Crearío — Creating MATSim Models for any place on earth

Nadine Rieser, Marcel Rieser
Simunto GmbH

Simunto

Creating MATSim Models is challenging

- A lot of different data is required
- Different countries have different data, and different data formats
- Data needs to be validated and cleaned,
  outliers removed, data errors fixed

Creating MATSim Models on global data

Is it possible to create (good) MATSim Models based on data available worldwide, instead of country-specific data?

- One generic data-processing pipeline
- Well-defined stages where local data can be integrated
- Highest possible re-use of best practices and lessons learned

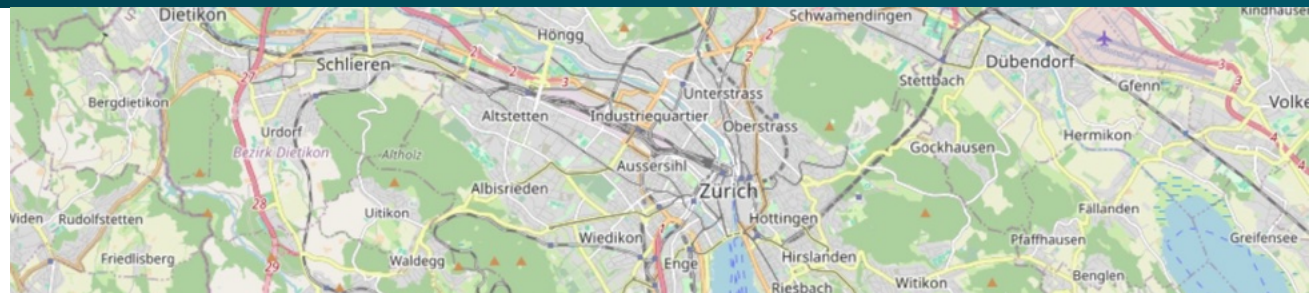
Crearios

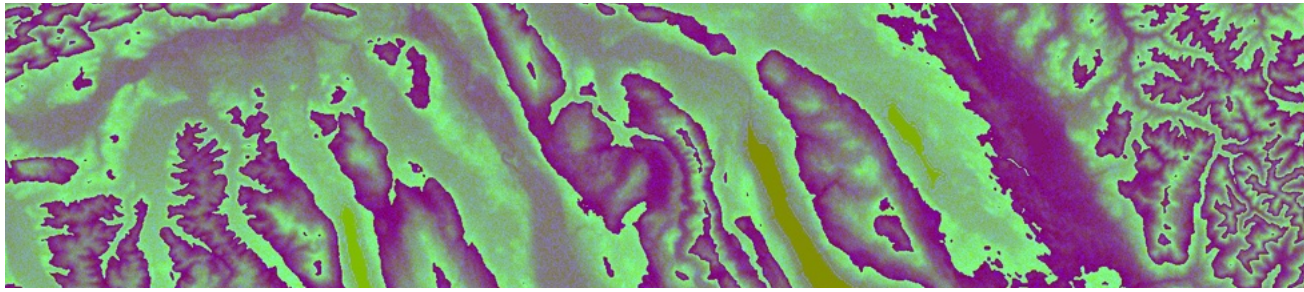
Provide users a MATSim scenario that is

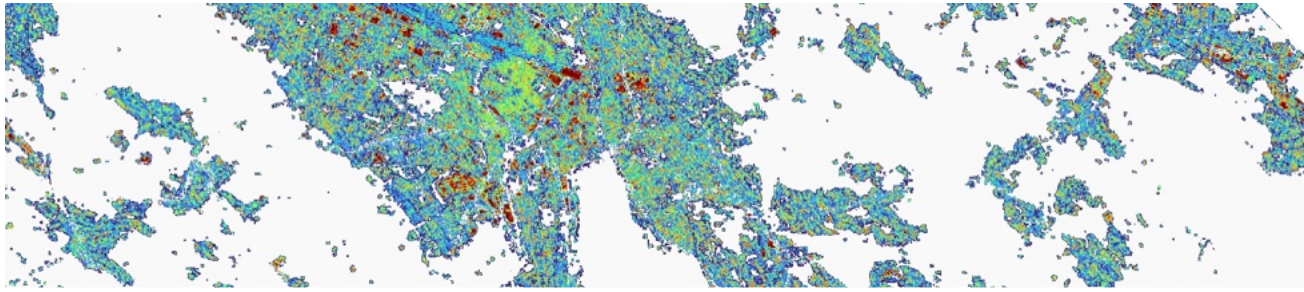
- Ready for simulation
- Based on worldwide publicly available data
- Integrates optional local user inputs
- Algorithmically validated
- Accompanied by an extensive validation report and further validation outputs

Network – Input data

- OpenStreetMap (main source)
- AWS Terrain Tiles for elevations
- Global Human Settlement Layers for location class
- Norms for capacities







