Fleet Sizing for Pooled (Automated) Vehicle Fleets

Milos Balac1, Sebastian Hörl1, and Kay W. Axhausen1

Abstract
This paper proposes an (automated) on-demand public transport service using different vehicle capacities to serve current car demand in cities. The service relies on space and time aggregation of passengers that have similar origins and destinations. It provides a point-to-point service with predefined pick-up and drop-off locations. In this way, detours to pick-up en-route passengers is avoided. The optimization problem that minimizes the fleet size along with limiting rebalancing distances is defined as a mixed-integer linear programming problem. Solving the problem for Zurich, Switzerland, yields, in the best case, a fleet size equal to 3.7% of the current fleet that could serve current car demand. Vehicle kilometers traveled could also be reduced by nearly 10%. Results also show that the speed of automated vehicles has a substantial effect on the necessary fleet size and free-flow speeds generally produce over-optimistic results.

Automated vehicles (AVs) are rarely seen on the streets today. Even when occasionally a glimpse of one is caught, there is always a person in the driver’s seat ready to take control in case of an emergency. However, plenty of research has been devoted in the last years to how AVs might affect the way people travel (1,2,3). Elimination of the driver can bring two important changes in how cars are used. First, people who cannot drive, because they do not possess a driver’s license or are otherwise unable to drive, will gain access to the freedom that the automobile provides. Second, a substantial cost element of taxi services will be eliminated by removing the driver from the equation, which shows the potential for shared vehicle fleets to thrive and potentially reduce private vehicle ownership.

Researchers focusing on the impact of shared AV fleets suggest that the proportion of vehicles needed to serve the demand in urban cores is about 10% of the current fleet (2,3). Although this promises to reduce the necessary parking space, it creates additional vehicle miles traveled, because of the need to re-position vehicles to efficiently serve the demand. Potential induced demand from people who currently do not use a car only worsens the picture. To mitigate this problem, ride-sharing is seen as a potential solution.

This paper investigates how either automated or conventional vehicles can be used in a dynamic public transport service in the city of Zurich, Switzerland. The focus here is on AVs, as they can bring about substantial reduction of costs; the methodology can also be applied to conventional vehicles. By pooling previous individual car travelers, the paper proposes a point-to-point public transport service. An optimization problem that aims to minimize vehicles needed to serve the demand and to minimize vehicle kilometers traveled (VKT) is formulated by looking at both the congested and free-flow speed cases.

This paper is structured as follows. Background information is provided on the current state of knowledge about the potential of shared AVs. This is followed by an explanation of the methodology for the case study presented. Results are then portrayed before discussion and concluding remarks.

Background
Considerable research effort has been devoted in recent years to investigate the level of disruption that AVs can cause in different environments. Many studies have focused on investigating the impact of AVs on road capacities, the necessary parking space, travel cost

1ETH Zurich, Zurich, Switzerland

Corresponding Author:
Milos Balac, milos.balac@jvt.baug.ethz.ch
reduction, increasing comfort, potential use as public transport feeder services, and decreasing vehicle car ownership by deploying shared vehicle fleets to serve the current demand.

How fleets of shared AVs can reduce vehicle ownership and the potential consequences have been studied in many countries and cities. One of the first large-scale simulations has been performed for Singapore (1): The study finds that the whole transport demand of the city could be covered by one-third of today’s vehicle fleet if it entirely consisted of automated single-occupancy vehicles.

Subsequently, a series of studies has been performed for Austin, Texas. In Fagnant and Kockelman’s study (2) a grid-based simulation for the city is introduced. For an artificial demand based on real-world trip generation rates and randomly assigned destinations, the study found that Austin’s demand for private car trips could be served by an AV fleet that reduces the necessary car fleet by 90%. However, owing to the need to relocate vehicles to pick-up passengers a considerable amount of additional VKT is introduced. The case is further extended in the study by Chen et al. (4), in which an electric charging infrastructure is assumed. Other studies introduce a more detailed demand for the scenario, based on static trips from the regional household travel survey (HTS). Levin et al. (5) introduce congestion to the simulation and find that this has a strong impact on fleet size. Liu et al. (6) apply a choice model, although in a post-processing step, and feed a discrete choice model with information about travel and waiting times to analyze potential mode shares in Austin. Finally, Fagnant and Kockelman (7) extend the Austin case with a ride-sharing component and find that it can reduce wait times for customers and mitigate the increase of VKT owing to empty rides.

