HOW DOES NETWORK COMPLEXITY IMPACT SHARED AUTONOMOUS VEHICLE FLEET OPERATIONS? COMPARING COMPLETE VS INCOMPLETE NETWORKS AND REAL VS AGGREGATED ADDRESSES

Kentaro Mori  
Graduate Research Assistant  
Department of Civil, Architectural and Environmental Engineering  
The University of Texas at Austin  
mori@utexas.edu

Krishna Murthy Gurumurthy, Ph.D.  
Argonne National Laboratory, Energy Transportation and Power Systems Division  
9700 S. Cass Avenue  
Lemont, IL 60439  
kgurumurthy@anl.gov

Fatemeh Fakhrmoosavi, Ph.D.  
Assistant Professor  
Department of Civil and Environmental Engineering  
University of Connecticut  
moosavi@uconn.edu

Pedro Camargo, Ph.D.  
Argonne National Laboratory, Transportation and Power Systems Division  
9700 S. Cass Avenue, Lemont, IL 60439  
pveigadecamargo@anl.gov

Kara M. Kockelman, Ph.D., P.E. (corresponding author)  
Dewitt Greer Centennial Professor in Engineering  
Department of Civil, Architectural, and Environmental Engineering  
The University of Texas at Austin  
kkockelm@mail.utexas.edu  
Tel: 512-471-0210

Submitted for presentation at BTR#6 and presented at TRB24.
ABSTRACT: Agent- and activity-based transportation models are increasingly used to simulate shared autonomous vehicle (SAV) fleet operations, enabling a growing understanding of SAVs’ operations, impacts, and opportunities. Such studies rely on a variety of assumptions and simplifications that can affect conclusions. This paper investigates the issue of network simplifications, since almost all studies have been conducted on coarsened networks, with many missing links and with aggregated addresses for passenger pickups and dropoffs (PUDOs). The issue of missing network is relevant for all travel demand modeling studies and especially pertinent for SAVs with dynamic ride-sharing enabled, door to door. This work compares fleet operations in Austin, Texas using two networks and two sets of addresses for travel and trip ends across the region’s 6 counties. These are the Capital Area Metropolitan Planning Organization’s (CAMPO’s) planning network with addresses highly aggregated, vs OpenStreetMap’s (OSM’s) real network with OpenAddresses (OA) actual addresses. The CAMPO network contains 40.6% of the OSM lane-miles, and there is just one aggregated address point for every 23 actual addresses. Agent-based mesoscopic traffic assignment results (using the POLARIS model) suggest that network simplifications greatly affect traffic conditions, which impact SAV fleet performance. SAVs are able to serve more requests with lower wait times and less vehicle-miles traveled (VMT), thanks to the expanded network. Interestingly, using realistic address points had little to no effect on network congestion and fleet operations. Overall, VMT per SAV and empty VMT (%eVMT) were the fleet metrics most affected by network complexity, but SAV fleet performance was surprisingly consistent relative to the variations in network congestion metrics across the six scenarios studied.

Keywords: Transportation Network Simplification, Shared Autonomous Vehicle Simulations, Agent-based Modeling, Spatial Aggregation

INTRODUCTION
Future transport, of persons and freight, represents an open frontier, characterized by the emergence of innovative solutions enabled by information technologies and vehicle automation, as well as interest in environmental protection with access-centered planning (Dominković et al., 2018; Hancock et al., 2019; Sumalee and Ho, 2018; von Schönfeld and Bertolini, 2017). In an effort to capture the complex interactions of emerging transportation systems, researchers have turned to agent-based modeling, which excels in simulating activity-based travel and operations of emerging mobility services. Agent-based modeling tools with dynamic traffic assignment (DTA) include MATSim (Axhausen et al., 2016), POLARIS (Auld et al., 2016), and SimMobility (Lu et al., 2015). Advances in these tools’ capabilities and computer performance enable researchers to conduct simulations with unprecedented levels of detail and scale. (For example, POLARIS can now handle 30+ million trips per 24-hr simulated day across realistic regional networks with endogenous mode and route choices, mesoscopic traffic assignment, and congestion feedbacks (Gurumuthy et al., 2021).)

