

The 10th International Workshop on Agent-based Mobility, Traffic and Transportation Models, (ABMTRANS) March 22 - 25, 2022, Porto, Portugal

# Creating an agent-based long-haul freight transport model for Germany

Chengqi Lu<sup>a,\*</sup>, Kai Martins-Turner<sup>a</sup>, Kai Nagel<sup>a</sup>

<sup>a</sup>*Technische Universität Berlin, Chair of Transport Systems Planning and Transport Telematics, Straße des 17. Juni 135, 10623 Berlin, Germany*

---

## Abstract

The freight traffic plays an important role in the agent-based transport simulation models. Compared to the commuter traffic, the freight traffic receives relatively less attention. This study aims to bring freight traffic into the spotlight of the traffic simulation. In this study, the German-wide long-haul freight traffic is generated based on reliable and open-source data in the context of an agent-based model. The outcome of this study can be applied to any of the simulation scenarios within Germany. In addition, it can also be used as a full scenario for studies on logistics and sustainable solutions for goods delivery.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the Conference Program Chairs.

*Keywords:* freight transport, long-haul, multi-agent transport model, MATSim

---

## 1. Introduction

Freight traffic is an important part of the transportation system. According to the statistics, 76.1% of the goods are transported by road (i.e., with trucks) within Europe in 2019 [4]. In Germany, heavy goods vehicle alone contributes to 17% of the traffic count on the freeways and highways[1]. Freight traffic also contributes to one third of CO2 emission from the road traffic[5]. Meanwhile, in the field of agent-based transport modeling, most of the focus up till now are on the car traffic in the cities. For example, in one widely used agent-based simulation platform, MATSim, there are already scenarios for cities all over the world, including Berlin [15], Paris [9] and Zurich [8]. In most of these scenarios, however, the focus is the daily commute trips inside the cities. The freight trips are either excluded or synthesized based on simple algorithms. In order to improve those models, it makes sense to include better modeled freight traffic. In addition, a well-modeled freight traffic is also essential for a more realistic analysis on the traffic congestion and the environmental impact of the traffic system.

---

\* Chenqgi Lu, Tel.: +49-30-314-29592 ; fax: +49-30-314-26269

*E-mail address:* [lu@vsp.tu-berlin.de](mailto:lu@vsp.tu-berlin.de)

There are mainly two types of the freight trips: long-haul trips and short-distance trips. The long-haul trips refer to the freight trips that run between different cities, regions or even nations. They are usually carried out by larger trucks and will travel from and to freight hubs located in the industrial or commercial area. The major part of the trips are on the freeway or major roads. The short-distance trips are the freight trips that travel within cities or regions. The composition of short-distance freight trips are usually more complex and there can be different kind of trips.

One can argue that the line that distinguish long-haul and short-distance trips is rather ambiguous. For example, a truck that carry goods for a supermarket may depart from a rural region and visit several supermarkets in a city. In this case, the trip has the characteristics for both long-haul and short-distance trips. In order to distinguish the two types of trips more systematically when generating the freight traffic in the agent-based simulation platform, we rely on the aggregation level of the data source. The trips that are generated based on data that focus on the amount of the good flow among different regions or cities are considered as long-haul trips. The freight traffic that are generated based on more detailed data within a given study area (e.g. cities or regions) belong to the short-distance trips.

In this study, we will focus on the long-haul freight trips for the whole German network. This is a good starting point, as long-haul freight trips are less complex than the short-distance trips. Meanwhile, it is also more general applicable to various scenarios within the scoop (in this case within Germany).

**Contribution:** A German-wide long-haul freight traffic based on open data is generated in the context of an agent-based transport simulation environment, namely MATSim [7]. The German-wide model can be used, possibly with a simple extraction process, on any of the scenarios within Germany. In addition, the model can also be used for the national-scale freight transport studies.

## 2. Generation of German-wide long-haul freight traffic data

### 2.1. About MATSim

The Multi-Agent Transport Simulation (MATSim) [7] is used in this project to materialize the freight traffic in an agent-based model. MATSim is a widely used open-source framework<sup>1</sup> for implementing large-scale agent-based transport simulations. It is capable of simulating very large network and population while maintaining a relatively high level of details. One iteration for a city wide (or even nationwide) network with tens of thousands of agents and different modes of transport can be simulated within hours. This characteristic makes MATSim a suitable environment to accommodate the German-wide freight traffic and relevant studies.