For the case of Berlin, Bischoff and Maciejewski (8) use a static travel demand from the regional HTS to create a MATSim (9) simulation with automated taxis. All car trips within the city boundaries are replaced by the service, leading to a scenario in which one-tenth of all vehicles can replace the current fleet if every agent in the simulation uses the service with acceptable wait times. In the same scenario, Bischoff and Maciejewski (10) find that also carrying public transport users leads to a linear increase in needed fleet size. Finally, Maciejewski and Bischoff (11) introduce congestion to the simulation showing that without significant gains in road capacity owing to automation a fleet of automated taxis serving all of the city’s demand will worsen congestion dramatically. On the other hand, AVs are found to be likely to mitigate problems of searching for parking in the city (12).

Heilig et al. (13) use an agent-based approach to investigate the impact of an on-demand shared automated fleet of vehicles that supports ride-sharing. They use a simplified approach to determine the minimum required fleet size without any rebalancing of vehicles during the day. They find that around half the number of vehicle trips can be saved through pooling but only about 20% of VKT for the city of Stuttgart. They also report that around 15% of vehicles at the time of their study will be needed to serve the car-travel demand.

Vosooghi et al. (14) use a dispatching ride-sharing algorithm proposed by Bischoff et al. (15) to estimate how pooling can benefit the metropolitan region of Rouen Normandie in France, using a predefined fleet size. They find that using four-seaters with shared rides is the best option among the tested scenarios.

For Zurich, Hörl et al. (16) show that under ideal flow conditions a fleet size of around 7,000 to 14,000 automated taxis will be able to serve the mobility demand of the city. Using a detailed agent-based daily travel demand from HTS data, it is shown that the dispatching strategy has a large impact on the performance of the fleet. The work by Hörl et al. (17) combines the simulation developed in (16) with a detailed model of costs for automated mobility (18). A choice model for conventional and automated modes of transport alike is estimated from a large-scale stated preference survey in the canton of Zurich and added dynamically to the simulation. The study constitutes the first simulation in which a closed cycle between simulation of demand and supply is able to estimate not only what fleet size would be able to serve a certain demand but also for what fleet size and service characteristics customers will be willing to pay. A similar study is available for the city of Paris where it is found that 25,000 AVs can serve 1.2 million trips with dynamically adjusted service costs (19).

Based on the current state of the art in AV research, and with the aim to fill some of the current research gap, this paper proposes:

- a point-to-point on-demand (automated) public transport service to replace car travel;
- aggregating travelers in space and time, at their origin and destination, thus avoiding de-tours; and
- an optimization algorithm that minimizes necessary fleet size and keeps additional VKT minimal with short relocations.

Methodology
This paper addresses the problem of minimal fleet sizing of a pooled vehicle fleet using a mixed-integer linear programming (MILP) model. A variant of a discrete minimum-cost flow problem is used.

The basic idea is that a specific service area is divided into a set of zones. In the example in this study, hexagons
as in Figure 2 are used. From previously obtained travel demand data, time can be divided into bins and the number of trips needed to serve each origin–destination (O-D) pair during every time bin can be tracked. The idea is that all trip departures can be shifted to the end of their respective time bins such that all the travel movements can be pooled.

To serve the trip demand in a certain time bin, vehicles need to be relocated. For that, how much time it takes for a vehicle to move from one zone to another, at any time of the day, needs to be known. Computationally, the demand flows are regarded as constraints that need to be fulfilled in any case. Vehicles also need to be moved empty to be present for when demand occurs for later O-D pairs. Without loss of generality, it is assumed that each vehicle has a capacity of serving one trip. The aim of the algorithm, which will be presented in the next section, is then to find the minimum number of vehicles needed to cover all the demand.

Problem Definition

Formally, assume a set of zones \( \mathcal{Z} \). Furthermore, assume a set of discrete time bins \( t \in T \). Each time bin has a specific end time \( \tau(t) \in \mathbb{R} \). The mandatory demand flows leaving from origin zone \( u \in \mathcal{Z} \) at the end of time bin \( s \in T \) and arriving at destination \( v \in \mathcal{Z} \) before the end of time bin \( t \) shall be denoted as \( d_{s,t,u,v} \in \mathbb{N} \) and be recorded from given data. Note that a discrete number of trips is counted.

