Impacts of downscaled inputs on the predicted performance of taxi fleets in agent-based scenarios including Mobility-as-a-Service

Conference Paper

Author(s):
Saprykin, Aleksandr; Chokani, Ndaona; Abhari, Reza S.

Publication date:
2022

Permanent link:
https://doi.org/10.3929/ethz-b-000544968

Rights / license:
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

Originally published in:
Procedia Computer Science 201, https://doi.org/10.1016/j.procs.2022.03.074
Impacts of downscaled inputs on the predicted performance of taxi fleets in agent-based scenarios including Mobility-as-a-Service

Aleksandr Saprykin, Ndaona Chokani, Reza S. Abhari

Keywords: large-scale modelling; agent-based simulation; sampling; downscaling; taxi fleet; maas

Abstract

The impacts of using downscaled inputs in mobility simulation involving taxi fleets have not been well studied. Downscaled inputs, that is, inputs with a fraction of the full-scale population, are often used for large-scale scenarios in order to keep run times short and to avoid using high-performance computing resources. In this paper, a large-scale multi-modal scenario of the Munich metropolitan region, with a taxi fleet that serves the first/last mile legs of public transit passengers, is used to systematically quantify the impacts of downsampling. The results show that the full-scale population is required to obtain accurate spatio-temporal estimations of the fleet’s performance. Sample sizes of at least 30% of the full-scale population can thus be used if only an approximate knowledge of spatially aggregated performance metrics (such as average waiting time, fleet utilization and empty mileage) is desired. However, with downscaled inputs, different characteristics during peak hours, such as a shift in the peak utilization, can result. Moreover, we find evidence that downsampling can affect the spatial distributions of trip densities of taxi fleets.

© 2022 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: large-scale modelling; agent-based simulation; sampling; downscaling; taxi fleet; maas

1. Introduction

The ongoing spread of emerging transportation modes, such as autonomous vehicles, coordinated taxi fleets for shared and private rides, and micro-mobility services like scooters, is changing the ways people move and use mobility infrastructure. With increased penetrations of battery-electric vehicles [8], the availability and reliability of the infrastructure become more critical, as, for example, the electricity distribution grids can be overloaded during periods of peak charging [11]. Traditional macroscopic modelling tools do not yield fine-grained and disaggregated details of transportation systems as these tools do not account for the complex behavioural patterns of human beings in sim-
ulated scenarios. To address this limitation, researchers and practitioners are shifting focus to agent-based models (ABMs) where each agent has his/her own rules of behaviour, and interacts with other agents and the environment.

However, the application of high-fidelity ABMs to large scale (that is, the extent of a metropolitan area with millions of residents and commuters) scenarios leads to long simulation run times and requires substantial computing resources. One approach to reduce the run times is to utilize high-performance computing hardware [12, 17]; typically, this requires additional investments in the hardware and that the simulation models are adapted to efficiently use the hardware. An alternative approach, which is widely employed, is to downscale the inputs of the scenario, and then upscale the simulation outputs. For example, one uses a random sample of 10% of the population, reduces the flow capacities of roads to 10% of their nominal capacities, and then multiplies the simulation outputs by a factor of 10. However, recent studies indicate that simulations with downscaled inputs and upscaled outputs produce different traffic dynamics and occupancy levels in public transit than simulations with full-scale inputs/outputs. While the impacts of downscaled inputs on simulations of car traffic and public transit systems have received some attention, the impacts on simulations of coordinated taxi fleets have not been addressed. Thus, the main contribution of this work is a systematic quantification of the impacts of using downsampling factors on the predicted performance of a fleet’s operation within the context of large-scale agent-based scenarios.