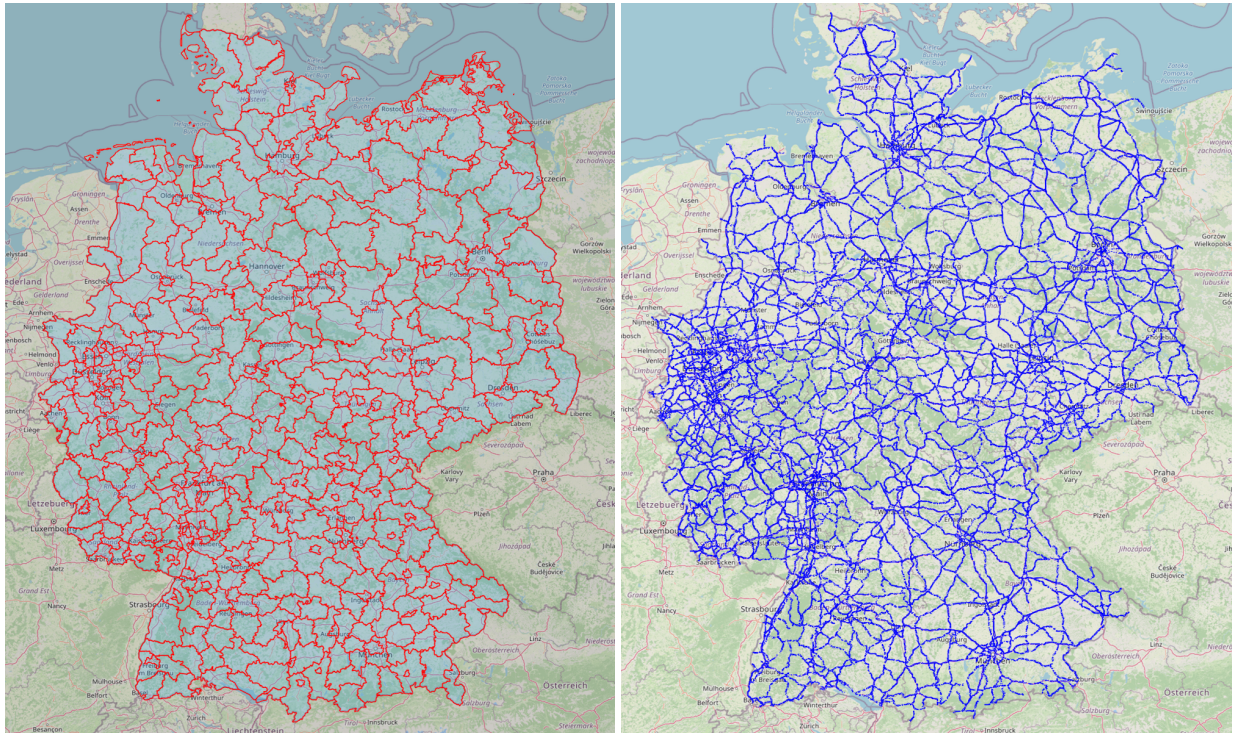
### 2.2. From open data to German-wide long-haul freight traffic in agent-based transport model

In order to generate German-wide long-haul freight traffic, we use the open data from the German Federal Ministry for Digital and Transport (BMVI) [2] and the German Federal Highway Research Institute (BAST). In addition, the Open Street Map [14] is used to generate the German-wide major road network. The data are both open and with high reliability. The freight traffic is first created based on the Verkehrsprognose 2030 [3] project data created by BMVI. The data quantify the amount of goods flow between any two regions inside Germany based on the Nomenclature of Territorial Units for Statistics Level 3 (in short, NUTS 3 regions, see figure 1a), as well as international freight trips that starts or terminates in Germany. We then map the data to the German-wide major road network (see figure 1b), which contains all the freeway (Autobahn) and Federal highway (Bundesstraße). Finally, the freight traffic is optimized against the traffic count data available on the BAST website [1].

