Submitted to *Transportation Science*
manuscript TS-2021-0530

“Online Supplement - ”Cooperative Learning for Smart Charging of Shared Autonomous Vehicle Fleets”

Appendix A: Agent-based Simulation

In this appendix, we explain the details of our multi-agent simulation for shared autonomous electric vehicles (SAEVs). Each module of the simulation is described in the following.

A.1. Trip Module

Figure A.1 shows the process of serving a trip. When a vehicle is assigned to a trip, it starts relocating to the trip’s origin. Meanwhile, the traveler might cancel her request due to long waiting time or urgent events. Thus, we consider the probability of trip cancellation, which is a function of the driving distance between the trip’s origin and the assigned vehicle. Once a traveler cancels her request, the assigned vehicle runs the parking module (explained later on). Otherwise, the vehicle picks up the traveler, drives to the destination, and drops it off (we assume no cancellation during serving time). After serving the trip, the vehicle updates its state and runs the operations module to take the best operational action.

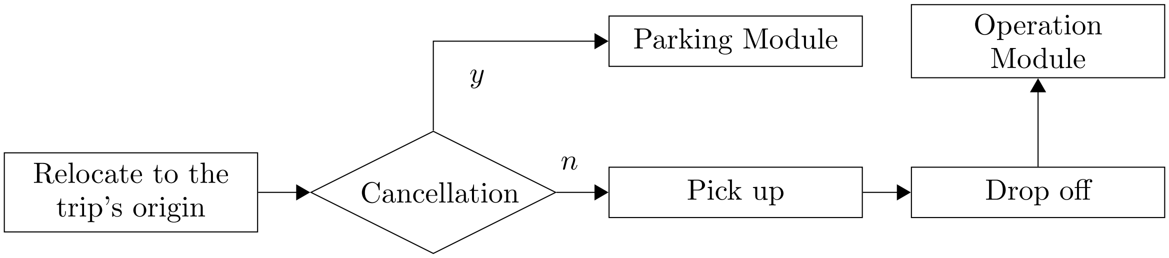

Figure A.1 **Trip Module of SAEVs: the Process of Serving a Trip after Matching to a Vehicle**

A.2. Operations Module

The operations module comprises three sub-modules: charging, repositioning, and parking; executed sequentially (see Figure A.2). First, the vehicle checks whether it should charge or not. If the decision is to charge, the vehicle must also choose a charging destination, relocate to the selected CS, wait in a queue (if there is any), and start charging. Instead of having a first-in-first-serve strategy, the CS prioritizes vehicles based on their SoC (the lowest SoC, the highest priority). We also assume that charging and waiting vehicles with enough SoC can be assigned to requests. We avoid interrupting charging/waiting processes if idle vehicles are available to serve all open requests, which is considered in our matching strategy (see A.3). If the vehicle does not interrupt charging, it finishes the task and follows the repositioning module. Also, when the vehicle decides not to charge, it checks the repositioning conditions.

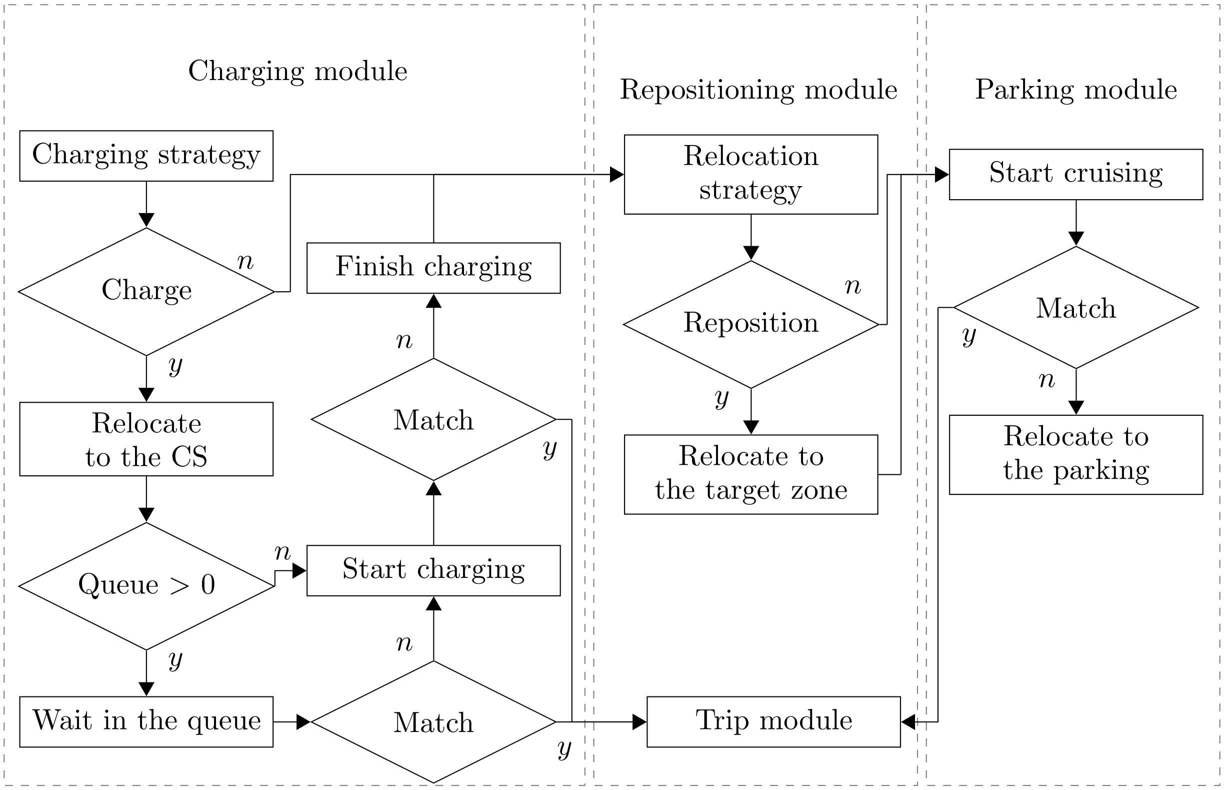

Figure A.2 **Operations Module of SAEVs: the Process of Taking and Executing Operational Decisions**