2. Related work

Prior literature [10, 1, 2] has mostly focused on the impacts of downscaled inputs in relation to car traffic, and one study [14] quantifies the impacts on both car traffic and public transit occupancy. The observations of these prior works are consistent, and suggest that for analysis of disaggregated performance metrics, population samples of at least 25%–30% are used, while only when analysis of highly aggregated simulation results is required, can small populations samples of 5%–10% be used. Bischoff and Maciejewski [3] evaluated the effects of scaling down a fleet of 100 000 vehicles that is used to replace private cars in the city of Berlin. Population samples of 10% and 100% were compared, and the fleet size was scaled down to 11% in order to keep constant the trips-per-vehicle ratio. The simulations with downscaled inputs did not show any substantial differences in terms of waiting times and fleet occupancy compared to the simulations with full-scale inputs. However, in the simulations with downscaled inputs, no background traffic congestion was simulated, but, instead, the per-link travel times from the simulations with full-scale inputs were used. Indeed, other studies [9, 5] consider that it is unrealistic to assume that the large-scale fleet does not impact congestion and vice versa.

3. Methodology

In this study, we use our in-house GPU-accelerated agent-based mobility simulator, GEMSIm [13]. The demand comprises the daily-activity plans of the agents, that is, a set of activities that the agents do and the legs travelled by the agents using an assigned mode between the locations of the activities. The supply comprises (i) a road network graph with uni-directed edges as street segments and nodes as intersections; and, (ii) a public transit schedule. A simulation is run iteratively, whereby agents can learn between iterations and adapt their behaviour based on previous experience. This iterative process converges to a Nash equilibrium when an agent cannot unilaterally improve his/her daily performance. The simulator can handle taxi fleets [15], and uses a load-balancing algorithm [3] to serve the requests of the clients. An operator schedules requests as follows:

- During periods of over-supply, each queued request is matched with the nearest available vehicle.
- During periods of over-demand, each idle vehicle is matched with the nearest queued request.
- All requests are served without pre-booking; there is no drop rule for long-waiting requests.
- A vehicle stays at its last location after drop-off.
- Initially, the vehicles are placed based on the known morning-time demand.

The scheduling is executed every 30 seconds of simulated time, such that the operator can aggregate requests. The matching of the nearest vehicle or request is done based on travel time and takes congestion into account.
The scenario used in this study covers the whole of the Munich metropolitan region, which includes the agglomeration areas of Munich, Augsburg, Ingolstadt, Landshut, Rosenheim, and Landsberg am Lech. The region’s total population is about 6 million inhabitants, and, as shown in Fig. 1, the region covers about 40% of the area of Bavaria. The region’s road network, which includes 289,893 nodes and 840,752 links, was generated from OpenStreetMap [7] data. The public transit schedule is taken from the OpenData ÖPNV portal [4], and includes all 1,555 lines and 34,206 stop facilities in the region. The initial daily-activity plans of the agents were generated based on the methodology described in [16] and using the Mobility in Germany [6] nationwide travel survey. The synthetic population includes 3,248,739 agents who either drive cars or use public transit; walking is considered to be a sub-mode of public transit. People who either stay at home or use other modes, such as bicycles, are not included in the simulation.

The simulations are conducted as follows. First, a simulation is run for 350 iterations, with 10% of the agents re-routed after each iteration, to converge to an equilibrium. Then in successive simulations, a new transport mode, Mobility-as-a-Service (MaaS), is available to the agents; the MaaS includes a combination of public transit and taxi rides for the first/last mile travel legs. The taxi trips are limited to a direct distance of up to 3 km, and the total trip length must be at least 6 km long to prevent the use of taxis for uni-modal travelling. The area of fleet operation is shown in red in the left plot of Fig. 1, and only agents who travel within this area can switch to the MaaS mode. The fleet size is 15,000 vehicles. The simulation is run again for 700 iterations: first, 350 iterations with 10% of the agents re-routed and 10% of the agents allowed to switch the mode between either car, public transit, or MaaS; and then, 350 iterations where agents can pick only one of the previously experienced plans. In the end, about 7.5%, or 252,389, of the agents switched to the MaaS mode. This converged scenario was used to assess the impacts of downsampling.