One especially exciting application of these tools is fleet operations of shared and fully automated or “autonomous” vehicles (SAVs). Along with surveys of potential SAV users, such simulations provide major insights into how SAVs will impact human systems. Past papers emphasize travel choices, operational strategies, and environmental impacts (Li et al., 2021). While several studies suggest that SAVs can improve transportation equity (Lee and Kockelman, 2022; Nahmias-Biran et al., 2021a; Nahmias-Biran et al., 2021b), uncontrolled implementation
can increase vehicle-miles traveled (VMT) and worsen network congestion (Bischoff and Maciejewski, 2016; Oh et al., 2021; Oh et al., 2020). Fortunately, dynamic ridesharing (DRS), geofencing, and roadway-demand management policies (like credit-based congestion pricing) can mitigate many negative impacts (Fagnant and Kockelman, 2018; Gurumurthy et al., 2020; Simoni et al., 2019). Researchers have also studied the ways in which SAVs can support public transit use (Basu et al., 2018; Gurumurthy et al., 2020b; Huang et al., 2021; Huang et al., 2022; Oke et al., 2020; Shen et al., 2018; Wen et al., 2018).

While significant effort has been put into modeling the complex features of SAVs, details of more fundamental aspects of transportation simulations that can influence the outcomes have been largely neglected. For example, many studies simulate only a small percentage of the total population and use simplified networks that only cover a fraction of the actual roads in a city. Although these simplifications save simulation/computer run time, their influence on fleet metrics (like number of person-trips served, wait times, average vehicle occupancies (AVOs), and shares of empty VMT (%eVMT)) has not been investigated, weakening our collective understanding and investigative conclusions (on fleet sizes needed, added VMT, emissions, profitability, etc.).

While network (and address) abstraction issues are relevant in any transportation simulation, this paper is focused on SAV operations. This work compares SAV fleet operations in Austin using two networks: the 6-county's typically modeled network (from the local metropolitan planning organization: CAMPO) and the region’s actual (OSM) network. The MPO’s network contains just 41% of the OSM’s lane-miles. POLARIS is run on both in order to compare results of multiple-fleet-size, network, and address-aggregation scenarios. Developed by Argonne National Laboratory, POLARIS enables simulations of SAV operations across realistic networks and complex regions (Auld et al., 2016). Similar to other agent-based models, POLARIS empowers users to carefully trace trajectories of individual vehicles and travelers across interconnected networks (involving roadways, walkways, and transitway links). To our knowledge, this study is the first to use two networks from distinct sources for the same region or city and compare the results. The effects of network complexity on SAV fleet operations are revealed through metrics, such as wait times, VMT, %eVMT, and vehicle occupancy, and the run times are compared.

The structure of the paper is as follows. We first describe the problem of network complexity in transportation modeling and examine simplifications taken in recent studies (<4 years) on agent-based simulations of SAVs. We then delve further into the problem of network complexity in the context of SAV simulations and illustrate what a complete network looks like for several popular cities in literature. Next, we describe the experiment using POLARIS traffic simulation tool and analyze the effects of network complexity on SAV fleet operations. Finally, we discuss the conclusions and future directions.

LITERATURE REVIEW
The origin of transportation modeling can be traced back to the 1950s with Chicago Area Transportation Study’s four-step travel model and Beckman’s solution to Wardrop’s user equilibrium (Weiner, 1997; Beckman, 1956). The second breakthrough in transportation modeling occurred in the 1990s with the advent of activity-based models and DTA. Since then,
various dynamic network modeling software, such as DYNASMART (Jayakrishnan, 1994) and DynaMIT (Ben-Akiva et al., 1998), have been introduced to be used in conjunction with demand models. Today, there are various commercial traffic assignment software available, such as Emme, Visum, and TransCAD, offering static traffic assignment, DTA, or both. Recent developments in transportation modeling have been focused on the integration of demand and supply models into one platform. One of the earliest efforts was TRANSIMS, developed by the Los Alamos National Laboratory (Smith et al., 1995). Some popular comprehensive agent-based activity-based modeling systems currently used for cutting-edge research are MATSim (Axhausen et al., 2016), POLARIS (Auld et al., 2016), and SimMobility (Lu et al., 2015).