Then, vehicle rebalancing flows \( r_{s,t,u,v} \in \mathbb{N} \) from zone \( u \) to zone \( v \) are considered, which start after the vehicles become idle during the time bin \( s \in T \) and arrive during time bin \( t \); that is, they arrive at the latest before the respective demand flow. Note that rebalancing flows cannot go back in time, but they can be performed during the same time bin if travel time allows it, that is,

\[
r_{s,t,u,v} = 0 \quad \forall \tau(t) < \tau(s) \tag{1}
\]

Two special artificial nodes are also introduced: “source” and “sink.” It is assumed that initially all vehicles reside in the source, although they all need to go to the sink eventually. When a vehicle moves to a certain zone from the source, it represents a vehicle being available at the zone at the beginning of the day; a vehicle going to the source represents a vehicle that goes into an idle state until the end of the day in its zone. All flows going to the source are zero:

\[
r_{s,t,\text{Source}, u} = 0 \quad \forall (s, t) \in T^2, \; u \in \mathcal{Z} \tag{2}
\]

and that all flows coming from the sink are zero, too:

\[
r_{s,t,\text{Sink}, v} = 0 \quad \forall (s, t) \in T^2, \; v \in \mathcal{Z} \tag{3}
\]

Each node (zone) also has to fulfill a flow conservation constraint:

\[
\sum_{s \in T} \sum_{u \in \mathcal{Z}} d_{s,t,u,v} + \sum_{s \in T} \sum_{v \in \mathcal{Z}} r_{s,t,u,v} = \sum_{s \in T} \sum_{v \in \mathcal{Z}} d_{t,q,u,v} + \sum_{s \in T} \sum_{v \in \mathcal{Z}} r_{t,q,u,v} \quad \forall l \in T, \; q \in \mathcal{Z}
\]

Consider Figure 1 for a more intuitive example. In Figure 1a, there are four time points (horizontal) and three zones A, B, and C (vertical). There is a demand of 44 trips from time point 1 to 2 from zone A to B, and there is another demand flow from time point 3 to 4 from zone C to A. These demand flows are fixed and need to be served. From the constraints mentioned above, it can be inferred that all points that are independent of any demand (like A2) require that there be no inflow and no outflow at all. However, the inflow constraint suggests that A1 needs to be served by 44 vehicles. As one cannot go back in time, the only way to fulfill this constraint is to create a flow of 44 vehicles from the source. The outflow constraint then demands that 44 vehicles need to be removed from B2. In Figure 1b, a flow of 44 vehicles is added from B2 to the sink. The same procedure applies for the demand from C3 to A4: a vehicle flow is created from the source to C3
to fulfill the inflow constraint and then create a flow of 21 vehicles from A4 to the sink.

However, Figure 1b is not optimal for vehicle count. In Figure 1c all constraints are fulfilled, but first there is a flow of 44 vehicles from the source to A1. Also, there is no option other than sending 21 vehicles from A4 to the source as these 21 demand trips end in A. However, in between there is another option. The constraints are that 44 vehicles must leave B2 and that 21 vehicles must enter C3. Therefore, 21 of those 44 vehicles at B2 are sent to C3 (which fulfills the inflow constraint of C3), and the remaining 23 vehicles are sent to the sink (which fulfills the outflow constraint of B2). Of the total number of vehicles that arrive in the sink, Figure 1c shows only 44 vehicles are registered whereas there are 65 in Figure 1b. As the source resembles the “end of the day” when all vehicles go into the “idle” state, this number also resembles the total number of vehicles in the system. The example in Figure 1 shows how different configurations of feasible vehicle flows influence the objective. Therefore, by rearranging vehicle flows the system can be optimized to yield the minimum number of vehicles.

