bicycles, including per-adult rates for motorized vehicles (cars and motorcycles) and per-capita rates for bicycles.
- **Land-use / facilities data:** Residential facilities (buildings) within each zone, including zone ID, facility/building ID, building footprint area, and a capacity measure (e.g., derived from footprint area multiplied by an assumed people-per-m^2, or a predefined capacity).

Processing Calculation

1. **Calculation 1 — Generate household sizes & allocate households to residential facilities**
   Goal: This step aims to create households and place them in residential buildings while respecting building capacity and matching population totals.
   How it works:
   - Household size sampling: For each new household, draw a household size using random sampling from a normal distribution or from an empirical discrete distribution.
   - Facility selection: Select a residential facility using capacity-weighted probabilistic sampling, where facilities with larger available capacity have a higher probability of being selected.
2. **Calculation 2 – Assign household income and age composition**
   Goal: Add socio-demographics consistent with aggregate census distributions, while remaining compatible with household size.
   How it works:
   - Income assignment: for each household, assign income using normal distribution sampling
   - Age-group target calculation: compute target totals for minors and elders using share-based scaling from the total population



Clear Households: ☑
Update Facilities: ☑
Min Area Facilities: 20.0
Max Area Facilities: 100000.0
Study Area: Entire study area
Input Data Type: Use aggregated population data
Average Living Density: 12.0 m^2/pers
Aggregated Population: 1000 persons
Population Scale: 1.0
Aggregated Minor Share: 0.13
Aggregated Major Share: 0.232
Household Sizing Method: Use share of household sizes
Aggregated Household Size Average: 2.4
Aggregated Household Size Stdev: 1.0
Aggregated Household Size Shares: 0.328,0.277,0.19,0.153,0.039,0.013
Household Income Estimation Method: Use average incomes for each size and stdev
Average Income: 31785.543 $
Average Income By Size: 22125.5463,38502.6789,46908.5676,51080.2128,55607.3691 $
Stdev Household Income: 1181 $
Cars Per Person: 0.76 cars/person
Motorcycle Per Person: 0.14 motos/person
Bikes Per Person: 0.9 bikes/person

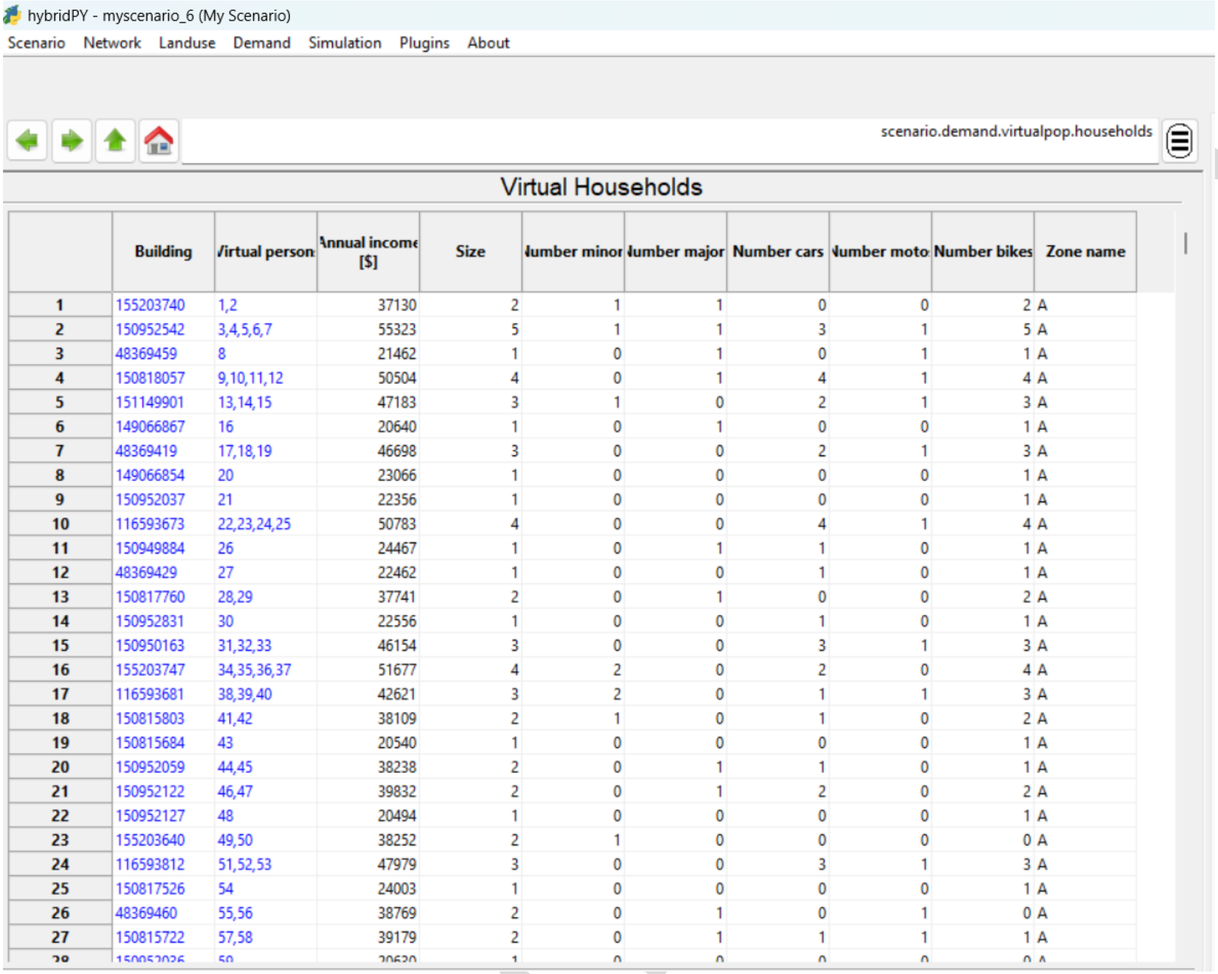
Figure 3. *Configuration menu for input data and parameters to generate facility-based households*

   - Age-group assignment to households: assign minors and elders using weighted probabilistic assignment until targets are met: The probability of having minors and elders generally increases with household size, and the likelihood of minors may also rise with income.

3. **Calculation 3 – Assign household vehicle ownership**
   Goal: Allocate vehicles to match city-level aggregate ownership rates while reflecting differences by household size and income.
   How it works:
   - Vehicle target calculation: Convert ownership rates into target vehicle counts.
   - Household ownership estimation: Motorized vehicles are allocated using income-and-size weighted probabilities, while bicycles are allocated using size-proportional probabilities.

Key Outputs

- A household list/table containing: household ID, zone ID, assigned residential facility/building ID, household income, household size, age compositions (e.g., number/presence of minors and elders), and vehicle ownership (cars, motorcycles, and bicycles).