Mainly due to computational constraints, early studies on transportation modeling relied on simple networks for validation before moving on to application in real networks. Some of these toy networks have become deeply engrained in the transportation research community as benchmarking networks. Perhaps the most notable example is the Sioux Falls network introduced by LeBlanc et al. (2975). Although there are several versions, the Sioux Falls network is generally comprised of 24 nodes and 76 directed links. Despite it being well-acknowledged that the Sioux Falls network is not a realistic network, many recent studies on SAVs have still relied only on the Sioux Falls network, especially research on the operation algorithms of SAVs (Hasanpour Jesri and Akbarpour Shirazi, 2022; Noruzoliaee and Zou et al., 2022; Rong et al., 2022; Tian et al., 2022; Xu et al., 2023; Zhou and Roncoli, 2022). The selection of the Sioux Falls network is often justified due to its well-established status in the transportation field and low computational load, which makes it a reasonable choice for conducting initial analyses. However, utilizing realistic networks in the same studies could enhance the comprehensiveness and practicality of their applications.

As agent-based activity-based modeling systems have been introduced, refined, and adopted over the past few years, many studies on SAVs have been conducted on such platforms using networks from various cities around the world. Table 1 shows a sample of recent studies on agent-based simulations of SAVs that contain at least some basic information about the network. Unfortunately, many studies fail to mention basic information about the simulation, such as the number of links and nodes in the network or the ratio of the simulated population. It should be noted that depending on the platform, links are either unidirectional or bidirectional, but the relative complexity of the network for the modeled area can still be understood. However, links can be of any length, so we recommend that authors provide the percentage of lane-miles to provide the best idea of network complexity. It can be seen that more researchers are starting to obtain more complex networks from OSM in recent years. To the best of our knowledge, only Kamijo et al. (2022) has used both a complete network and population. However, their study was conducted in a small city with a population of only 47,000. Information on the number of address points or possible origins and destinations is even more scarce. This information is only included by Hunter et al. (2023), whose network of Austin contained 39,638 address points. Although some other studies mention the number of TAZs, it is unclear if the simulations utilize more disaggregate possible origin and destination (OD) points instead of TAZ centroids. From the examination of recent literature, it is clear that 1) basic simulation properties in different studies are important in generalizing conclusions and should be
stated, 2) there is a gap to conduct simulations with complete networks, addresses, and populations, and 3) the effects of using incomplete demand or supply need to be understood.

A few studies have been conducted on the effects of population downscaling in agent-based transportation simulations. Ben-Dor et al. (2021) investigated this point using a more detailed version of the Sioux Falls network with 334 links and 282 nodes in MATSim. They looked at statistics, such as average travel distance and duration, trip counts, and traffic volume. They found that a simulated population percentage of 25% or higher was required to maintain the full population outcome. Ratios between 10% and 25% led to biases in some statistics, and results from ratios below 10% proved to be unusable. Also using MATSim, Kagho et al. (2022) studied the effects of population downscaling for a ride-hailing service with DRS in Zurich. They focused on AVO and wait, travel, and detour times. They found that AVO does not asymptotically converge when increasing the simulated population ratio and recommended using the full population or estimating the bias of results when using a partial population. Similarly, Kamijo et al. (2022) examined the effects of population downscaling on simulations with SAVs in the regional city of Numata, Japan, reporting the minimum required simulated population ratio to keep biases to less than 10%. They examined key statistics for describing SAV service, such as wait time, number of trips served, VMT, eVMT%, and profit margin. They found that while SAVs without DRS can be reliably simulated with a population ratio of 10%, the required simulated population ratio is much higher for SAVs with DRS at 60%. The results of Kamijo et al. (2022) and Kagho et al. (2022) match the findings of Fagnant and Kockelman (2018) and Zwick et al. (2021) in that the performance of a fleet with DRS improves with higher demand density. These studies imply that results of SAV simulations with population downscaling, which is usually around 10%, are likely unreliable.
TABLE 1 Recent studies on agent-based simulations of SAVs

