HOW DO NETWORK COMPLETENESS & OD ADDRESS DETAILS IMPACT TRAFFIC AND SAV FLEET OPERATIONS? INPUT DATA RESOLUTIONS IN AGENT-BASED TRANSPORTATION MODELS

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

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

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

Pedro Camargo, Ph.D.
Transportation and Power Systems Division
Argonne National Laboratory
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 publication in *Computers, Environment and Urban Systems* (September 2024)

ABSTRACT: Agent-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). This issue is relevant for all travel demand modeling studies and especially pertinent for SAVs with dynamic ridesharing 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 actual addresses from OpenAddresses. 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 dynamic traffic assignment results (using the POLARIS model) suggest that while omitting links significantly affects network congestion, address aggregation does not have a substantial impact, provided the addresses are detailed enough at the level of census tract centroids supplemented with business establishment information. Relative to the variations in network congestion observed across the scenarios studied, SAV fleet performance remained surprisingly consistent, though VMT per SAV and empty VMT (%eVMT) were the fleet metrics most affected by omitted links in the network.

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 activity scheduling and 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, 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; Karolemeas et al., 2024). While several studies suggest that SAVs can improve transportation equity (Lee and Kockelman, 2022; Nahmias-Biran et al., 2021a & 2021b), research also suggests that uncontrolled implementation can increase vehicle-miles traveled (VMT) and worsen network congestion (Bischoff and Maciejewski, 2016; Oh et al., 2020 & 2021). 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 & 2022; Oke et al., 2020; Shen et al., 2018; Wen et al., 2018).

While significant effort has been devoted to modeling the complex features of SAVs, more fundamental aspects of transportation simulations, which can influence outcomes, have been largely overlooked. 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 research conclusions (on fleet sizes needed, added VMT, emissions, profitability, etc.).

While network (and origin-destination (OD)/address) abstraction issues are relevant in any transportation simulation, this paper focus on SAV operations while also examining network traffic patterns. This work compares simultaneous DTA and SAV fleet operation simulations in Austin using two different networks: the 6-county's typically modeled network (from the local metropolitan planning organization: CAMPO) and the region's actual (OpenStreetMap) network (OpenStreetMap, 2024). It also applies two different sets of addresses: aggregated addresses from past studies and real addresses from OpenAddresses (2024). The simulation experiments are conducted using the POLARIS Transportation Simulation Tool. 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 (and addresses) from distinct sources for the same region or city and compare the results. The effects of network completeness and OD address details 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 completeness and OD address details in transportation modeling and examine simplifications taken in recent (<5 years) studies on agent-based simulations of SAVs. We then delve further into the problem of network completeness and OD address details 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 completeness and OD address details 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, such as Emme, Visum, and TransCAD, offering static traffic assignment, DTA, or both. Recent developments in transportation modeling have been focused on integrating 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). Popular comprehensive agent-based activity-based modeling systems used in cutting-edge research include MATSim (Axhausen et al., 2016), POLARIS (Auld et al., 2016), and SimMobility (Lu et al., 2015).

Early studies on transportation modeling relied on simple networks for validation before moving on to 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. (1975). Although several versions exist, the Sioux Falls network is generally comprised of just 24 nodes and 76 directed links. Although it is well-acknowledged that the Sioux Falls network is not a realistic representation, many recent studies on SAVs, particularly those on SAV operation algorithms, have still relied solely on this network (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, adding realistic networks to 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 several years, numerous studies on SAVs have been conducted 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 used. 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 can be either unidirectional or bidirectional, but the relative complexity of the network for the modeled area can still be understood from link and node counts. 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 completeness. It can be observed from Table 1 that more researchers are obtaining complex networks from OpenStreetMap (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 rural city with a population of only 47,000. Information on the number of address points or possible ODs is even more scarce. Among the works in Table 1, this information is provided only by Hunter et al. (2023) and Fakhrmoosavi et al. (2024), both using a network of Austin with 39,638 OD points for a population of approximately 2 million. Although some other studies mention the number of traffic analysis zones (TAZs), it is unclear whether the simulations use TAZ centroids or more disaggregate points as ODs. From the examination of recent literature, it is evident that 1) basic simulation properties in different studies are important in generalizing conclusions and should be stated, 2) there is a gap in conducting simulations with complete networks, addresses, and populations, and 3) the effects of using incomplete demand or supply need to be understood.

Few studies have investigated the effects of population downscaling in agent-based transportation simulations. Ben-Dor et al. (2021) studied this effect using a more detailed version of the Sioux Falls network with 334 links and 282 nodes in MATSim. By comparing average travel distances and durations, trip counts, and traffic volumes, they found that a minimum simulated population percentage of 25% is 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, Switzerland, focusing on AVO and wait, travel, and detour times. They found that AVO does not asymptotically converge with respect to simulated population ratio and recommends using a 100% population or estimating the bias of results if using a partial population. Kamijo et al. (2022) conducted a similar study in Numata, Japan, and reported the simulated population ratio needed to keep biases of SAV operation metrics (including wait time, number of trips served, VMT, eVMT%, and profit margin) to less than 10%. They found that while SAVs without DRS can be reliably simulated with a population ratio of just 10%, the 60% is required for simulations with DRS. The results of Kagho et al. (2022) and Kamijo et al. (2022) are consistent with the findings of Fagnant and Kockelman (2018) and Zwick et al. (2021) in that the performance of a vehicle fleet with DRS improves with higher demand density. Kuehnel et al. (2023) proposed a solution to this problem called “flow-inflated selective sampling,” which calls for fully simulating modes sensitive to demand density, while downscaling demand for other modes and inflating their capacity consumption in the network. These studies suggest that results of SAV simulations with population downscaling, typically around 10%, are likely unreliable.