Virtual Households

	Building	Virtual person	Annual income [$]	Size	Number minor	Number major	Number cars	Number moto	Number bikes	Zone name
1	155203740	1,2	37130	2	1	1	0	0	2	A
2	150952542	3,4,5,6,7	55323	5	1	1	3	1	5	A
3	48369459	8	21462	1	0	1	0	1	1	A
4	150818057	9,10,11,12	50504	4	0	1	4	1	4	A
5	151149901	13,14,15	47183	3	1	0	2	1	3	A
6	149066867	16	20640	1	0	1	0	0	1	A
7	48369419	17,18,19	46698	3	0	0	2	1	3	A
8	149066854	20	23066	1	0	0	0	0	1	A
9	150952037	21	22356	1	0	0	0	0	1	A
10	116593673	22,23,24,25	50783	4	0	0	4	1	4	A
11	150949884	26	24467	1	0	1	1	0	1	A
12	48369429	27	22462	1	0	0	1	0	1	A
13	150817760	28,29	37741	2	0	1	0	0	2	A
14	150952831	30	22556	1	0	0	1	0	1	A
15	150950163	31,32,33	46154	3	0	0	3	1	3	A
16	155203747	34,35,36,37	51677	4	2	0	2	0	4	A
17	116593681	38,39,40	42621	3	2	0	1	1	3	A
18	150815803	41,42	38109	2	1	0	1	0	2	A
19	150815684	43	20540	1	0	0	0	0	1	A
20	150952059	44,45	38238	2	0	1	1	0	1	A
21	150952122	46,47	39832	2	0	1	2	0	2	A
22	150952127	48	20494	1	0	0	0	0	1	A
23	155203640	49,50	38252	2	1	0	0	0	0	A
24	116593812	51,52,53	47979	3	0	0	3	1	3	A
25	150817526	54	24003	1	0	0	0	0	1	A
26	48369460	55,56	38769	2	0	1	0	1	0	A
27	150815722	57,58	39179	2	0	1	1	1	1	A

Figure 4. *Household generation result*

Validation checks

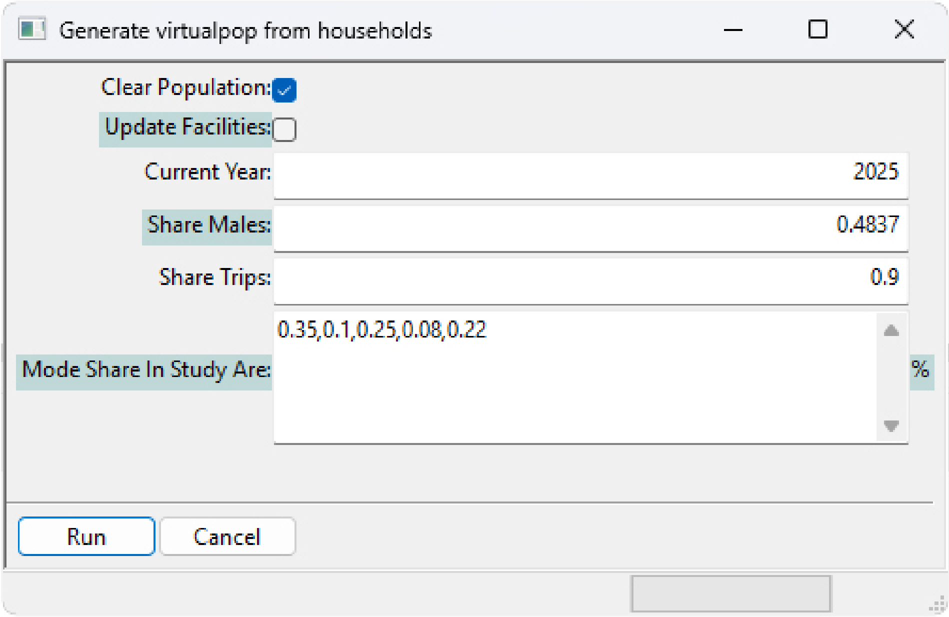
- Household size distribution vs input targets;
- Income distribution vs input targets;
- Minors/elders totals vs targets;
- Vehicle totals vs ownership-rate targets.

3.2.4 Generating the Virtual Population (Step 10)

This step reconstructs virtual population for each synthetic household. Using aggregate census statistics, it probabilistically generates individual demographics (gender, age), then allocates household vehicles to eligible members while respecting household vehicle ownership quotas and eligibility rules. Finally, it assigns each person a preferred travel mode based on their available modes (given vehicle access) and the observed citywide mode-share targets.

Key Inputs

- **Household inputs (from the previous step):** Household ID, household size, household age composition (number of minors/adults/elders); household vehicle quotas by type (cars, motorcycles, bicycles); and household income.



Generate virtualpop from households

Clear Population: ☑

Update Facilities: ☐

Current Year: 2025

Share Males: 0.4837

Share Trips: 0.9

Mode Share In Study Are: 0.35,0.1,0.25,0.08,0.22 %

Figure 5. *Configuration menu for input data and parameters for generating the virtual population*

- **Citywide demographic shares:** Male and female shares.
- **Age-group definitions:** Age ranges for minors, adults, and elders.
- **Vehicle eligibility rules:** Minimum eligible age thresholds by vehicle type.
- **Mode choice configuration:** Set of available modes; citywide mode shares.
- **Transit accessibility (optional):** Bus stop locations and an access distance threshold from home (e.g., 500 m), to determine whether bus is an available option.

Processing Calculation

1. **Calculation 4 – Generate individual demographics within each household**
   Goal: Generate basic individual demographics for each household member consistent with citywide demographic shares and household composition.
   How it works:
    - Create ID of individuals per household, where number of individuals equals the household size.
    - Assign gender to each person using probabilistic sampling based on citywide male and female shares.
    - Assign individuals to age groups (minors/adults/elders) to match the household composition.
    - Draw an exact age for each person using conditional discrete sampling within the corresponding age-group range.
2. **Calculation 5 – Allocate household vehicles to eligible members**
   Goal: Allocate household vehicles to individuals while respecting vehicle quotas and eligibility constraints.
   How it works:
    - Identify eligible members for each vehicle type based on age thresholds.
    - Allocate household vehicles to individuals using quota-based sequential assignment: allocate cars first (typically to adults), then motorcycles and bi-

cycles with priority given to unassigned eligible members; if needed, backfill from already-assigned eligible members to allow multiple vehicles per person.

3. **Calculation 6 – Assign preferred travel mode for each person**
   Goal: Assign a preferred mode per person consistent with individual mode availability and observed citywide mode shares.
   How it works:
   - Determine each person's available mode from previous steps: private modes available if the person owns the corresponding vehicle, bus available if the household within the bus-stop threshold distance, walking is available for every people.
   - Compute citywide mode-assignment targets based on observed mode shares.
   - Assign preferred mode using a rule-based assignment with priority ranking: If only one non-walk option is available, assign it directly. If multiple options are available, assign the highest-priority mode according to a configurable ranking (e.g., by relative speed/utility). Otherwise, assign walking.

Key Outputs

- Virtual population table: Person ID, household ID, gender, age group, age, individual vehicle ownership and preferred mode.