<table>
<thead>
<tr>
<th>Paper</th>
<th>City</th>
<th>Simulation platform</th>
<th>Network source</th>
<th>#Links</th>
<th>#Nodes</th>
<th>Population %</th>
<th>Persons</th>
<th>Trips</th>
<th>Area (sq miles)</th>
<th>Time period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hunter et al. (2023)</td>
<td>Austin, USA</td>
<td>POLARIS MPO*</td>
<td>16,059</td>
<td>10,435</td>
<td>100%</td>
<td>5,300</td>
<td>–</td>
<td>–</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Dean et al. (2022)</td>
<td>Austin, USA</td>
<td>POLARIS MPO</td>
<td>16,100</td>
<td>10,400</td>
<td>100%</td>
<td>5,300</td>
<td>–</td>
<td>–</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Gurumurthy and Kockelman (2022)</td>
<td>Bloomington, IL, USA</td>
<td>POLARIS</td>
<td>–</td>
<td>4,000</td>
<td>2,500</td>
<td>100% to 2500%</td>
<td>120,000</td>
<td>74</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Gurumurthy et al. (2021)</td>
<td>Chicago, USA</td>
<td>POLARIS MPO</td>
<td>31,000</td>
<td>19,000</td>
<td>100%</td>
<td>11M</td>
<td>30M</td>
<td>11,246</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2020)</td>
<td>Hague, Netherlands</td>
<td>Anylogic</td>
<td>–</td>
<td>836</td>
<td>510</td>
<td>–</td>
<td>–</td>
<td>27,452</td>
<td>5:30-10:00 AM</td>
<td></td>
</tr>
<tr>
<td>Alisoltani et al. (2020)</td>
<td>Lyon, France</td>
<td>OSM</td>
<td>27,000</td>
<td>11,310</td>
<td>–</td>
<td>480,000</td>
<td>31</td>
<td>6:00-10:00 AM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shafiei et al. (2023)</td>
<td>Melbourne, Australia</td>
<td>DynaMel</td>
<td>–</td>
<td>55,719</td>
<td>24,502</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>6:00-10:00 AM</td>
<td></td>
</tr>
<tr>
<td>Yan et al. (2020)</td>
<td>Minneapolis–Saint Paul, USA (7-counties)</td>
<td>MATSim OSM</td>
<td>42,485</td>
<td>20,746</td>
<td>2%, 5%</td>
<td>180,000 and 457,000</td>
<td>–</td>
<td>6,364</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minneapolis–Saint Paul, USA (Twin Cities)</td>
<td>MATSim OSM</td>
<td>–</td>
<td>–</td>
<td>20%</td>
<td>487,000</td>
<td>–</td>
<td>245</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kamijo et al. (2022)</td>
<td>Numata, Japan</td>
<td>MATSim OSM</td>
<td>–</td>
<td>–</td>
<td>2% to 100%</td>
<td>–</td>
<td>–</td>
<td>360</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Nguyen-Phuoc et al. (2023)</td>
<td>Singapore</td>
<td>SimMobility</td>
<td>–</td>
<td>15,128</td>
<td>6,375</td>
<td>100%</td>
<td>6.7M</td>
<td>283**</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Oh et al. (2021)</td>
<td>Singapore</td>
<td>SimMobility</td>
<td>–</td>
<td>15,128</td>
<td>6,375</td>
<td>100%</td>
<td>6.7M</td>
<td>283**</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Oh et al. (2020)</td>
<td>Singapore</td>
<td>SimMobility</td>
<td>–</td>
<td>15,128</td>
<td>6,375</td>
<td>100%</td>
<td>5.2M and 6.7M</td>
<td>283**</td>
<td>24-h</td>
<td></td>
</tr>
<tr>
<td>Ishibashi and Akiyama (2022)</td>
<td>Tokyo, Japan</td>
<td>MATSim OSM</td>
<td>338,652</td>
<td>134,112</td>
<td>10%***</td>
<td>200,000</td>
<td>–</td>
<td>208**</td>
<td>AM commute trips</td>
<td></td>
</tr>
</tbody>
</table>
Mori et al.