Formally, the objective $J \in \mathbb{N}$ can be written as:

$$ J = \sum_{s \in T} \sum_{t \in T} \sum_{u \in Z} r_{s,t,u,\text{Sink}} \quad (4) $$

Therefore, the optimization problem becomes:

$$ \text{minimize } r_{s,t,u,v} = \sum_{s \in T} \sum_{t \in T} \sum_{u \in Z} r_{s,t,u,\text{Sink}} $$

subject to

$$ \sum_{t \in T} \sum_{u \in Z} d_{s,t,u,q} + \sum_{u \in Z} \sum_{t \in T} r_{s,t,u,q} = \sum_{t \in T} \sum_{v \in Z} r_{t,v,q,u} \forall t \in T, \quad q \in Z $$

$$ r_{s,t,u,v} = 0 \quad \forall \tau(t) \leq \tau(s) $$

$$ r_{s,t,u,\text{Source}} = 0 \quad \forall (s,t) \in T^2, \quad u \in Z $$

$$ r_{s,t,u,\text{Sink}} = 0 \quad \forall (s,t) \in T^2, \quad v \in Z \quad (5) $$

**Travel Time Constraints and Maximum Distance**

Although the above-mentioned problem is the basic version of the algorithm used in this paper, some additional refinements can be added. So far, travel time between zones is not incorporated into the model formulation. Let $\tau_{s,u,v}$ define the travel time (with respect to free-flow speeds or congested speeds) between two zones $u$ and $v$ during time bin $s$. It may then happen that zone $v$ cannot be reached from zone $u$ within one time bin, because the travel time is too long. However, it is always possible that a vehicle departs at some time at $u$ to arrive on time for a specific time index at $v$. This constraint can be formalized as follows:

$$ r_{s,t,u,v} = 0 \quad \text{if } \tau(s) + \tau_{s,u,v} > \tau(t) \quad (6) $$

Equation 6 indicates that a flow cannot exist if the sum of the departure time and the travel time is greater
than the end time of a certain destination time bin. Looking at Figure 1, this means that certain edges that violate the travel time constraint can be removed. If, for instance, C3 cannot be reached from B2 anymore, because traveling from B to C at this time of day takes too long, the example in Figure 1c will not be a feasible solution and that in Figure 1b will be the optimum.

A second adjustment to the basic model is vehicle movements being only allowed within a certain distance of a specific zone $u$. For instance, Figure 2 shows an origin node in blue. In the hexagon grid in the figure, a “maximum distance of one zone” means that vehicle flows can only happen to the current zone itself or to its direct neighbors. A distance of two means that vehicles are allowed to go to the next further ring of hexagons and so forth. Structurally, this also translates to forcing certain flows to be zero (or removing edges from a graph as shown in Figure 1). This constraint can be formulated as follows:

$$r_{s,t,u,v} = 0 \text{ if } v \text{ not neighboring } u \quad (7)$$

**Vehicle Sizes**

One major the methodological contribution of this paper is to use the MILP model to test the performance of fleets of differently sized vehicles. So far, one vehicle flow unit is assumed to serve one demand flow unit in the above-discussed model. It is easy to simulate the case of a ten-seater vehicle by aggregating the demand into packages. If the demand shown in Figure 1 is to be served by a fleet of ten-seater vehicles, all flows would be rounded up to multiples of ten. The flow from A1 to B2 would become 50, and, therefore, a new demand flow of five units would be required; the flow from B3 to A4 would become 30 and a new flow of three units would be required. Then again, the MILP can be solved and the minimum fleet size of ten-seaters can be found to serve this demand.

In a more advanced-use case, this procedure can be performed in a hierarchical way to find the optimal vehicle size mix. Such a case would start with ten-seater vehicles, but round all flows to the lower bound. In Figure 1, four ten-seaters would be required on A1 to B2 and two ten-seaters on B3 to A1. Then, the minimum number of ten-seaters can be calculated. However, some flow units would be left from the initial problem. On the first trip, four trips that were served would be missing and one trip on the second O-D pair would be missing. Then, an attempt can be made to serve those trips with five-seater or even smaller vehicles.

This procedure is summarized in Algorithm 1.

**Algorithm 1: Finding the fleet mix**

**Input:** Initial demand flow $d_{s,t,u,v}$

**For each** $n$ in {10, 5, 2, 1} (FleetSize $n$)

$$d_{s,t,u,v} = \lfloor d_{s,t,u,v} / n \rfloor$$

FleetSize ($n$) = Solve MILP with $d_{s,t,u,v}$

$$d_{s,t,u,v} = d_{s,t,u,v} - d_{s,t,u,v} \cdot n$$

**Continue**

**Return** FleetSize

**Post-Processing**

After solving the optimization problem in a next step, total VKT is calculated as a sum of all rebalancing and passenger trip movements. In Figure 1c, this would be $(44 \cdot distance_{A1 \to B2} + 44 \cdot distance_{B2 \to C3} + 21 \cdot distance_{C3 \to A4})$

$(44 distance_{A1 \to B2} + 44 distance_{B2 \to C3} + 21 distance_{C3 \to A4})$. 

