Transportation (2019) 46:1975–1996
https://doi.org/10.1007/s11116-019-10007-9

Evaluating the impact of spatio-temporal demand forecast aggregation on the operational performance of shared autonomous mobility fleets

Florian Dandl[1] · Michael Hyland[2] · Klaus Bogenberger[1] · Hani S. Mahmassani[3]

Published online: 4 May 2019

Abstract

Fleet operators rely on forecasts of future user requests to reposition empty vehicles and efficiently operate their vehicle fleets. In the context of an on-demand shared-use autonomous vehicle (AV) mobility service (SAMS), this study analyzes the trade-off that arises when selecting a spatio-temporal demand forecast aggregation level to support the operation of a SAMS fleet. In general, when short-term forecasts of user requests are intended for a finer space–time discretization, they tend to become less reliable. However, holding reliability constant, more disaggregate forecasts provide more valuable information to fleet operators. To explore this trade-off, this study presents a flexible methodological framework to evaluate and quantify the impact of spatio-temporal demand forecast aggregation on the operational efficiency of a SAMS fleet. At the core of the methodological framework is an agent-based simulation that requires a demand forecasting method and a SAMS fleet operational strategy. This study employs an offline demand forecasting method, and an online joint AV-user assignment and empty AV repositioning strategy. Using this forecasting method and fleet operational strategy, as well as Manhattan, NY taxi data, this study simulates the operations of a SAMS fleet across various spatio-temporal aggregation levels. Results indicate that as demand forecasts (and subregions) become more spatially disaggregate, fleet performance improves, in terms of user wait time *and* empty fleet miles. This finding comes despite demand forecast quality decreasing as subregions become more spatially disaggregate. Additionally, results indicate the SAMS fleet significantly benefits from higher quality demand forecasts, especially at more disaggregate levels.

Keywords Autonomous vehicles · Mobility service · Autonomous mobility on-demand · Demand forecasting · Fleet operations

✉ Hani S. Mahmassani
masmah@northwestern.edu

[1] Institute for Intelligent Transportation Systems, Bundeswehr University Munich, Werner-Heisenberg-Weg 39, 85779 Neubiberg, Germany

[2] Department of Civil and Environmental Engineering, Institute of Transportation Studies, University of California-Irvine, 4000 Anteater Instruction and Research Bldg (AIRB), Irvine, CA 92697-3600, USA

[3] William A. Patterson Distinguished Chair in Transportation, Northwestern University, 215 Chambers Hall, 600 Foster Street, Evanston, IL 60208, USA

Introduction

The growth of mobility services (e.g. Uber and Lyft) and the availability of large data sources (e.g. taxi and carsharing data) has prompted significant research in the transportation literature. The advent of fully-autonomous vehicles (AVs) and their expected inclusion in mobility service fleets has further motivated research relating to the operation and management of shared-use AV mobility services (SAMSs). In terms of SAMS fleet operations, two research areas have proliferated, namely, forecasting demand/user requests (i.e. modeling arrival processes) and developing operational policies/strategies to efficiently operate a SAMS fleet dynamically.

The existing literature largely treats these two SAMS problems independently. To address the forecasting problem, researchers are developing and comparing demand forecasting methods (Sayarshad and Chow 2016). To address the problem of operating SAMS fleets efficiently, researchers are developing strategies to assign AVs to user requests (Alonso-Mora et al. 2017; Hyland and Mahmassani 2018; Maciejewski et al. 2016) and reposition empty AVs (Dandl and Bogenberger 2018; Fagnant and Kockelman 2014; Hörl et al. 2018; Pavone et al. 2012; Sayarshad and Chow 2017; Spieser et al. 2016). The repositioning strategies rely on forecasts of future demand to reposition empty AVs; hence, these two SAMS subproblems are inherently interconnected.

Figure 1 shows the relationship between demand forecasting (i.e. predictive analytics) and SAMS operational decision-making (i.e. prescriptive analytics). Predictive analytics methods convert 'raw' data into (demand) forecasts; whereas, prescriptive analytics (i.e. optimization) methods rely on these demand forecasts to make informed (operational) decisions (IBM 2017). Hence, the efficient operation of a SAMS fleet requires reliable demand forecasts. Moreover, forecasts only provide real value to a SAMS provider if they improve decision making and fleet performance.

Motivated by the emergence of SAMSs and the inherent interconnection between demand forecasts and the operational performance of SAMS fleets, this study aims to connect these two research areas. This is not simply an academic exercise; mobility service providers need to consider both problems (jointly). Specifically, this study aims to evaluate and quantify the impact of spatio-temporal demand forecast aggregation on the operational performance of a SAMS fleet. Given the sizeable market share of existing on-demand mobility services that do not allow shared rides (e.g. UberX, traditional Lyft, and taxi services), this study analyzes an on-demand SAMS with no shared rides, defined in Hyland and Mahmassani (2018).

The remainder of this paper is structured as follows. The next section provides an overview of the research problem and the study's hypothesis. The following section briefly discusses the relevant literature. The next section presents the methodological framework employed to address the research problem. The following two sections present the experimental design and computational results, respectively. The final section concludes the paper and presents limitations of the analysis along with future research directions.

Fig. 1 Schematic of process to convert data into better decisions

Research problem and hypothesis

The purpose of this study is to evaluate and *quantify* the impact of spatio-temporal demand forecast aggregation on the operational performance of SAMS fleets. There is an inherent trade-off in the selection of a spatio-temporal aggregation level. From an operational standpoint, holding forecast reliability constant, more disaggregate spatio-temporal forecasts—for smaller subregions—provide the fleet more valuable information. For example, knowing three users will request rides in a 100 m^2 area between 9:00 a.m. and 9:05 a.m. is more valuable than knowing three users will request rides in a 1000 m^2 area between 9:00 a.m. and 9:30 a.m. However, it is likely that short-term SAMS demand forecast errors will increase as forecasts become very disaggregated in space and time due to the law of large numbers (statistical variability increases as the number of items to forecast decreases) and the underlying demand generation process (Makridakis 1988). Given the inherent trade-off associated with choosing a spatio-temporal aggregation level, this study's working hypothesis is that:

SAMS fleet performance will initially increase as forecasts (and subregions) become more disaggregate; however, eventually, SAMS fleet performance will decrease, or at least stagnate, as forecasts and subregions become progressively more disaggregate.

This study also aims to determine the optimal spatio-temporal demand forecast aggregation level to most efficiently operate a SAMS fleet. However, the optimal spatio-temporal aggregation level depends on a multitude of factors including the demand forecasting method, the SAMS operational strategy, and even the characteristics (e.g. density of user requests) of the service area. Hence, this study introduces a flexible methodological framework that other researchers and mobility service providers can employ to determine the optimal spatio-temporal aggregation level for their own forecasting method, fleet operational strategy, and service area.

To obtain an upper bound on the operational performance of the SAMS fleet, the study runs experiments where the fleet has perfect demand forecasts, across all spatio-temporal aggregation levels. To obtain a lower bound, the study runs experiments where the fleet has no information about future demand forecasts.

Literature review

A SAMS fleet serving travelers who request rides and want service immediately represents a highly-dynamic and stochastic operational problem (Powell et al. 2012). Although, exact problem formulations vary based on the mobility provider's business model and service offerings (Hyland and Mahmassani 2017), if the fleet provides on-demand service, the underlying problem is a stochastic dynamic vehicle routing problem.

Travel behavior research indicates that short user wait times are likely to be a key factor in the success of an on-demand SAMS (Krueger et al. 2016). Motivated by this finding, researchers addressing fleet management and fleet operational problems are trying to determine the necessary fleet size to provide high-quality service (Boesch et al. 2016; Brownell and Kornhauser 2014; Fagnant et al. 2015; Spieser et al. 2014; Vazifeh et al. 2018) and develop operational strategies to efficiently assign AVs to open user requests (Alonso-Mora et al. 2017; Hyland and Mahmassani 2018; Maciejewski et al. 2016), respectively. Hyland and Mahmassani (2018) present and compare AV-user assignment strategies for an on-demand

SAMS without shared rides. Alonso-Mora et al. (2017) use an anytime optimization solution approach in which they enumerate feasible user-to-user sharing opportunities, as well as feasible matches between vehicles and groups of feasible users and then solve a bipartite matching problem.