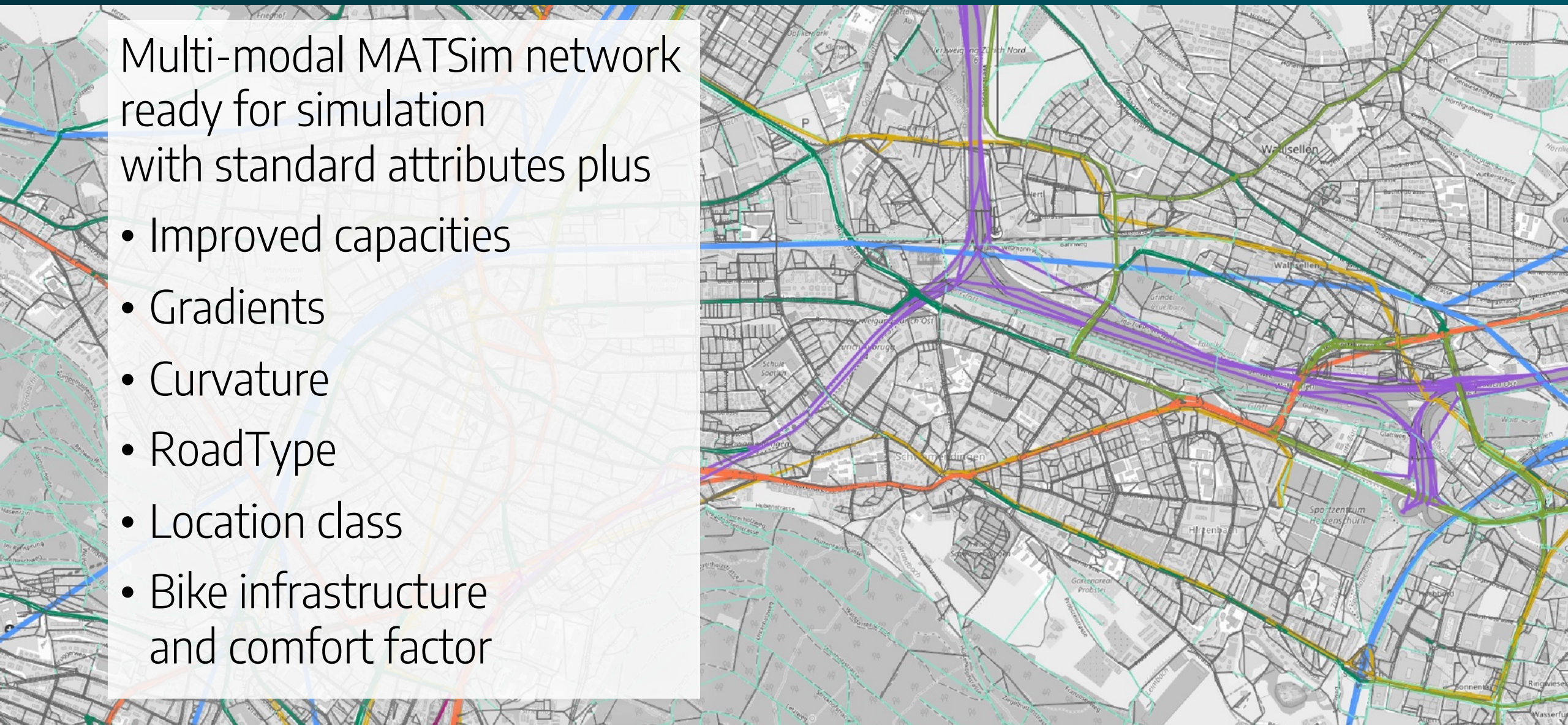
Capacité C [véh/h] d'une autoroute à deux voies de circulation en fonction de la vitesse maximale autorisée V_{aut}, du pourcentage de véhicules lourds SVA et de la rampe selon [15]

Schwerverkehrsanteil SVA *Pourcentage de véhicules lourds SVA*	V_{zul} / V_{aut}								
	120 km/h			100 km/h			80 km/h		
	Steigung / *Rampe*								
	< 2%	2…4%	> 4%	< 2%	2…4%	> 4%	< 2%	2…4%	> 4%
≤ 5%	4000	3800	3550	4000	3800	3600	4000	3800	3650
> 5…15%	3800	3500	3150	3800	3600	3350	3800	3700	3450

Network – Deliverables

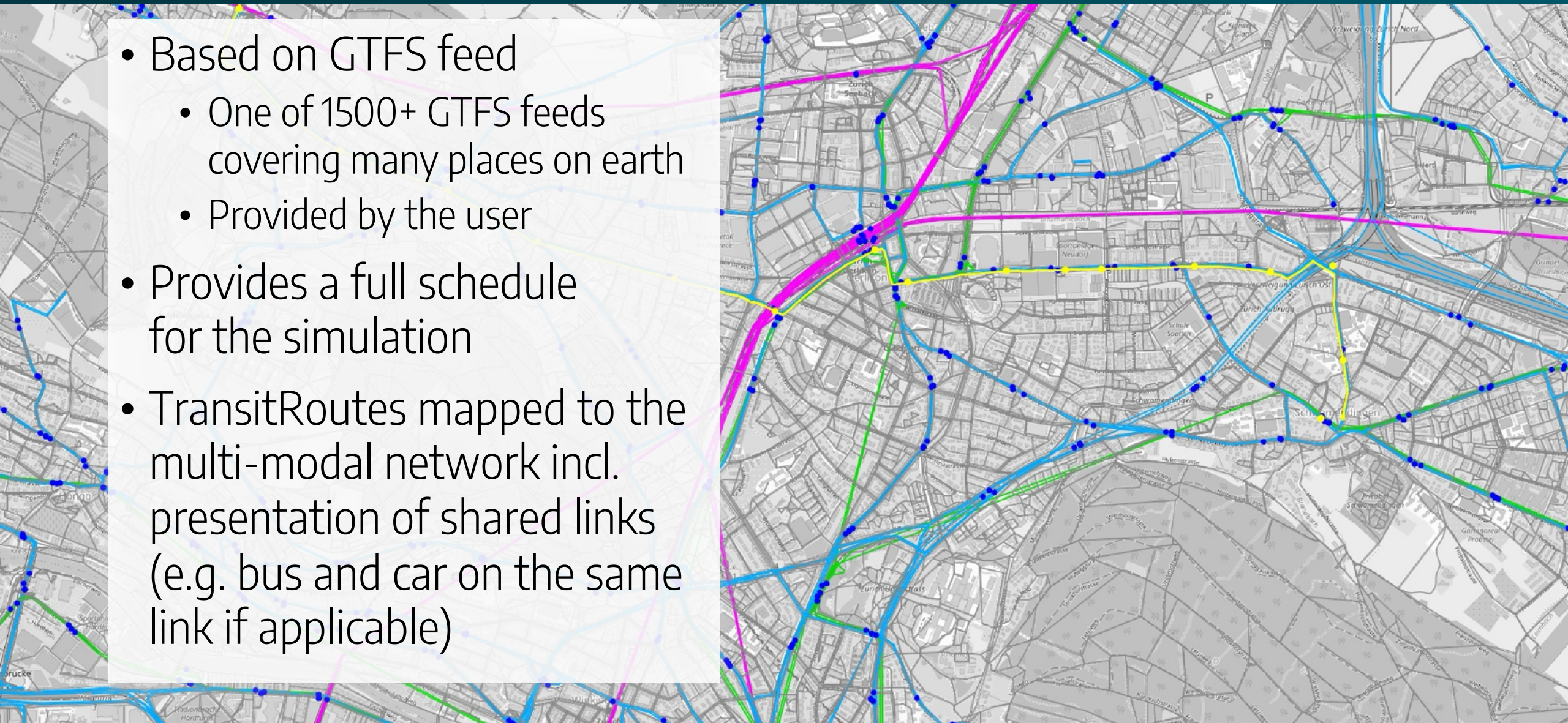
Multi-modal MATSim network
ready for simulation
with standard attributes plus

- Improved capacities
- Gradients
- Curvature
- RoadType
- Location class
- Bike infrastructure
  and comfort factor