Validation Checks

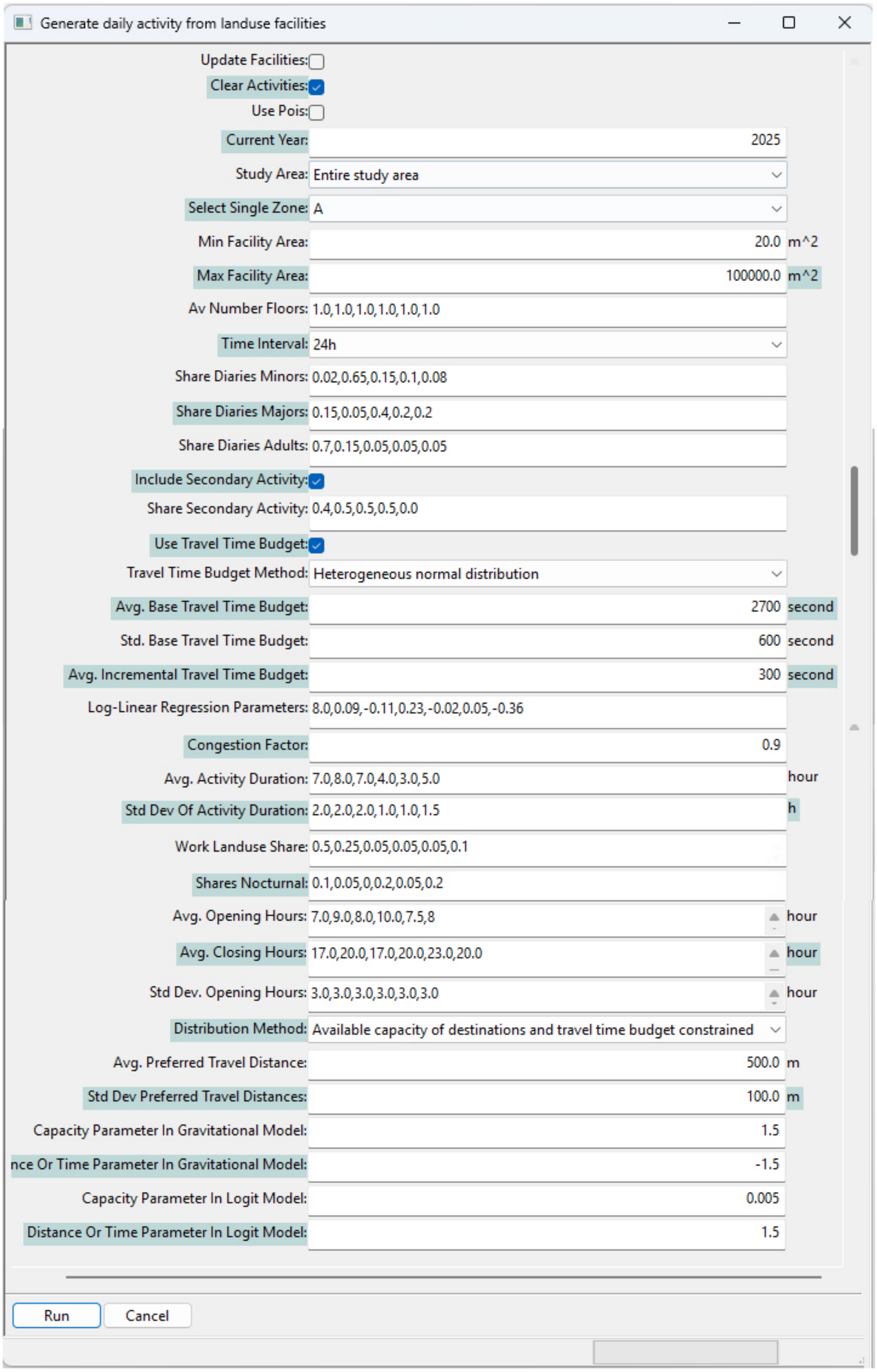
- Checks of gender shares, age-group totals, household vehicle quota satisfaction, and preferred-mode distribution versus citywide targets.

3.2.5 Generating Daily Activities and Schedules (Step 11)

This sub-module reconstructs daily activity diaries and travel plans for the virtual population by assigning activity chains (primary and optional secondary activities), selecting land-use-consistent activity locations under travel-time constraints (travel time budget - TTB), and scheduling start times, durations, and trips across a full 24-hour day. The output is a complete daily plan for each person (activity types, locations, travel times, and time schedule), which is then used for downstream activity-based modeling and microscopic traffic simulation.

Key Inputs

- **Virtual population data:** Person ID, household ID, home location (residential building), age group (minor/adult/elder), and preferred mode (defined in previous steps).
- **Activity diary configuration:**
  - Defined activity types (home, work, education, shopping, leisure, other).
  - Age-group-specific shares for primary diary chains.
  - Diary-specific shares for optional secondary activities (e.g., shopping/leisure extensions).
- **Land-use / facility database (from OSM or equivalent):**



Generate daily activity from landuse facilities

Update Facilities: ☐
Clear Activities: ☑
Use Pois: ☐
Current Year: 2025
Study Area: Entire study area
Select Single Zone: A
Min Facility Area: 20.0 m^2
Max Facility Area: 100000.0 m^2
Av Number Floors: 1.0,1.0,1.0,1.0,1.0,1.0
Time Interval: 24h
Share Diaries Minors: 0.02,0.65,0.15,0.1,0.08
Share Diaries Majors: 0.15,0.05,0.4,0.2,0.2
Share Diaries Adults: 0.7,0.15,0.05,0.05,0.05
Include Secondary Activity: ☑
Share Secondary Activity: 0.4,0.5,0.5,0.5,0.0
Use Travel Time Budget: ☑
Travel Time Budget Method: Heterogeneous normal distribution
Avg. Base Travel Time Budget: 2700 second
Std. Base Travel Time Budget: 600 second
Avg. Incremental Travel Time Budget: 300 second
Log-Linear Regression Parameters: 8.0,0.09,-0.11,0.23,-0.02,0.05,-0.36
Congestion Factor: 0.9
Avg. Activity Duration: 7.0,8.0,7.0,4.0,3.0,5.0 hour
Std Dev Of Activity Duration: 2.0,2.0,2.0,1.0,1.0,1.5 h
Work Landuse Share: 0.5,0.25,0.05,0.05,0.05,0.1
Shares Nocturnal: 0.1,0.05,0,0.2,0.05,0.2
Avg. Opening Hours: 7.0,9.0,8.0,10.0,7.5,8 hour
Avg. Closing Hours: 17.0,20.0,17.0,20.0,23.0,20.0 hour
Std Dev. Opening Hours: 3.0,3.0,3.0,3.0,3.0,3.0 hour
Distribution Method: Available capacity of destinations and travel time budget constrained
Avg. Preferred Travel Distance: 500.0 m
Std Dev Preferred Travel Distances: 100.0 m
Capacity Parameter In Gravitational Model: 1.5
nce Or Time Parameter In Gravitational Model: -1.5
Capacity Parameter In Logit Model: 0.005
Distance Or Time Parameter In Logit Model: 1.5

Figure 6. *Configuration menu for input data and parameters for generating daily activities and schedules*

  - Buildings/facilities classified by functional land-use types and linked to activity purposes.
  - Facility attributes (e.g., location, capacity, opening and closing hours).
- **Transport network data:**
  - Road / path network with edge travel times by mode and connectivity (adjacency).
  - Edge-to-facility mapping (nearest network edge for each building/facility).