*Reading the data.* The data from the Verkehrsprognose 2030 project contains detailed information on the domestic and international freight. The amount of goods transported (in tons) between different regions, the type of the goods and the mode of transport are all included. Each relation is divided into 3 different runs: pre-run (Vorlauf), main run (Hauptlauf) and post-run (Nachlauf). While the main run is always present in any of the relations, the pre-run and the post-run are optional. The division makes it possible to consider the multi-modal freight trip as one single trip. For our project, however, we treat the sub-trips as independent trips, as we only focus on the freight transported with road

---

<sup>1</sup> <https://github.com/matsim-org/matsim-libs/> and <https://matsim.org/>



(a) Nomenclature of Territorial Units for Statistics Level 3 (NUTS 3) regions within Germany (b) The German-wide major road network (including freeway and highway)

Fig. 1: Creating the German-wide freight long-haul transport

vehicles (i.e., trucks). After re-structuring the trips in the data, we then extract all the freight trips that are carried out by road vehicles.

For the domestic freight flows, the Verkehrsprognose 2030 data is based on the NUTS 3 level regions. We can calculate the average daily number of trucks traveling between any two regions based on the amount of the goods transported between those two regions per year. Dividing the amount of goods transported between two regions by number of working days and the average load of a truck, we can get the daily average freight trips that are made between those two regions. The average number of working days are relatively easy to determine. The average load of a truck, on the other hand, is not as straight forward. First, there are different type of goods with different density. Second, the actual load on a truck can vary significantly. The trucks can be fully loaded, partially loaded or even empty. These two factors will lead to very different weights even with the same type of the truck. To make the problem even more complicated, the types of trucks running on the roads also varies. As a result, it is difficult to come up with very accurate values based on the available data. Because of this, we will not distinguish the type of goods or vehicle types. Instead, we treat all the goods and vehicles as uniform type. That means, for any two regions, we will sum up the total amount of goods flow and then divide it with a single average load of vehicle. The traffic count data will be used to adjust the value of average load, such that the difference between the generated data and the actual count data is minimized.

*Determine the departure and arrival location with land use data.* After the conversion from the goods flow to the number of trucks is set up, we need to determine the exact departure and arrival location within a zone of each freight trip for the agent-based model. The land use data is used to determine the departure and arrival location of the freight trips within the NUTS 3 level regions. The land use data is generated based on Open Street Map with an open-source tool Geofabrik [6]. For the freight trips, we choose “industrial”, “commercial” and “retail” as the relevant land use types. The freight trips will start or terminate inside the area with relevant land use type. By doing so, the freight trips that originate or terminate in a residential area or in the middle of nowhere can be avoided. As this model is meant to

be open-sourced, the level of detail is restricted by the open-source data. When more detailed information about the origins and destinations of the freight trips is available (e.g., location of the freight terminals), the model can be easily adapted by replacing the open source land use data with a more detailed data.

*Adapting international trips.* In addition to the freight trips that travel within Germany, there are also international freight trips. The Verkehrsprognose 2030 data contain the information for all the cross-border trips that originate or terminate in Germany. The origin or destination location within Germany are also based on NUTS 3 level region. The locations outside Germany are based on region system with lower resolution (e.g., NUTS1, NUTS 2). The farther a region is away from Germany, the lower the resolution will be. As the network we used for the model is a German-wide network, we will modify the international trips by extracting the trip segments that are within Germany. For an international freight trip, we assign the border-crossing point based on the orientation of the origin/destination region outside Germany. The trip segment between the border-crossing point and the destination/origin region inside Germany will be kept. The rest part of the trip will be trimmed. For example, Spain is to the southwest of Germany. The trip coming from or going to there is likely to enter or exit Germany on the southwestern border. Then, we keep the trip segment between the southwestern border of Germany and the destination or origin inside Germany and use this trip segment to represent the trip traveling from or to Spain. This determination of the orientation of international regions towards Germany is carried out manually and the summarized data is openly available.

*Calibration against traffic count data.* By the steps described so far, all the data entries in the Verkehrsprognose 2030 can be converted to freight trips in an agent-based model. As mentioned above, we use the traffic count data from the BAST is used to determine the number of trips. From the BAST website, the complete counting data is available. We extract the locations of all the counting stations on the freeway and Federal highway and map them to the German-wide network. Since we are generating long-haul freight trips, which are mostly carried out with heavy trucks, we will use the heavy goods vehicle (HGV) count in the counting data. For trips traveling between two given regions, a route can be calculated on the German-wide network. The route will pass through the counting stations and will contribute to the traffic count. By varying the average load of a truck, the number of trucks that travel between the two regions will vary. This will then change the traffic count for each of the counting stations along the route in the model.