The downscaling is performed by sampling, uniformly at random, a fraction of agents from the reference scenario and adjusting the flow capacities of the road graph following [14]. The fleet size is scaled down in proportion to the size of the population sample. Each downscaled scenario is run for 150 iterations to converge: 100 iterations with 10% re-routing and 50 iterations using plans from the agents’ memories. Sample sizes of 1%, 2%, 5% and 10% to 90% with increments of 10% were used. The memory of agents is always set for the five last chosen plans.

4. Results

The impacts of sample size on the average waiting time of clients, that is the time from when a request is sent to the time when a taxi arrives, is shown in Fig. 2. For sample sizes less than 5%, the evening peak hour has substantially longer waiting times of about 1 hour compared to about 20 minutes for the larger sample sizes. The waiting times in the morning for smaller sample sizes are also longer and range from about 30 to 50 minutes, compared to the waiting times of 10–20 minutes for sample sizes larger than 10%. The substantially longer waiting times of small sample sizes...
in the evening peak hour can be explained by the fact that the initial placement of vehicles is based on the known morning-time demand; hence the waiting times are less affected in the morning. The larger sample sizes of 20%—50% yield quite similar results over the whole day, with some differences observed during the evening. As in prior works [2, 14], we found that for sample sizes larger than 30% there are no substantial differences in the waiting times. However, for the full-scale population, the longest waiting times occur 3 hours later than with the smaller samples.

Fig. 2: Impact of sample size on the average waiting times of agents.

Fig. 3 compares the impact of sample size on fleet utilization. During the morning peak hour, the fleet utilization differs substantially over the range of 60%–100% for sample sizes of 1% to 10%, while sample sizes of 30% to 50% yield fleet utilization in the range of 70%–80%. Sample sizes larger than 60% consistently yield fleet utilization around 60% in the morning. All sample sizes have a fleet utilization of 100% in the evening peak hour. However, the distribution of the fleet utilization over the evening differs with sample size: for the smaller sized samples, the distribution tends to be more flatter and wider than for the larger sized samples; only the full-scale population has a narrow distribution with a sharp peak. The flatter and wider distributions of fleet utilization with the smaller sized samples indicate that for longer portions of the evening peak hours, a fleet operator finds it challenging to schedule the increased number of requests. The most probable reason for this is the spatial imbalance between supply and demand that results when downscaled samples are used; this imbalance is most pronounced in municipalities with low population densities. It is worthwhile to note that the narrow and sharp distribution of fleet utilization in the full-scale population case explains the shift in the evening peak of waiting times compared to the downscaled sample sizes.

Fig. 3: Impact of sample size on fleet utilization.

The impact of sample size on the fleet empty mileage is compared in Fig. 4. It is evident that the fleet tends to drive longer distances with downscaled sample sizes. Similar to the fleet utilization, for sample sizes of 1% to 10% the fleet empty mileage differs substantially, especially in the morning peak hour. This is so as the small sample sizes are more affected by the imbalance between supply and demand. The sample sizes of 30% to 50% differ less across the day, and have a smaller empty mileage in the morning of 40% compared to 60% in the sample sizes of 1% to 10%. The larger sample sizes of 60% to 100% yield empty mileage of 30% in the morning. The magnitudes of the empty mileage in
the evening are similar across all sample sizes. However, similar to the fleet utilization, only the full-scale population case has a narrow and sharp distribution of the fleet empty mileage.

![Fig. 4: Impact of sample size on fleet empty mileage.](image)

The distributions of the per-vehicle daily driven distances for sample sizes of 1%, 10% and 100% are compared in Fig. 5. It can be seen that the distribution of distances changes with sample size. For sample sizes of 10% and less, the distributions are close to a normal distribution, with the mean per-vehicle daily driven distances of 175 km and 250 km, for the sample sizes of 1% and 10% respectively. As the sample size increases, the distributions tend towards a bimodal distribution, for the 100% sample size, with peaks around 100 km and 270 km, and an overall mean per-vehicle daily driven distance of 215 km. The bimodal distribution can be explained by the fact that the more vehicles and clients are in the system, the higher is the probability that a nearby vehicle can be found for each request. This is especially relevant for low-density areas which have fewer vehicles and clients.