TABLE 1 Recent studies on agent-based simulations of SAVs

Paper	City	Simulation Platform	Network Source	#Links	#Nodes	Population %	Area (sq. miles)	Time Period
Fakhrmoosavi et al. (2024)	Austin, USA	POLARIS	MPO*	16,059	10,435	100%	5,300	24-h
Huang et al. (2024)	Austin, USA	POLARIS	–	22,863	17,231	100%	5,480	24-h
Hunter et al. (2023)	Austin, USA	POLARIS	MPO	16,059	10,435	100%	5,300	24-h
Dean et al. (2022)	Austin, USA	POLARIS	MPO	16,100	10,400	100%	5,300	24-h
Gurumurthy and Kockelman (2022)	Bloomington, IL, USA	POLARIS	–	4,000	2,500	100% to 2500%	74	24-h
Li et al. (2024)	Brussels, Belgium	MATSIM	OSM	–	–	10%	–	24-h
Dean et al. (2024)	Chicago, USA	POLARIS	MPO	48,400	35,800	–	–	24-h
Gurumurthy et al. (2021)	Chicago, USA	POLARIS	MPO	31,000	19,000	100%	11,246	24-h
Wang et al. (2020)	Hague, Netherlands	Anylogic	–	836	510	–	–	5:30-10:00 AM
Ben-Dor et al. (2022)	Jerusalem, Israel	MATSim	–	8500	–	30%	695	24-h
Alisoltani et al. (2020)	Lyon, France	–	OSM	27,000	11,310	–	31	6:00-10:00 AM
Shafiei et al. (2023)	Melbourne, Australia	DynaMel	–	55,719	24,502	–	–	6:00-10:00 AM
Yan et al. (2020)	Minneapolis–Saint Paul, USA (7-counties)	MATSim	OSM	42,485	20,746	2% & 5%	6,364	24-h
	Minneapolis–Saint Paul, USA (Twin Cities)					20%	245	
Kamijo et al. (2022)	Numata, Japan	MATSim	OSM	–	–	2% to 100%	360	24-h
Nguyen-Phuoc et al. (2023)	Singapore	SimMobility	–	15,128	6,375	100%	283**	24-h
Oh et al. (2021)	Singapore	SimMobility	–	15,128	6,375	100%	283**	24-h
Oh et al. (2020)	Singapore	SimMobility	–	15,128	6,375	100%	283**	24-h
Ishibashi and Akiyama (2022)	Tokyo, Japan	MATSim	OSM	338,652	134,112	10%	208**	AM commute trips
Müller et al. (2021)	Vienna, Austria	MATSim	OSM	156,000	71,000	12.5%	1,610	24-h
Peer et al. (2024)	Vienna, Austria	MATSim	OSM	156,000	–	12.5%	1,583	24-h
Hörl et al. (2021)	Zurich, Switzerland	MATSim	–	150,000	–	10%	–	24-h

* Metropolitan planning organization

** Estimated based on information given in the cited paper

NETWORK COMPLETENESS AND OD ADDRESS DETAILS 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 simulation results. However, it is also plausible that the omitting of links artificially restricts routes, thereby increasing congestion. In this case, the resulting network condition can negatively impact SAV fleet operations. Figures 1 through 4 show the difference between simulated/coded and actual networks, for a range of examples repeatedly used in the published literature.

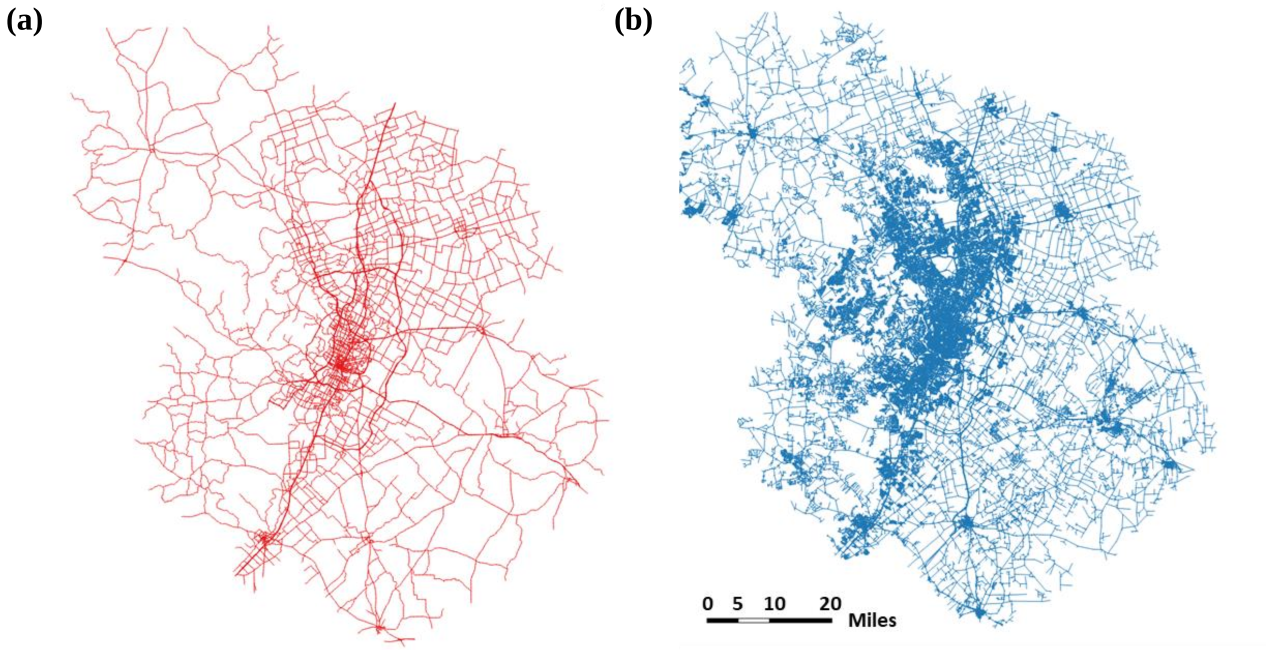

Figure 1 (a) CAMPO's Austin's 6-County Network and (b) the Corresponding OSM Network

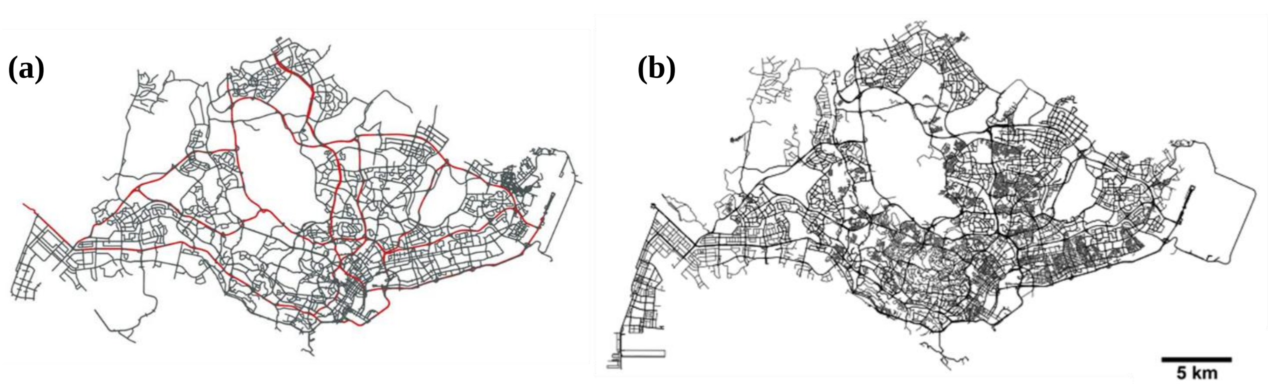

Figure 2 (a) Alho et al.'s (2021) Singapore Network and (b) the Corresponding OSM Network