The utilization of advanced (deterministic or stochastic) information can improve the operation of mobility services. In the context of goods transport, Yang et al. (2004) present the generic real-time truckload pickup and delivery problem and present computational results as a function of advanced information about demand requests. Tjokroamidjojo et al. (2006) and Jaillet and Wagner (2006) quantify the value of advanced deterministic information (i.e. known future requests) in dynamic freight routing problems. The on-demand SAMS modeled in this study does not allow users to request rides in advance; therefore, the SAMS fleet cannot obtain advanced (deterministic) information. However, predictive analytics methods and big data can help SAMS operators forecast demand and reposition vehicles based on demand predictions, thereby reducing user wait times.

In the context of carsharing, Weikl and Bogenberger (2015) introduce an algorithm to relocate vehicles, based on forecasts of future demand, in order to maximize profit. In goods transport, Ichoua et al. (2006) use demand forecasts to decide whether a vehicle should wait in its current position for a future demand before continuing its planned tour. Some SAMS studies introduce empty vehicle repositioning strategies (Dandl and Bogenberger 2018; Fagnant and Kockelman 2014; Hörl et al. 2018; Pavone et al. 2012; Sayarshad and Chow 2017; Spieser et al. 2016); however, these studies do not focus on the implications of demand forecast aggregation and/or quality on fleet performance.

The existing literature includes short-term demand forecasting studies related to carsharing (Müller and Bogenberger 2015), taxi (Ihler et al. 2006; Moreira-Matias et al. 2013), and public transportation (Zhong et al. 2016). In their survey and comparative analysis of taxi user arrival process models, Sayarshad and Chow (2016) categorize forecast methods into offline models and online models. Offline models rely entirely on historic data; whereas, online models utilize real-time data. Sayarshad and Chow (2016) evaluate the prediction quality of two offline and three online forecast models using New York taxicab data. Recent demand forecasting research incorporates new features from other data sources (e.g. social media) to further improve the quality of online models (Chaniotakis et al. 2016; Tong et al. 2017). In a more general analysis (i.e. broader than transportation), Zotteri et al. (2005) present an in-depth analysis of the impact of aggregation level on forecasting performance.

The current study makes several scientific contributions to the existing literature. First, as far as the authors are aware, this is the first study in the passenger or freight transportation literature to analyze the impacts of spatio-temporal demand forecast aggregation on the performance of vehicle fleets, where forecast quality varies across aggregation level. Second, the research methodology introduced in this study provides a flexible approach that can easily be adapted by other researchers and mobility service providers to determine the optimal spatio-temporal aggregation level with their own demand forecasting methods and fleet operational strategies for a particular service area.

Research methodology

To perform the computational analysis and test the hypothesis, this study employs an agent-based simulation tool. The simulation tool models the operations of an AV fleet, employs an algorithm to efficiently assign AVs to open user requests, and uses demand forecasts to proactively reposition AVs to serve future user requests. After providing an overview of the simulation framework, this section details the user requests, the demand forecast model, and the SAMS fleet operational strategy employed in this study. The agent-based simulation model in this study follows the general three-component framework for modeling SAMSs that includes a demand (i.e. traveler request) generator, an SAMS fleet operator/dispatcher, and some representation of the transportation network (Levin et al. 2017; Rigole 2014).

Agent-based simulation framework

Figure 2 displays a flowchart of the agent-based simulation tool. The simulation tool is time-driven and updates the position and status of AVs and users every time step. To initialize the simulation, the current time, τ, is set to zero, the AVs are positioned throughout the service region, and the statuses of all AVs are set to idle.

The simulation first updates the current time ($\tau \leftarrow \tau + \Delta\tau$) by the simulation time step ($\Delta\tau$) and then checks if τ is less than the length of the simulation period T. If $\tau \geq T$, the simulation ends, otherwise the simulation moves en-route pickup AVs, en-route drop-off AVs, and en-route repositioning AVs one step ($\Delta\tau \times v$, where v is vehicle speed) closer to their assigned destination.

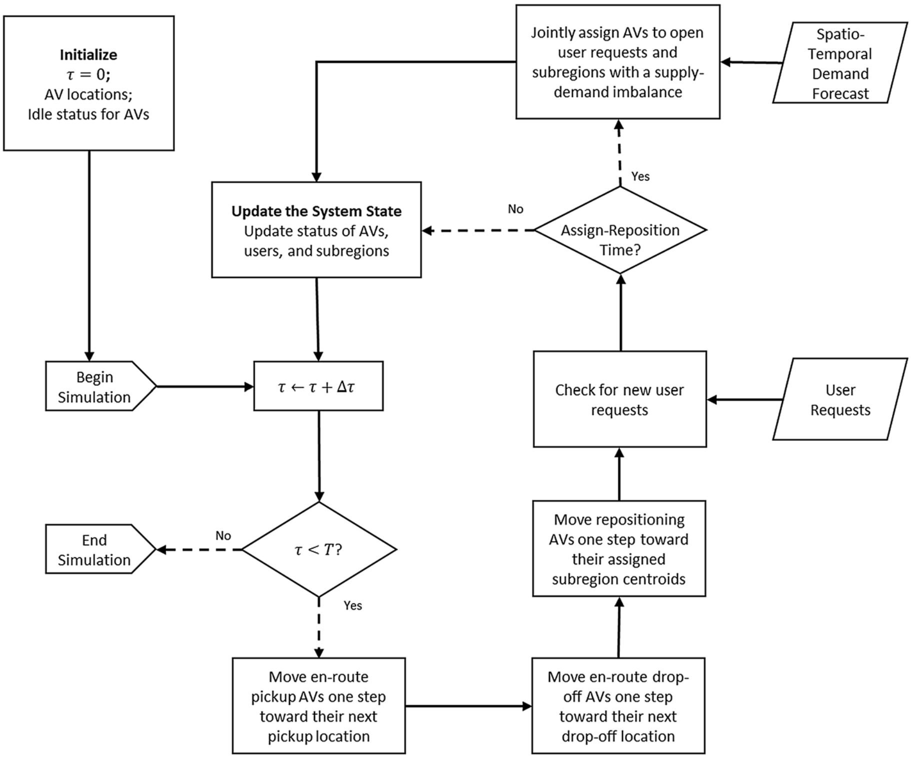


Fig. 2 Simulation framework

After the simulation moves the vehicles, it checks for new user requests with a request time $r_i = \tau$, where r_i is the request time of user i, and C is the set of user requests ($i \in C$). Then the simulation checks to see if it is time to assign AVs to open user requests and reposition AVs to different subregions. Every I^D, the inter-decision time interval length, the fleet simultaneously assigns and repositions AVs. Figure 2 shows that the joint assignment-repositioning strategy requires spatio-temporal demand forecasts, which are a key input in the operational strategy, as they determine how many, when, and where AVs should reposition.

The fleet only assigns and repositions AVs every I^D for strategic reasons and due to computational constraints. Strategically, it is often beneficial to allow user requests to queue before assigning AVs to them, especially, if AV-user assignments are final (i.e. if AV diversions and user reassignments are not allowed). The constraint comes from the fact that it can take more than a few seconds to solve a decision problem that involves assigning and repositioning large numbers of AVs.

After moving the AVs, checking for new user requests, and assigning AVs to user requests and subregions, the simulation updates the system state via changing the statuses of AVs, users, and subregions, if necessary. For example, if an AV reaches its drop-off point, the simulation changes the status of the AV from *en-route drop-off* to *idle*.

The simulation ends when $\tau = T$, even if AVs are still active and users are still unserved. The simulation can output metrics for individual AVs, users, and subregions, such as wait time and (empty) vehicle miles. As this study focuses on the performance of the SAMS fleet across different spatio-temporal aggregation levels, the computational analysis section presents performance statistics at the system level, such as average user wait time and empty fleet miles.

User requests

The main input to the agent-based simulation model is the set of user requests. Each user request i includes an origin (o_i), destination (d_i), and request time (r_i). In the simulation, the fleet becomes aware of each user and her origin and destination, at her request time (i.e. when the user requests a ride on her smartphone). This study assumes that the SAMS serves every user request within the service region; moreover, it assumes users will wait indefinitely to be served.

Demand forecasts

As described previously, fleet repositioning algorithms require forecasts of future demands. This study employs two sets of demand forecasts to analyze the impact of spatio-temporal aggregation on SAMS fleet performance. The first set of forecasts come from a simple time-varying Poisson forecast model based on historical demand (Ihler et al. 2006; Moreira-Matias et al. 2013; Sayarshad and Chow 2016; Tong et al. 2017), whereas, the second set of forecasts are *perfect* demand forecasts.