**Scenario Setup**

The presented Algorithm 1 is applied to a case for the city of Zurich, Switzerland. As stated, one major input to the algorithm is a time-varying O-D matrix. To obtain such a matrix for Zurich, results from a detailed agent-based transport simulation are used.

The transport simulation is based on the eqasim framework (20). It combines the well-known agent-and activity-based transport simulation framework MATSim (9) with capabilities to use discrete mode choice models inside of the simulation (21, 22). In the present work, the output of such a simulation is used, namely, the detailed mobility plans of the full artificial agent population of Zurich. It allows the extraction of origin location, destination location, departure time, and arrival time for each car trip in the city. Although for the present work only this output is relevant, the interested reader can refer to the sources cited above for information on the general framework, to another study by Hörl et al. (17) for details about the specific implementation of the large-scale model of Switzerland and Zurich, and to a variety of case studies in which the modeling framework has successfully been applied to AVs (19), car-sharing (23) and urban air mobility (24).

The detailed trip data are aggregated into time bins and a hexagonal grid that covers the city of Zurich. In the present study, two time-bin sizes, 7.5 and 15 min, are used. The hexagonal grid with a zone radius of 500 m can be seen in Figure 2. Besides the 500-m grid, the 350-m grid is considered in this study.

Time bins of 15 min are interesting to investigate as more than 80% of car trips within the city have a
duration of less than 15 min in a free-flow speed case (Figure 3). Furthermore, 7.5-min time bins are representative for bus and tram headways at peak times in Zurich. In the case of 15-min time bins, departures from every second zone are moved 7.5 min forward to allow for better scheduling of the vehicles. Distances of 350 and 500 m are well in the range of reasonable walking distance from any departure point to public transport facilities.

The study area is slightly larger than the political boundaries of the city of Zurich, as it was designed to include all areas with considerable population density. Figure 4 shows a distribution of start times of car trips that are entirely contained in this area. To find the minimum number of required vehicles but to keep the optimization problem tractable as well, the focus here is on the period of the day that has the highest demand, namely, from 4:00 to 6:15 p.m.

In the status quo scenario, 162,648 cars in total are used in the study area, and they are driven on average 1,779,764 km per day. During the afternoon peak time, these vehicles travel for 386,043 km.

To estimate travel times between zones, routing between all zones’ centroids for each time bin is performed. This routing is based on either free-flow speeds or congested network speeds from the simulation. In this way, fleet performance based on free-flow travel times and congested travel times is investigated. This comparison is interesting because research suggests that AVs have the potential to use road infrastructure much more efficiently. In such a case, their actual performance will probably lie between the free-flow and congested scenarios.

Results

First, the problem presented in the previous section is solved for AVs of passenger capacity equal to one. The results can be seen in Table 1. In the best case the number of vehicles can be reduced to 9,018, but that case does not provide the minimum VKT. Minimum VKT can be observed for the case of the 7.5-min time bin and 350-m radius, in which the increase in VKT as a result of vehicles needing relocation is only 3%.

These results also show that the number of one-seater vehicles required is in the range of 11–18 times smaller than the current fleet size. This reduction is in the range of previous studies (2,3), although with less additional
Table 1. Number of Vehicles Required to Serve the Demand with One-Seaters

<table>
<thead>
<tr>
<th>Time bin (min)</th>
<th>Radius (m)</th>
<th>Congested</th>
<th>One-seaters</th>
<th>VKT (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>350</td>
<td>No</td>
<td>9,333</td>
<td>401,763 (+3.8)</td>
</tr>
<tr>
<td>7.5</td>
<td>500</td>
<td>No</td>
<td>9,018</td>
<td>452,735 (+17.0)</td>
</tr>
<tr>
<td>7.5</td>
<td>350</td>
<td>Yes</td>
<td>10,607</td>
<td>399,830 (+3.5)</td>
</tr>
<tr>
<td>7.5</td>
<td>500</td>
<td>Yes</td>
<td>10,614</td>
<td>448,726 (+16.0)</td>
</tr>
<tr>
<td>15</td>
<td>350</td>
<td>No</td>
<td>10,268</td>
<td>426,322 (+10.4)</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>No</td>
<td>9,797</td>
<td>480,810 (+24.5)</td>
</tr>
<tr>
<td>15</td>
<td>350</td>
<td>Yes</td>
<td>11,649</td>
<td>424,324 (+9.9)</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>Yes</td>
<td>11,567</td>
<td>476,728 (+23.5)</td>
</tr>
</tbody>
</table>

Note: VKT = vehicle kilometers traveled.