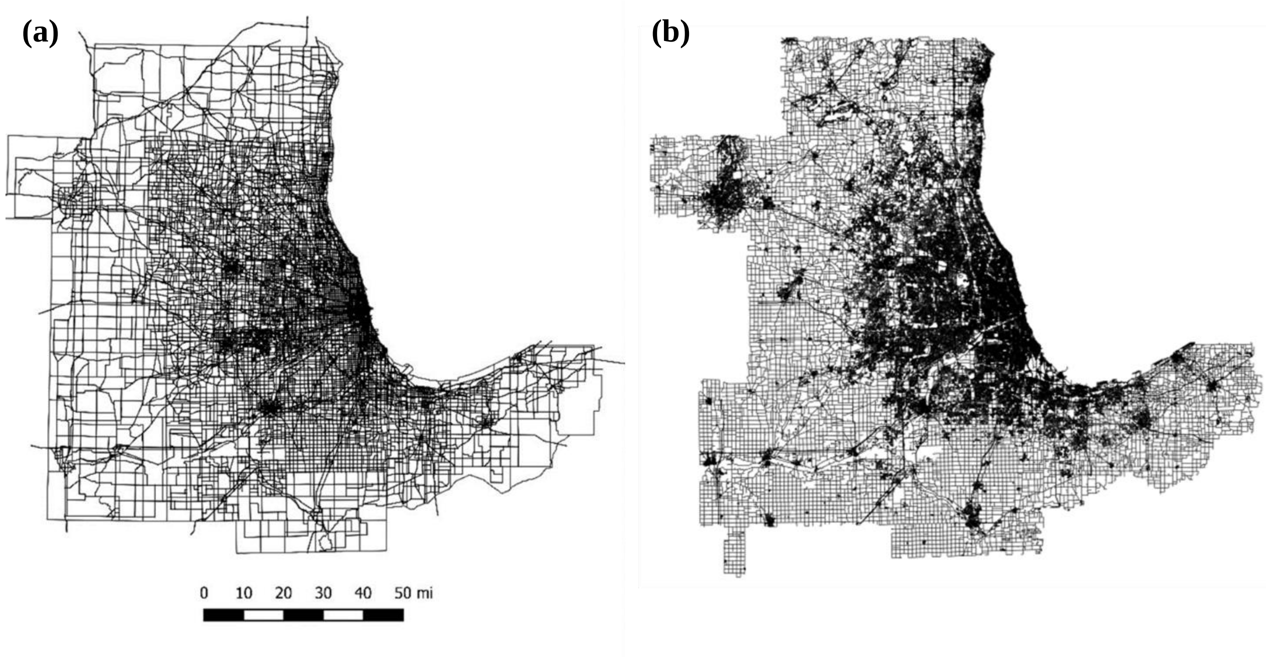

Figure 3 (a) Gurumurthy et al.'s (2020a) 20-County Chicago Network and (b) the Corresponding OSM Network

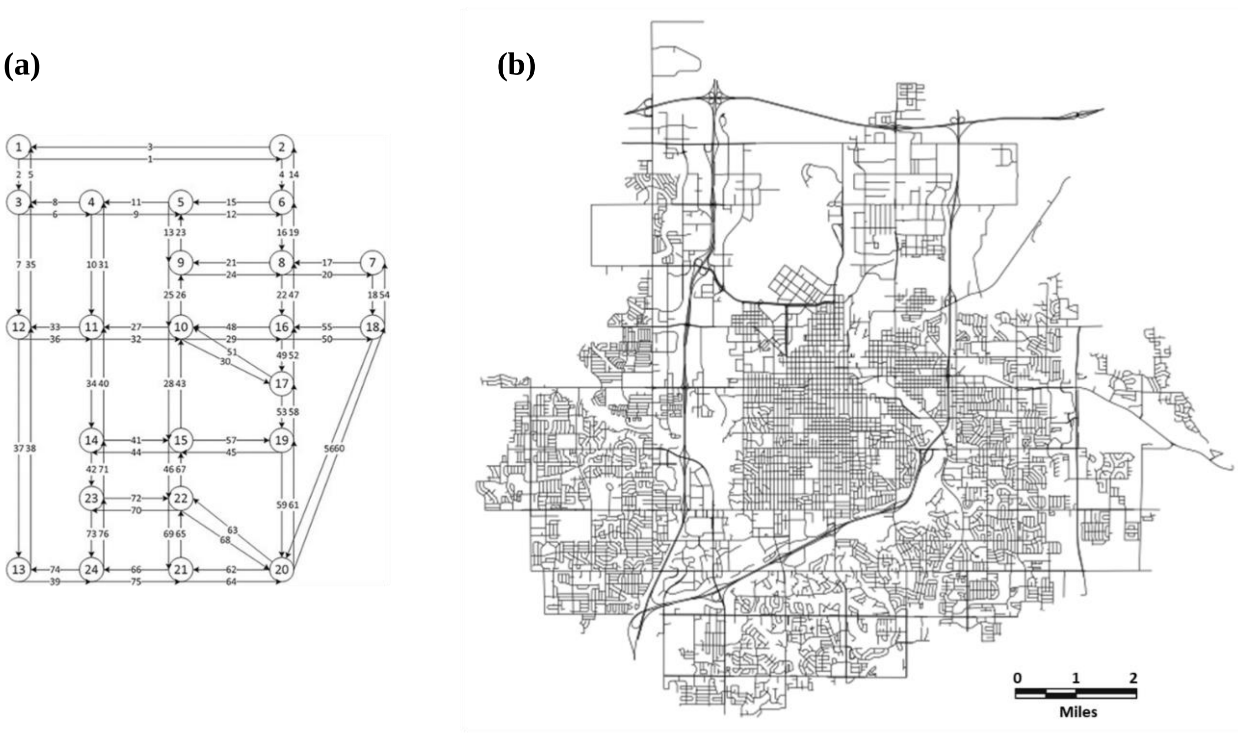

Figure 4 (a) The Classic Sioux Falls Network (Figure from Xu et al., 2023) and (b) the Corresponding OSM Network

The degree of OD aggregation is another important consideration. While traditional travel demand models have long relied on TAZs as the smallest spatial unit of analysis and used zone centroids as ODs (Miller, 2021), agent-based simulation tools like POLARIS offer a more granular approach by tracking trips to individual “locations” or “addresses” serving as ODs (Hunter et al., 2023). While this is a significant improvement over TAZ centroids, the simulated addresses can still act as aggregated PUDO points depending on the level of detail, impacting SAV simulation results. Figure 5 shows the 39,638 location points that have used in Austin’s 6-county network, which are comprised of non-residential locations from CoStar (CoStar, 2024) and census block centroids for residential locations, versus the 905,925 actual addresses available on OpenAddresses (OA). Figures 6 and 7 zoom into the network as two contrasting examples on the extent of aggregation. Locations are well-covered in some parts of the downtown area (Figure 6), whereas in many areas, particularly on the periphery of the network, the number of generated addresses is markedly lower than the actual number (Figure 7).