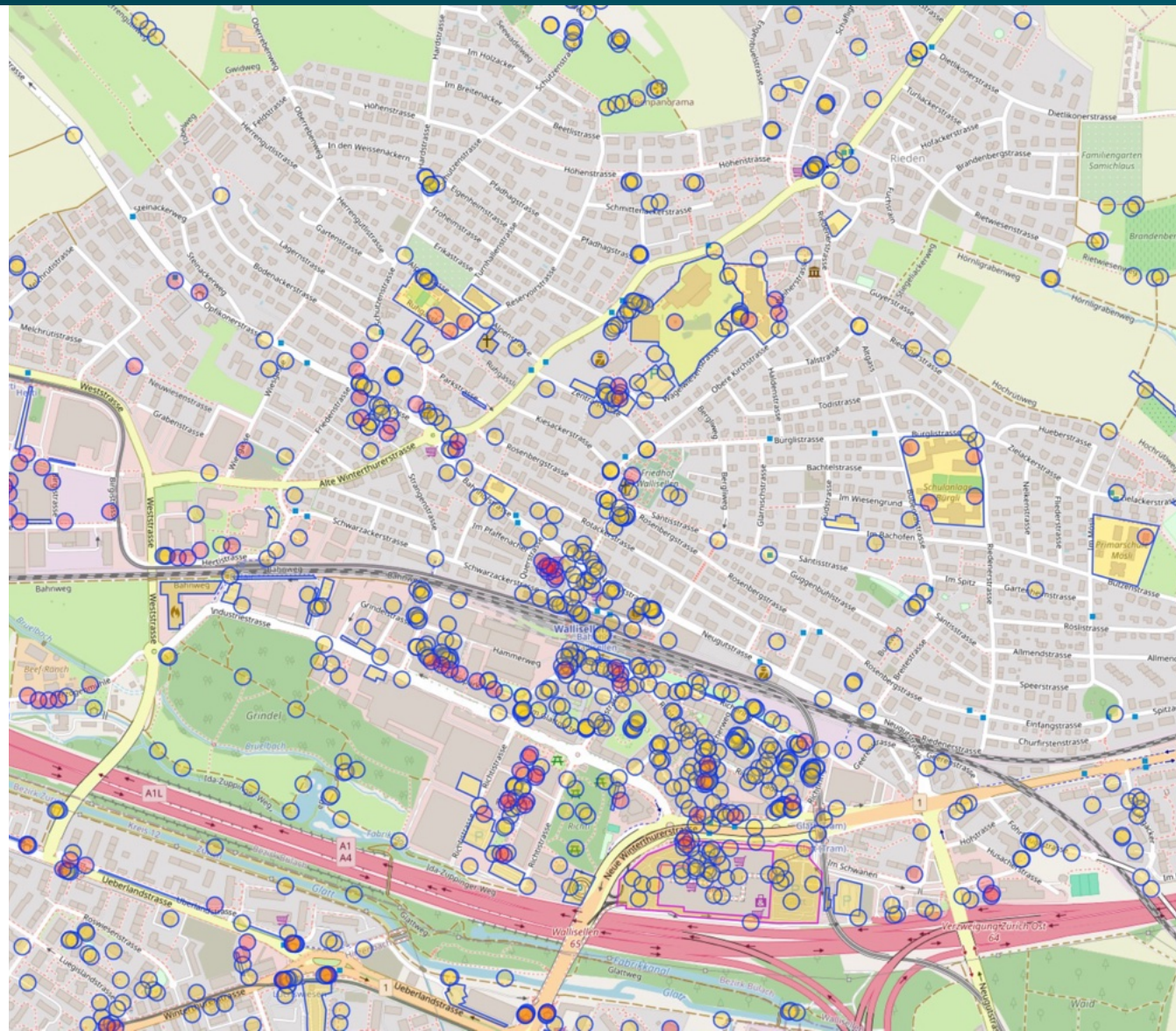
Schedule

- Based on GTFS feed
  - One of 1500+ GTFS feeds covering many places on earth
  - Provided by the user
- Provides a full schedule for the simulation
- TransitRoutes mapped to the multi-modal network incl. presentation of shared links (e.g. bus and car on the same link if applicable)

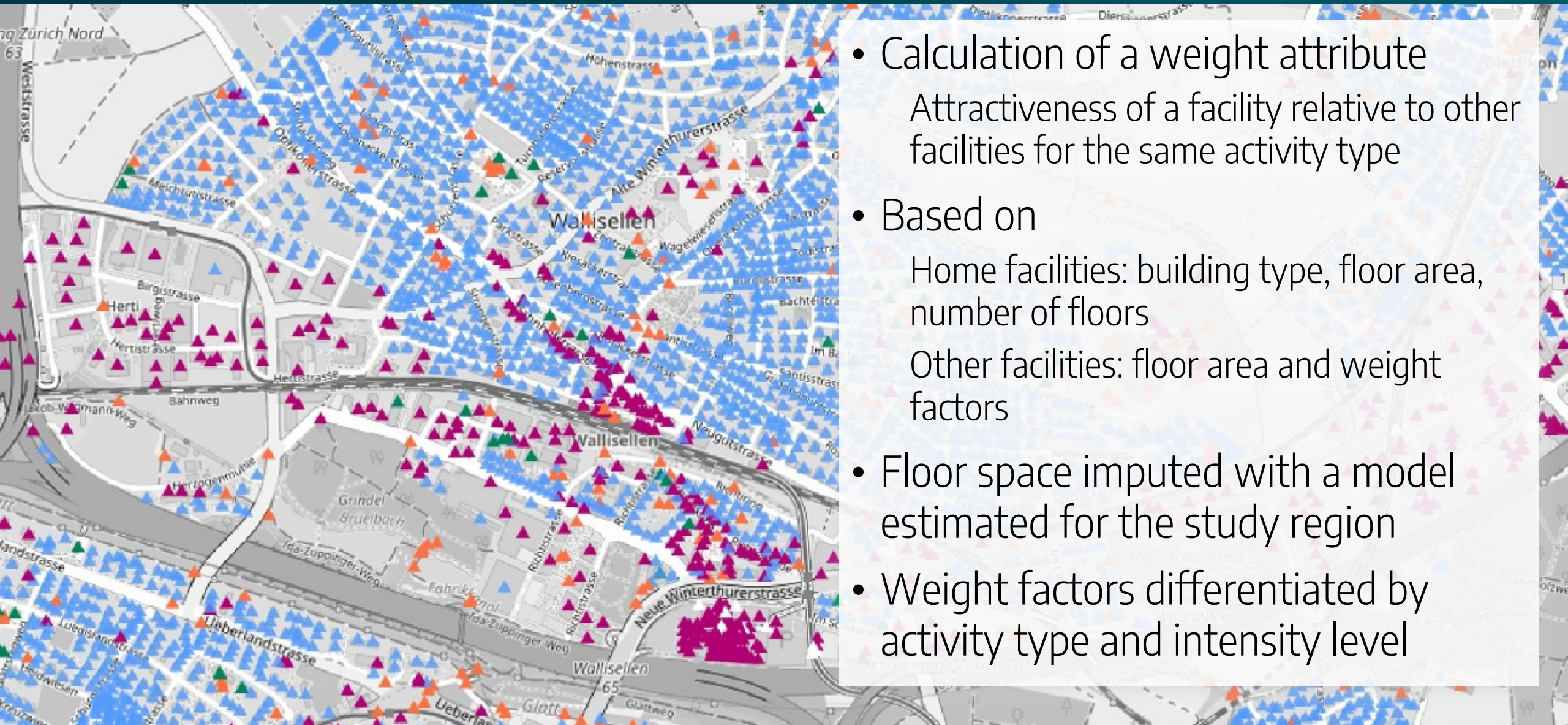


Facilities – Input data

- Data source: OpenStreetMap
- Creating facilities for POIs and Buildings,
- Land-use areas used when no POIs or buildings available
- Using a large variety of tags (e.g. amenity, shop, office, leisure, healthcare)
- Opening times included if tagged in OSM