Table 2. Number of Vehicles Required to Serve the Demand with Two-Seaters

<table>
<thead>
<tr>
<th>Time bin (min)</th>
<th>Radius (m)</th>
<th>Congested</th>
<th>Two-seaters</th>
<th>VKT (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>350</td>
<td>No</td>
<td>8,347</td>
<td>378,341 (-2.0)</td>
</tr>
<tr>
<td>7.5</td>
<td>500</td>
<td>No</td>
<td>7,342</td>
<td>389,585 (+0.9)</td>
</tr>
<tr>
<td>7.5</td>
<td>350</td>
<td>Yes</td>
<td>9,568</td>
<td>376,612 (-2.5)</td>
</tr>
<tr>
<td>7.5</td>
<td>500</td>
<td>Yes</td>
<td>8,692</td>
<td>386,642 (+0.1)</td>
</tr>
<tr>
<td>15</td>
<td>350</td>
<td>No</td>
<td>8,556</td>
<td>385,038 (-0.3)</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>No</td>
<td>7,302</td>
<td>384,548 (-0.4)</td>
</tr>
<tr>
<td>15</td>
<td>350</td>
<td>Yes</td>
<td>9,834</td>
<td>383,232 (-0.7)</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>Yes</td>
<td>8,697</td>
<td>381,403 (-1.2)</td>
</tr>
</tbody>
</table>

Note: VKT = vehicle kilometers traveled.

Table 3. Number of Vehicles Required to Serve the Demand with Free-Flow Speeds

<table>
<thead>
<tr>
<th>Time bin (min)</th>
<th>Radius (m)</th>
<th>Two-seaters</th>
<th>Five-seaters</th>
<th>Ten-seaters</th>
<th>VKT (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>350</td>
<td>7,928</td>
<td>224</td>
<td>9</td>
<td>375,983 (-2.6)</td>
</tr>
<tr>
<td>7.5</td>
<td>500</td>
<td>6,216</td>
<td>429</td>
<td>55</td>
<td>374,035 (-3.1)</td>
</tr>
<tr>
<td>15</td>
<td>350</td>
<td>7,628</td>
<td>409</td>
<td>31</td>
<td>377,518 (-2.2)</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>5,363</td>
<td>605</td>
<td>146</td>
<td>351,088 (-9.1)</td>
</tr>
</tbody>
</table>

Note: VKT = vehicle kilometers traveled.

Table 4. Number of Vehicles Required to Serve the Demand with Congested Speeds

<table>
<thead>
<tr>
<th>Time bin (min)</th>
<th>Radius (m)</th>
<th>Two-seaters</th>
<th>Five-seaters</th>
<th>Ten-seaters</th>
<th>VKT (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>350</td>
<td>9,152</td>
<td>229</td>
<td>9</td>
<td>373,852 (-3.2)</td>
</tr>
<tr>
<td>7.5</td>
<td>500</td>
<td>7,476</td>
<td>477</td>
<td>61</td>
<td>370,736 (-4.0)</td>
</tr>
<tr>
<td>15</td>
<td>350</td>
<td>8,857</td>
<td>432</td>
<td>31</td>
<td>375,976 (-2.6)</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>6,574</td>
<td>717</td>
<td>154</td>
<td>348,174 (-9.8)</td>
</tr>
</tbody>
</table>

Note: VKT = vehicle kilometers traveled.

empty vehicle kilometers than reported by Hörnl et al. (16).

Going one step further, Table 2 shows the results when travelers are pooled, if possible, into two-seater AVs. The size of the zones now starts to play an important role and the number of required AVs is reduced to 7.302 in the best case, which is only 4.5% of the number of currently privately owned cars. VKT is very similar to the status quo scenario, meaning that pooling of individual travelers cancels out the empty distance driven, with the maximum reduction of VKT of 2.5% in the case of the 7.5-min time bin, 350-m zone radius, and congested speed. It is clear that aggregation to larger time bins also starts to have an effect, which for one-seaters had a negative effect on the fleet size and VKT.