The repositioning module distributes available vehicles according to supply and demand. We use a simple decentralized heuristic approach to reposition vehicles. Each vehicle first compares the supply and demand for its current zone and neighbors. If the supply exceeds the forecasted demand, the vehicle decides to relocate to the target zone; otherwise, it keeps its position. The target zone is the closest among low-supply zones (for each, we determine a net supply by considering all available vehicles and open requests within a certain radius of its center). If the output of the repositioning module is no, the vehicle follows the parking module; otherwise, it first relocates to the target zone and then follows the parking module. The parking strategy in our framework is straightforward. Vehicles first cruise for a while, during which they can be assigned to a trip. The algorithm calls the trip module for paired vehicle and trip if a match occurs. Otherwise, it sends the vehicle to the closest free parking.

A.3. Matching Strategy

In this work, we consider a static system-centric matching strategy. A basic trick to tackle stochastic demand and supply is to use time windows to pool vehicles and requests and match them simultaneously. Long time windows could decrease pick-up distances but increase waiting times for assignments. Thus, we quantify the length of time windows to two minutes to avoid prolonged waiting times while reducing pick-up distances. After pooling all available vehicles and open requests, we generate our dispatching policy by solving a mixed-integer linear programming model assigning vehicles to trips. We define the sets, parameters, and decisions to do so. The set of available vehicles and the set of open requests are indicated by $\mathcal{J}_S$ and $\mathcal{I}_O$, respectively. The term $\beta^{Energy}D_{i,j}$ represents energy that vehicle j needs to serve request i, $C_{i,j}$ is the cost of assigning vehicle j to trip i (a function of distance $D_{i,j}$ and the vehicle's mode), $D_{i,j}$ is the distance between vehicle j and the origin of trip i, P_i is the price of serving request i, SoC_j is the state of charge of vehicle j, δ is safety energy after serving requests. Finally, the decision variable $(x_{i,j})$ is a binary variable set to one if vehicle j is assigned to request i.

$$Max_x \sum_{j\in\mathcal{J}_S}\sum_{i\in\mathcal{I}_O} x_{i,j}(P_i - C_{i,j}). \tag{1}$$

The objective is to match vehicles with trips while maximizing total profits. Each trip's price varies according to its length and duration in addition to a base fair. Assigning a vehicle to a trip entails different costs, depending on the vehicle mode (e.g., charging, parked) and the distance to the trip. A higher cost is designated to charging vehicles to allocate them with trips only when there is a lack of supply and to reduce charging interruptions. The assignment costs are set to be less than trip profits to only distinguish vehicles according to their distances with trips and avoid affecting the acceptance rate (i.e., not rejecting trips if there is enough supply).

$$\sum_{j\in\mathcal{J}_S} x_{i,j} \leq 1, \;\; \forall i \in \mathcal{I}_o, \tag{2}$$

$$\sum_{i\in\mathcal{I}_O} x_{i,j} \leq 1, \;\; \forall j \in \mathcal{J}_S, \tag{3}$$

$$x_{i,j}(\beta^{Energy}D_{i,j} - SoC_j\kappa_j^{Battery} - \delta) \leq 0, \;\; \forall j \in \mathcal{J}_S, i \in \mathcal{I}_O, \tag{4}$$

$$x_{i,j}D_{i,j} \leq \Delta, \;\; \forall j \in \mathcal{J}_S, i \in \mathcal{I}_O, \tag{5}$$

$$x_{i,j} \in \{0,1\}, \;\; \forall j \in \mathcal{J}_S, i \in \mathcal{I}_O. \tag{6}$$

As we match vehicles and open requests simultaneously, the model must ensure to assign each vehicle to at most one trip and each trip to at most one vehicle; guaranteed by Constraints (2) and (3), respectively. Constraint (4) ensures that a vehicle can be assigned to a trip only if it has enough energy to pick up the traveler, take her to the destination, and reach a CS. As a simplification, we consider the same safety threshold throughout the business area, assuming that CSs are evenly distributed. Moreover, to avoid prolonged waiting time after assignments, we add Constraint (5) to only allow matching when driving distances to trips' origin is not too large. Although we aim to serve all trips in the earliest time window, sometimes demands exceed supplies, and unassigned trips shift to the next time window.

A.4. Revenue and Costs

To calculate trip revenues, we use a base fare (2 USD) plus a variable fare, which is a linear function of trip distance and duration. Other factors, such as the time of day and traffic congestion, are excluded for simplicity. Each trip also has a travel cost that must be taken into account. We show the profit calculation for a trip in Equation (7), where ω is the revenue per distance (1.3 USD/km),

ϕ is the revenue per duration (0.35 USD/min), and ψ is travel cost per distance (0.53 USD/km). Note that the minimum charge for each trip is 5 USD (Taxi costs in Berlin 2022).

$$TripProfit := min(BaseFare + \omega * Distance + \phi * Duration, MinimumFare) - \psi * Distance. \quad (7)$$

Regarding other costs, we calculate the charging cost based on an hourly energy fee:

$$ChargingCost = \sum_{t \in \mathcal{T}_{charging}} EnergyCharged_t * EnergyPrice_t, \quad (8)$$

Where $\mathcal{T}_{charging}$ is the set of periods (minutes) during a charging session. It means that charging costs are independent of charging rates, which are aligned with current electricity tariffs for large-scale EV fleets (Lee et al. 2019). Moreover, the driving cost is a linear function of the distance ($DrivingCost = \psi * Distance$), and the waiting cost is also based on a fixed ratio per time ($WaitingCost = \Omega * WaitingTime$), where Ω is the parking cost per hour (3 USD/hr).