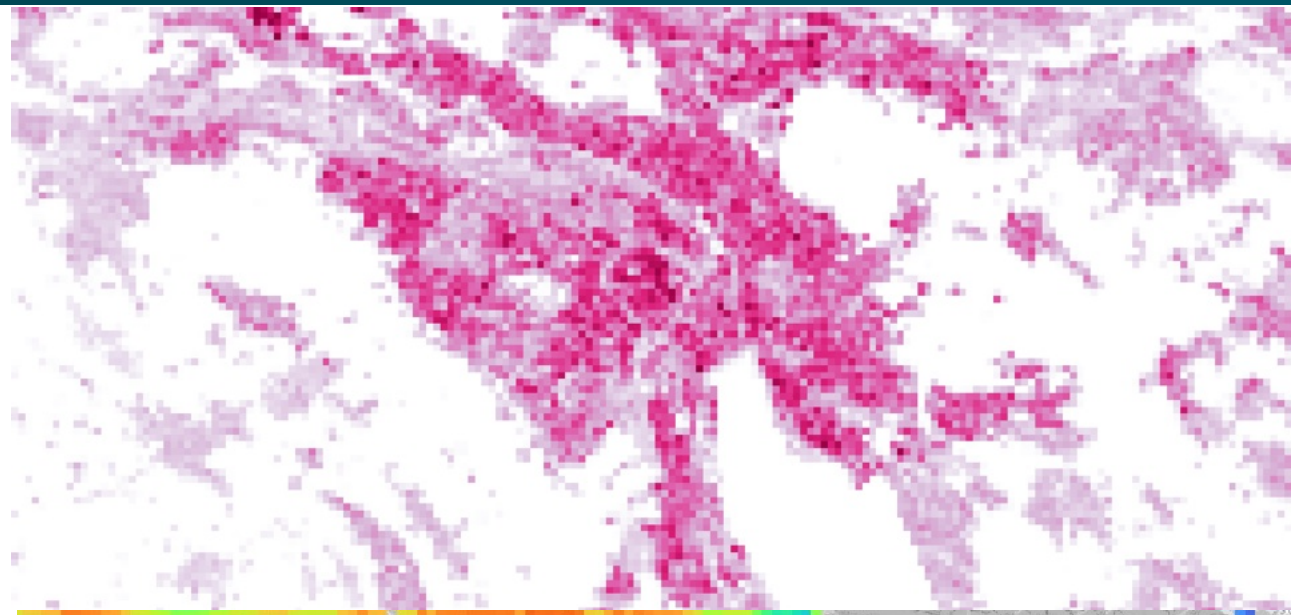
Facilities – Deliverables

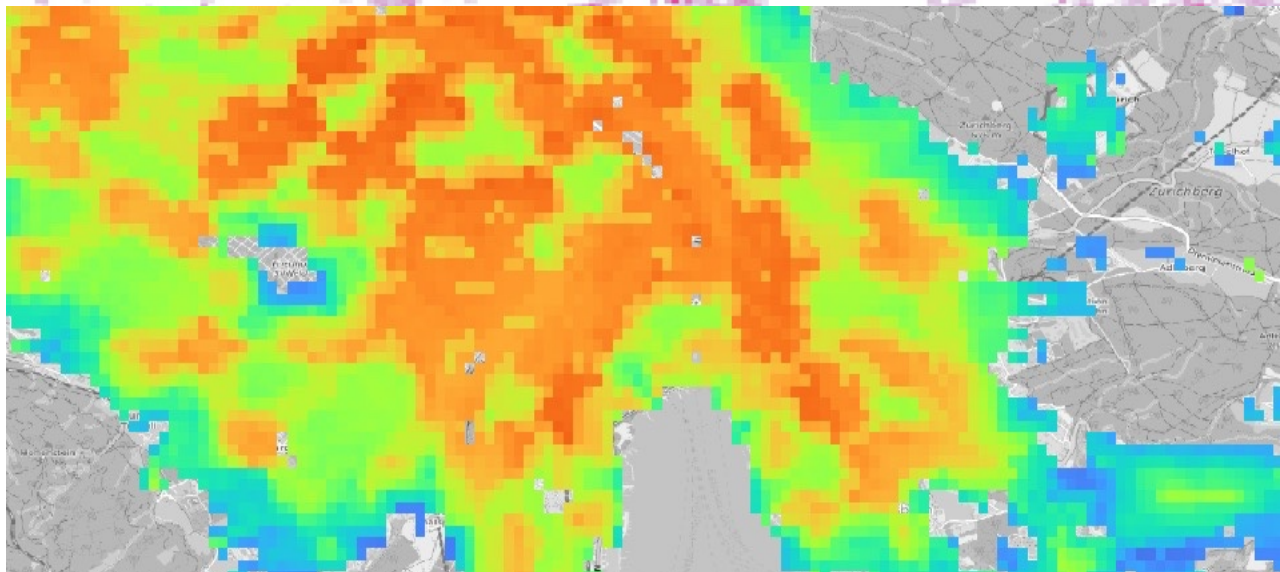


- Calculation of a weight attribute
  - Attractiveness of a facility relative to other facilities for the same activity type
- Based on
  - Home facilities: building type, floor area, number of floors
  - Other facilities: floor area and weight factors
- Floor space imputed with a model estimated for the study region
- Weight factors differentiated by activity type and intensity level

Population – Input data

- Global Human Settlement Layer for number of inhabitants
- Population data by Worldpop for age and gender distribution
- Users can upload own population statistics (incl. age and gender if available)