<table>
<thead>
<tr>
<th>Study</th>
<th>Location</th>
<th>Model</th>
<th>OSM</th>
<th>Trip Generation</th>
<th>O-D Demand</th>
<th>VMT</th>
<th>O-D Generation</th>
<th>VMT</th>
<th>OD Entries</th>
<th>VMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Müller et al. (2021)</td>
<td>Vienna, Austria</td>
<td>MATSim</td>
<td>156,000</td>
<td>71,000</td>
<td>200,000</td>
<td>–</td>
<td>1,610</td>
<td>24-h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hörl et al. (2021)</td>
<td>Zurich, Switzerland</td>
<td>MATSim</td>
<td>150,000</td>
<td>–</td>
<td>220,000</td>
<td>–</td>
<td>–</td>
<td>24-h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Metropolitan planning organization
** estimated by us based on information given in the cited paper
*** of corresponding entries in OD table from the Tokyo Person Trip Survey
NETWORK COMPLEXITY IN SAV SIMULATIONS

The use of incomplete networks is a key concern in generalizing the output of SAV simulations because most studies assume that SAVs will provide door-to-door (address point to address point) service, which means first-mile and last-mile travel on missing (or uncoded) minor roads/links. Additionally, simplified networks have the effect of aggregating pickup and drop-off (PUDO) positions of travelers, as it restricts PUDOs to the coded links, resulting in easier ridesharing. Both these features are likely to optimistically bias SAV and non-SAV simulation results. However, it is also plausible that omitting links leads to restricted path sets, thereby increasing congestion and possibly VMT. In such case, the resulting network condition can negatively impact SAV operations. Figures 1 through 4 show the difference between simulated/coded and actual networks, for a range of examples used in the published literature.

Figure 1 (a) Austin’s 6-county network used in several published works and (b) the OSM network of the Austin 6-county area

Figure 2 (a) Singapore network (figure from Alho et al., 2021) used in several published works and (b) the OpenStreetMap network of Singapore
Figure 3 (a) Chicago’s 20-county network (figure from Gurumurth et al., 2020a) used in several published works and (b) the OSM network of the 20-county Chicago region

Figure 4 (a) The classic Sioux Falls network (figure from Xu et al., 2023) and (b) the OSM network of Sioux Falls
Another aspect of network complexity is the impact of using actual addresses. While traditional travel demand models have long relied on traffic analysis zones (TAZs) as the smallest spatial unit of analysis and used zone centroids as the possible origins and destinations (Miller, 2021), agent-based simulation tools like POLARIS offer a more granular approach by tracking travel at the level of individual “locations” or “addresses” serving as origins and destinations (Hunter et al., 2023). While this approach is an improvement over relying on TAZ centroids, the level of detail provided by the simulated addresses can still pose an issue and have a similar effect on SAV simulations as missing links, because simulated addresses essentially act as aggregated PUDO points. Figure 5 shows the location points used in Austin’s 6-county network and actual addresses obtained from OpenAddresses (OA). Austin’s 6-county network contains 39,638 addresses, while 905,925 addressed were obtained from OA for the same area. Figures 6 and 7 zoom into the network as two contrasting examples on the extent of aggregation. It can be seen in these figures that locations are covered very well in some parts of the downtown area (Figure 6), while there are many parts of the network, especially the periphery of the network, that generated addresses in Austin’s 6-county network are much less than the actual ones (Figure 7).

Figure 5. (a) Addresses used in Austin’s 6-county network and (b) actual addresses obtained from OA
Figure 6. Simulated addresses vs actual ones in downtown Austin (blue dots are actual addresses and the red dots are the ones in Austin’s 6-county network)

Figure 7. Example of a residential area on the periphery of network (blue dots are actual addresses and the red dots are the ones in Austin’s 6-county network)
On the demand side, using a smaller synthesized population is a common simplification. Since SAV performance is influenced by population density, simulated population ratio is also a concern. Many studies simulate approximately 10% of the total population in order to save computational cost. To this end, Kamijo et al. (2022) and Kagho et al. (2022) simulated on-demand vehicle services with DRS using various population ratios and recommends using population ratios that are much higher than what has typically been used in literature. Thanks to their work, the effects of simulated population rates on SAV simulation outcomes has been illuminated. However, a similar comparative study looking at the effects of the level of network completeness has not been conducted. While simulated population ratios are simply a matter of computational cost, networks require additional data and effort to alter and are often used repeatedly in consecutive studies, so understanding the effect of network complexity is paramount.