- **Travel-time budget parameters:**
    - Daily travel time budget distribution (e.g., normal distribution) with mean adjustments by demographics (age, gender).
    - Rules to split daily budget into primary and (optional) secondary components.
- **Scheduling parameters:**
    - Activity duration distributions by activity type (e.g., normal distribution).
    - Facility opening-, closing-time distributions.

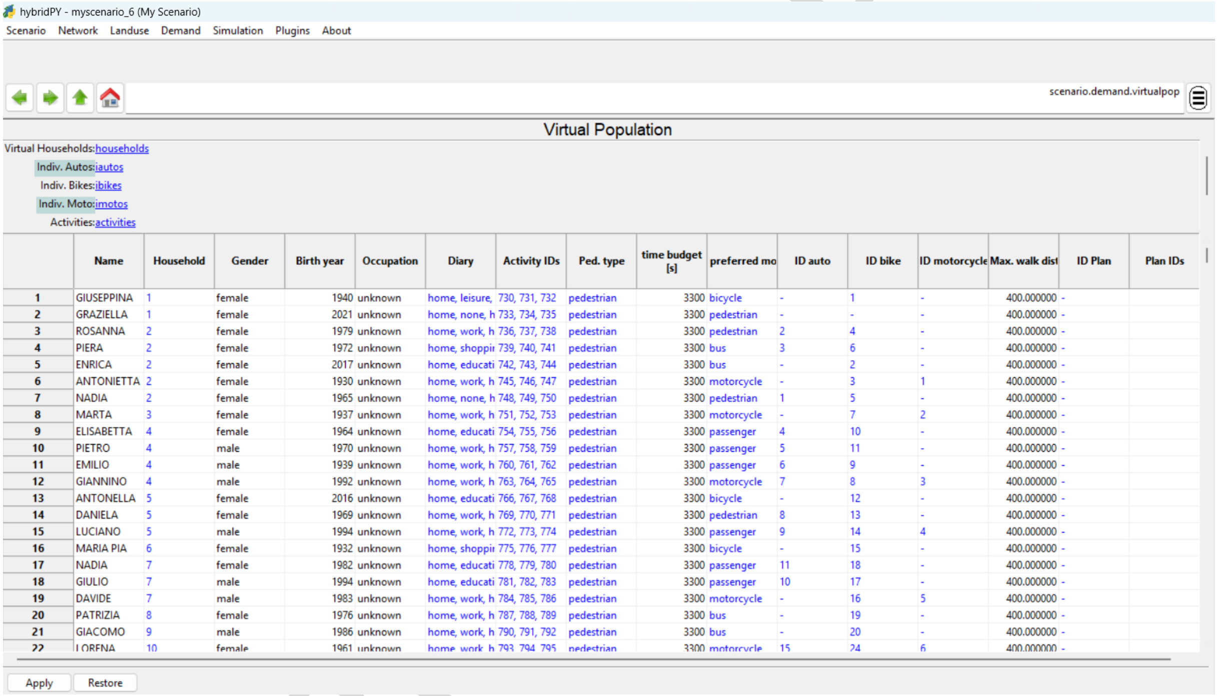
Processing Calculation

1. **Calculation 7 – Schedule daily activities (activity diary type)**
   Goal: Schedule a daily activity diary for each person by assigning a home-based activity chain with primary and (optional) secondary activities.
   How it works:
    - Define the set of activity types and classify them as primary (essential) and secondary (discretionary).
    - Construct home-based chains that start and end at home (e.g., home – primary – [secondary] – home).
    - Assign a primary diary chain based on age-group-specific shares (minors, adults, elders).
    - Optionally add one secondary activity using diary-specific probabilities (e.g., add shopping or leisure after the primary activity).
2. **Calculation 8 – Select activity locations (destination choice)**
   Goal: Select activity locations (facilities) that are consistent with activity purposes and feasible under each person's travel-time budget.
   How it works:
    - Map each activity purpose to eligible land-use types (e.g., education activities to education facilities).
    - Assign a daily travel time budget to each person using probabilistic sampling (e.g., normal distribution) with demographic adjustments.
    - Generate the isochrone (reachable area) by computing all reachable network edges and associated facilities within the travel time budget using shortest-path forward search.
    - Build the set of reachable facilities for the activity and compute travel time to each facility from the origin (home or previous activity).
    - Select a destination using an attractiveness-based probabilistic choice model that combines (i) access time proximity to the TTB and (ii) facility dynamic capacity.
    - Repeat destination selection for each activity in the chain (primary first, then secondary if present).
3. **Calculation 9 – Schedule activity times and trips within a 24-hour day**
   Goal: Generate a temporally feasible 24-hour schedule by assigning activity durations, start times, and trip times between consecutive activities.
   How it works:
    - Assign activity durations for primary and secondary activities using probabilistic sampling (e.g., normal distributions by activity type).
    - Determine feasible start times using temporal constraints from facility opening and closing hours (opening times may be sampled from distributions or if data available for each facility).

  - Compute departure time from home by subtracting the travel time to the first activity from its start time.
  - Schedule secondary activities sequentially after the primary activity by checking inter-activity travel time and opening-hour of secondary facilities.
  - Compute return time home and evening-at-home duration as the remaining time after all out-of-home activities and travel times, ensuring a complete 24-hour schedule.

Key Outputs

- Daily travel plan for each person: Daily activity chain, selected facility IDs and coordinates, preferred mode, trip travel times between consecutive activities, and a complete 24-hour schedule (departure time, activity start times, durations, return time)



Virtual Households:households
Indiv. Autos:iautos
Indiv. Bikes:ibikes
Indiv. Moto:imotos
Activities:activities