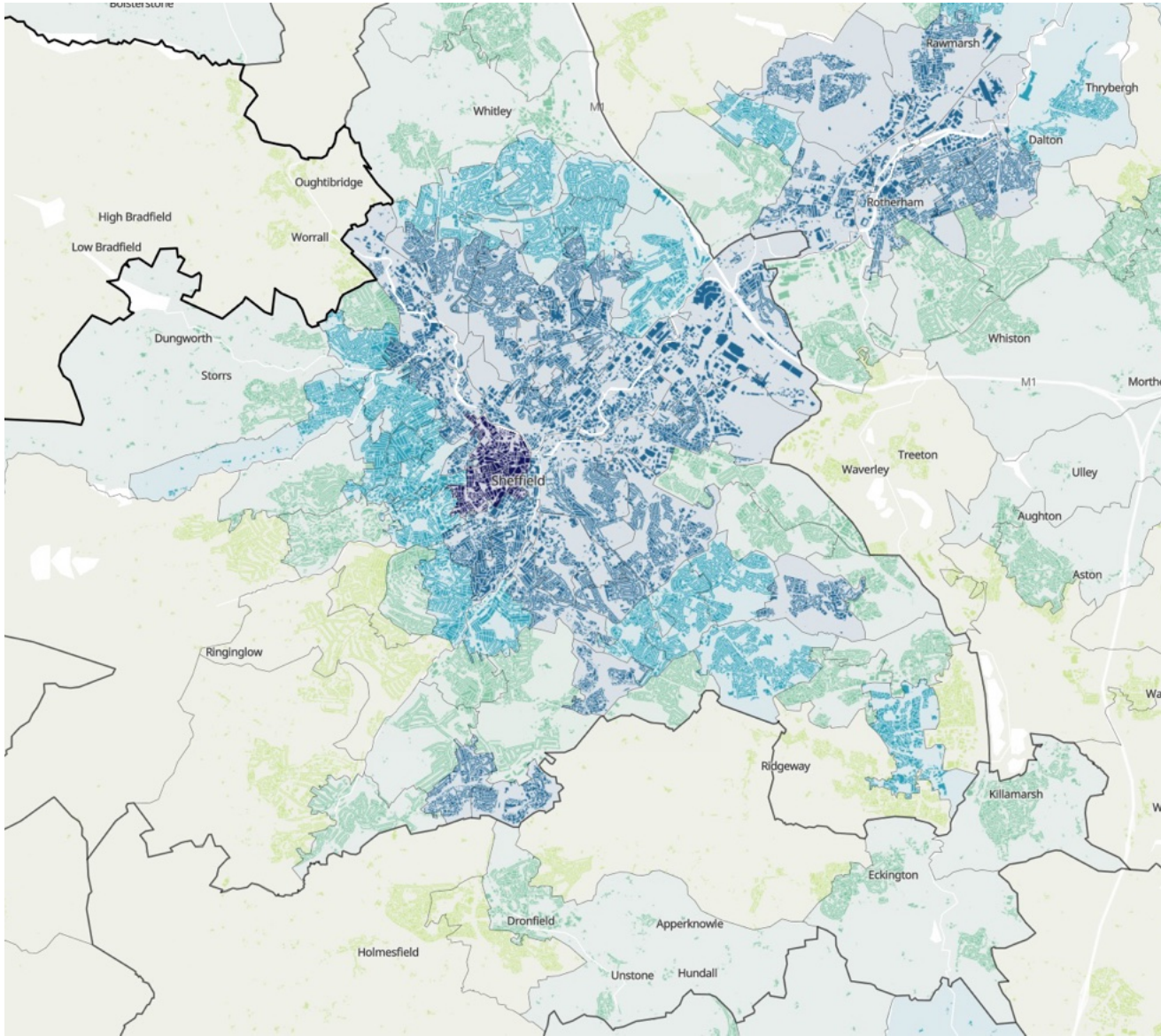




Population – Input data

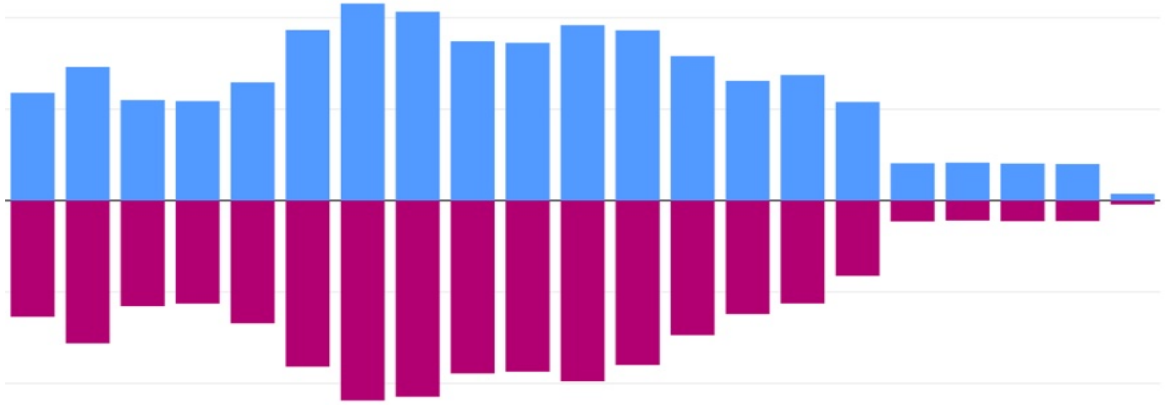
Additional attributes for better behaviour choice:

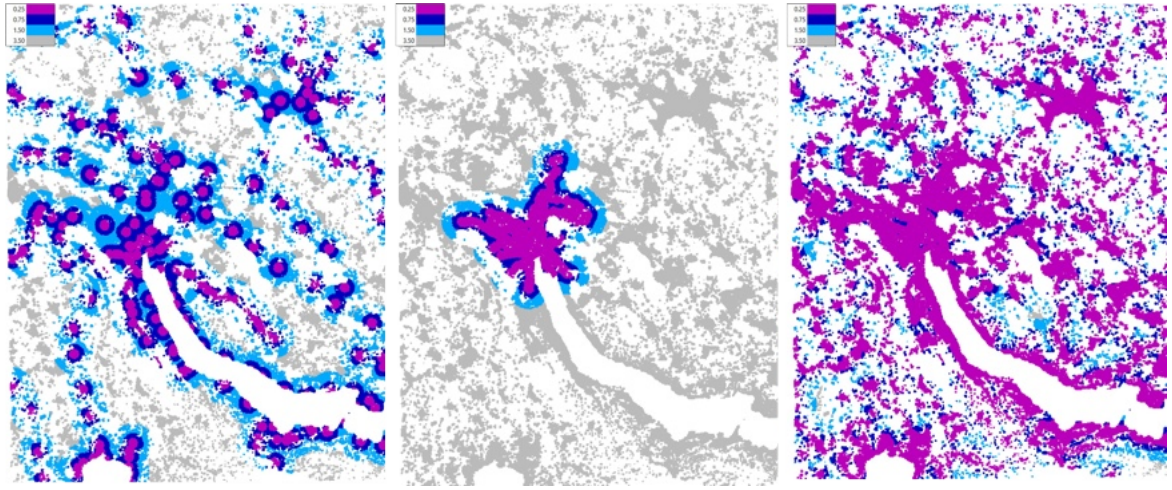
- Global Human Settlement Layer for degree of urbanisation
- GTFS feed or OSM for public transport attributes
- UK census maps for vehicle ownership model