![Fig. 5: Distribution of per-vehicle daily driven distance.](image)

The spatial distribution of daily average waiting times is presented in Fig. 6. Here, the locations of requests are aggregated in hexagons having an edge size of 500 m. It can be seen that the 1% sample size has a substantially different spatial distribution than for the full-scale case. One reason is that downscaling significantly reduces the number of requests in the low-density areas outside Munich. Hence, fleet vehicles are mostly concentrated in the city, where the vehicles generate and encounter more congestion, and thus for requests coming from low-density areas, there is a lower likelihood of finding a nearby vehicle. Another reason for the difference is that traffic dynamics are more changed with smaller samples [14] causing different travel times. For the 10% sample size, the spatial distribution in the city is more similar to the full-scale case, while low-density areas outside of the city are still highly under-sampled.

5. Discussion and conclusions

The impacts of downscaled inputs on the predicted performance of a coordinated taxi fleet that is used for MaaS, which combines public transit with taxi trips for the first/last mile legs, is systematically quantified. In comparison
to the full-scale population, sample sizes of 30% or more predict fleet performance metrics that are of comparable, not exact, magnitudes. These sample sizes of 30% or more can thus be used if only an approximate knowledge of spatially aggregated externalities of the fleet, such as noise or pollution, are desired. However, as the predicted performance using downscaled inputs can have quite different characteristics during peak hours, such as a shift in the peak utilization, the use of the full-scale population is recommended if accurate spatial and temporal estimates are required. For example, with a sample size of 50%, the mean per-vehicle daily driven distance is 235 km, while with the full-scale population, the mean driven distance of 215 km is 8.5% less. Thus, for an all-electric fleet, based on the current electricity price in Germany of EUR 150/MWh and the average electric vehicle energy consumption of 0.2 kWh/km, the annual difference in cost for charging the fleet is about EUR 3.3 million. Similarly, the required charging infrastructure and corresponding investments will be overestimated when downscaled inputs are used in the mobility simulations. On the other hand, in low-density areas the charging infrastructure would be underestimated. Based on our assessment of the impacts of downscaled inputs, we do not recommend using a population sample size of less than 30% as the estimates of fleet performance are then highly biased.

The shift of the average waiting time during the evening peak hour, when the sample size is reduced from 100% to 90%, was examined in detail. In the current case, it seems that this shift arises due to a combination of multiple factors, including the spatial distributions of fleet and requests, the congestion patterns (in particular, the highly congested city centre and uncongested low-density surroundings), as well as the fleet size and served demand. The experiment was repeated three times using different sets of samples in order to exclude substantial stochastic variations in the simulation outcome; however, the outputs of simulations were observed to be similar for each set of samples. Further analysis of the case with the full-scale population showed that the equilibrium, which is finally reached, is relatively fragile, and even insignificant fluctuations in the non-taxi demand shifted the evening peak to earlier times, for example, as in the case of the 90% sample. Therefore, even downsampling the scenario by 10% can lead to shifts of peak utilization of the fleet. That means that scenarios with fully utilized fleets must be very carefully evaluated during the downsampling process.

In future work, it would be interesting to further quantify the impacts of trip density when the inputs are downscaled. The results of the present study differ from the results of [3] where a 10% sample was found to be sufficient to yield consistent outputs. However, as the spatial distributions of demand are not the same in the two studies, this may impact the results when downscaled inputs are used. As we can see in the present study, the predicted performance in the city of Munich is not adversely affected when downscaled inputs are used, and even with a sample size of 10%, very similar waiting times are predicted. On the other hand, in the low-density areas around the city, the fleet’s overall performance is poorly predicted as these areas are under-sampled. Another interesting aspect for future research would be to investigate the impacts of different network topologies in other areas on fleet performance; for example, a grid network may affect the results distinctly.
References