Therefore, in this study, we compared the results of simulations involving SAVs in POLARIS using Austin 6-county region’s network from CAMPO and a network obtained from OSM (Figure 1). The CAMPO network is a well-established network that has been used in numerous previous studies of various topics (Fagnant and Kockelman, 2018; Hunter et al., 2023; Dean et al., 2022; Duthie and Unnikrishnan, 2014; Hu et al., 2017; Perrine et al., 2015; Zhao and Kockelman; 2018). The OSM network of Austin has previously been used by Gurumurthy et al. (2019), Huang et al. (2021), and Liu et al. (2017), but Gurumurthy et al. (2019) simulated only 5% of the population and the analysis by Liu et al. (2017) was limited to an 18 sq. mile central region. To our knowledge, this study is the first to use two networks from distinct sources for the same city. We also compared the results of simulations using aggregated CAMPO addresses and real addresses obtained from OA (Figure 5). The CAMPO addresses have been used by any study using the CAMPO network in an agent-based simulation setting. On the other hand, no study has used addresses from OA as possible OD points.

**SIMULATION FRAMEWORK**

This study uses the POLARIS transportation system simulation tool (Auld et al., 2016) to simulate SAV fleet operations in the Austin 6-county region. POLARIS is an agent-based activity-based modeling framework. It is able to create a synthetic population, simulate a full day of activities for that entire population, and track the movements of individual agents routed through a time-dependent dynamic traffic assignment model. POLARIS contains many features for simulating traffic network company (TNC) vehicles, including SAVs, and is able to output key metrics for understanding the performance of fleet operations, such as wait times, VMT, %eVMT, and vehicle occupancy. In POLARIS, TNC vehicles are centrally controlled by a TNC operator, which gives assignment, operation, and repositioning instructions while considering network congestion. Based on the instructions from the fleet operator, each TNC vehicle maintains a task list, calculates optimal paths from the current location to the next operation location for the certain task, and records trip details (Gurumurthy et al., 2020a). Additionally, POLARIS is the able to simulate DRS, a process by which a single vehicle can be arranged (in real time) to concurrently satisfy the travel demands of multiple travelers (Fagnant and Kockelman, 2018). In this model structure, incomplete networks might lead to missing some of the links taken by vehicles or altered trajectories, which needs to be investigated.
CASE STUDY
In this study, we evaluated the effects of network complexity on the performance of an SAV fleet with 15,000 vehicles (1 SAV per 125 residents) using two different networks and two different sets of addresses. For the base scenario, we used the CAMPO network (Figure 1a) of the Austin 6-county region. For the more complex or complete network, we used the network of the same region obtained from OSM (Figure 1b). The raw OSM network data was converted to the general modeling network specification (GMNS) format using an open-source Python package osm2gmns (Lu and Zhou, 2022), demonstrating the ease of obtaining a more complete network for any region. The GMNS network was then converted to the native POLARIS format and combined with other existing supply inputs for Austin (e.g., zones, locations, transit network) using a Python software developed in this study. Furthermore, in order to eliminate differences in the coding of the network (lanes, capacities, free flow speeds, intersection geometry, etc.) and isolate the effects of the network topology, the corresponding links of the CAMPO network were extracted from the OSM network (Figure 8). The extracted network contains 16,176 lane-miles, which is 40.6% of that of the OSM network. From this point on, this trimmed OSM network is referred to as the CAMPO network, and any mention of the original CAMPO network is explicitly stated as “original CAMPO.” Additionally, the study examined the combinations of the two networks and two sets of addresses obtained from CAMPO and OA (Figure 5), producing four main scenarios. The OA database for the Austin 6-county region contains nearly 1 million addresses, which is roughly 23 times that of the CAMPO addresses. Because a fixed trip table (from Dean et al. (2023) with 100% population) was used in this study instead of endogenously modeling travel demand to maintain the focus on SAV operations, each OA address was mapped to the nearest CAMPO address, as shown in Figure 9, for randomly reassigning trips ODs. In POLARIS, each location can be connected to up to four links to avoid network loadings that would create artificially high levels of congestion when using a simplified network. However, to avoid unequal effects of distance thresholds in different networks (based on link density), locations were only connected to the nearest link, except in one scenario, as no gridlocks were observed. To examine the effect multiple location links, the scenario with CAMPO network and CAMPO locations was run with single location links as well as multiple location links with a distance threshold of 300 meters. The scenarios were identical in all other aspects not mentioned.
RESULTS