A.5. Linearizing Queue Management of the Reoptimization Benchmark Policy

To linearize the waiting constraint of the reoptimization benchmark, we use a big M trick. Therefore, we can replace the nonlinear constraint as follows.

$$t_{j,c}^{Queue} \geq max\{\frac{N_{j,c} + \sum_{j' \in \mathcal{J}_C} y_{j',c}(1 - z_{j,j'}) - \kappa_c^{Charging}}{\kappa_c^{Charging}} E/\kappa_c^{Power}, 0\} \quad , \forall j \in \mathcal{J}_C, c \in \mathcal{C}, \quad (9)$$

$$t_{j,c}^{Queue} \leq M^{Queue} y_{j,c} \quad , \forall j \in \mathcal{J}_C, c \in \mathcal{C}. \quad (10)$$

This pair of constraints guarantees that the waiting time of vehicle j in charging station c only could get a non-zero value if vehicle j is allocated to CS c (i.e., $y_{j,c} = 1$).

Appendix B: Configuration Data

In this section, we quantify the mobility environment parameters in addition to the charging agents' hyperparameters.

B.1. Hyperparameters

Here, we specify the hyperparameters of smart charging agents and additional details regarding the training and implementation of MARL and HMARL agents. Hyperparameters, chosen based on different experiments, are given in Table B.1. The HMARL agent has two separate networks (high-level and low-level), and the MARL agent has one network with the same structure and hyperparameters as the high-level network of the HMARL agent.

Parameter	**High-level**	**Low-level**
Optimizer (learning rate)	Adam(0.001)	Adam(0.001)
Loss function	MSE	MSE
Discount factor γ	0.99	0.99
Memory capacity	1,000,000	1,000,000
Steps prior to learning	1000	500
Training frequency	50	20
Batch size	32	32
Initial ϵ	0.6	0.6
Final ϵ	0.01	0.05
Target network update frequency	1000	1000
DQN activation functions	ReLU	ReLU
Number of hidden layers (nodes)	2 (256, 512)	2 (256, 512)

Table B.1 **Agents' Hyperparameters**

B.2. Energy Price

We use an actual EV-focused time-of-use (ToU) electricity tariff (Tariff 2 in Table B.2), that is used by Southern California Edison (SCE) and in recent related works (e.g., Lee et al. 2019). In addition, we test two other tariffs in this paper. The first is a flat tariff (Tariff 1), where electricity prices are the same for all hours of the day. The second one (Tariff 3) is similar to Tariff 2 with a larger price variation, where the price in super on-Peak hours is costly.

	Tariff 1	**Tariff 2**	**Tariff 3**
Super Off-Peak (9pm-6am)	0.23 USD/kWh	0.08 USD/kWh	0.08 USD/kWh
Off-Peak (6am-4pm)	0.23 USD/kWh	0.08 USD/kWh	0.23 USD/kWh
On-Peak (4pm to 6pm)	0.23 USD/kWh	0.23 USD/kWh	0.23 USD/kWh
Super on-Peak (6pm to 9pm)	0.23 USD/kWh	0.23 USD/kWh	0.50 USD/kWh

Table B.2 **Time-of-use Tariff for Large-scale EV Charging Customers**

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
SoC threshold % (morning)	0.45	0.60	0.65	0.62	0.58	0.55	0.52	0.50	0.40	0.40	0.40	0.40
SoC threshold % (afternoon)	0.38	0.35	0.32	0.25	0.25	0.20	0.20	0.25	0.27	0.35	0.35	0.40

Table B.3 **Hourly Charging Threshold for the Reoptimization Benchmark Policy**

B.3. SoC Thresholds of the Reoptimization Benchmark Model

To make charging decisions using REOPT, we determine vehicles needing to charge using an hourly SoC threshold. Whenever a vehicle's energy level falls below this threshold, it is marked as a charging demand. We consider an hourly threshold with the opposite pattern of mobility demands;

i.e., when the demand is low (e.g., during the night), the threshold is high, meaning that more vehicles will charge. The values are provided in Table B.3.

B.4. Details of the Upper Bound Scenario

For the upper bound scenario, we use the Audi 3 with a fuel efficiency of 7.29 liter/km Fuel economy guide (2020), and consider the average price of gasoline in Germany in 2021 ($1.8) Global petrol prices (2021) to compute energy costs (i.e., the driving cost of petrol vehicles equals $0.13/km).

B.5. Modification for Tabular Solutions

We compare our proposed model (MADC) with a tabular version (Q-learning) to ensure that deep learning does not compromise the learned policies. Therefore, we discretize the state space; hexagons and hours represent locations and time, respectively, SoC is divided into ten levels (10%, 20%, ..., 100%), and local supply and demand are categorized into three levels (low, medium, and high). Low/medium/high supply means less than 3/between 3 and 10/more than 10 available SAEVs within the coverage area of the vehicle (10 km radius). Low/medium/high demand means that there are less than 5/between 5 and 10/more than 10 open requests within the coverage area of the vehicle (10 km radius). We also identify the CSs with at least one free spot using binary variables and restrict the action space to only limited charging destinations.