Finally, to investigate the full potential of pooling passengers with similar O-Ds, vehicles with capacities of two, five, and ten are used as part of a mixed vehicle fleet. Tables 3 and 4 show the results for each of these scenarios, for the free-flow and congested speed cases, respectively. In all cases, the number of vehicles needed to serve the demand is reduced along with the total VKT. In the scenario with a 15-min time bin, 500-m radius zone, and free-flow speed, only 3.7% of the original fleet is needed to serve the demand. If one assumes that the congestion will stay at the same level as at the time of writing this share rises to 4.6%. Reduction of VKT is between 2.6% and 9.8%.

Discussion

The results presented in the previous section show that pooling car travelers can not only substantially reduce the number of vehicles needed, with reductions of up to 96%, but can also reduce VKT, which was previously highlighted as one of the drawbacks of shared AV taxi fleets (3,16). Although travelers are delivered close to their destination without making detours on the way for other passengers, they are expected to walk a certain distance from their origin to a pick-up location and from the drop-off location to their final destination. The average walking distances are rather small, between 250 and 350 m depending on the zone size (note that for free-floating car-sharing, customers are willing to walk up to 500 m to rent a car [25]). The average distance translates to 3–5 min of additional access and egress walk times. As maximum delay to a person’s departure time is the size of the time bin, the actual waiting time can be considered to be low.

It is interesting to note that pooling travelers in 15-min bins is more efficient than pooling them in 7.5-min bins from the operator’s perspective, though not substantially. From the user’s perspective, one might consider that 7.5 min is more acceptable to the
users; therefore, in the long run that might be a better option.

Another important finding of this paper is that if AVs drive in the same congested roads as at the time of writing, the number of required resources to serve the demand may rise as high as 20%. This puts the findings of Hörl et al. (16) into perspective, which considers only a free-flow speed case.

Limitations

Although the methodology introduced in this paper can provide important insights into the potentials of an (automated) public transport (AdPT) service, this study has made several simplifications that should be discussed.

Only car trips starting and ending in the city of Zurich are considered as potential trips for the AdPT service. Therefore, commuters surrounding the study area that arrive to the city of Zurich with their cars are still allowed to do that. However, later if they need to conduct a trip within the city of Zurich, they have to switch to an AdPT service. Although this can be regarded as a policy constraint, not allowing previous travelers, who used to walk, cycle, or take public transport, to switch to this new service is a limitation. As a next step it will be important to see what happens if everyone is allowed to use this service, with respect to VKT, required fleet size, and service in general.

Demand is considered to be static in this study. However, as shown by Hörl et al. (16,19), the cost of a shared AV service depends on the fleet size, which, in turn, affects the demand. Therefore, it will be interesting to understand how this new service can be priced to efficiently serve all potential users.

Determining who would switch to an AdPT service if the usage of cars in the city is forbidden is important. However, it is maybe even more crucial to understand what kind of new demand will form. This is beyond the scope of this paper, but an important topic that surrounds all work around AVs.

It is assumed in this paper that the demand is known a priori. However, some of the trips cannot be predicted and scheduled in advance. One of the future steps would be to investigate how different levels of spontaneous trips can affect the needed fleet size and what kind of dispatching strategy will work best in this case.

Conclusion

This study provides insights into how a pooled (automated) vehicle fleet with a point-to-point service can operate in the city of Zurich and what impact it can have on the number of vehicles needed and VKT. Besides the methodology presented, several key findings are also a contribution of this paper. Congestion has a substantial effect on the necessary fleet size to serve the demand and should not be neglected. Pooling travelers not only helps to reduce the vehicles needed but also reduces VKT. Moreover, vehicles of different passenger capacity provide further benefits and should be considered when designing a shared automated fleet service. Aggregating passengers in larger time bins increases the positive effects of the service from the operator perspective, but the benefits might not be large enough to overcome the reduced frequency of the service from the user’s perspective.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: M. Balac and S. Hörl; data collection: M. Balac and S. Hörl; analysis and interpretation of results: M. Balac and S. Hörl; draft manuscript preparation: M. Balac, S. Hörl and K.W. Axhausen. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