The results of the simulations are summarized in Table 2, showcasing key insights into the performance of SAVs in the two networks. Additionally, the network loading curves showing the number of vehicles in the network in each minute of the simulation are shown in Figure 10. We conducted these simulations on the Texas Advanced Computing Center's Lonestar6 high-
performance computing system, utilizing 2 nodes equipped with 2x AMD EPYC 7763 64-Core Processors and 256 GB (3200 MT/s) DDR4 RAM. The run times for the CAMPO network were approximately 6 hours, while the OSM scenarios took around 12 hours. The number of addresses did not affect run time. Each scenario was run for 10 iterations to ensure convergence.

As shown in Figure 10a, the additional links of the OSM network (with CAMPO addresses) had the effect of noticeable network decongestion, with 8.0% and 13.7% reductions in VMT and vehicle-hours traveled (VHT), respectively. This suggests that streets deemed negligible by CAMPO provide attractive alternative paths according to the time-dependent A* routing algorithm used in POLARIS (Verbas et al., 2018). Interestingly, the original CAMPO network experienced substantially higher levels of congestion compared to the CAMPO network trimmed out of the OSM network, with 4.7% and 22.3% higher VMT and VHT, respectively. This is despite the original CAMPO network containing 65.7% more lane miles than the OSM-based CAMPO network, as a result of inflations in the number of lanes during calibration of static traffic assignment by CAMPO. This suggests that there may be some bottlenecks in the original CAMPO network for dynamic traffic assignment. The network conditions reflected directly on SAV fleet performance. Consequently, SAV fleet performance was more optimistic and efficient in the OSM network, serving 0.7 more trips per SAV while VMT/SAV fell by the same margin as the network-wide VMT. %eVMT and median wait time fell by 3.6 percentage points and 1.32 minutes, respectively. No large change in average trip distance or AVO was observed.

The effects of aggregating addresses proved to be minor. As shown in Figure 10b, using real addresses with the CAMPO network led to slight decongestion, with no change in VMT and 6.7% reduction in VHT. It was also revealed that using multiple location links can mimic the effect of using disaggregate addresses, as both has the effect of diversifying network loading points. By using multiple location links with the CAMPO network, VMT and VHT fell by 3.1% and 3.3%, respectively, compared to connecting each address to the closest link. SAV fleet performance was similar between the two sets of addresses for the CAMPO network. Network conditions were nearly identical in the OSM network regardless of what addresses were used, as demonstrated by Figure 10c. SAV fleet performance was also similar between the two sets of addresses for the OSM network, but interestingly, average trip distance fell by 0.5 miles when using OA addresses.

Overall, SAV fleet performance was surprisingly consistent relative to the variations in the overall traffic conditions observed across the six scenarios. SAVs were able to satisfy at least 94% of all requests, on average completing 28 to 29 trips. Median wait times were roughly 4±1.5 minutes. VMT/SAV and %eVMT saw the most variations across the scenarios, ranging between 270-310 miles and 17-23%, respectively.
TABLE 2 Summary of SAV simulation results