Appendix C: Results and Sensitivity Analysis of Strategic Factors

This section provides detailed results of the scenario analyses using the proposed models.

C.1. The Impact of Fast Charging

We resolve SAEV-CC assuming there is ”no” fast charging technology. The results show that without fast charging, the fleet performance decreases by 14% and 24% using MADC and REOPT, respectively. We display vehicles' SoC in Figure C.1. The disparity between MADC and REOPT is wider when there is ”no” fast charger, which can be seen from the gap between the vehicles' SoC. Comparing the results with the fast charging scenario, vehicle batteries have lower energy since they need more time to charge, during which they might interrupt the charging process to meet urgent mobility demands.

Regarding the charging strategy, Figure C.2 shows that there is no longer a significant gap between the fast CSs and standard CSs occupancy (CSs locations cause the difference). Another change occurs in the hourly utilization of CSs, which is visualized in Figure C.3. As can be seen, although both strategies follow the same pattern, the average utilization is higher compared to the fast charging scenario due to longer charging sessions.

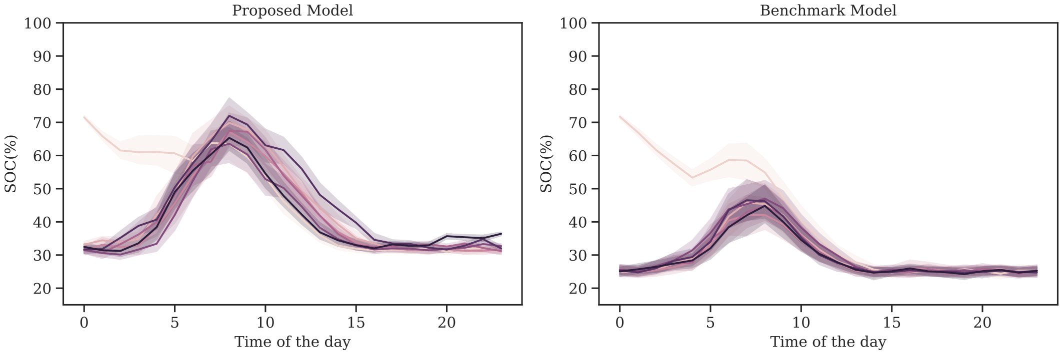

Figure C.1 **No Fast Charging Scenario: Average SoC of Vehicles for the Proposed and Benchmark Models**

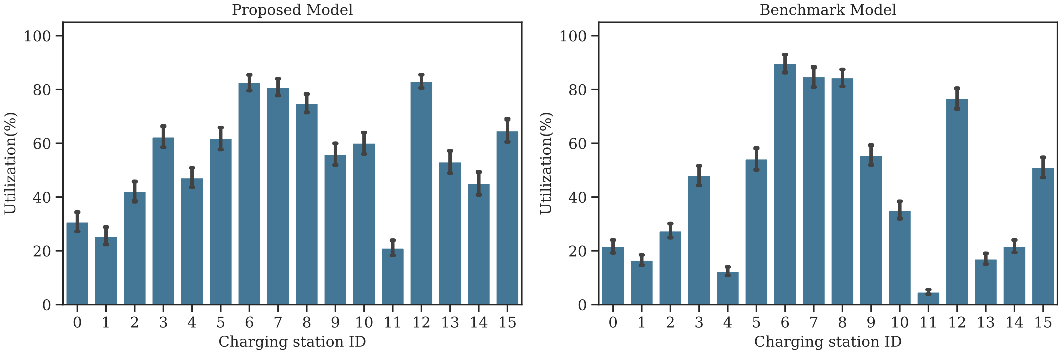

Figure C.2 **No Fast Charging Scenario: Individual CS Utilization for the Proposed and Benchmark Models**

C.2. The Impact of Charging Infrastructure Capacity

Another influential charging-related factor is the size of each CS. The more charging capacity, the higher fleet performance since vehicles have more resources to charge their batteries. Results show that with half of the charging capacity, service quality reduces by 5% and 12% for MADC and REOPT, respectively, while vehicle *SoC* patterns do not change significantly (see Figure C.4). The shape of SoC curves is the same. However, the peaks drop and occur later compared to

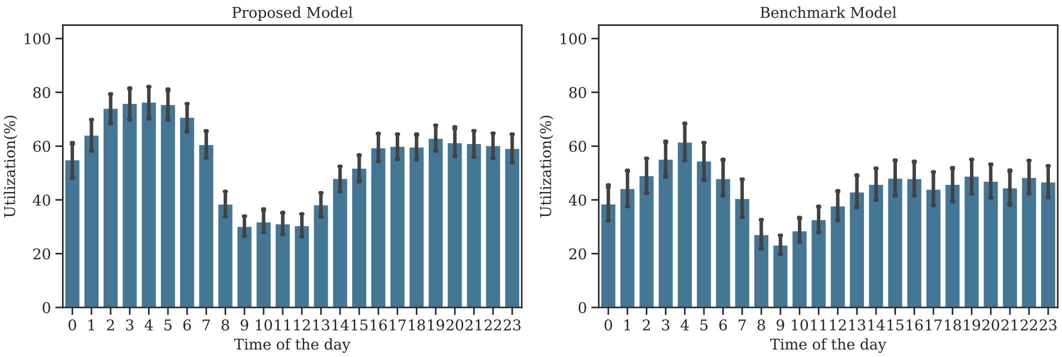

Figure C.3 **No Fast Charging Scenario: Hourly CSs Utilization for the Proposed and Benchmark Models**

the full charging capacity scenario. We also plot the utilization of CSs in Figure C.5, where still fast chargers are more attractive for both charging strategies. Figure C.6 depicts a very similar occupancy pattern for both strategies while the values are almost doubled.