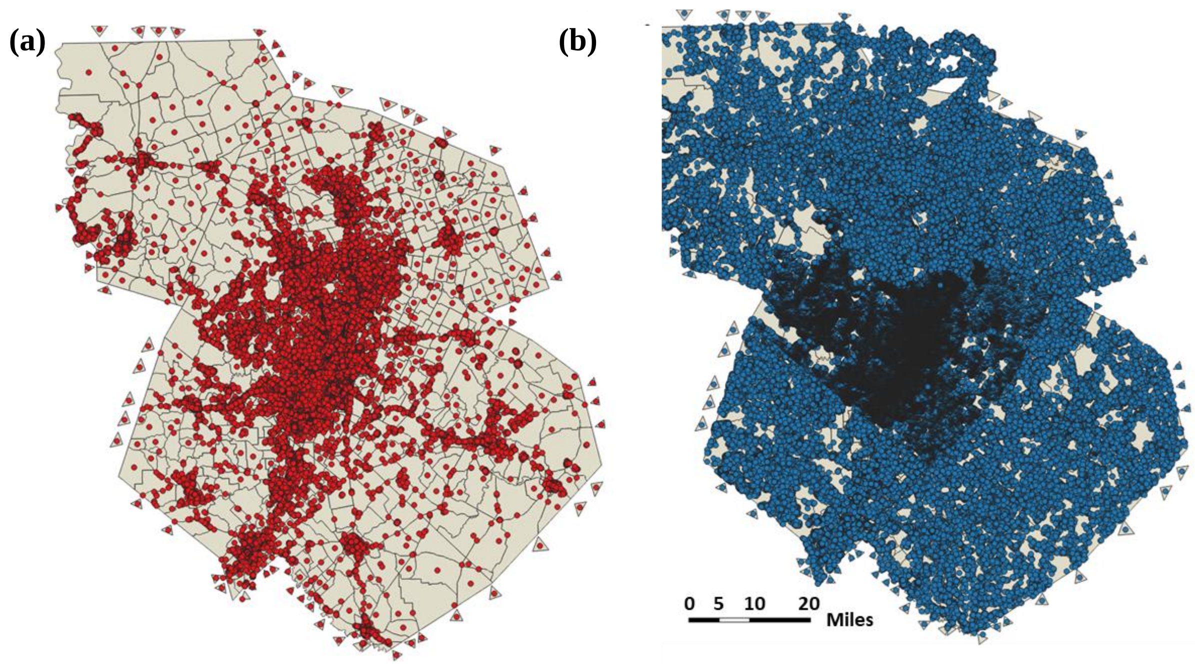

Figure 5. (a) Addresses used in Austin’s 6-county network and (b) actual addresses obtained from OpenAddresses (OA)

Figure 6. Simulated (red) vs Actual (blue) Addresses in Downtown Austin

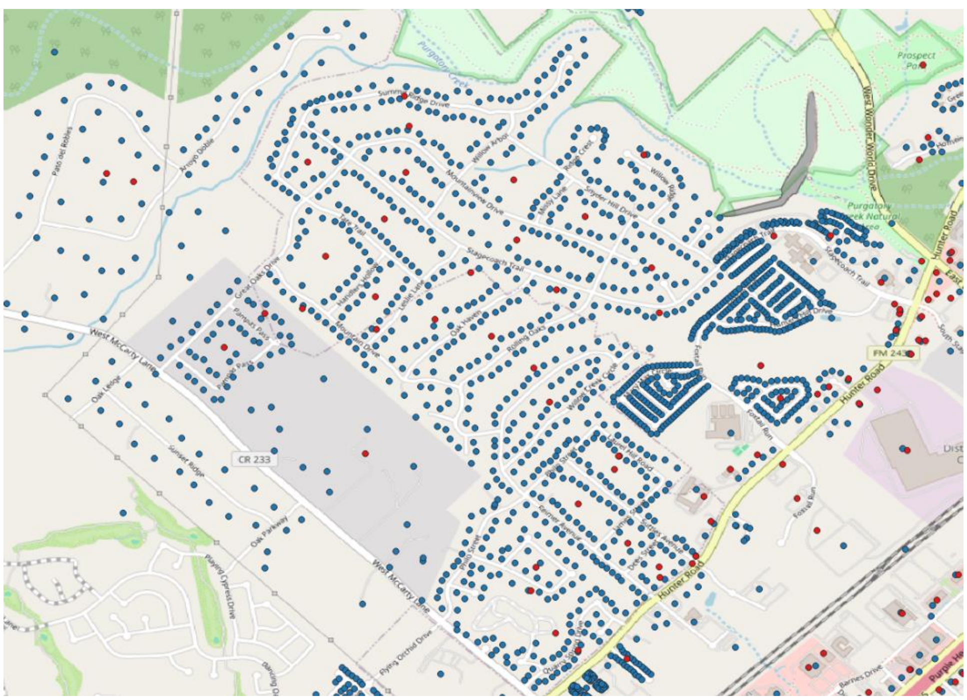

Figure 7. Simulated (red) vs Actual (blue) Addresses a Residential Area on the Periphery of the Network

As mentioned, using a simulating a smaller population is a common simplification on the demand side. To this end, Kagho et al. (2022) and Kamijo et al. (2022) has illuminated on the effects of simulated population rates on SAV fleet simulation outcomes, which is sensitive to demand density. However, a similar comparative study on the effects of the level of network completeness or address aggregation has not yet been conducted. Furthermore, while simulated population ratios can easily be changed with one parameter, networks require additional data and effort to alter and are often reused in multiple studies, so understanding the effect of network completeness and OD address details is paramount.

Therefore, in this study, we compare 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 (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). 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 addresses created from CoStar and census block data and real addresses obtained from OA (Figure 5). The aggregated addresses have been used by all existing studies of Austin using POLARIS. 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 fleets of 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 AVO. 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, 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.

Typically, POLARIS simulations are run with a fully endogenous activity-based travel demand model that is calibrated to match known aggregate statistics in the base year. However, this study uses a fixed trip table across all scenarios to maintain the focus on SAV fleet operations and traffic simulations. To this end, a DTA iteration procedure is used for route choice convergence, where route choices of agents with satisfactory experienced travel times (using thresholds determined by the method of successive averages) are preserved between iterations - instead of rerouting them through the network using the time-dependent A* shortest path algorithm (Verbas et al., 2018).

CASE STUDY

In this study, we evaluated the effects of network completeness and OD address details on the performance of an SAV fleet with 15,000 vehicles (approximately 1 SAV per 125 residents) using combinations of different networks and sets of addresses. The raw OSM network data for the Austin 6-county region 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 package developed in this study (Argonne National Laboratory, 2024). Signalized intersections were added to the network using information from OSM, and traffic signs were added using rule-based procedure.

Furthermore, in order to eliminate differences in local attributes of the CAMPO and OSM networks (lane counts, capacities, free-flow speeds, intersection/interchange geometries and controls, etc.), links in the OSM network that are in the *original CAMPO* network were extracted to create the *OSM-based CAMPO network* (Figure 8). Figure 8 shows the three networks compared in this study. Table 3 shows the numbers of links and nodes, lane-miles by facility type, and numbers of sign-controlled and signalized intersections for the three networks studied. Although the *original* and *OSM-based CAMPO networks* contain the same roads, as illustrated in Figure 8, there are significant discrepancies in lane-miles. The original CAMPO network has less freeway and expressway lane-miles (due to removing of auxiliary lanes) and less controlled intersections (since missing links are not considered in controls placement) than the OSM-based CAMPO network, but more lane-miles for other arterials because the MPO added lanes to adjust for missing links in their static model. Links from OSM are shorter, resulting in higher link and node counts in the OSM-based CAMPO network compared to the original CAMPO network. Future work using OSM networks could consider combining links to reduce network loading computations during DTA. The OSM-based CAMPO network contains 16,176 lane-miles, which is 40.6% of that of the OSM network. While the CAMPO network contains nearly all expressways and arterials, it excludes approximately 80% of local roads.