<table>
<thead>
<tr>
<th>Network</th>
<th>Addresses</th>
<th>Trips/SAV per day</th>
<th>% requests met</th>
<th>VMT/SAV per day</th>
<th>% eVMT</th>
<th>Avg trip distance (miles)</th>
<th>Revenue trip AVO</th>
<th>Revenue distance AVO</th>
<th>Median wait time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAMPO</td>
<td>CAMPO</td>
<td>28.5 trips</td>
<td>97.3%</td>
<td>305 mi</td>
<td>20.7%</td>
<td>6.93 mi</td>
<td>1.69</td>
<td>1.51</td>
<td>4.00 min</td>
</tr>
<tr>
<td></td>
<td>OA</td>
<td>28.6</td>
<td>97.6</td>
<td>309</td>
<td>21.6%</td>
<td>6.77</td>
<td>1.71</td>
<td>1.52</td>
<td>4.1</td>
</tr>
<tr>
<td>OSM</td>
<td>CAMPO</td>
<td>29.2</td>
<td>99.5</td>
<td>278</td>
<td>17.1%</td>
<td>6.76</td>
<td>1.66</td>
<td>1.45</td>
<td>2.68</td>
</tr>
<tr>
<td></td>
<td>OA</td>
<td>29.2</td>
<td>99.5</td>
<td>272</td>
<td>17.2%</td>
<td>6.29</td>
<td>1.74</td>
<td>1.54</td>
<td>2.97</td>
</tr>
<tr>
<td>Original CAMPO</td>
<td>CAMPO</td>
<td>27.8</td>
<td>94.7</td>
<td>298</td>
<td>23.1%</td>
<td>6.34</td>
<td>1.73</td>
<td>1.58</td>
<td>5.87</td>
</tr>
<tr>
<td>CAMPO (multiple location links)</td>
<td>CAMPO</td>
<td>28.7</td>
<td>97.8</td>
<td>295</td>
<td>19.8%</td>
<td>6.85</td>
<td>1.69</td>
<td>1.49</td>
<td>3.17</td>
</tr>
</tbody>
</table>
Figure 10. Network loading curves
CONCLUSIONS

In this study, we evaluated the influence of network complexity on SAV fleet operations using the POLARIS agent-based modeling platform. We simulated SAV fleet operations in Austin using two networks: Austin 6-county region’s network from CAMPO and a network extracted from OSM, and two sets of addresses: aggregated address points from CAMPO and real addresses from OA. The CAMPO network consists 40.6% of the OSM network’s lane-miles, while each aggregated CAMPO address point represents 23 actual addresses, although there is significant spatial variation. The results showed that omission/addition of links affects the network-wide traffic condition, which reflects on the SAV fleet performance. Consequently, SAV fleet performance was more optimistic and efficient in the OSM network, serving more requests and reducing wait times while traveling less cumulative distance. On the other hand, the number of address points had minor to negligible effects on both the overall network state and SAV fleet operations. Overall, VMT/SAV and %eVMT were the metrics most affected by network complexity, but SAV fleet performance was surprisingly consistent relative to the variations in the overall traffic conditions observed across the six scenarios studied. Therefore, researchers can rest assured that their (i.e., on the level of networks used by metropolitan planning organizations). Although adding missing links is an easy way to increase realism in agent-based simulations, adding all links to the network was not worthwhile in this case, as simulation run time doubled. While additional address points did not add to the computational load, the assumptions for trip distribution (like the one made in this paper) or additional land use data needed for trip generation may not be preferable. Therefore, we recommend connecting each address point to multiple nearby links for loading, as this seems to have a similar effect to adding address points.

Although this study provides insights into the impact of network simplifications on SAV operations, there are certain limitations that warrant consideration. In this study, we did not calibrate the networks against traffic flow counts. Doing so would likely narrow the gap between the traffic conditions of networks of varying complexity. However, this step was forgone in this study, as supply calibration is not always performed and also to identify the effects of missing links that theoretically should have minimal effects on routing. Additionally, as mentioned above, the random redistribution of trip ODs from aggregated addresses to real ones may not accurately reflect the trip demand that would be generated considering the land use. However, this was done instead of generating trips so that the same fixed trip table could be used across all scenarios for the best comparison. Finally, only one fleet size was tested. The differences across the scenarios may be more pronounced with a smaller fleet size, as fleet performance deteriorates rapidly if fleet size is not adequate.

ACKNOWLEDGEMENTS

The work done in this paper was sponsored by the U.S. Department of Energy (DOE) Vehicle Technologies Office (VTO) under the Systems and Modeling for Accelerated Research in Transportation (SMART) Mobility Laboratory Consortium, an initiative of the Energy Efficient Mobility Systems (EEMS) Program. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

AUTHOR CONTRIBUTIONS
Kentaro Mori: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. Krishna Murthy Gurumurthy: Methodology, Resources, Writing – review and editing, Funding acquisition. Fatemeh Fakhrmoosavi: Conceptualization, Methodology, Data curation, Writing – review and editing. Pedro Camargo: Software, Writing – review and editing. Kara M. Kockelman: Conceptualization, Methodology, Resources, Writing – review and editing, Project administration, Funding acquisition.

REFERENCES