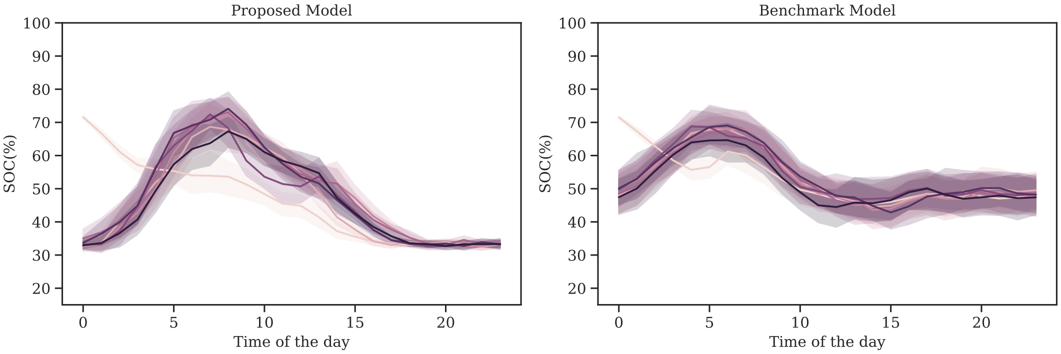

Figure C.4 **Half Charging Capacity Scenario: Average SoC of Vehicles for the Proposed and Benchmark Models**

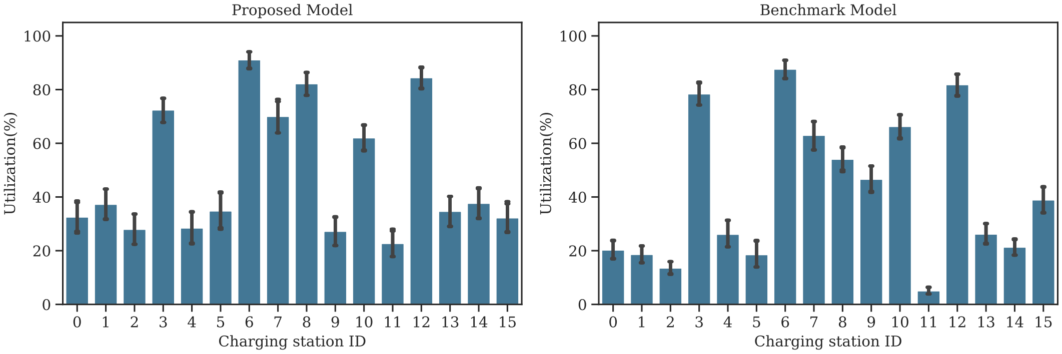

Figure C.5 **Half Charging Capacity Scenario: Individual CS Utilization for the Proposed and Benchmark Models**

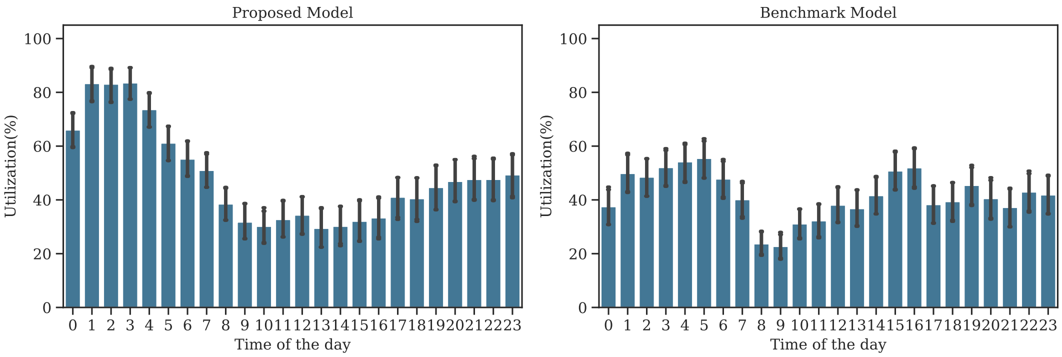

Figure C.6 **Half Charging Capacity Scenario: Hourly CSs Utilization for the Proposed and Benchmark Models**

C.3. The Impact of Fleet Size

One critical factor in shared autonomous fleet performance is the number of SAEVs. It gains more importance when the fleet is electric, and vehicles need more time to refuel their batteries. We test our proposed charging strategy (from now on, we will only analyze results using MADC and exclude REOPT) for four different fleet sizes (150, 200, 250, and 300 SAEVs). By increasing the number of SAEVs, the service quality increases, but not linearly (71%, 90%, 98%, and more than 99%, respectively, for the mentioned fleet sizes). Finding the optimal fleet size is beyond the scope of this paper; we merely aim to show its impacts on performance and charging policies. Therefore, if the goal is maximizing fleet performance, it makes no sense to facilitate more than 250 SAEVs to cover demands for this case. In more detail, vehicles' SoC is illustrated in Figure C.7 and what is clear is that by increasing the number of vehicles, the SoC distribution gets flatter (i.e., there is no such a gap between the vehicles' SoC in the early morning and evening when there are 300 SAEVs), which is caused by the low utilization of vehicles.