Population – Deliverables

- Fitted to match statistics provided by the user (optional)
- Home facility
- Age
- Gender
- Land-use type of the home location
- Public transport quality class
- Distance to public transport stops
- Vehicle ownership
- Initial daily plan

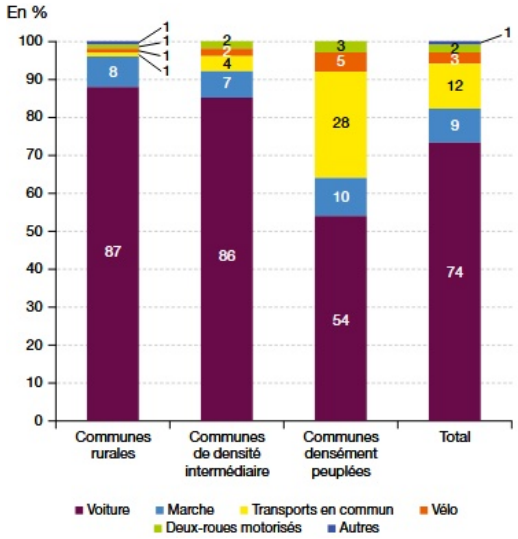


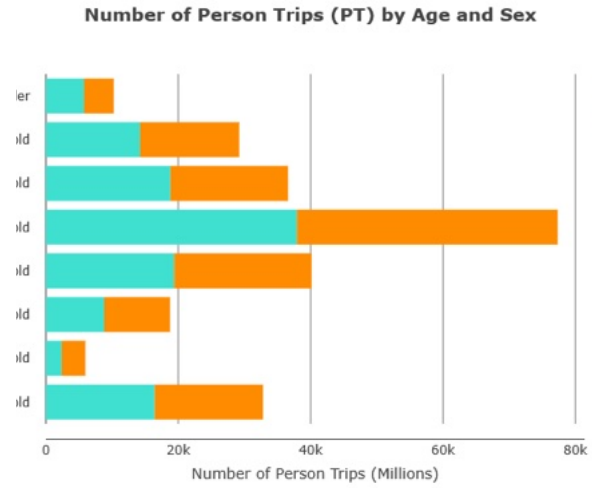


Plans – Input data

Travel diaries

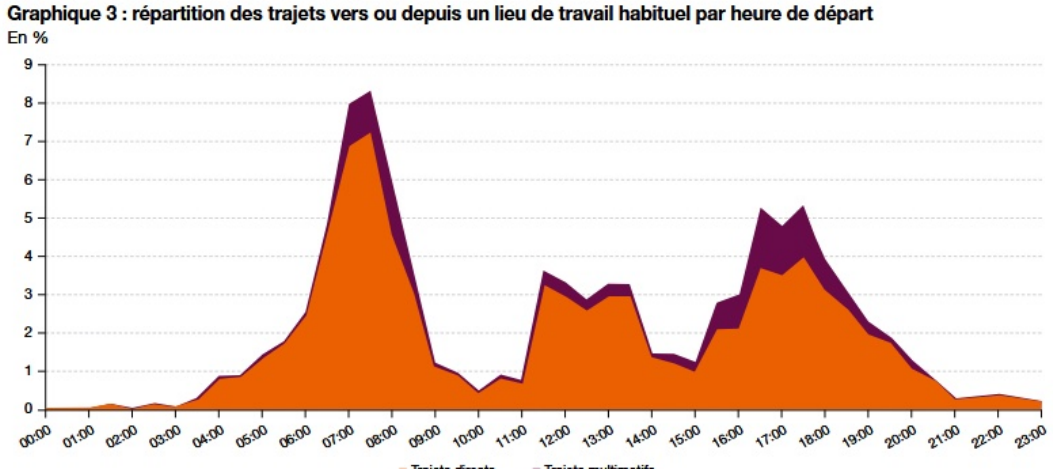
- French National Household Survey 2019
- US National Household Survey 2022





Using houshold and person characteristics to assign plans to agents

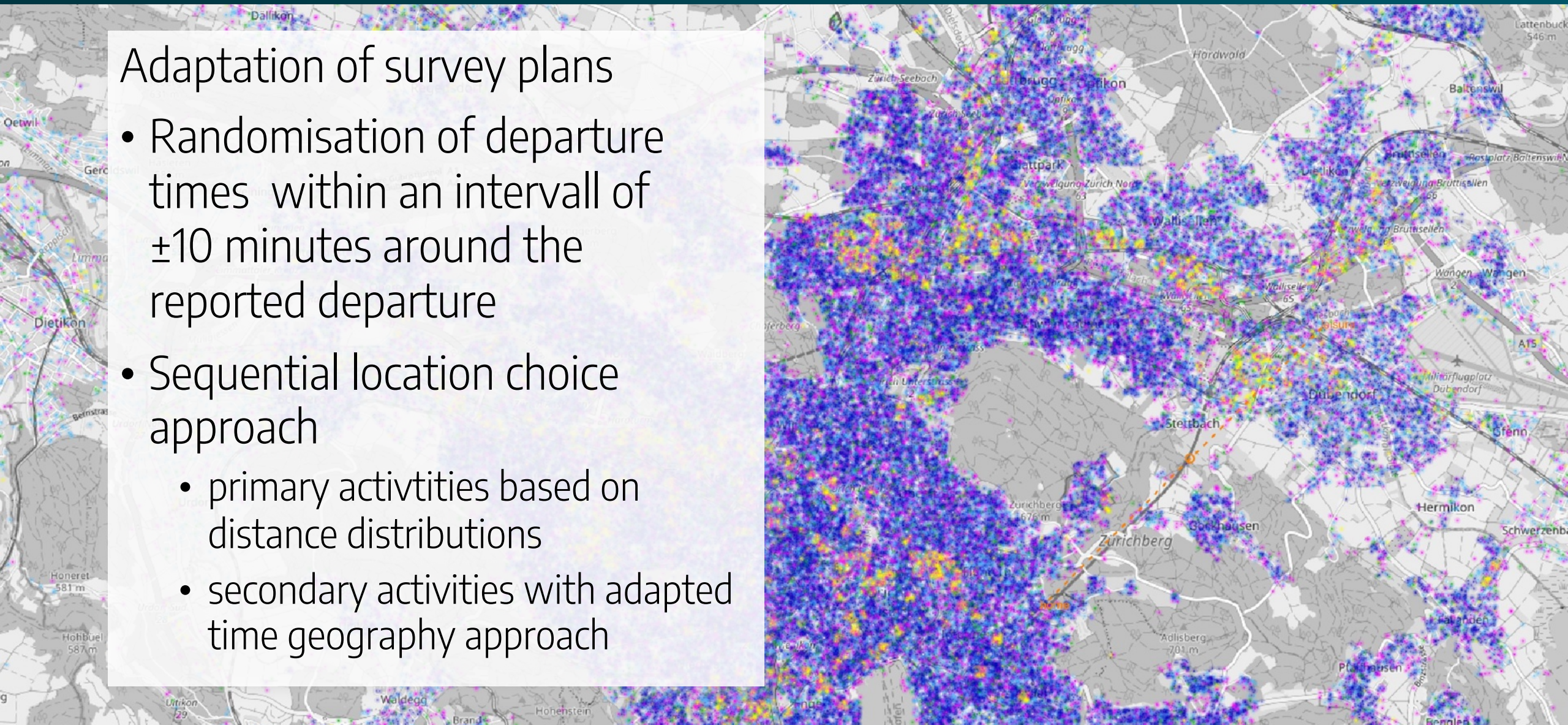
- Household: number of vehicles, distance to public transport stops, land-use type of the home location
- Person: Age, Gender