	Name	Household	Gender	Birth year	Occupation	Diary	Activity IDs	Ped. type	time budget [s]	preferred mo	ID auto	ID bike	ID motorcycle	Max. walk dist	ID Plan	Plan IDs
1	GIUSEPPINA	1	female	1940	unknown	home, leisure,	730, 731, 732	pedestrian	3300	bicycle	-	1	-	400.000000	-	
2	GRAZIELLA	1	female	2021	unknown	home, none, h	733, 734, 735	pedestrian	3300	pedestrian	-	-	-	400.000000	-	
3	ROSANNA	2	female	1979	unknown	home, work, h	736, 737, 738	pedestrian	3300	pedestrian	2	4	-	400.000000	-	
4	PIERA	2	female	1972	unknown	home, shoppi	739, 740, 741	pedestrian	3300	bus	3	6	-	400.000000	-	
5	ENRICA	2	female	2017	unknown	home, educati	742, 743, 744	pedestrian	3300	bus	-	2	-	400.000000	-	
6	ANTONIETTA	2	female	1930	unknown	home, work, h	745, 746, 747	pedestrian	3300	motorcycle	-	3	1	400.000000	-	
7	NADIA	2	female	1965	unknown	home, none, h	748, 749, 750	pedestrian	3300	pedestrian	1	5	-	400.000000	-	
8	MARTA	3	female	1937	unknown	home, work, h	751, 752, 753	pedestrian	3300	motorcycle	-	7	2	400.000000	-	
9	ELISABETTA	4	female	1964	unknown	home, educati	754, 755, 756	pedestrian	3300	passenger	4	10	-	400.000000	-	
10	PIETRO	4	male	1970	unknown	home, work, h	757, 758, 759	pedestrian	3300	passenger	5	11	-	400.000000	-	
11	EMILIO	4	male	1939	unknown	home, work, h	760, 761, 762	pedestrian	3300	passenger	6	9	-	400.000000	-	
12	GIANNINO	4	male	1992	unknown	home, work, h	763, 764, 765	pedestrian	3300	motorcycle	7	8	3	400.000000	-	
13	ANTONELLA	5	female	2016	unknown	home, educati	766, 767, 768	pedestrian	3300	bicycle	-	12	-	400.000000	-	
14	DANIELA	5	female	1969	unknown	home, work, h	769, 770, 771	pedestrian	3300	pedestrian	8	13	-	400.000000	-	
15	LUCIANO	5	male	1994	unknown	home, work, h	772, 773, 774	pedestrian	3300	passenger	9	14	4	400.000000	-	
16	MARIA PIA	6	female	1932	unknown	home, shoppi	775, 776, 777	pedestrian	3300	bicycle	-	15	-	400.000000	-	
17	NADIA	7	female	1982	unknown	home, educati	778, 779, 780	pedestrian	3300	passenger	11	18	-	400.000000	-	
18	GIULIO	7	male	1994	unknown	home, educati	781, 782, 783	pedestrian	3300	passenger	10	17	-	400.000000	-	
19	DAVIDE	7	male	1983	unknown	home, work, h	784, 785, 786	pedestrian	3300	motorcycle	-	16	5	400.000000	-	
20	PATRIZIA	8	female	1976	unknown	home, work, h	787, 788, 789	pedestrian	3300	bus	-	19	-	400.000000	-	
21	GIACOMO	9	male	1986	unknown	home, work, h	790, 791, 792	pedestrian	3300	bus	-	20	-	400.000000	-	
22	LORENA	10	female	1961	unknown	home, work, h	793, 794, 795	pedestrian	3300	motorcycle	15	24	6	400.000000	-	

Figure 7. Virtual population attributes and activity generation results

Validation checks

- Distributions of diary types by age group, travel time budgets, destination distributions by land-use type, and facility occupancy/capacity diagnostics.

4. Case Study

4.1 Introduction and key model inputs

The proposed framework is demonstrated using Bologna (Italy) as a practical test case for generating daily activity-based travel demand. The core simulation area is represented by 2,060 zones and a total population of 381,574 residents, where the virtual population and daily activity schedules are generated.

Table 4. *Key data inputs and parameters for generating the Bologna daily ABM*

Parameters	**Values**
Traffic analysis zones and population	2,060 zones; population: 381,574
Household composition	1 (27.4%), 2 (13.6%), 3 (7.1%), 4 (4.1%), 5 (1.0%), 6+ (0.4%)
Gender distribution	48% male; 52% female
Household size	Mean size: 2.4 persons/household
Vehicle ownership per 100 people	63 cars; 14 motorcycles; 90 bicycles
City mode share (%)	35.6% car; 10.6% motorcycle; 25.6% bus ; 6.9% bicycle ; 21.0% walk
Mean travel time budget (min)	45 min (+5 min for workers and males)
Minor diary shares (%)	2% h–w–h; 65% h–e–h; 15% h–l–h; 10% h–s–h; 8% h–o–h
Elderly diary shares (%)	15% h–w–h; 5% h–e–h; 40% h–l–h; 20% h–s–h; 20% h–o–h
Adult diary shares (%)	70% h–w–h; 15% h–e–h; 5% h–l–h; 5% h–s–h; 5% h–o–h
Secondary activity probability (%)	40% h–w–h; 50% h–e–h; 50% h–l–h; 50% h–s–h; 0% h–o–h
Mean activity duration (h)	Home 7; work 8; education 7; leisure 4; shopping 3; other 5
Land-use facilities (count; area km^2)	Res. 72,625 (18.0); Com. 456 (2.0); Ind. 6,409 (7.8); Edu. 281 (0.2); Lei. 1,278 (6.5); Mixed 3,516 (2.0)
Typical opening–closing hours	Ind. 7:00–17:00; Com. 9:00–20:00; Edu. 8:00–17:00; Lei. 10:00–20:00; Res. 7:00–23:00; Mixed 8:00–20:00

Note: h = home, w = work, e = education, l = leisure, s = shopping, o = other.

Data inputs combine aggregated population, travel, and land-use data from multiple open sources (city-level, national statistics, or adapted regional datasets). Key socio-demographic assumptions include an average household size of 2.4 persons, a gender split of 48% male and 52% female [26]. and vehicle ownership of 63 cars per 100 inhabitants [27]. The daily travel time budget of 45 minutes is adopted from Italian statistics [28] with group-based adjustments.

Land-use data are extracted from OSM and structured into 84,565 facilities with type, coordinates, and capacity. Facility capacity is estimated from building footprint areas. Opening hours are assigned based on typical operating patterns inferred from OSM POI attributes.

4.2 Modeling Results and Discussions

Generating household sizes and allocating households to residential facilities (Step 9) yields a Bologna synthetic population across 2,060 zones comprising 160,788 households and 376,747 virtual persons (average household size: 2.34), dominated by 52,497 single-person households, followed by 44,286 two-person, 30,753 three-person, and 24,930 four-person households, with relatively few larger households (6,233 five-person and 2,089 six-or-more-person households).

Vehicle ownership rises with both household income and size, and the synthetic fleet matches citywide baseline rates of about 90 bicycles, 63 cars, and 14 motorcycles per 100 residents.

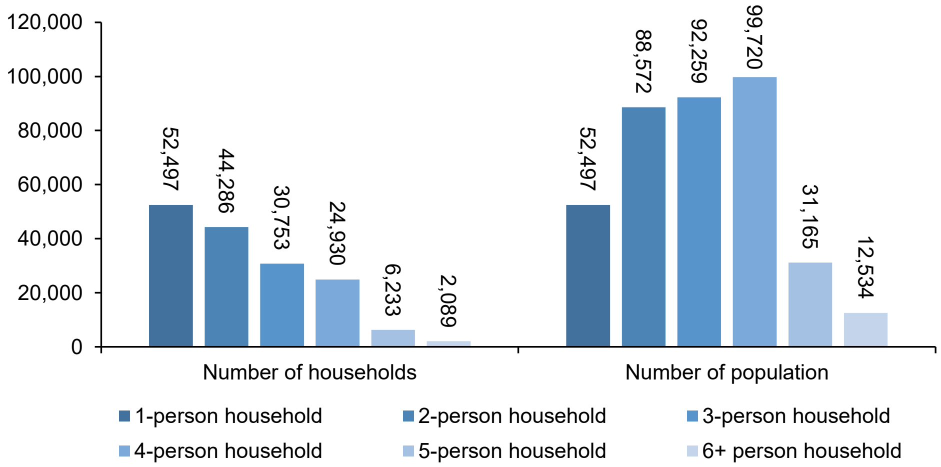


Figure 8. *Distribution of generated households and population by household size*