More importantly, different fleet sizes alter the learned charging strategies. Figure C.8 demonstrates that the smaller fleet sizes, the more fast charging demands. Indeed, when the number of vehicles is low, MADC takes the fast CSs as the best destination to ensure enough SoC for providing high service quality. On the other hand, the utilization distribution gets flat when there are many SAEVs (i.e., there is no need for urgent charging sessions). Regarding the hourly utilization, a rise in the number of vehicles creates a peak during midday hours (Figure C.9). This needs to be

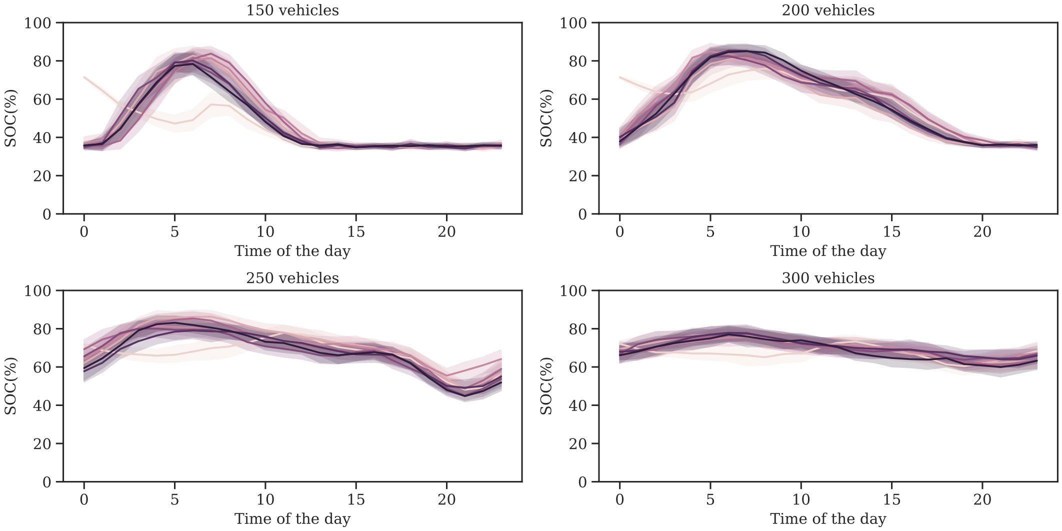

Figure C.7 Different Fleet Sizes Scenario: Average SoC of Vehicles for Four using the Proposed Model

interpreted cautiously as many factors might be influential, like electricity price, mobility demand, and charging capacity.

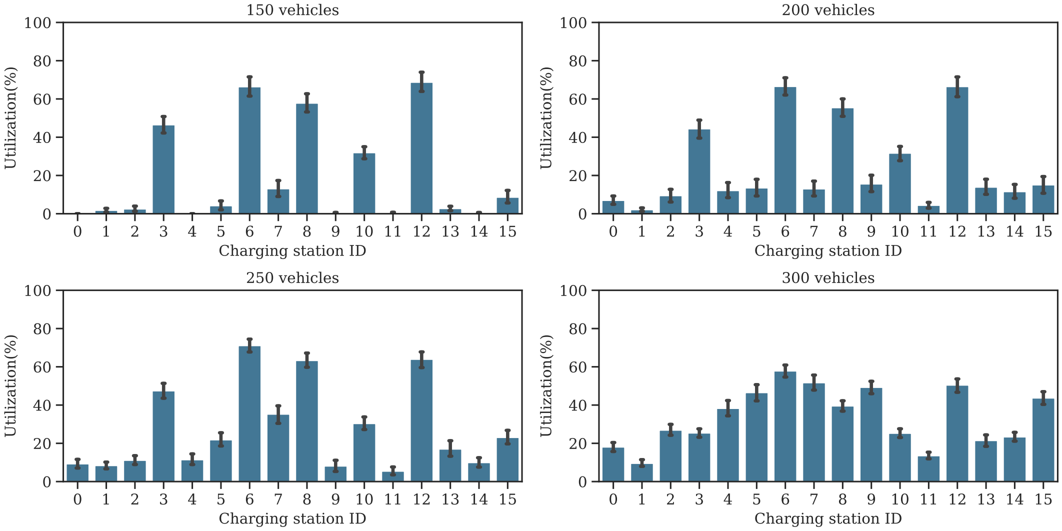

Figure C.8 Different Fleet Sizes Scenario: Individual CS Utilization using the Proposed Model

C.4. The Impact of Battery Capacity

Another impactful fleet configuration is the vehicles' battery capacity. We test MADC for three different battery sizes (50 kWh, 75 kWh, and 100 kWh), yielding service levels of 90%, 92.5%, and 94%, respectively. It shows that a bigger battery size could increase fleet performance and is worth