In order to determine the freight traffic that is most representative to the reality, we try to minimize the total absolute error between the traffic count in our model and the traffic count data from the BAST. Equation 1 shows the objective function that we want to minimize:

$$\min_x \sum_i |c_i - \sum_j (n_j \cdot \alpha_{ij})| \quad \text{for } \forall i \in \text{Counting Stations}, j \in \text{OD pair} \quad (1)$$

$c_i$  is the HGV count at counting station  $i$  and  $n_j$  is the number of freight trips that travel between the origin-destination pair (OD pair), which includes the modified international trips. The term  $\alpha_{ij} \in \{0, 1\}$  is a decision variable that determines whether the trips of OD pair  $j$  contribute to the traffic count at counting station  $i$ . For each OD pair  $j$ , we can calculate the route on the German-wide network. The route can be considered as a set of road segments. If the road segment where counting station  $i$  locates is contained by the route for the OD pair  $j$ , then the route will contribute one unit of traffic count for each vehicle that travels along the route (i.e.,  $\alpha_{ij} = 1$ ). Otherwise, no traffic count will be contributed (i.e.,  $\alpha_{ij} = 0$ ). The value of  $n_j$  is determined by equation 2, where  $x$  is the average load of the truck and  $L_j$  is the total amount of good flow for the OD pair based on the data.

$$\text{s.t. } n_j = L_j/x \quad (2)$$

By solving this optimization problem, we can find the optimal  $x$  (average load of trucks) that minimize the difference between generated traffic and the reality. Note that, in this optimization process, we simplify the problem by assigning all the trips that travel between two given regions (i.e., an OD pair) a same route. This is possible because the traffic counting stations are on the freeway and highways, and the exact locations of the origin and destination of the trips within the respective regions do not impact the traffic count significantly. Therefore, we can approximate the routes of all the trips traveling between the two given regions with the route that connects the road segments that are closest to the centroid of the origin and designation region respectively.

*Generating freight trip in agent-based model.* In MATSim, an agent performs a sequence of activities during the simulation time horizon (usually a day) based on his/her plan. If two subsequent activities are performed at different locations, the need to be connected by a trip. The departure time of the trip is the ending time of the former activity. The departure location is the location of the former activity, and the destination of the trip is the location of the later activity.

The agent can be created to represent a real person in the reality or it can be used simply as a trip creator. In this project, we will treat the agent as the trip creator, because the truck driver's behavior and the operational model of the freight company require more detailed information, which we do not have. In our case, to generate one trip in the agent-based model, we just need 2 dummy activities and a leg in between for each agent. The first dummy activity is happening at the departure location of the freight trip and the second dummy activity is happening at the destination location of the freight trip. For the departure time of a trip, a random time during the day is assigned since there is no information in the open data. When a more detailed data is used, the departure time of the trips can be adjusted based on a proper distribution.

The conversion from the annual amount of goods to daily number of trucks can lead to non-integer values and values below one. To adapt these non-integer values into the agent-based model, we use the probability interpretation. For the non-integer part of the value, we interpret it as the probability of having one (additional) trip. For example, from region A to region B, there are 4.3 trips per day on average. Then we will first generate 4 trips from region A to region B. Then there is a 30% of probability of having another trip from region A to region B. This also applies to the OD pair with number of trips below 1.

### 2.3. *Extracting data for a given scenario*

The plans file for the German-wide long-haul freight trips can be added directly to the MATSim scenario where a German-wide network is used. For many MATSim scenarios in Germany, however, a smaller but denser network is used (for example in Berlin scenario). In that case, the plans need to be extracted before adding it to the scenario.

To extract plans, we need the network for the scenario where the long-haul freight trip is to be added and the study area (in the format of shape file) of the scenario. First, we overlay the study area onto the German-wide network from the freight model. Then we can categorize the freight trips into four different categories: (1) both origin and destination within the study area, (2) either origin or destination within the study area, (3) the route passing through the study area, (4) not touching the study area at all. The first three types of freight trips are relevant to the scenario, and we will extract them from the freight trips model. For the trips with both origin and destination within the study area, no adjustment is required, and the trips can be directly added to the scenario. For the trips with segments outside the study area (i.e., type 2 and type 3), we will perform the trip trimming process, which is similar to the international trip trimming.

To trim the trips with segments outside the study area, we will first calculate the route for the trips on the German-wide network. The intersection between the route and the boundary of the study area will be the location where the trip will enter and/or exit the study area. The trip can then be divided into different parts. We will take the trip segment within the study area and discard the rest of the trip segment(s). In addition to the location of the trips, the departure time of the trips whose starting point is changed to the boundary of the study area will also be adjusted. The departure time of such trips is delayed by the time it takes to travel from the original starting point of the trip to the enter point of the study area.