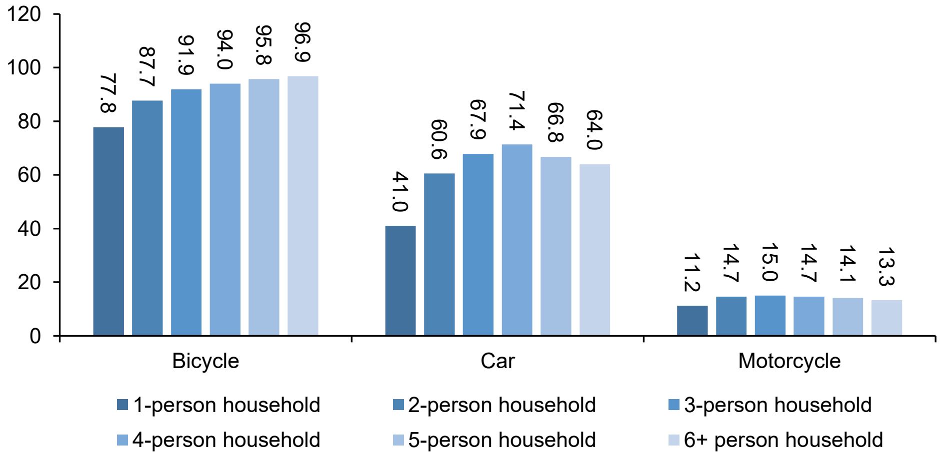


Figure 9. *Distribution of generated vehicle ownership rates by household size*

For computational efficiency, we simulated a 50% demand scenario comprising 188,014 individuals. Among 188,014 agents, activity chains are dominated by work, followed by education, while leisure and shopping account for smaller shares. About half of agents include a secondary stop—typically shopping or leisure—before returning home, indicating realistic tour extensions beyond the primary purpose and aligning with observed statistics.

Simulated activity durations for all agents differ significantly between tours with and without secondary activities (Figure 11). Single-purpose tours align with the input mean durations (work-related: 7.71 hours; education-related: 6.93 hours; shopping-related: 2.81 hours; leisure-related: 3.77 hours), while adding secondary activities consistently extends durations (work-related: 11.13 hours; education-related: 10.62 hours; shopping-related: 6.87 hours; leisure-related: 6.86 hours). This increase is expected, since inserting discretionary activities between the main activities lengthens the overall tour.

For each agent, we modeled time use over a 24-hour day (00:00–24:00), covering both activity participation and travel between locations. Figure 12 shows the time-use profiles of 165,042 agents ordered by age. It is noted that the results represent the allocation of total activity time over 24 hours; the time spent at home in the morning

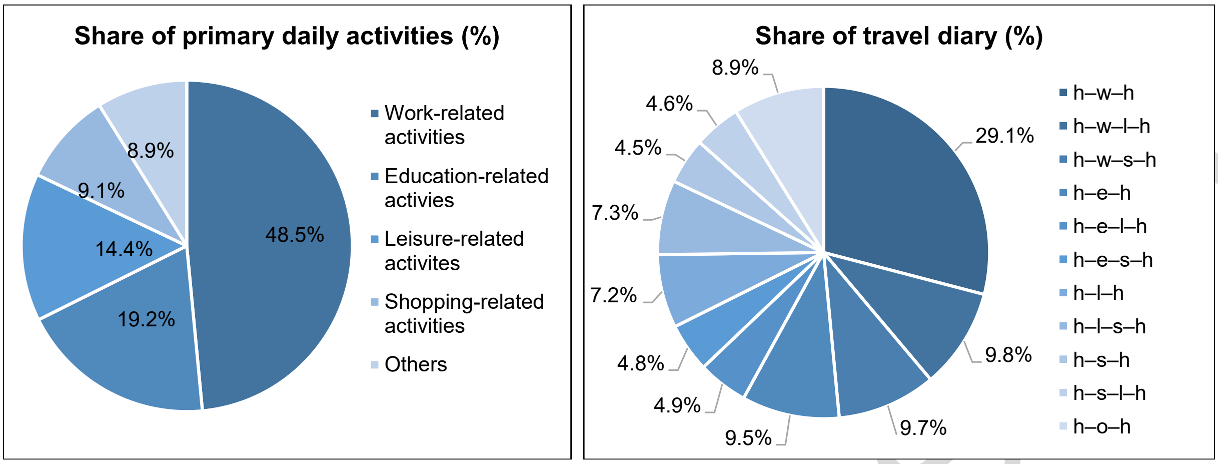
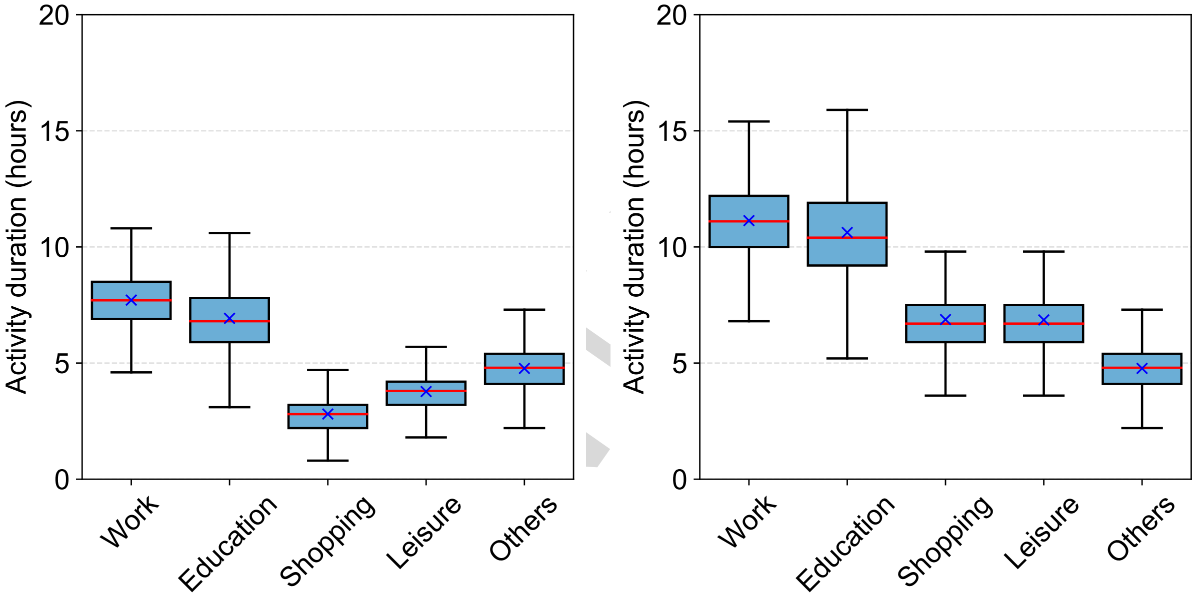


Figure 10. *Reconstructed primary daily activities (left) and daily travel diaries (right) in the Bologna daily ABM model*



(a) Only primary activities.