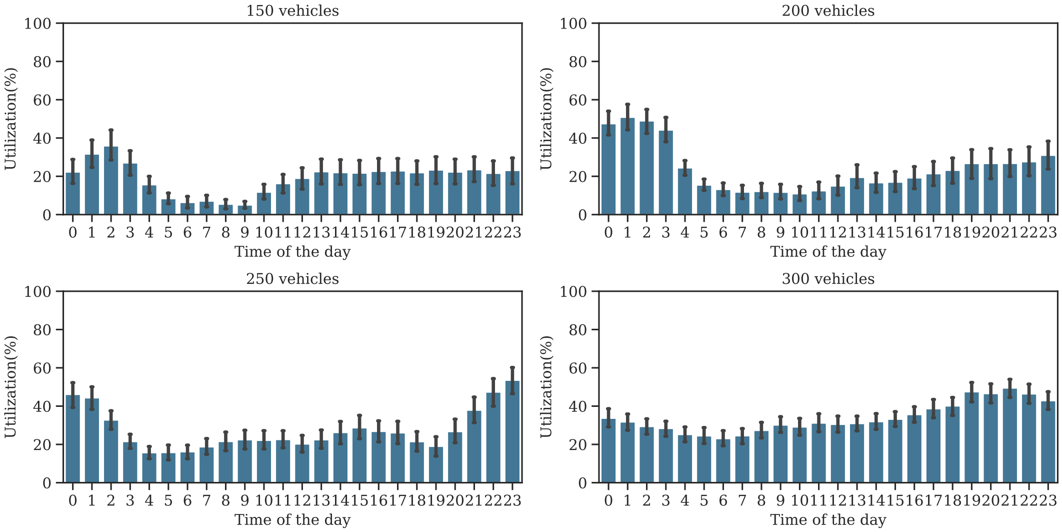

Figure C.9 **Different Fleet Sizes Scenario: Hourly Utilization of CSs using the Proposed Model**

considering (further analysis and the cross effects with fleet size could also be interesting, which is beyond the scope of this study). Figure C.10 shows that the SoC of vehicles with smaller batteries reaches a higher peak occurring earlier than the case of larger batteries (due to the shorter charging time). Figure C.11 displays that fast CSs are still more popular destinations for different battery sizes. Thus, larger batteries do not remove the need for fast chargers in our case. Hourly patterns are also similar. However, increasing the battery capacity decreases the reduction rate (from the peak during night hours until the valley during day hours) as it takes longer to charge the batteries (see Figure C.12).

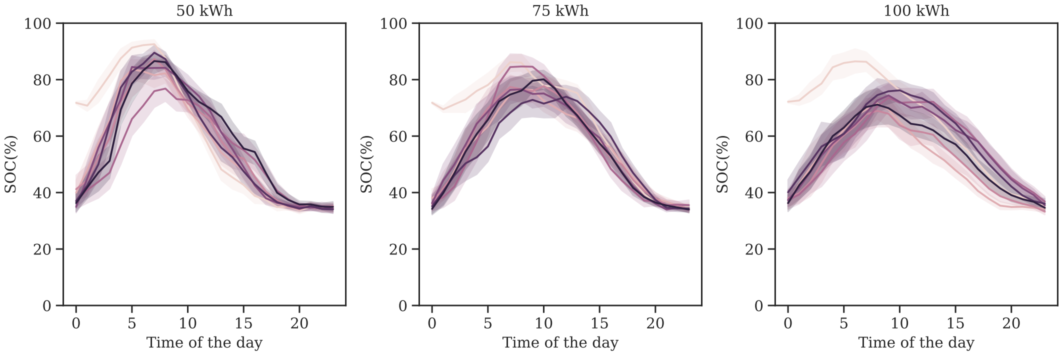

Figure C.10 **Different Battery Capacity Scenario: Average SoC of Vehicles using the Proposed Model**

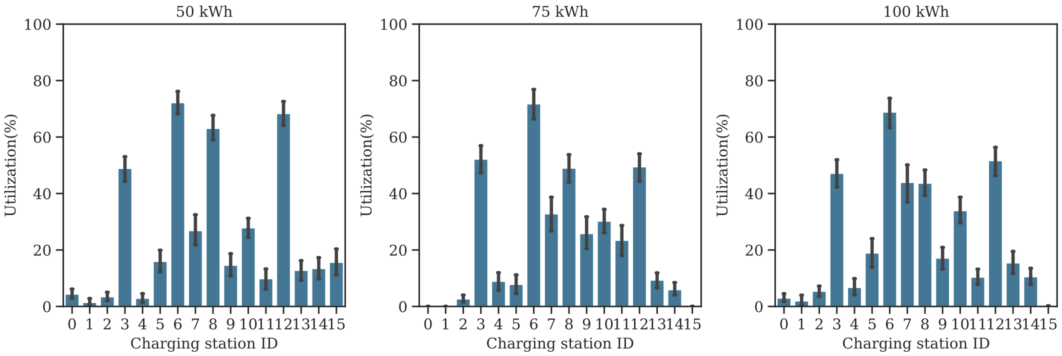

Figure C.11 Different Battery Capacity Scenario: Individual CS Utilization using the Proposed Model

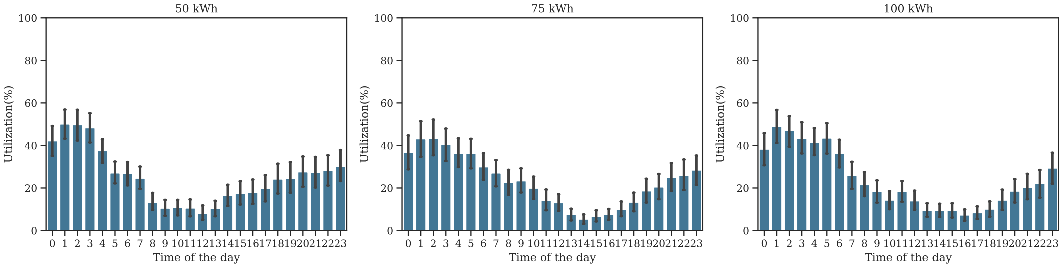

Figure C.12 Different Battery Capacity Scenario: Hourly Utilization of CSs using the Proposed Model

C.5. The Impact of Electricity Tariff

We check three different tariffs to understand how electricity prices affect charging behaviors (See Table B.2). We consider 250 vehicles for this comparison to avoid the bias of supply scarcity (i.e., with lower supply than demand, the charging behavior only follows the mobility demand and disregards the electricity price). Figure C.13 illustrates that vehicles increase their SoC during the night using the flat tariff (Tariff 1), and that it will drop relatively until the end of the peak hours. On the other hand, vehicles' SoC for two other tariffs (peak and off-peak prices) have periodic levels. Vehicles recharge mainly during the night (due to the low electricity price and mobility demand) and exactly before high price hours. This is more visible for Tariff 3, which has higher prices during super-peak hours (18-21).