The OA database for the Austin 6-county region contains nearly 1 million addresses, which is roughly 23 times that of the aggregated addresses used in past works (Figure 5). A fixed trip table (from Dean et al. (2023) but with 25% population) was used in this study instead of endogenously modeling travel demand, so each OA address was mapped to the nearest aggregated address, as illustrated in Figure 9, to randomly reassign trips ODs. In POLARIS, each location can be connected to up to four nearby links to avoid network entries 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, as no gridlocks were observed. A total of five scenarios were examined in this study: the OSM-based CAMPO and full OSM networks, each tested with both aggregated and OA addresses, as well as the original CAMPO network with aggregated addresses.

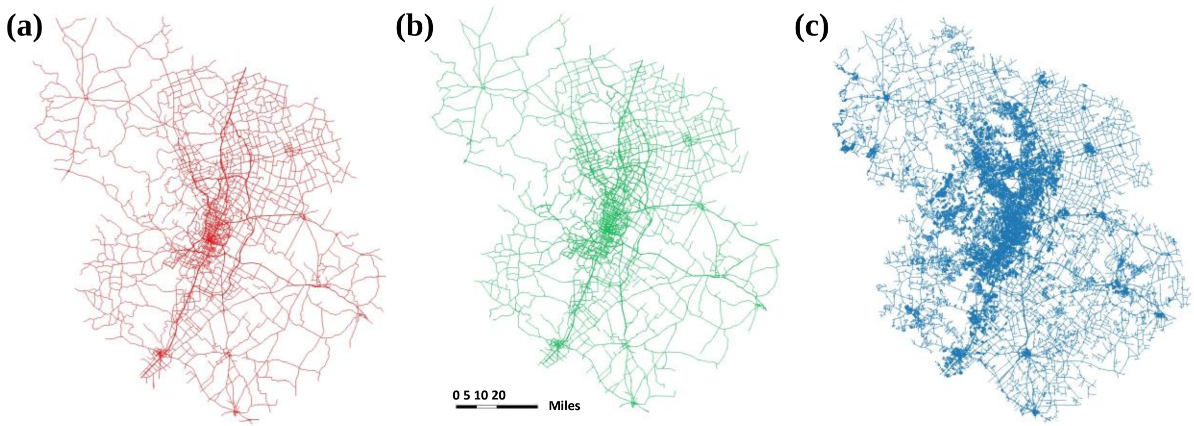

Figure 8. (a) Original CAMPO Network, (b) OSM-Based CAMPO Network, and (c) Full OSM Network

TABLE 2 Summary of Networks

Network	# Links*	# Nodes	Total Lane-Miles	Freeway + Expressway lane-miles	Other Arterial Lane-Miles	Collector + Local Road Lane-Miles	# Sign-Controlled Intersections	# Signalized Intersections
Original CAMPO	15,833	13,177	28,609	1,383	19,367	6,059	1,823	1,150
OSM-Based CAMPO	39,650	33,052	16,176	2,159	8,106	5,910	5,802	1,856
Full OSM	146,176	111,704	39,800	2,269	8,260	29,271	64,820	2,361

* Links in POLARIS are bidirectional.

Figure 9. OA Addresses (blue) Mapped to the Nearest Aggregated Address (red) for Reassigning Trip ODs

RESULTS

The results of the simulations are summarized in Table 3, showcasing key insights into traffic in the network, performance of SAVs, and run time. Additionally, the in-network 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 24 threads on 1 node equipped with 2x AMD EPYC 7763 64-Core Processors and 256 GB (3200 MT/s) DDR4 RAM. The OSM-based CAMPO network had approximately 25% longer run times compared to the original CAMPO network, because the shorter OSM links and higher link and node counts mean more computation is required for traffic simulations. Adding the missing links to the OSM-based CAMPO network more than doubled the run time. Meanwhile, the number of addresses did not affect the run times. Each scenario was run until satisfactory route convergence was reached using the process outlined in the Simulation Framework section.

Compared to the OSM-based CAMPO network, the original CAMPO network had 7% lower vehicle-hours traveled (VHT) but similar VMT, due to inflated lane counts and fewer controlled intersections. In particular, the PM peak in Figure 10a is 25% higher in the OSM-based CAMPO network compared to the original CAMPO network. On the other hand, the full OSM network exhibits a 3% lower VMT but a 3% higher VHT compared to the OSM-based CAMPO network. However, VMT and VHT on links that are present in the OSM-based CAMPO network decreased by 11% and 16%, respectively, indicating decongestion due to alternative paths. From Figure 10a, a significant portion the travel on the missing links appears to occur during the mid-day period between 10:30 AM and 5:00 PM. Despite differences in network congestions, SAV fleet performances were similar across the board. The SAV fleet performed slightly better in the full OSM network than in the OSM-based CAMPO network due to decongestion, with VMT per SAV, %eVMT, and median wait time falling 5%, 2.3 percentage-points, and 0.5 minutes, respectively. Interestingly, %eVMT was the highest in the original CAMPO network, while trips were nearly 1 mile shorter compared to the OSM-derived networks.

The effects of aggregating addresses proved to be minor. As shown in Figure 10b, using OA addresses resulted in virtually no change in the network loading characteristics. Consequently, SAV fleet performances were nearly identical as well. This suggests that the aggregated addresses derived from CoStar and census block data are sufficiently detailed for simulating passenger travel and SAV fleet operations.

Overall, SAV fleet performance remained surprisingly consistent relative to the variations in the network congestion (due to network differences but not address differences) observed across the five scenarios studied. Each SAV was able to complete, on average, 28 trips with median wait times of 5 to 6 minutes. VMT per SAV and %eVMT exhibited the most variation across the scenarios, ranging roughly from 325 to 350 miles and 22% to 27%, respectively.

TABLE 3 Summary of SAV simulation results

Scenario		Network Metrics		SAV Fleet Metrics						Average Iteration Run Time (hh:mm)
Network	Addresses	VMT	VHT	Trips/SAV per day	VMT/SAV per day	% eVMT	Avg Trip Distance (miles)	Revenue Distance AVO	Median Wait Time (min)	
	Aggregated	43.9 M	1.55 M	28.3 trips	348 mi	24.5%	7.8 mi	1.41 pax	5.6 min	4:18

OSM-Based CAMPO	OA	43.9 M	1.58 M	28.1	347	24.9	7.7	1.43	5.7	4:20
Full OSM	Aggregated	42.4 M	1.60 M	28.5	331	22.2	7.9	1.37	5.1	9:15
	OA	42.7 M	1.59 M	28.2	326	22.3	7.7	1.39	5.2	9:23
Original CAMPO	Aggregated	43.5 M	1.44 M	28.2	330	26.7	7.0	1.38	5.4	3:31