Plans – Deliverables

Adaptation of survey plans

- Randomisation of departure times within an intervall of ±10 minutes around the reported departure
- Sequential location choice approach
  - primary activtities based on distance distributions
  - secondary activities with adapted time geography approach



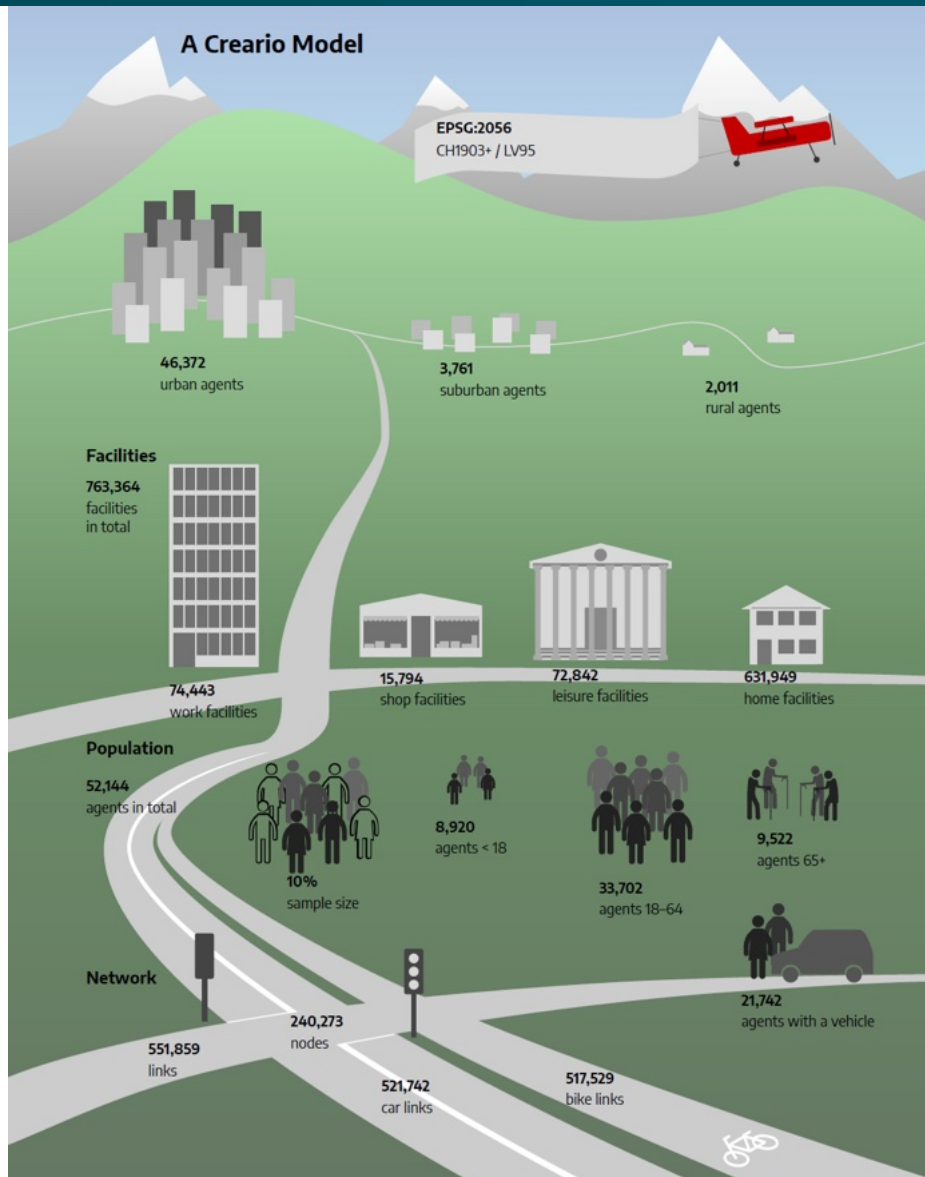
Validation report

The validation report assists the user in evaluating the deliverables provided by Creario.

It provides network plots, statistical charts and (cross) tables that allow the user to