Demand forecast model

Like Moreira-Matias et al. (2013), this study uses a simplified version of the time-varying Poisson model in Ihler et al. (2006), which exploits weekly periodicity in demand to make forecasts for future days. The model in this paper does not include seasonality terms because it only uses three months of data. Historical request data ($\{o_i, d_i, r_i\}_{\forall i \in C}$) are aggregated into spatio-temporal bins. The underlying assumption is that (for example) the requests on Sunday between 5:00 p.m. and 5:30 p.m. will be similar to the historic average of requests on previous Sundays between 5:00 p.m. and 5:30 p.m. Hence, the forecasted trip origin count for subregion k during period h on day-of-the-week d is based on the historical average of trip counts in subregion k during period h on day-of-the-week d. Although more advanced methods tailored to specific problem instances can likely produce better results, this study employs the 'historical average' model or time-varying Poisson model because of its wide-use in practice and in the literature due to its ease of implementation.

The Poisson distribution is defined in Eq. 1, where λ is the rate of new user requests entering the system, n is the number of new user requests, and $P(n;\lambda)$ is the probability of exactly n new user requests entering the system over a specified time period, given rate λ.

$$P(n;\lambda) = \frac{e^{-\lambda}\lambda^n}{n!} \tag{1}$$

However, the rate λ is not time-invariant and space-invariant in real-world shared-use mobility services; rather, it varies across space and time. Similar to the model in Moreira-Matias et al. (2013), this study assumes the time- and space-variant rate $\lambda_k(t)$ is a function of the day of the week $d(t)$, the period of the day $h(t)$, and the subregion k. This functional relationship is displayed in Eq. 2, where $\lambda_{k,0}$ is the average rate over the week in subregion k, $\delta_{k,d(t)}$ is the relative change for day-of-the-week $d(t)$ in subregion k, and $\eta_{k,d(t),h(t)}$ is the relative change for period $h(t)$ on day-of-the-week $d(t)$ in subregion k.

$$\lambda_k(t) = \lambda_{k,0} \times \delta_{k,d(t)} \times \eta_{k,d(t),h(t)} \tag{2}$$

This study varies the size of the period $h(t)$ and the size of each subregion k in order to determine the impact of temporal aggregation and spatial aggregation, respectively, on demand forecast quality and SAMS fleet performance. The parameters in Eq. 2 are calibrated using historical trip request data for various period $h(t)$ sizes and subregion k sizes. Multiplying $\lambda_k(t)$ by the size of the period $h(t)$ gives the expected number of new user requests in subregion k, on day $d(t)$, during period $h(t)$ that is used by the SAMS fleet operator.

Perfect demand forecasts

Obtaining perfect demand forecasts requires aggregating the actual request data into different spatio-temporal bins. If spatial aggregation is set at the census tract level and temporal aggregation is set at the one-hour level, then, with perfect forecasts, the SAMS fleet knows the exact number of users who will request service originating at each census tract every hour of the day. However, the SAMS fleet does not know the exact location within the census tract, nor does it know the exact time within the one-hour interval the requests will occur. Hence, more disaggregate subregions and time intervals provide the SAMS fleet more valuable information.

SAMS fleet operational strategy

This section describes a SAMS fleet operational strategy that jointly assigns AVs to open user requests and repositions AVs between subregions.

Let V denote the set AVs in the SAMS fleet and let $j \in V$ denote an AV in the fleet. Moreover, let V^I, V^P, V^D, and V^R be the set of idle, en-route pickup, en-route drop-off, and en-route repositioning AVs respectively; $V = \{V^I, V^P, V^D, V^R\}$.

Similarly, let C denote the set of open user requests (meaning, they have not been assigned to an AV yet) and let $i \in C$ denote an open user request. If τ is the current time and r_i is the request time of user i, then user i's elapsed wait time (w_i) is $w_i = \tau - r_i$.

Additionally, let R denote the set of subregions in the service area, and let $k \in R$ denote a subregion. The expected imbalance between AVs and open user requests in subregion $k \in R$ over the prediction horizon h^p is denoted I_k. The expected imbalance I_k is determined by taking the difference between expected future demand and planned future supply in subregion k between the current time τ and the end of the prediction horizon $\tau + h^p$. The expected future demand is the sum of:

- the number of open user requests currently in subregion k
- the expected number of future requests in subregion k over the prediction horizon h^P (this value comes from the demand forecasts).

The planned future supply is the sum of:

- the number of repositioning AVs and idle AVs currently in subregion k
- the number of en-route drop-off and en-route pickup AVs assigned to users who have destinations in subregion k (the AVs must drop off their users in subregion k within the prediction horizon h^p)

The current distance between AV j and open user request i is denoted d_{ij}. The distance between AV j and the demand-weighted centroid of subregion k is denoted d_{jk}.

To solve the stochastic dynamic problem of operating a SAMS fleet, this study employs a rolling-horizon solution approach, where every I^D the fleet solves an optimization problem. In this study, the fleet can only control the AVs that are currently idle V^I or repositioning V^R. In fact, from the perspective of the fleet, at the decision epoch (every I^D), there is no difference between AVs that are currently repositioning and AVs that are currently idle, both sets of AVs can be assigned to travelers or assigned to reposition. Hence, let $V' = \{V^I, V^R\}$ denote the subset of AVs the fleet controller can (i) assign to open user requests, (ii) reposition to subregions, or (iii) choose to be idle and remain in its current position.

To model these decisions mathematically, let x_{ij} equal one if AV j is assigned to pick up user i, and zero otherwise. Moreover, let r_{jk} equal one if AV j is assigned to reposition to subregion k, and zero otherwise. Equation 3 displays the objective function driving the fleet operator's decisions. Equation 4–7 constrain the decision set.

$$\begin{aligned}\min_{x_{ij}, r_{jk}} c^{ED} \sum_{i \in C} \sum_{j \in V'} x_{ij} d_{ij} - r^{asgn} \sum_{i \in C} \sum_{j \in V'} x_{ij} + c^I \max\left(0, \sum_{k \in R} \left(I_k - \sum_{j \in V'} r_{jk}\right) - I^{min}\right) \\ + c^{ED} \sum_{j \in V'} \sum_{k \in R} r_{jk} d_{jk} - c^{VOT} \sum_{i \in C} w_i \sum_{j \in V'} x_{ij}\end{aligned} \tag{3}$$

$$\sum_{i \in C} x_{ij} + \sum_{k \in R} r_{jk} \leq 1 \quad \forall j \in V' \tag{4}$$

$$\sum_{j \in V'} x_{ij} \leq 1 \quad \forall i \in C \tag{5}$$

$$x_{ij} \in \{0, 1\} \quad \forall i \in C, \forall j \in V' \tag{6}$$

$$r_{jk} \in \{0, 1\} \quad \forall j \in V', k \in R \tag{7}$$

The objective function contains five separate terms that are associated with a penalty or a reward. The first term is a penalty term that denotes the cumulative distance between each newly assigned AV j and the user i it will pick up. The second term rewards the fleet for assigning an AV j to an open user request i. The third term penalizes the fleet for allowing an imbalance, greater than the minimum imbalance parameter I^{min}, in subregion k. The fourth term is a cost term that denotes the cumulative distance between each AV j and the demand-weighted centroid of subregion k it is assigned. The fifth term further rewards the fleet for assigning AVs to user requests with a long elapsed wait time.

The parameters set $\left(c^{ED}, r^{asgn}, c^{I}, c^{VOT}\right)$ convert units of empty vehicle distance, passengers assigned, expected subregion imbalances, and elapsed wait time into monetary units. The objective function implicitly makes trade-offs between assigning AVs to open requests now, reducing subregion imbalances now, and waiting until later (when other AVs become available) to assign AVs to open requests or balance subregions.

The constraint in Eq. 4 ensures that each AV j is assigned to at most one open user request or subregion k. The constraint in Eq. 5 ensures that no more than one AV is assigned to a single open user request. The constraints in Eqs. 6–7 ensure the two sets of decision variables take on binary values.