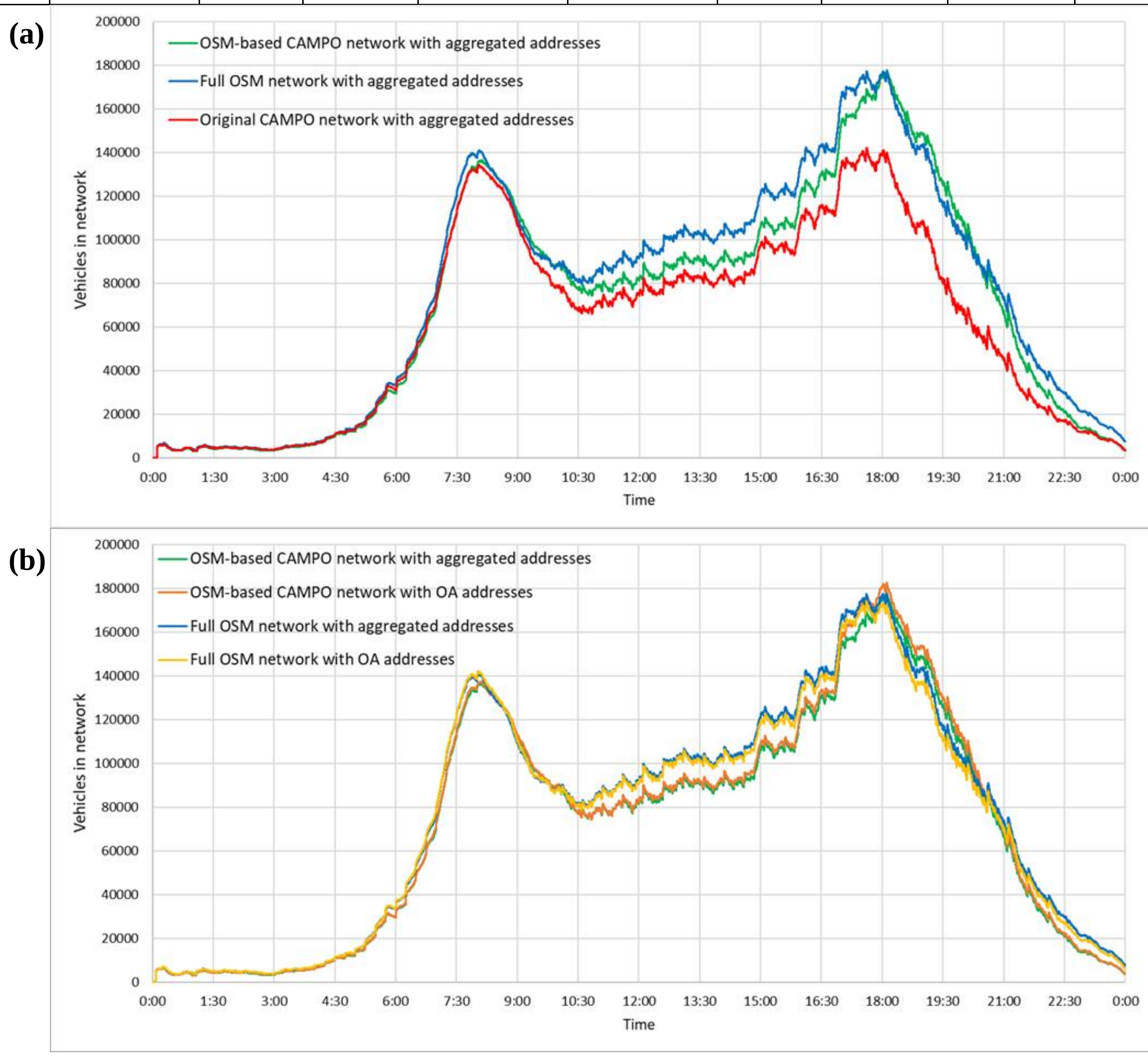

Figure 10. Number of Vehicles Active/Moving in the Network by Time of Day: (a) Across the 3 Distinct Networks (with aggregated addresses) and (b) Using Different Address Sets (for the OSM-based CAMPO and full OSM networks)

CONCLUSIONS

In this study, we evaluated the influence of network completeness and OD address details 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 past studies and real addresses from OA. The CAMPO network, once standardized to OSM link attributes, comprises 40.6% of the lane-miles in the OSM network, while each aggregated address point represents, on average, 23 actual addresses, though there is significant spatial variation. The results show that omission/addition of links affects the network-wide traffic condition, while the set of addresses used had negligible impact. In contrast to the variations in the network congestion observed across scenarios, SAV fleet performance remained relatively consistent. Therefore, researchers can rest assured that their findings from SAV simulations hold as long as a network of reasonable realism is used (e.g., networks used by MPOs). 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.

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 observed traffic flow counts. However, differences in route choices revealed in this study when comparing an incomplete versus a complete network call into question the validity of calibrating to observed traffic flow counts when links are missing. 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 examined in this study. 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

This manuscript and the work described were sponsored by the U.S. Department of Energy (DOE) Vehicle Technologies Office (VTO) under the Pathways to Net-Zero Regional Mobility, 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. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC and database resources that have contributed to the research results reported within the study.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Alho, A.R., Sakai, T., Oh, S., Cheng, C., Seshadri, R., Chong, W.H., Hara, Y., Caravias, J., Cheah, L., and Ben-Akiva, M. (2021) A Simulation-Based Evaluation of a Cargo-Hitching Service for E-Commerce Using Mobility-on-Demand Vehicles. *Future Transportation,* 1, 656.

Alisoltani, N., Zargayouna, M., and Leclercq, L. (2020) Real-Time Autonomous Taxi Service: An Agent-Based Simulation. *Agents and Multi-Agent Systems: Technologies and Applications 20,* , pp. 199-207.

Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., and Zhang, K. (2016) POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. *Transportation Research Part C: Emerging Technologies,* 64, 101-116.

Argonne National Laboratory. (2024) polaris-studio. *POLARIS 24.03 Documentation*, 2024.

Axhausen, K.W., Horni, A., and Nagel, K. (2016) *The Multi-Agent Transport Simulation MATSim.* Ubiquity Press.

Basu, R., Araldo, A., Arun, P.A., Bat Hen, N.B., Basak, K., Seshadri, R., Deshmukh, N., Kumar, N., Carlos, L.A., and Moshe Ben-Akiva. (2018) Automated Mobility-on-Demand vs. Mass Transit: A Multi-Modal Activity-Driven Agent-Based Simulation Approach. *Transportation Research Record,* 2672, 608-618.

Beckmann, M.J., McGuire, C.B., and Winston, C.B. (1956) *Studies in the Economics of Transportation.* Yale University Press, New Haven, CT.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., and Mishalani, R. (February, 1998) DynaMIT: a simulation-based system for traffic prediction. *DACCORD Short Term Forecasting Workshop*.

Ben-Dor, G., Ben-Elia, E., and Benenson, I. (2021) Population downscaling in multi-agent transportation simulations: A review and case study. *Simulation Modelling Practice and Theory,* 108, 102233.

Ben-Dor, G., Ogulenko, A., Klein, I., and Benenson, I. (2022) Modal Shift and Shared Automated Demand-Responsive Transport: A Case Study of Jerusalem. *Procedia Computer Science,* 201, 581–586.

Bischoff, J. and Maciejewski, M. (2016) Simulation of City-wide Replacement of Private Cars with Autonomous Taxis in Berlin. *Procedia Computer Science,* 83, 237-244.