## 3. Results

After applying the whole process, we reach 654,396 long-haul freight trips throughout the Germany for a normal working day. This is based on an average truck load of 16 tons, which is determined by the calibration process mentioned in the section 2.2. The freight plans, the scripts as well as the open-source raw data are all openly available and can be easily adapted to different studies or reproduced based on custom data<sup>2</sup>.

---

<sup>2</sup> <https://svn.vsp.tu-berlin.de/repos/public-svn/matsim/scenarios/countries/de/german-wide-freight/>

In figure 2a, the freight trips is shown on the German-wide network. Each semi-transparent line in the figure represent one single trip. For a better visualisation effect, we show 25% of the total freight trips in the figures. To demonstrate the extraction process, a geometry (shape file) representing the dummy study area is fed to the script and the extracted freight trips are shown in figure 2b. Note that, the departure and arrival locations of all the trips are within or on the boundary of the given geometry. This means the trips can be easily mapped to a smaller and denser network for the study area.

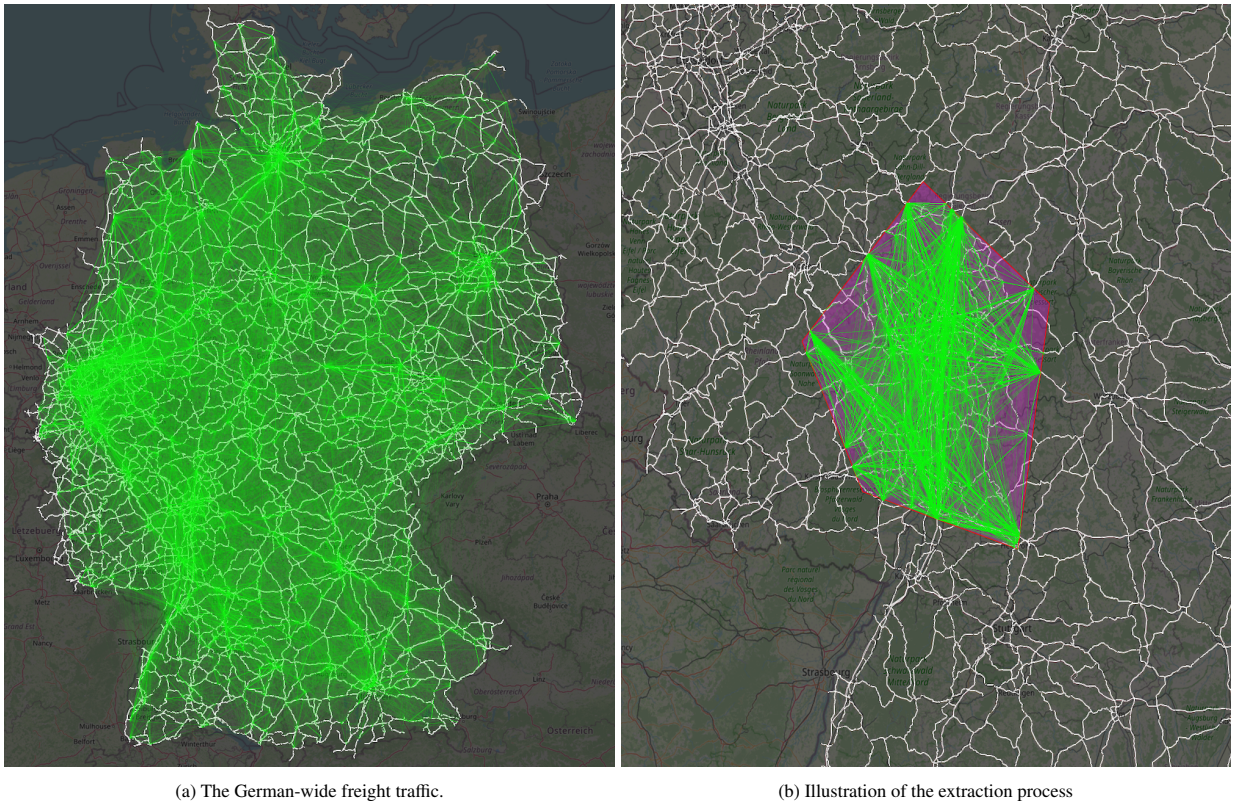


Fig. 2: Results

The German-wide long-haul freight model has already been applied to various open scenario in Germany. The Düsseldorf scenario [12] (based on the KoMoDnext project [11]) and Kelheim scenario [13] (based on KelRide project [10]) are two examples. The added long-haul freight trips help to increase the representativeness of the scenarios to the reality.