The third term in the objective function with the max() term is nonlinear. Fortunately, it is easy to convert this term and the mathematical program in Eqs. 3–7 into a linear integer programming problem. The term Z_k in Eq. 8 replaces the max() term in Eq. 3. The constraints in Eqs. 9 and 10 ensure that Z_k takes a value greater than or equal to the original value in the max() term of Eq. 3, and zero, respectively. The constraints in Eqs. 4–7 remain.

$$\min_{x_{ij}, r_{jk}} c^{ED} \sum_{i \in C} \sum_{j \in V'} x_{ij} d_{ij} - r^{asgn} \sum_{i \in C} \sum_{j \in V'} x_{ij} + c^{I} \sum_{k \in R} z_k + c^{ED} \sum_{j \in V'} \sum_{k \in R} r_{jk} d_{jk} - c^{VOT} \sum_{i \in C} w_i \sum_{j \in V'} x_{ij} \tag{8}$$

$$I_k - \sum_{j \in V} r_{jk} - I^{min} \leq z_k \quad \forall k \in R \tag{9}$$

$$z_k \geq 0 \quad \forall k \in R \tag{10}$$

$$\text{Constraints Eqn.}4 - 7$$

Fortunately, the constraint matrix (Eqs. 4–7 and 9–10) is totally unimodular; therefore, the linear relaxation of the integer program in Eqs. 8–10 and 4–7 always produces integer solutions. Hence, even for large instances of this problem, solutions can be obtained in a reasonable amount of time. This is quite beneficial as the fleet needs to repeatedly resolve the problem every I^{D}.

Fleet strategy with no demand forecasts

As mentioned previously, this study aims to create a lower-bound on fleet performance via testing scenarios that only allow myopic operational strategies that do not consider demand forecasts, to replicate the case in which short term demand forecasts are not available. Using the variable definitions described above, Eqs. 11–14 define the myopic user assignment strategy with no AV repositioning. This math program parallels the formulation in Eqs. 3–7, except it does not include the repositioning terms and constraints.

$$\min_{x_{ij}} c^{ED} \sum_{i \in C} \sum_{j \in V^I} x_{ij} d_{ij} - r^{asgn} \sum_{i \in C} \sum_{j \in V^I} x_{ij} - c^{VOT} \sum_{i \in C} w_i \sum_{j \in V^I} x_{ij} \tag{11}$$

$$\sum_{i \in C} x_{ij} \leq 1 \quad \forall j \in V^I \tag{12}$$

$$\sum_{j \in V^I} x_{ij} \leq 1 \quad \forall i \in C \tag{13}$$

$$x_{ij} \in \{0, 1\} \quad \forall i \in C, \forall j \in V^I \tag{14}$$

Experimental design

NYC taxi data

This study utilizes taxi data from New York City provided by the NYC Taxi and Limousine Commission (2017). The yellow taxi data was filtered for trips starting *and* ending in Manhattan since this simplifies the process of aggregating data into subregions of different sizes. The simulation treats the recorded taxi trip start times in the NYC taxi data as the users' request times.

Averaging over all days in April 2016, the number of trips per hour varies between 2500 trips per hour between 4:00 a.m. and 5:00 a.m. and more than 21,300 trips per hour between 06:00 p.m. and 08:00 p.m. For trips per day, the mean is 314,796 trips and the standard deviation is 69,122 trips. The mean taxi trip length is 2.8 km (1.7 mi) with a standard deviation of 2.0 km (1.2 mi).

We transform the origins and destinations of all taxi records into a metric system and create a minimum bounding rectangle. To create the largest forecast subregions, we cut the short edge of the minimum bounding rectangle in two and the long edge in eight pieces to create approximately square areas. To generate more disaggregate subregions, the edges of each subregion are cut in half. Figure 3 displays the resulting forecast subregions for four spatial aggregation levels. This method produces large differences in the number of trips per zone (i.e. a high coefficient of variation for daily trips per subregion), but provides an efficient means to test different spatial aggregation levels.

Since the simulation framework only allows movements along the x-axis and y-axis, the coordinate system for Manhattan, along with user's origins and destinations, were rotated to align with the gridded street network.

Table 1 Parameter Settings that do not vary across scenarios

Parameter	Math notation	Value	Units
Fleet size		5000	Vehicles
Vehicle speed	v	5	Meters/sec.
Drop-off time		15	Sec.
Pickup time		45	Sec.
Simulation length	T	21	Hours
Simulation time step	$\Delta\tau$	1	Sec.
Inter-decision interval	I^D	30	Sec.
Value of time	c^{VOT}	21.6	\$/hr.
Empty distance cost rate	C^{EDCR}	0.3	\$/km
Assignment reward	r^{asgn}	2.1	\$/user
Imbalance penalty	c^I	1.5	\$/user
Minimum imbalance	I^{min}	1	User
Prediction horizon	h^p	30	Minutes

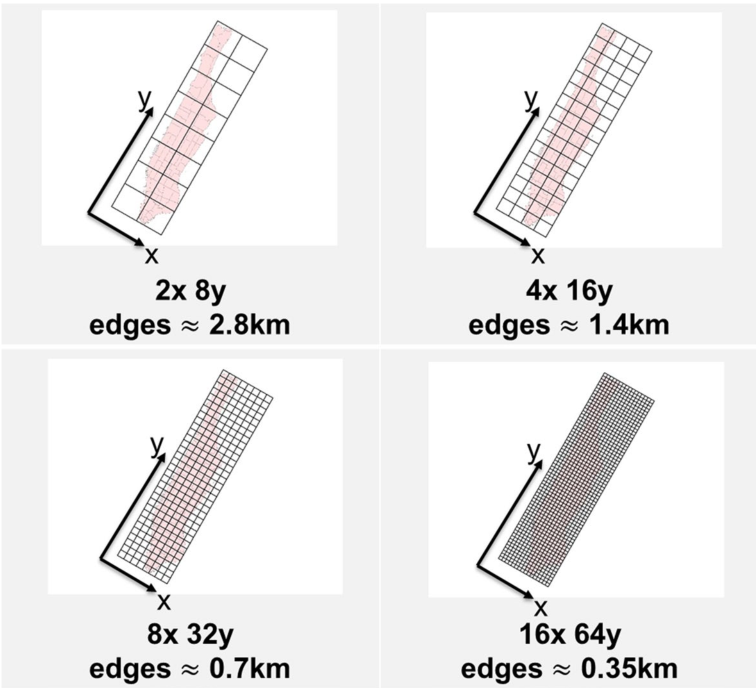
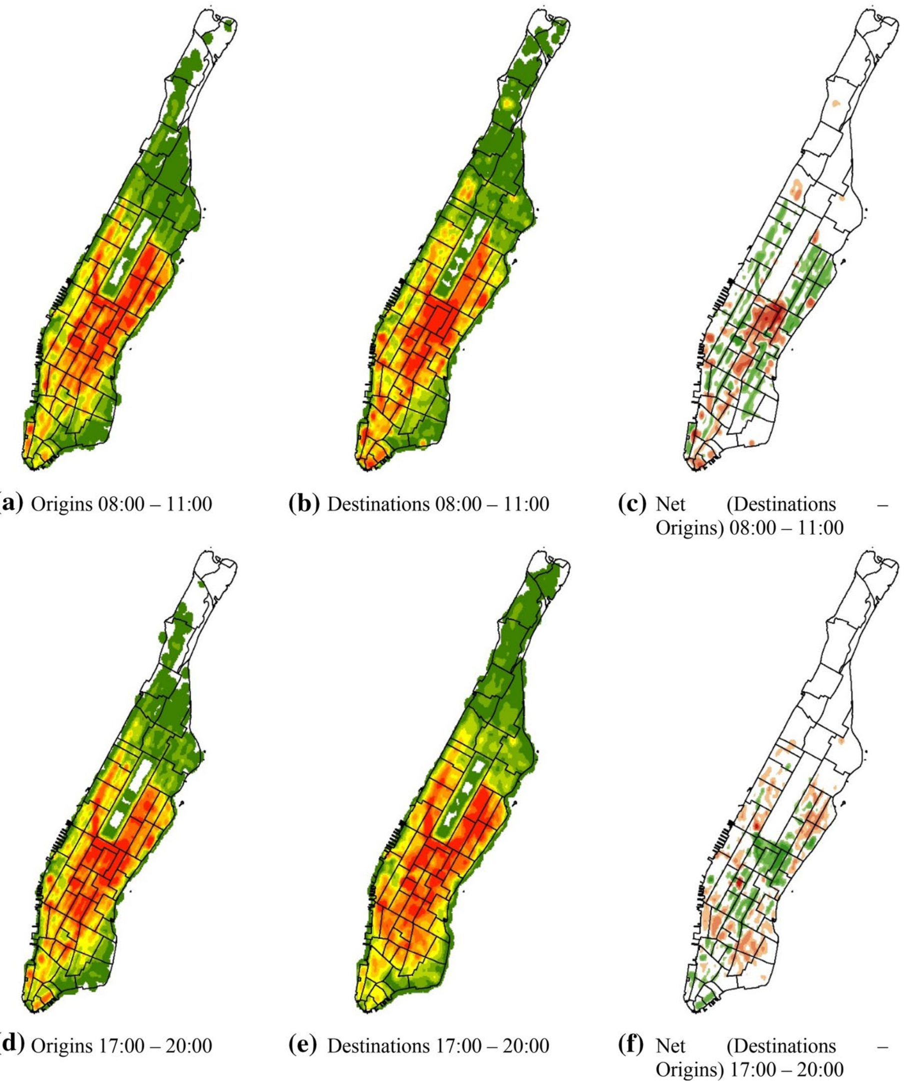


Fig. 3 Subregion layout for different spatial aggregation levels (official NYC taxi subregions for Manhattan are drawn in the background)

Simulating a full day and using a realistic spatio-temporal demand distribution adds practical value to the results presented in the next section. Moreover, given the natural spatio-temporal fluctuations in demand throughout a typical day, the SAMS fleet relies on demand forecasts to reposition AVs in advance of future demand surges. The demand forecast model was calibrated based on three months of historical data.