CoStar. (2024) CoStar. *CoStar,* 2024.

Dean, M.D., de Souza, F., Gurumurthy, K.M., and Kockelman, K.M. (2023) Multi-stage charging and discharging of electric vehicle fleets. *Transportation Research Part D: Transport and Environment,* 118, 103691.

Dean, M.D., Gurumurthy, K.M., de Souza, F., Auld, J., and Kockelman, K.M. (2022) Synergies between repositioning and charging strategies for shared autonomous electric vehicle fleets. *Transportation Research Part D: Transport and Environment,* 108, 103314.

Dean, M.D., Gurumurthy, K.M., Zhou, Z., Verbas, O., Cokyasar, T., and Kockelman, K. (2024) An integrated transportation-power system model for a decarbonizing world. *Transportation Planning and Technology*, 1–37.

Dominković, D.F., Bačeković, I., Pedersen, A.S., and Krajačić, G. (2018) The future of transportation in sustainable energy systems: Opportunities and barriers in a clean energy transition. *Renewable and Sustainable Energy Reviews*, 82, 1823-1838.

Duthie, J. and Unnikrishnan, A. (2014) Optimization Framework for Bicycle Network Design. *Journal of Transportation Engineering,* 140, 04014028.

Fagnant, D.J. and Kockelman, K.M. (2018) Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas. *Transportation,* 45, 143-158.

Fakhrmoosavi, F., Gurumurthy, K.M., Kockelman, K.M., Hunter, C.B., and Dean, M.D. (2024) Parking Strategies and Outcomes for Shared Autonomous Vehicle Fleet Operations. *Journal of Transportation Engineering, Part A: Systems*, 150, 04024009.

Gurumurthy, K.M., de Souza, F., Enam, A., and Auld, J. (2020a) Integrating Supply and Demand Perspectives for a Large-Scale Simulation of Shared Autonomous Vehicles. *Transportation Research Record,* 2674, 181-192.

Gurumurthy, K.M. and Kockelman, K.M. (2022) Dynamic ride-sharing impacts of greater trip demand and aggregation at stops in shared autonomous vehicle systems. *Transportation Research Part A: Policy and Practice,* 160, 114-125.

Gurumurthy, K.M., Kockelman, K.M., and Auld, J. (2021) A system of shared autonomous vehicles for Chicago; Understanding the effects of geofencing the service. *Journal of Transport and Land Use,* 14, 933-948.

Gurumurthy, K.M., Kockelman, K.M., and Simoni, M.D. (2019) Benefits and Costs of Ride-Sharing in Shared Automated Vehicles across Austin, Texas: Opportunities for Congestion Pricing. *Transportation Research Record,* 2673, 548-556.

Gurumurthy, K.M., Kockelman, K.M., and Zuniga-Garcia, N. (2020b) First-Mile-Last-Mile Collector-Distributor System using Shared Autonomous Mobility. *Transportation Research Record,* 2674, 638-647.

Hancock, P.A., Nourbakhsh, I., and Stewartd, J. (2019) On the future of transportation in an era of automated and autonomous vehicles. *Proceedings of the National Academy of Sciences - PNAS,* 116, 7684-7691.

Hasanpour Jesri, S.O. and Akbarpour Shirazi, M. (2022) Bi Objective Peer-to-Peer Ridesharing Model for Balancing Passengers Time and Costs. *Sustainability,* 14.

Hörl, S., Becker, F., and Axhausen, K.W. (2021) Simulation of price, customer behaviour and system impact for a cost-covering automated taxi system in Zurich. *Transportation Research Part C: Emerging Technologies,* 123, 102974.

Hu, X., Chiu, Y., and Shelton, J. (2017) Development of a behaviorally induced system optimal travel demand management system. *Journal of Intelligent Transportation Systems,* 21, 12-25.

Huang, Y., Kockelman, K.M., and Garikapati, V. (2022) Shared automated vehicle fleet operations for first-mile last-mile transit connections with dynamic pooling. *Computers, Environment and Urban Systems,* 92, 101730.

Huang, Y., Kockelman, K.M., Garikapati, V., Zhu, L., and Young, S. (2021) Use of Shared Automated Vehicles for First-Mile Last-Mile Service: Micro-Simulation of Rail-Transit Connections in Austin, Texas. *Transportation Research Record,* 2675, 135-149.

Huang, Y., Kockelman, K.M., and Gurumurthy, K.M. (2024) Agent-based simulations of shared automated vehicle operations: reflecting travel-party size, season and day-of-week demand variations. *Transportation,* .

Hunter, C.B., Kockelman, K.M., and Djavadian, S. Curb Allocation and Pick-Up Drop-Off Aggregation for a Shared Autonomous Vehicle Fleet. *International Regional Science Review,* , 01600176231160498.

Ishibashi, Y. and Akiyama, E. (2022) Predicting the Impact of Shared Autonomous Vehicles on Tokyo Transportation Using MATSim. - *2022 IEEE International Conference on Big Data (Big Data),* pp. 3258-3265.

Jayakrishnan, R., Mahmassani, H.S., and Hu, T. (1994) An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C: Emerging Technologies,* 2, 129-147.

Kagho, G.O., Meli, J., Walser, D., and Balac, M. (2022) Effects of population sampling on agent-based transport simulation of on-demand services. *Procedia Computer Science,* 201, 305-312.

Kamijo, Y., Parady, G., and Takami, K. (2022) Required Simulated Population Ratios for Valid Assessment of Shared Autonomous Vehicles’ Impact Using Agent-Based Models. *SSRN Electronic Journal*.

Karolemeas, C., Tsigdinos, S., Moschou, E., and Kepaptsoglou, K. (2024) Shared autonomous vehicles and agent based models: a review of methods and impacts. *European Transport Research Review,* 16, 25.

Kuehnel, N., Rewald, H., Axer, S., Zwick, F., and Findeisen, R. (2024) Flow-Inflated Selective Sampling for Efficient Agent-Based Dynamic Ride-Pooling Simulations. *Transportation Research Record,* 2678, 229–244.

LeBlanc, L.J., Morlok, E.K., and Pierskalla, W.P. (1975) An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research,* 9, 309-318.

Lee, J. and Kockelman, K.M. (2022) Access Benefits of Shared Autonomous Vehicle Fleets: Focus on Vulnerable Populations. *Transportation Research Record,* 2676, 568-582.

Li, J., Rombaut, E., and Vanhaverbeke, L. (2021) A systematic review of agent-based models for autonomous vehicles in urban mobility and logistics: Possibilities for integrated simulation models. *Computers, Environment and Urban Systems,* 89, 101686.

Li, J., Rombaut, E., and Vanhaverbeke, L. (2024) How far are we towards sustainable Carfree cities combining shared autonomous vehicles with park-and-ride: An agent-based simulation assessment for Brussels. *Computers, Environment and Urban Systems*, 112, 102148.

Liu, J., Kockelman, K.M., Boesch, P.M., and Ciari, F. (2017) Tracking a system of shared autonomous vehicles across the Austin, Texas network using agent-based simulation. *Transportation*, 44, 1261-1278.