(b) Secondary activities included.

Figure 11. *Comparison of simulated daily activity durations with and without secondary activities*

and evening is aggregated. Home time dominates the day, particularly for older agents who mainly undertake short out-of-home activities (e.g., leisure and shopping), whereas working-age adults and children devote most non-home time to work and education, respectively.

Figure 13 compares the simulated and observed traffic volumes from 605 detectors in the core simulation area on typical weekdays in October 2022. The simulation reproduces the observed diurnal pattern across peak and off-peak periods: morning peak shares align closely, and the afternoon peak is only slightly underestimated (about 1–1.5%), a negligible difference.

The above results indicate that the model can reconstruct households, synthesize a virtual population, and generate daily plans consistent with observed statistics, while minimizing reliance on costly survey- and interview-based travel diary data, highlighting its potential as a powerful and practical tool for large-scale demand modeling.

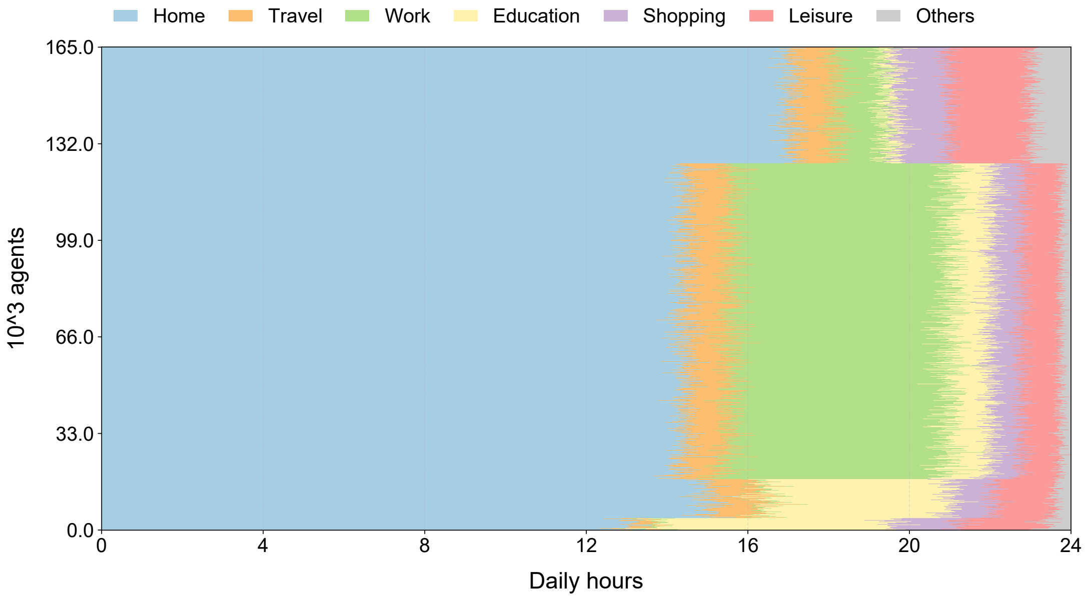


Figure 12. *Daily activity-time allocation for modeled agents*

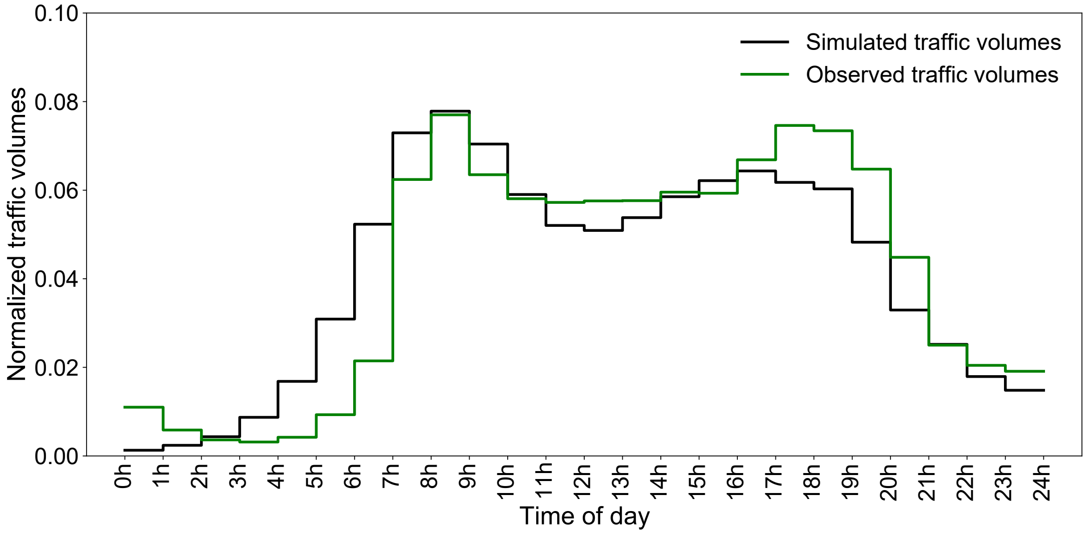


Figure 13. *Comparison of normalized simulated and observed traffic volumes across the Bologna study area over a 24-hour period*

5. Conclusions

The daily activity-based demand generation framework is integrated into hybridPy/SUMO and implemented through three sequential sub-modules. First, it reconstructs a synthetic household population by sampling household sizes and assigning demographic and socioeconomic attributes to match census distributions. Second, these attributes are disaggregated to individuals, and vehicle ownership and preferred travel modes are allocated to reproduce citywide ownership rates and modal shares while preserving heterogeneity across income, household size, age, and gender. Third, the model generates complete daily activity-travel diaries, including primary activities and optional secondary activities by accounting for facility capacity, individual travel time budgets, and facility opening/closing hours. The resulting agent-based travel demand can then be used as input to dynamic traffic simulation step.

A key contribution of this work is the step-by-step implementation guide of a daily activity-based travel demand module in hybridPy/SUMO, which clearly specifies the required datasets, key input parameters, computational procedures, and expected outputs at each stage. This structured workflow enables users to independently reproduce the full demand generation pipeline and develop baseline and "what-if" scenarios with minimal manual intervention.

The framework outputs detailed individual daily travel plans (e.g, home and activity locations, activity start times and durations, travel modes, and tour structures) suitable for large-scale microscopic traffic simulation in hybridPy. This allows traffic researchers and planners to efficiently leverage open-source data to support large-scale scenario assessment and urban digital twin applications through a reproducible workflow.

Author contributions

Conceptualization, J.S. and N.A.N.; methodology, N.A.N. and C.P; software, N.A.N. and C.P; validation, N.A.N.; formal analysis, N.A.N.; investigation, N.A.N.; resources, N.A.N. and J.S; data curation, N.A.N.; writing—original draft preparation, N.A.N., and J.S.; writing—review and editing, J.S. and F.R; visualization, N.A.N.; supervision, J.S. and F.R; project administration, J.S; and funding acquisition, J.S; All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

This research has been funded by the Italian PNRR program.