A preliminary analysis of the NYC taxi trip request data supports the choice of a historical average forecast model that segments the data by day-of-the-week. In the case of hourly demand forecasts for the entire borough of Manhattan, not segmenting by day-of-the-week results in a coefficient of variation (CV) for taxi trip count of 20%. The CV is between 3% (Wednesdays) and 7% (Saturdays) when the data are segmented by day-of-the-week.



(a) Origins 08:00 – 11:00 **(b)** Destinations 08:00 – 11:00 **(c)** Net (Destinations – Origins) 08:00 – 11:00

(d) Origins 17:00 – 20:00 **(e)** Destinations 17:00 – 20:00 **(f)** Net (Destinations – Origins) 17:00 – 20:00

Fig. 4 Taxi trip density during the morning (**a–c**) and evening (**d–f**) on Wednesday, 2016-04-06 for trip origins (**a** and **d**), trip destinations (**b** and **e**), and net trips (**c** and **f**) where red (green) areas represent more (fewer) trip origins, trip destinations, and net trips, respectively. (Color figure online)

Figure 4 displays demand density across Manhattan during the morning (8:00 a.m.–11:00 a.m.) and evening (5:00 p.m.–8:00 p.m.) for taxi users. These plots were created using the ArcGIS 'kernel plot' function. The density plots on the left-side and in the middle display the density of trip origins, and destinations, respectively. The density plots on the right-side display the net difference between trip destinations and trip origins. In the density plots on the right-side, areas in red (green) denote areas where there are more (fewer) trips terminating than originating. As the fleet serves demand throughout the day, without repositioning AVs, red (green) areas are likely to have a surplus (deficit) of AVs. Hence, repositioning empty AVs, from surplus areas to deficit areas, should improve the operational performance of SAMS fleets.

Parameter settings

Table 1 displays the parameter settings in the computational analysis that do not vary across scenarios. Test simulations with 3500 to 6000 AVs indicate fleet size is a crucial parameter for this study. On the one hand, very small fleet sizes essentially preclude repositioning trips because all vehicles are continuously busy serving a growing queue of open user requests. On the other hand, very large fleet sizes allow the fleet to easily serve all user requests without any subregions ever experiencing a deficit of AVs. Test simulations indicated that 5000 AVs was a reasonable fleet size to both operate an on-demand SAMS with no shared rides, and to answer the research problem in this study.

The 5000 AVs travel at a fixed rate of 5 m/s (11 mph) because this is approximately the average taxi speed in Manhattan. The simulation assumes AVs take 15 s to drop off a user and 45 s to pick up a user.

Each simulation runs from 3:00 a.m. to 11:59 pm (i.e. $T=21$ h); however, user requests only enter the system between 3:00 a.m. and 10:30 p.m. The AVs finish picking up and dropping off users between 10:30 p.m. and 11:59 p.m. This procedure ensures the fleet can serve all requests in all scenarios. In all scenarios, all the AVs are initially located (at 3:00 a.m.) in one location. This forces the AVs to reposition before the morning peak period.

The simulation time step $\Delta\tau$ is one second. The inter-decision interval I^D is 30 s, which is long enough to solve the optimization problem instances in this study. Allowing user requests to queue over 30 s also allows the fleet controller to make efficient AV-user assignments. The impact of I^D on the operational performance is another interesting research area; however, it is beyond the scope of this study.

The value of wait time c^{VOT} and the empty distance cost rate c^{ED} used in the objective function are \$21.6/hr. and \$0.3/km (\$0.48/mi), respectively. These values coincide with estimates of value of time in the literature and the U.S. governmental mileage rate (Internal Revenue Service 2018). The reward for assigning an AV to open user request r^{asgn} is \$2.1/user. Given the empty distance cost rate c^{ED} of \$0.3/km, AVs will not be assigned to new user requests if the AVs are more than 7.0 km (minus $c^{VOT}\times w_i$) away from the new request. A smaller assignment reward value would increase user wait times but decrease empty fleet kilometers. The imbalance penalty c^I is \$1.5/user, indicating an empty AV will only be considered for repositioning to an imbalanced subregion, if the AV and subregion are less than 5.0 km away from each other. The parameter values chosen by an SAMS fleet operator will likely depend on how they want to position themselves. If the SAMS fleet is concerned with user wait times they can choose larger values for value of wait time c^{VOT}, and the assignment reward r^{asgn}. If they are more concerned with offering low prices through keeping their operational costs down they can decrease the assignment reward r^{asgn} and the imbalance penalty c^{I} to avoid empty miles.

In this study, the minimum imbalance parameter I^{min} is set to 1 vehicle. This parameter allows the fleet to control the aggressiveness of empty AV repositioning. Larger values of I^{min} should decrease empty repositioning miles while increasing average user wait times; however, this analysis is beyond the scope of this study. The prediction horizon h^{p} is 30 min.

Scenarios

Given the parameter values listed in the previous subsection, the computational analysis involves simulating the performance of a SAMS fleet under a variety of scenarios. The scenarios vary:

- Forecast type: Perfect and model forecasts
- Spatial partition for demand forecasts (side 1 length| side 2 length| area):
    - 2.83 km | 2.65 km | 7.49 km^2
    - 1.41 km | 1.32 km | 1.87 km^2
    - 0.71 km | 0.66 km | 0.47 km^2
    - 0.35 km | 0.33 km | 0.12 km^2
- Temporal aggregation for demand forecasts: 5-min, 30-min, and 60-min
- Request data: 30 days of taxi data from April 2016

In combination, this represents 2×4x3x30=720 simulations/scenarios. The analysis also includes 30 experiments for the no forecast/no AV repositioning case. In the scenarios with no repositioning, the fleet solves the math program in Eqs. 11–14 every I^{D}.

Results

Demand forecast results

This subsection presents statistical error measurements for demand forecasts across different spatio-temporal aggregation levels. The following two metrics are used to measure statistical error:

$$RMSRE = \frac{1}{N_T}\frac{1}{N_{nz}}\sum_{h=1}^{N_T}\sum_{k\in Z_{nz}}\sqrt{\left(\frac{X_k^h - Z_k^h}{Z_k^h}\right)^2},$$

$$sMAPE = \frac{1}{N_T}\frac{1}{N_{nz}}\sum_{h=1}^{N_T}\sum_{k\in Z_{nz}}\frac{\left|X_k^h - Z_k^h\right|}{X_k^h + Z_k^h + 1},$$

where N_T is the number of time intervals and Z_{nz} is the set of subregions containing demand ($N_{nz} = \left|Z_{nz}\right|$). X_k^h and Z_k^h are the forecasted number of requests in subregion k during time-interval h from the demand model and the perfect forecast, respectively.