Lu, J. and Zhou, X.S. (2022) Modeling partially schedulable connected and automated mobility systems on layered virtual-track networks: modeling framework and open-source tools.

Lu, Y., Adnan, M., Basak, K., Pereira, F.C., Carrion, C., Saber, V.H., Loganathan, H., and Ben-Akiva, M.E. (2015) SimMobility Mid-Term Simulator: A State of the Art Integrated Agent Based Demand and Supply Model. *94th Annual Meeting of the Transportation Research Board*.

Miller, E.J. (2021) *Traffic Analysis Zone Definition: Issues & Guidance.*

Müller, J., Straub, M., Naqvi, A., Richter, G., Peer, S., and Rudloff, C. (2021) MATSim Model Vienna: Analyzing the Socioeconomic Impacts for Different Fleet Sizes and Pricing Schemes of Shared Autonomous Electric Vehicles. *Transportation Research Board 100th Annual Meeting*.

Nahmias-Biran, B., Oke, J.B., Kumar, N., Lima Azevedo, C., and Ben-Akiva, M. (2021) Evaluating the impacts of shared automated mobility on-demand services: an activity-based accessibility approach. *Transportation*, 48, 1613-1638.

Nahmias-Biran, B., Oke, J.B., and Kumar, N. (2021) Who benefits from AVs? Equity implications of automated vehicles policies in full-scale prototype cities. *Transportation Research Part A: Policy and Practice*, 154, 92-107.

Nguyen-Phuoc, D.Q., Zhou, M., Hong Chua, M., Romano Alho, A., Oh, S., Seshadri, R., and Le, D. (2023) Examining the effects of Automated Mobility-on-Demand services on public transport systems using an agent-based simulation approach. *Transportation Research Part A: Policy and Practice*, 169, 103583.

Noruzoliaee, M. and Zou, B. (2022) One-to-many matching and section-based formulation of autonomous ridesharing equilibrium. *Transportation Research Part B: Methodological*, 155, 72-100.

Oh, S., Lentzakis, A.F., Seshadri, R., and Ben-Akiva, M. (2021) Impacts of Automated Mobility-on-Demand on traffic dynamics, energy and emissions: A case study of Singapore. *Simulation Modelling Practice and Theory*, 110, 102327.

Oh, S., Seshadri, R., Azevedo, C.L., Kumar, N., Basak, K., and Ben-Akiva, M. (2020) Assessing the impacts of automated mobility-on-demand through agent-based simulation: A study of Singapore. *Transportation Research Part A: Policy and Practice*, 138, 367-388.

Oke, J.B., Akkinepally, A.P., Chen, S., Xie, Y., Aboutaleb, Y.M., Azevedo, C.L., Zegras, P.C., Ferreira, J., and Ben-Akiva, M. (2020) Evaluating the systemic effects of automated mobility-on-demand services via large-scale agent-based simulation of auto-dependent prototype cities. *Transportation Research Part A: Policy and Practice*, 140, 98-126.

OpenAddresses. (2024) OpenAddresses.

OpenStreetMap. (2024) OpenStreetMap.

Peer, S., Müller, J., Naqvi, A., and Straub, M. (2024) Introducing shared, electric, autonomous vehicles (SAEVs) in sub-urban zones: Simulating the case of Vienna. *Transport Policy,* 147, 232–243.

Perrine, K., Khani, A., and Natalia Ruiz-Juri. (2015) Map-Matching Algorithm for Applications in Multimodal Transportation Network Modeling. *Transportation Research Record,* 2537, 62-70.

S. Rong, C. Tang, and M. Chen. (2022) Dynamic shared autonomous vehicles dispatching algorithm. - *ISCTT 2022; 7th International Conference on Information Science, Computer Technology and Transportation,* , pp. 1-5.

Shafiei, S., Gu, Z., Grzybowska, H., and Cai, C. (2023) Impact of self-parking autonomous vehicles on urban traffic congestion. *Transportation,* 50, 183-203.

Shen, Y., Zhang, H., and Zhao, J. (2018) Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in Singapore. *Transportation Research Part A: Policy and Practice,* 113, 125-136.

Simoni, M.D., Kockelman, K.M., Gurumurthy, K.M., and Bischoff, J. (2019) Congestion pricing in a world of self-driving vehicles: An analysis of different strategies in alternative future scenarios. *Transportation Research Part C: Emerging Technologies,* 98, 167-185.

Smith, L., Beckman, R., Baggerly, K., Anson, D., and Williams, M. (1995) *TRANSIMS: Transportation Analysis and Simulation System.* , Los Alamos, NM.

Sumalee, A. and Ho, H.W. (2018) Smarter and more connected: Future intelligent transportation system. *IATSS Research,* 42, 67-71.

Tian, J., Jia, H., Wang, G., Lin, Y., Wu, R., and Lv, A. (2022) A Long-Term Shared Autonomous Vehicle System Design Problem considering Relocation and Pricing. *Journal of Advanced Transportation,* 2022, 1905526.

Verbas, Ö, Auld, J., Ley, H., Weimer, R., and Driscoll, S. (2018) Time-Dependent Intermodal A* Algorithm: Methodology and Implementation on a Large-Scale Network. *Transportation Research Record,* 2672, 219-230.

von Schönfeld, K.C. and Bertolini, L. (2017) Urban streets: Epitomes of planning challenges and opportunities at the interface of public space and mobility. *Cities,* 68, 48-55.

Wang, S., Almeida Correia, G.H.d., and Lin, H.X. (2020) Effects of Coordinated Formation of Vehicle Platooning in a Fleet of Shared Automated Vehicles: An Agent-based model. *Transportation Research Procedia,* 47, 377-384.

Weiner, E. (1997) Urban transportation planning in the United States: an historical overview.

Wen, J., Chen, Y.X., Nassir, N., and Zhao, J. (2018) Transit-oriented autonomous vehicle operation with integrated demand-supply interaction. *Transportation Research Part C: Emerging Technologies*, 97, 216-234.

Xu, T., Cieniawski, M., and Levin, M.W. (2023) FMS-dispatch: a fast maximum stability dispatch policy for shared autonomous vehicles including exiting passengers under stochastic travel demand. *Transportmetrica A: Transport Science,* , 1-39.

Yan, H., Kockelman, K.M., and Gurumurthy, K.M. (2020) Shared autonomous vehicle fleet performance: Impacts of trip densities and parking limitations. *Transportation Research Part D: Transport and Environment,* 89, 102577.

Zhao, Y. and Kockelman, K.M. (2018) Anticipating the Regional Impacts of Connected and Automated Vehicle Travel in Austin, Texas. *Journal of Urban Planning and Development,* 144.

Zhou, Z. and Roncoli, C. (2022) A scalable vehicle assignment and routing strategy for real-time on-demand ridesharing considering endogenous congestion. *Transportation Research Part C: Emerging Technologies,* 139, 103658.

Zwick, F., Kuehnel, N., Moeckel, R., and Axhausen, K.W. (2021) Ride-Pooling Efficiency in Large, Medium-Sized and Small Towns - Simulation Assessment in the Munich Metropolitan Region. *Procedia Computer Science,* 184, 662-667.