Acknowledgements

The authors are thankful for useful discussions with the SUMO team.

References

[1] E. Cascetta, *Transportation Systems Engineering: Theory and Methods*. Boston, MA: Springer US, 2001, ISBN: 978-0-7923-6792-4. DOI: 10.1007/978-1-4757-6873-2.

[2] M. Mladenovic and A. Trifunovic, "The shortcomings of the conventional four step travel demand forecasting process," *Journal of Road and Traffic Engineering*, vol. 60, no. 1, pp. 5–12, 2014.

[3] T. Adler and M. Ben-Akiva, "A theoretical and empirical model of trip chaining behavior," *Transportation Research Part B: Methodological*, vol. 13, no. 3, pp. 243–257, 1979. DOI: 10.1016/0191-2615(79)90016-X.

[4] R. Moeckel, N. Kuehnel, C. Llorca, A. T. Moreno, and H. Rayaprolu, "Agent-based simulation to improve policy sensitivity of trip-based models," *Journal of Advanced Transportation*, vol. 2020, no. 1, p. 1 902 162, 2020. DOI: 10.1155/2020/1902162.

[5] F. F. Bastarianto, T. O. Hancock, C. F. Choudhury, and E. Manley, “Agent-based models in urban transportation: Review, challenges, and opportunities,” *European Transport Research Review*, vol. 15, no. 1, p. 19, 2023. DOI: 10.1186/s12544-023-00590-5.

[6] N. A. Nguyen, J. Schweizer, and F. Rupi, “Large-scale activity-based demand generation modeling: A literature review and exploration of potential approaches,” *Transportation Engineering*, vol. 20, p. 100 329, 2025, ISSN: 2666-691X. DOI: https://doi.org/10.1016/j.treng.2025.100329.

[7] M. Balmer, K. W. Axhausen, and K. Nagel, “Agent-based demand-modeling framework for large-scale microsimulations,” *Transportation Research Record*, vol. 1985, no. 1, pp. 125–134, 2006. DOI: 10.3141/1985-14.

[8] M. Zilske, A. Neumann, and K. Nagel, “Eqasim: An activity-based equilibrium simulation framework,” in *Proceedings of the Transport Research Arena (TRA)*, Warsaw, Poland, 2016.

[9] A. Neumann and M. Zilske, “Eqasim: An open-source framework for activity-based demand modeling integrated with matsim,” Technische Universität Berlin, Berlin, Germany, Tech. Rep., 2018.

[10] A. Horni, K. Nagel, and K. W. Axhausen, “The multi-agent transport simulation matsim,” *Ubiquity Press*, 2016. DOI: 10.5334/baw.

[11] ActivitySim Consortium, *Activitysim: An open platform for activity-based travel modeling*, https://activitysim.github.io, Open-source software documentation and model framework, 2018.

[12] B. M. Paul, J. Doyle, B. Stabler, J. Freedman, and A. Bettinardi, “Multi-level population synthesis using entropy maximization-based simultaneous list balancing,” Tech. Rep., 2018.

[13] W. Davidson, P. Vovsha, J. Freedman, and R. Donnelly, “Ct-ramp family of activity-based models,” in *Proceedings of the 33rd Australasian transport research forum (ATRF)*, vol. 29, 2010, p. 29.

[14] H. Laarabi, Z. Needell, R. Waraich, C. Poliziani, and T. P. Wenzel, “A modeling framework for behavior, energy, autonomy and mobility (beam),” Tech. Rep., Feb. 2024. [Online]. Available: https://eta-publications.lbl.gov/publications/behavior-energy-autonomy-mobility.

[15] C. A. Spurlock, M. A. Bouzaghrane, A. Brooker, *et al.*, “Behavior, energy, autonomy & mobility comprehensive regional evaluator: Overview, calibration and validation summary of an agent-based integrated regional transportation modeling workflow,” Berkeley, Tech. Rep., Feb. 2024. [Online]. Available: https://eta-publications.lbl.gov/publications/behavior-energy-autonomy-mobility.

[16] F. Hülsmann, R. Gerike, B. Kickhöfer, and K. Nagel, “Tapas: An activity-based travel demand model for integrated transport and emission simulation,” in *Proceedings of the Transport Research Arena (TRA)*, European Commission, Paris, France, 2014.

[17] K. Nagel, F. Hülsmann, and B. Kickhöfer, “Integrated transport simulation with tapas and microscopic traffic models,” German Aerospace Center (DLR), Berlin, Germany, Tech. Rep., 2016.

[18] R. Moeckel, W.-C. Huang, J. Ji, *et al.*, “The activity-based model abit: Modeling 24 hours, 7 days a week,” *Transportation Research Procedia*, vol. 78, pp. 499–506, 2024, 25th Euro Working Group on Transportation Meeting, ISSN: 2352-1465. DOI: 10.1016/j.trpro.2024.02.062.

[19] R. Moeckel, “Constraints in household relocation: Modeling land use/transport interactions that respect time and monetary budgets,” *Journal of Transport and Land Use*, vol. 10, no. 1, pp. 211–228, 2017. DOI: 10.5198/jtlu.2015.810.

[20] A. Horni, K. Nagel, and K. Axhausen, *The multi-agent transport simulation title of book: The multi-agent transport simulation matsim subtitle positioned below*, 2016. DOI: 10.5334/baw.

[21] Eclipse SUMO Project, *Saga: Sumo activity generation application*, https://sumo.dlr.de, Accessed documentation and software reference, 2020.

[22] OpenStreetMap, *Openstreetmap (osm)*, Accessed 10 February 2026, 2023. [Online]. Available: https://www.openstreetmap.org.

[23] Eclipse, *Sumo netconvert*, Accessed 10 February 2026, 2023. [Online]. Available: https://sumo.dlr.de/docs/netconvert.html.

[24] E. SUMO, *Sumo netedit*, Accessed 10 February 2026, 2023. [Online]. Available: https://sumo.dlr.de/docs/Netedit/index.html.

[25] MobilityData, *General transit feed specification – schedule reference*, Accessed 10 February 2026, 2025. [Online]. Available: https://gtfs.org/documentation/schedule/reference/.

[26] Urbistat. “Demographic balance, population and familiy trends, age classes and average age, civil status and foreigners in bologna.” Accessed: 20 October 2025. (2023), [Online]. Available: https://ugeo.urbistat.com/AdminStat/en/it/demografia/eta/bologna/37/3.

[27] Servizio Studi e Statistica. “Parco veicolare nella città metropolitana di bologna – anno 2023.” Accessed: 2025-10-29. (2024), [Online]. Available: https://inumeridibolognametropolitana.it/studi-e-ricerche/parco-veicolare-nella-citta-metropolitana-di-bologna-anno-2023.

[28] J. Armoogum, S. Borgato, D. Fiorello, *et al.*, “Study on new mobility patterns in european cities: Eu-wide passenger mobility survey,” IFSTTAR – Institut Français des Sciences et Technologies des Transports, de l’Aménagement et des Réseaux, Tech. Rep., 2022, HAL Id: hal-03894185, p. 166. DOI: 10.2832/728583.