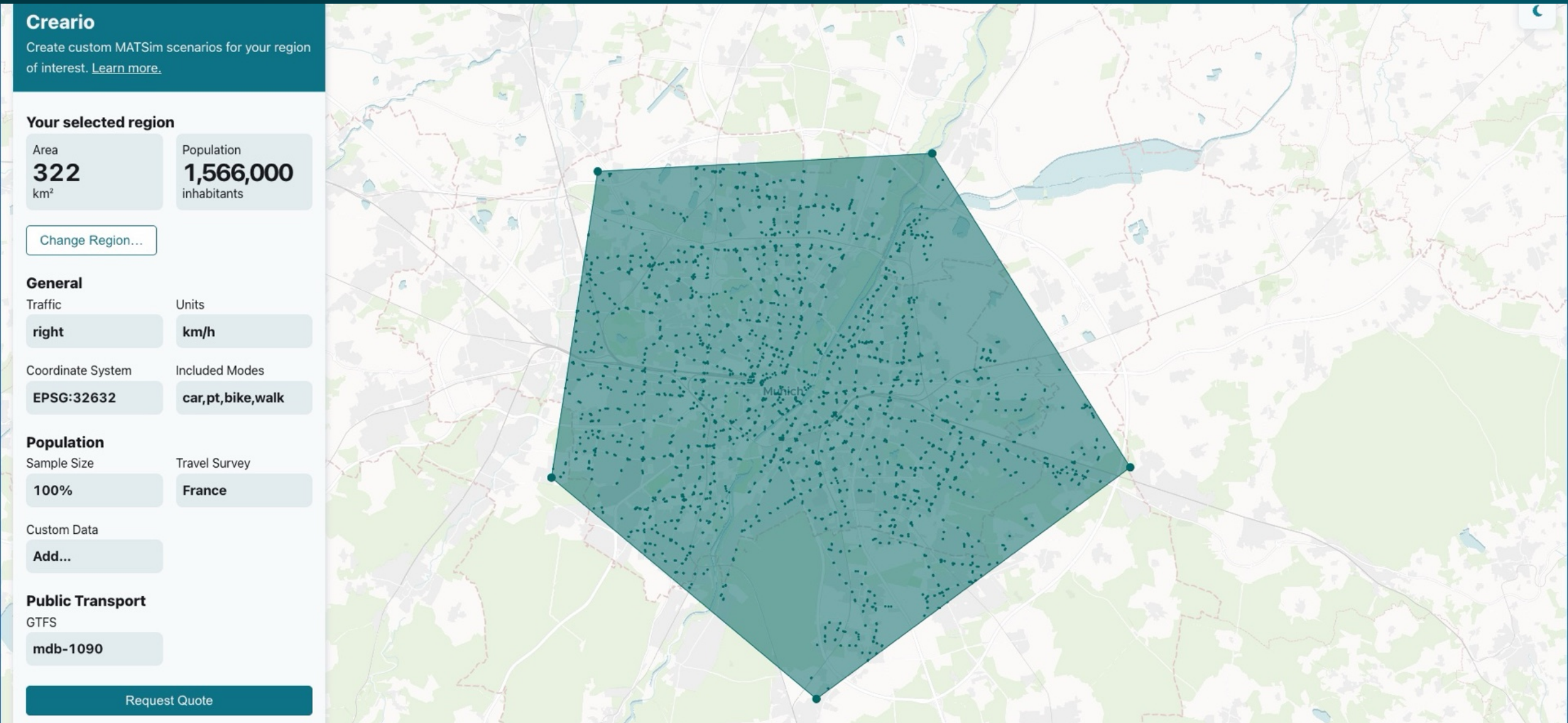
- verify the distribution of attribute values
- identify outliers and implausible values that should be investigated further

An interpretation guide helps the user with reading the analyses, judging the quality of the data provided and identifying aspects that need further individual attention.

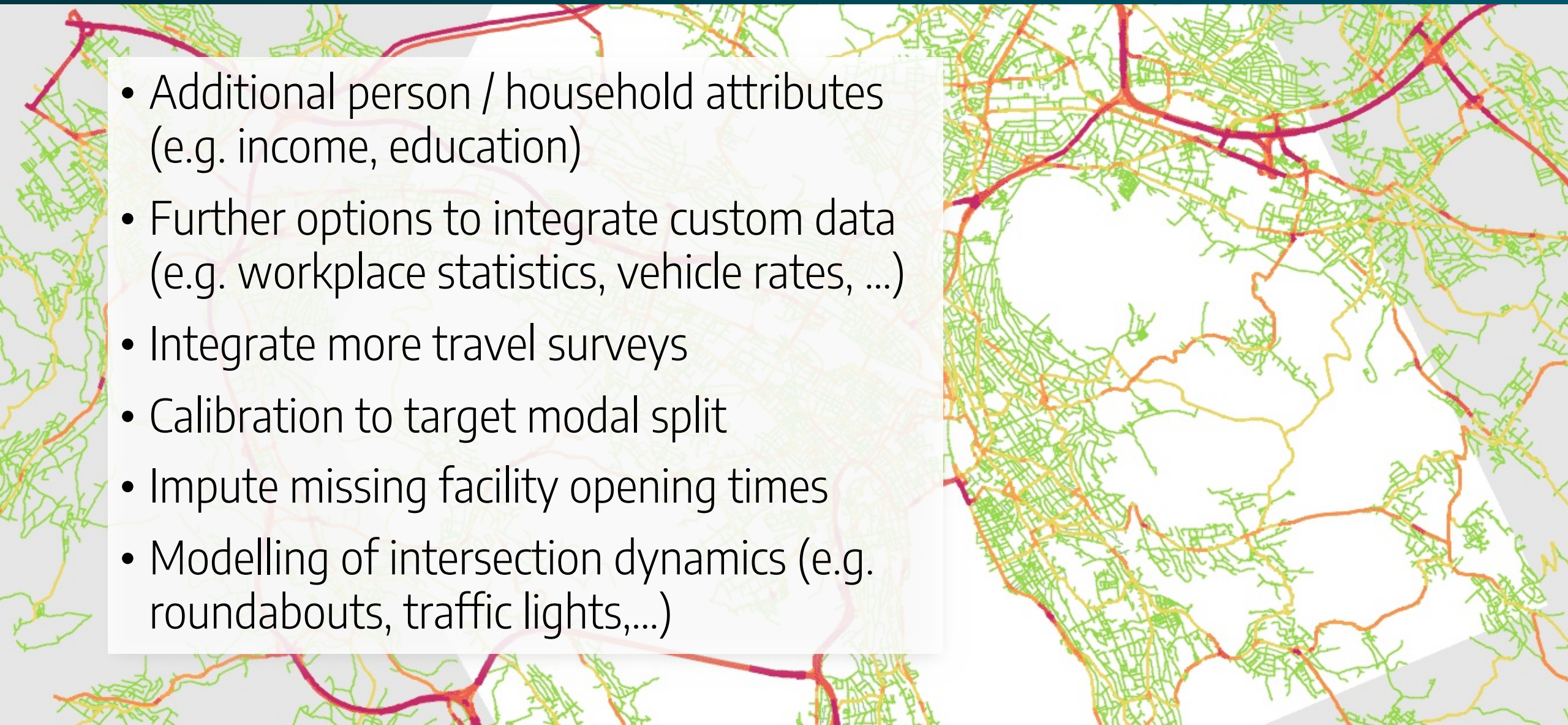


Demo

creario.simunto.com



What's next?



- Additional person / household attributes (e.g. income, education)
- Further options to integrate custom data (e.g. workplace statistics, vehicle rates, ...)
- Integrate more travel surveys
- Calibration to target modal split
- Impute missing facility opening times
- Modelling of intersection dynamics (e.g. roundabouts, traffic lights,...)