Figure 5 displays the average statistical error values for all days in April 2016 for different aggregation levels. The figure shows that errors increase with both shorter time intervals and smaller subregions. This finding is consistent across the two error measures employed in this study—root mean squared relative error (RMSRE) and symmetric mean absolute percentage error (sMAPE). In general, the RMSE metric penalizes large individual errors between actual and observed demand ($X_k^h - Z_k^h$) more severely than the sMAPE

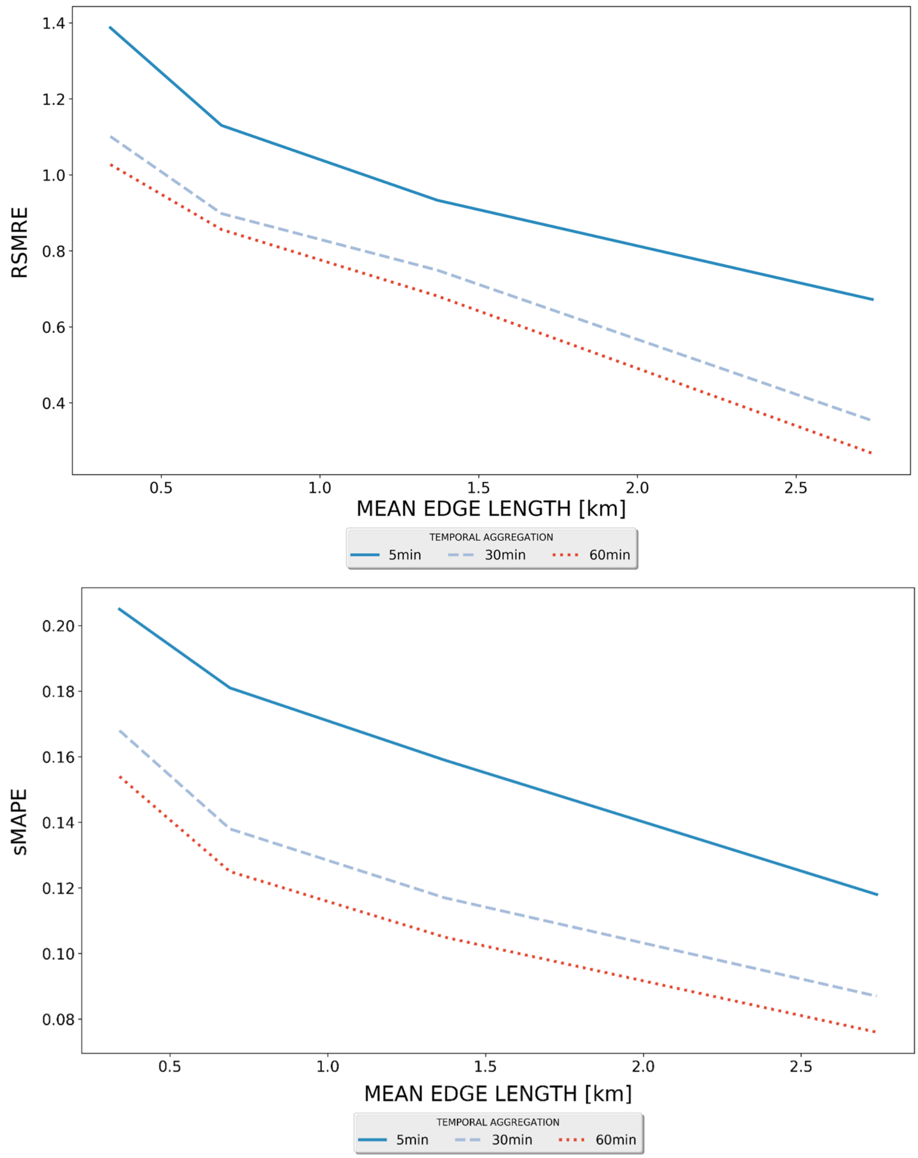


Fig. 5 Average demand forecast errors for different spatial (x-axis) and temporal (line type) aggregation levels according to RSMRE (top) and sMAPE (bottom) error measures

metric. However, the trends in Fig. 5 are similar for both error measures. The relationship between temporal aggregation as well as spatial aggregation and statistical forecast error appears to be non-linear. A rigorous analysis of these relationships requires more data points and is beyond the scope of this paper.

SAMS fleet performance results

This subsection presents the results of the computational analysis that was designed to evaluate and quantify the impact of spatio-temporal demand forecast aggregation on SAMS fleet performance. It includes two key performance metrics: average user wait times and the share of empty fleet miles.

Figure 6 displays average user wait time as a function of spatial aggregation and forecast type (the temporal aggregation level is 5 min in Fig. 6). Each point on the model and perfect forecast lines represents the mean of thirty separate experiments (i.e. the 30 days of April 2016) for a single spatial aggregation level. As the scenarios with no repositioning do not depend on forecast aggregation, the dotted line represents the mean of one set of 30 experiments.

There are several important findings displayed in Fig. 6. First, average user wait time increases significantly with spatial aggregation; i.e. using more disaggregate demand forecasts and smaller subregions significantly improve fleet performance in terms of average user wait time. Second, it is not until demand forecasts are spatially highly-disaggregate and subregions are small that a fleet using perfect demand forecasts begins to significantly outperform a fleet using model forecasts.

These first two findings provide some evidence to reject (and support) the study's hypothesis. Although Fig. 6 shows continued improvement at progressively more disaggregate levels, a comparison of the SAMS performance under perfect forecasts and model

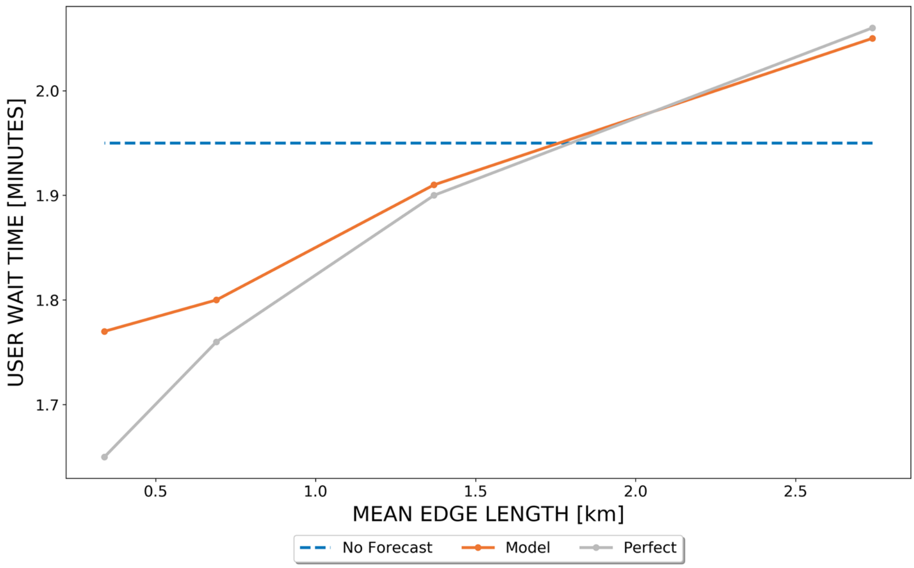


Fig. 6 Average user wait time as a function of subregion edge length (x-axis) and forecast type (line color)

forecasts suggests that it becomes progressively more difficult to improve fleet performance at highly-disaggregate levels using model forecasts with significant errors. This indicates that SAMS providers should benefit from improving demand forecast methods, especially at more spatially disaggregate levels.

Third, a fleet using no information about advanced requests and no empty AV repositioning outperforms a fleet using perfect demand forecasts in terms of average user wait times, when subregions are large in this study. This finding suggests that the SAMS operational strategy employed in this study is suboptimal in general, but especially when the service region is divided into large subregions. One potential method to improve the SAMS operational strategy includes making sure the AVs do not cluster in subregion centroids or at the edge of subregions. The operational strategy could force the available/empty AVs to spread out within their current subregions. This would decrease the distance between new user requests and the available AVs in their subregion. This improvement would likely have the biggest positive impact when subregions are large. Additionally, adjusting the parameters in Eq. 9 to emphasize reducing wait times and de-emphasize reducing empty fleet miles, may improve the average user wait times for large subregions.

Figure 7 displays the percentage of fleet miles that are empty across different spatial aggregation levels. The solid lines at the top of the figure represent *total* empty miles; whereas, the small vertical dash lines in the middle represent empty *pickup* miles and the horizontal dashed lines at the bottom represent empty *repositioning* miles. *Total* empty miles are the summation of empty *pickup* miles and empty *repositioning* miles.

The results in Fig. 7 are quite interesting, especially in the context of Fig. 6. Once again, fleet performance (measured in total fleet miles) improves with more disaggregate demand forecasts and smaller subregions. This finding suggests that there is not a trade-off in terms of operational costs and service quality when choosing a spatial aggregation level; rather, more disaggregate forecasts and smaller subregions perform better across both metrics.