#### 4. Conclusion and Outlook

In this project, a German-wide long-haul freight traffic is created in the context of agent-based traffic model based on reliable and publicly available data. The generated trip data as well as the scripts are all open-sourced and can be easily adapted to different scenarios within Germany. When more detailed information or project-specific data is available, users can also adapt the scripts to the data and re-generate the freight traffic according to their need. The German-wide freight model in this project is also a very suitable testbed for studies on logistics, environment impact of freight traffic as well as nation-wide sustainable freight system.

As this is the first version of the German-wide long-haul freight trip model, there are also several aspects that can be improved or further developed. For example, the departure time of the freight trips, the freight vehicles composition and loading condition of the freight vehicles for different types of goods are all simplified in the current model. In addition, the operational aspects of the freight traffic are also not included in this model. For example, the exact

locations of freight terminals or hubs and the driver-vehicle rotation can also have impact on the pattern of the freight trips in the network. As a more complex model will require more detailed information and data-processing, this calls for further studies. Meanwhile, thanks to the open-source design, users of the current freight traffic model, who are in need of a more sophisticated model, can feed in the data they have and sacrifice the open-source label in exchange for a more detailed model.

## Acknowledgement

This work was partly funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 398051144 and 323900421.

## References

- [1] Bundesanstalt für Straßenwesen, accessed in December 2021. Automatische zählstellen 2020. [https://www.bast.de/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/Aktuell/zaehl\\_aktuell\\_node.html](https://www.bast.de/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/Aktuell/zaehl_aktuell_node.html).
- [2] Bundesministerium für Verkehr und digitale Infrastruktur, 2016a. Bundesverkehrswegeplan 2030. <http://www.bmvi.de/DE/Themen/Mobilitaet/Infrastrukturplanung-Investitionen/Bundesverkehrswegeplan-2030/bundesverkehrswegeplan-2030.html>.
- [3] Bundesministerium für Verkehr und digitale Infrastruktur, 2016b. Verkehrsprognose 2030. <https://daten.clearingstelle-verkehr.de/276/>.
- [4] Eurostat, 2019. Road freight transport statistics. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Freight\\_transport\\_statistics\\_-\\_modal\\_split](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Freight_transport_statistics_-_modal_split).
- [5] Forschungs Informations System, 2021. Luft- und klimabelastung durch güterverkehr. <https://www.forschungsinformationssystem.de/servlet/is/39787/>.
- [6] Geofabrik tool, accessed in December 2021. <https://www.geofabrik.de/>.
- [7] Horni, A., Nagel, K., Axhausen, K.W. (Eds.), 2016. The Multi-Agent Transport Simulation MATSim. Ubiquity, London. doi:10.5334/baw.
- [8] Horni, A., Vitins, B., Axhausen, K., 2011. The zurich scenario: A technical overview. Institut für Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.
- [9] Hörl, S., Balac, M., 2021. Synthetic population and travel demand for paris and Île-de-france based on open and publicly available data. Transportation Research Part C: Emerging Technologies 130, 103291. doi:10.1016/j.trc.2021.103291.
- [10] KelRide: Weather-proof smart shuttle, accessed in December 2021. <https://www.vsp.tu-berlin.de/menue/forschung/projects/2021/kelride>.
- [11] KoMoDnext: Automatisiertes Fahren im digitalen Testfeld Dusseldorf, accessed in December 2021. <https://www.vsp.tu-berlin.de/menue/forschung/projects/2020/komodnext>.
- [12] MATSim open Dusseldorf Scenario, accessed in December 2021. <https://github.com/matsim-scenarios/matsim-duesseldorf>.
- [13] MATSim open Kelheim Scenario, accessed in December 2021. <https://github.com/matsim-scenarios/matsim-kelheim>.
- [14] OpenStreetMap, accessed 2021-06-23. <http://www.openstreetmap.org>.
- [15] Ziemke, D., Kaddoura, I., Nagel, K., 2019. The matsim open berlin scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. Procedia Computer Science 151, 870–877. doi:10.1016/j.procs.2019.04.120.