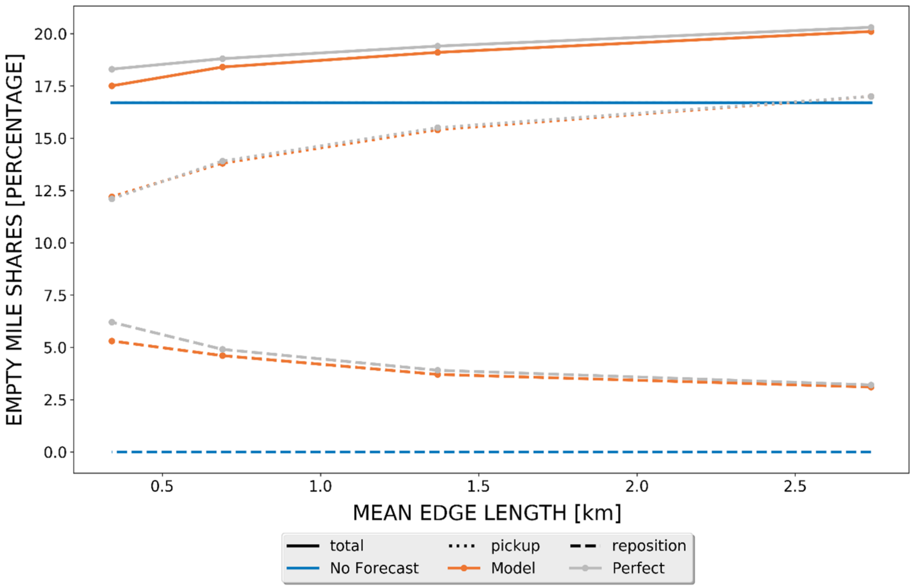


Fig. 7 Empty fleet miles as a function of subregion edge length (x-axis), type of empty miles (line type), and forecast type (line color)

Additionally, Fig. 7 indicates why/how more disaggregate forecasts and smaller subregions produce the shorter wait times in Fig. 6. Empty *pickup* miles significantly decrease for smaller subregions, meaning AVs are positioned closer to new user requests when subregions are smaller, effectively decreasing user wait times. This significant decrease in empty *pickup* miles more than offsets the increase in empty *repositioning* miles for smaller subregions.

Table 2 shows the computational results in tabular form for all three temporal aggregation levels. The table indicates that for the SAMS operational strategy in this study (i) temporal aggregation level had minimal impact on fleet performance, and (ii) the relationship between spatial aggregation and fleet performance holds across temporal aggregation levels. This first result likely does not hold in general. In fact, it suggests that the SAMS operational strategy employed in this study fails to effectively use higher-resolution temporal forecasts to improve operational performance. The authors are working to improve

Table 2 Complete SAMS fleet performance results as a function of forecast type, spatial aggregation, and temporal aggregation

Forecast type	Edge length (km)	Avg. user wait time (min)	Empty pickup miles share (%)	Empty reposition miles share (%)	Total empty miles share (%)
None	NA	1.95	16.7	0.0	16.7
5-min temporal aggregation					
Model	2.7	2.05	17.0	3.1	20.1
	1.35	1.91	15.4	3.7	19.1
	0.68	1.80	13.8	4.6	18.4
	0.34	1.77	12.2	5.3	17.5
Perfect	2.7	2.06	17.0	3.2	20.3
	1.35	1.90	15.5	3.9	19.4
	0.68	1.76	13.9	4.9	18.8
	0.34	1.65	12.1	6.2	18.3
30-min temporal aggregation					
Model	2.7	2.05	17.0	3.1	20.1
	1.35	1.91	15.4	3.7	19.1
	0.68	1.80	13.8	4.6	18.4
	0.34	1.78	12.2	5.3	17.5
Perfect	2.7	2.06	17.0	3.2	20.2
	1.35	1.91	15.5	3.8	19.3
	0.68	1.79	13.9	4.7	18.6
	0.34	1.74	12.2	5.6	17.7
60-min temporal aggregation					
Model	2.7	2.05	17.0	3.2	20.1
	1.35	1.91	15.4	3.8	19.1
	0.68	1.80	13.8	4.6	18.4
	0.34	1.77	12.2	5.4	17.5
Perfect	2.7	2.06	17.0	3.2	20.3
	1.35	1.91	15.5	3.8	19.4
	0.68	1.78	14.0	4.8	18.7
	0.34	1.69	12.2	5.8	18.0

the SAMS operational strategies in this paper to properly capture the temporal aspects of demand forecasts.

Conclusion

Summary and implications

This study evaluates and quantifies the impact of spatio-temporal demand forecast aggregation on the operational performance of an on-demand SAMS with no shared rides. This research problem combines two timely research areas, namely, forecasting demand for mobility services and developing strategies to dynamically operate SAMSs efficiently. The existing literature largely treats these problems independently despite their inherent interconnection. Hence, the research problem and methodological framework presented in this paper have significant practical value to SAMS providers who need to forecast demand in order to efficiently operate their AV fleets.

The computational analysis illustrates that more disaggregate forecasts significantly improve SAMS fleet performance in terms of empty fleet miles *and* user wait times. As forecasts become more disaggregate, the SAMS fleet more effectively repositions AVs into smaller subregions, thereby decreasing average user wait times and empty *pickup* miles. The decrease in empty *pickup* miles more than offsets the increase in empty *repositioning* miles. Additionally, the results indicate that while demand forecast quality has little impact on fleet performance when spatial aggregation is high, as demand forecasts become more disaggregate, forecast quality begins to significantly impact operational performance.

These findings suggest that (i) there are significant benefits associated with dividing service areas into smaller subregions to forecast demand and reposition AVs, and (ii) improvements in demand forecasting methods, particularly for disaggregate spatial scales, can produce significant value to on-demand SAMSs in terms of operational performance.

Limitations and future work

This study presented a variety of challenges in terms of conducting a truly scientific analysis to test the study's hypothesis. The study design clearly defines the demand forecasting model, the SAMS operational strategy, the NYC taxi data, and the modeling assumptions. Nevertheless, the SAMS operational strategy (i.e. the assignment and repositioning algorithm) is not an optimal strategy because it is highly unlikely that an optimal strategy exists for such a highly-dynamic, stochastic, and large problem. Hence, there is no way to guarantee results will hold across SAMS operational strategies. Moreover, there is no way to guarantee the results will hold across different demand forecasting methods, and in different service areas.

This limitation suggests the research problem presented in this paper along with the flexible methodological framework represent more significant scientific contributions than the results for one demand forecasting method, one set of taxi data, and one SAMS operational strategy. Future research should employ the methodological framework presented in this study, but use different SAMS operational strategies, demand forecasting methods, and different user request data to further test this study's hypothesis.

The 0.34-km edge length is the smallest spatial scale presented in this study due to computational constraints. Smaller edge lengths increase the computational time to solve the joint assignment-repositioning problem. The authors are working to improve computational performance and test more disaggregate spatial subregions. The authors are also working to improve the SAMS operational strategy and its exploitation of the demand forecast output.

Another future research area of interest is the impact of demand forecast errors on SAMS operational performance. The authors plan to test different demand forecast models, which will undoubtedly vary in terms of their demand forecast errors, to determine the relationship between demand forecast errors and SAMS operational performance. It is also possible to employ a single model, or perfect forecasts, and systematically create errors in the forecasts to answer this research question. In this study, forecast errors are also a function of the different demand data; i.e. the different days in the taxi data. A study design, which only varies forecast errors while keeping demand and aggregation level constant, could also highlight if any forecast error measure correlates better with fleet performance results.

Finally, the authors are working to more effectively handle temporal components of the short-term demand forecasts within the repositioning strategy.

Acknowledgements Partial funding for the first author is provided by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety through the project "City2Share". Partial funding for this work is provided through an Eisenhower Fellowship awarded by the US Department of Transportation to the second author. Additional funding is provided by the Northwestern University Transportation Center. The authors remain responsible for all findings and opinions presented in the paper.

Author contribution All authors contributed to all aspects of the study from conception and design, to data collection, to analysis and interpretation of results, and manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D.: On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proc. Natl. Acad. Sci. U. S. A. **114**, 462–467 (2017). https://doi.org/10.1073/pnas.1611675114

Boesch, P.M., Ciari, F., Axhausen, K.W.: Autonomous vehicle fleet sizes required to serve different levels of demand. Transp. Res. Rec. J. Transp. Res. Board **2542**, 111–119 (2016). https://doi.org/10.3141/2542-13

Brownell, C., Kornhauser, A.: A driverless alternative: fleet size and cost requirements for a statewide autonomous taxi network in New Jersey. Transp. Res. Rec. J. Transp. Res. Board **2416**, 73–81 (2014). https://doi.org/10.3141/2416-09

Chaniotakis, E., Antoniou, C., Pereira, F.: Mapping social media for transportation studies. IEEE Intell. Syst. **31**, 64–70 (2016). https://doi.org/10.1109/MIS.2016.98

Dandl, F., Bogenberger, K.: Comparing future autonomous electric taxis with an existing free-floating carsharing system. IEEE Trans. Intell. Transp. Syst. **7**, 1–11 (2018)

Fagnant, D.J., Kockelman, K.M.: The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. Transp. Res. Part C Emerg. Technol. **40**, 1–13 (2014). https://doi.org/10.1016/j.trc.2013.12.001

Fagnant, D.J., Kockelman, K.M., Bansal, P.: Operations of shared autonomous vehicle fleet for Austin, Texas. Market. Transp. Res. Rec. J. Transp. Res. Board **2536**, 98–106 (2015). https://doi.org/10.3141/2536-12

Hörl, S., Ruch, C., Becker, F., Frazzoli, E., Axhausen, K.W.: Fleet control algorithms for automated mobility: a simulation assessment for Zurich. In: TRB 2018 Annual Meeting (2017)
Hyland, M., Mahmassani, H.S.: Dynamic autonomous vehicle fleet operations: optimization-based strategies to assign AVs to immediate traveler demand requests. Transp. Res. Part C Emerg. Technol. **92**, 278–297 (2018). https://doi.org/10.1016/j.trc.2018.05.003
Hyland, M.F., Mahmassani, H.S.: Taxonomy of shared autonomous vehicle fleet management problems to inform future transportation mobility. Transp. Res. Rec. J. Transp. Res. Board **2653**, 26–34 (2017). https://doi.org/10.3141/2653-04
IBM: Descriptive, predictive, prescriptive: transforming asset and facilities management with analytics. IBM Corporation, Somers, NY (2017). https://www.ibm.com/downloads/cas/3V9AA9Y5
Ichoua, S., Gendreau, M., Potvin, J.-Y.: Exploiting knowledge about future demands for real-time vehicle dispatching. Transp. Sci. **40**, 211–225 (2006). https://doi.org/10.1287/trsc.1050.0114
Ihler, A., Hutchins, J., Smyth, P.: Adaptive event detection with time-varying poisson processes. Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. KDD'06 (2006). https://doi.org/10.1145/1150402.1150428
Internal Revenue Service: Standard mileage rates (2018). https://www.irs.gov/tax-professionals/standard-mileage-rates. Accessed 17 Feb 2019
Jaillet, P., Wagner, M.R.: Online routing problems: value of advanced information as improved competitive ratios. Transp. Sci. **40**, 200–210 (2006)
Krueger, R., Rashidi, T.H., Rose, J.M.: Preferences for shared autonomous vehicles. Transp. Res. Part C Emerg. Technol. **69**, 343–355 (2016). https://doi.org/10.1016/j.trc.2016.06.015
Levin, M.W., Kockelman, K.M., Boyles, S.D., Li, T.: A general framework for modeling shared autonomous vehicles with dynamic network-loading and dynamic ride-sharing application. Comput. Environ. Urban Syst. **64**, 373–383 (2017). https://doi.org/10.1016/j.compenvurbsys.2017.04.006
Maciejewski, M., Bischoff, J., Nagel, K.: An assignment-based approach to efficient real-time city-scale taxi dispatching. IEEE Intell. Syst. **31**, 68–77 (2016). https://doi.org/10.1109/MIS.2016.2
Makridakis, S.: Metaforecasting: ways of improving forecasting accuracy and usefulness. Int. J. Forecast. **4**, 467–491 (1988)
Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., Damas, L.: Predicting taxi-passenger demand using streaming data. IEEE Trans. Intell. Transp. Syst. **14**, 1393–1402 (2013). https://doi.org/10.1109/TITS.2013.2262376
Müller, J., Bogenberger, K.: Time series analysis of booking data of a free-floating carsharing system in Berlin. Transp. Res. Procedia **10**, 345–354 (2015). https://doi.org/10.1016/j.trpro.2015.09.084
NYC Taxi & Limousine Commission: TLC trip record data (2017). http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml. Accessed 17 Feb 2019
Pavone, M., Smith, S.L., Frazzoli, E., Rus, D.: Robotic load balancing for mobility-on-demand systems. Int. J. Rob. Res. **31**, 839–854 (2012). https://doi.org/10.1177/0278364912444766
Powell, W.B., Simao, H.P., Bouzaiene-Ayari, B.: Approximate dynamic programming in transportation and logistics: a unified framework. EURO J. Transp. Logist. **1**, 237–284 (2012). https://doi.org/10.1007/s13676-012-0015-8
Rigole, P.-J.: Study of a shared autonomous vehicles based mobility solution in Stockholm (2014). http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A746893&dswid=1021. Accessed 17 Feb 2019
Sayarshad, H.R., Chow, J.Y.J.: Survey and empirical evaluation of nonhomogeneous arrival process models with taxi data. J. Adv. Transp. **50**, 1275–1294 (2016). https://doi.org/10.1002/atr.1401
Sayarshad, H.R., Chow, J.Y.J.: Non-myopic relocation of idle mobility-on-demand vehicles as a dynamic location-allocation-queueing problem. Transp. Res. Part E Logist. Transp. Rev. **106**, 60–77 (2017). https://doi.org/10.1016/j.tre.2017.08.003
Spieser, K., Samaranayake, S., Gruel, W., Frazzoli, E.: Shared-vehicle mobility-on-demand systems: a fleet operator's guide to rebalancing empty vehicles. In: 95th Annual Meeting of the Transportation Research Board., Washington D.C. (2016)
Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Pavone, M.: Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: a case study in Singapore. In: Meyer, G., Beiker, S. (eds.) Road Vehicle Automation, pp. 229–245. Springer, Berlin (2014)
Tjokroamidjojo, D., Kutanoglu, E., Taylor, G.D.: Quantifying the value of advance load information in truckload trucking. Transp. Res. Part E Logist. Transp. Rev. **42**, 340–357 (2006). https://doi.org/10.1016/j.tre.2005.01.001
Tong, Y., Chen, Y., Zhou, Z., Chen, L., Wang, J., Yang, Q., Ye, J., Lv, W.: The simpler the better. Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. KDD'17 (2017). https://doi.org/10.1145/3097983.3098018

Vazifeh, M.M., Santi, P., Resta, G., Strogatz, S.H., Ratti, C.: Addressing the minimum fleet problem in on-demand urban mobility. Nature (2018). https://doi.org/10.1038/s41586-018-0095-1

Weikl, S., Bogenberger, K.: A practice-ready relocation model for free-floating carsharing systems with electric vehicles - Mesoscopic approach and field trial results. Transp. Res. Part C Emerg. Technol. **57**, 206–223 (2015). https://doi.org/10.1016/j.trc.2015.06.024

Yang, J., Jaillet, P., Mahmassani, H.: Real-time multivehicle truckload pickup and delivery problems. Transp. Sci. **38**, 135–148 (2004)

Zhong, C., Batty, M., Manley, E., Wang, J., Wang, Z., Chen, F., Schmitt, G.: Variability in regularity: mining temporal mobility patterns in London, Singapore and Beijing using smart-card data. PLoS ONE (2016). https://doi.org/10.1371/journal.pone.0149222

Zotteri, G., Kalchschmidt, M., Caniato, F.: The impact of aggregation level on forecasting performance. Int. J. Prod. Econ. **93–94**, 479–491 (2005). https://doi.org/10.1016/J.IJPE.2004.06.044

Florian Dandl is a doctoral researcher at the Institute for Intelligent Transportation Systems at Bundeswehr University, Munich. His research interests include shared mobility services and autonomous vehicle fleet planning and operations.

Michael Hyland is an Assistant Professor in Civil and Environmental Engineering at the University of California, Irvine, which he joined after completing his PhD in transportation Systems at Northwestern University. His research interests include emerging transportation systems such as bikesharing, ridesharing, and shared-use autonomous mobility services, as well as the integration of these emerging systems with existing transit networks.

Klaus Bogenberger holds the Chair of Intelligent Transportation Systems at Bundeswehr University, Munich. His research interests include transportation engineering and control systems Engineering.

Hani S. Mahmassani is the William A. Patterson Distinguished Chair in Transportation and Director of the Northwestern University Transportation Center. His research interests include network models, transportation science, travel behavior and dynamic decision models in transportation and logistics systems.