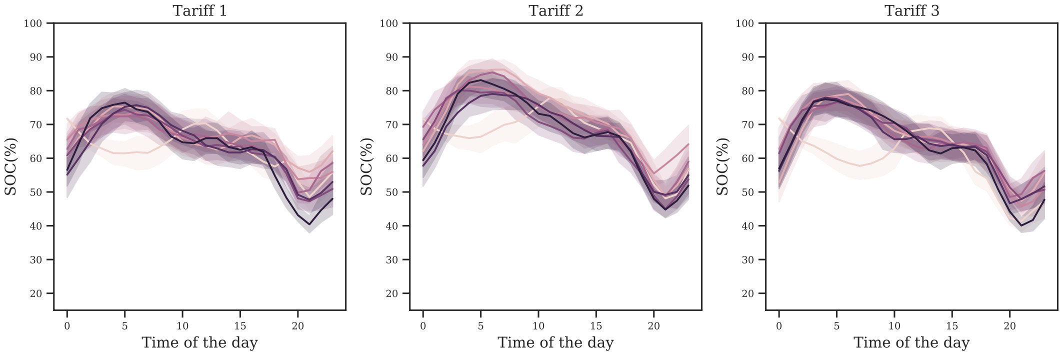

Figure C.13 Different Electricity Tariff Scenario: Average SoC of Vehicles using the Proposed Model

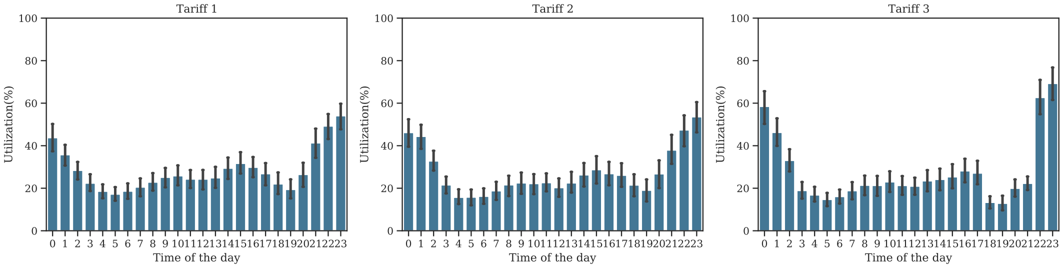

Figure C.14 Different Electricity Tariff Scenario: Hourly Utilization of CSs using the Proposed Model

References

Fuel economy guide (2020) `https://www.fueleconomy.gov/feg/pdfs/guides/FEG2020.pdf`.

Global petrol prices (2021) `https://www.globalpetrolprices.com/Germany/gasoline_prices`.

Lee ZJ, Li T, Low SH (2019) ACN-Data: Analysis and applications of an open EV charging dataset. *e-Energy 2019 - Proceedings of the 10th ACM International Conference on Future Energy Systems* 139–149.

Taxi costs in Berlin (2022) `http://taxihowmuch.com/location/berlin-de`.

Symbol	**Description**	**Unit**
Sets & Spaces		
$\mathcal{T}$	Set of operations time with index t	set
Γ	Set of hexagonal zones within the service region with index z	set
$\mathcal{H}$	Set of tasks while executing an operational decision with index h	set
$\mathcal{C}$	Set of charging stations with index c	set
$\mathcal{I}$	Set of all trip requests over the operations horizon with index i	set
$\mathcal{I}_O$	Set of open trip requests at the current time with index i	set
$\mathcal{J}$	Set of all vehicles (vehicle agents) with index j	set
$\mathcal{J}_S$	Set of available vehicles for serving trip requests with index j	set
$\mathcal{J}_C$	Set of vehicles with the need for charge with index j	set
$\mathcal{N}_j$	Set of neighboring vehicle agents to vehicle j	set
$\mathcal{S}$	The state space of the whole system with index s	space
$\mathcal{O}_j$	The observation space of of agent j with index o_j	space
$\mathcal{A}$	The joint action space of all agents with index a	space
$\mathcal{A}_j$	The action space of agent j with index a_j	space
Parameters		
α	Learning rate of Q-learning	ratio
γ	Reward discount rate	ratio
ϵ	Probability of taking a random action using an ϵ_{greedy} policy	ratio
θ	Parameter of action-value function approximations (neural network)	float
δ	Minimum safety energy threshold for reaching the closest CS	kWh
δ^{idle}	The maximum waiting time for vehicles before checking their status	minutes
Δ	The maximum coverage area of trip request or SAEVs	km
E	Average charging demand of vehicles	kWh
$\rho^{Charging}$	Penalty for an unassigned charging vehicle to CSs	USD
β^{Energy}	Energy consumption per driving distance	kWh/km
β^{Time}	Driving time per driving distance	minute/km
SoC_j	State of charge of vehicle j	%
N_c	Number of present vehicles in CS c	unit
$N_{j,c}$	Number of vehicles in CS c with lower SoC than vehicle j	unit
$\kappa_c^{Power}/\kappa_c^{Parking}$	Maximum charging power/number of parking spots of CS c	kW/unit
$\kappa_c^{Charging}$	Maximum number of charging docks in CS c	unit
$\kappa_j^{Battery}$	Maximum battery capacity of vehicle j	kWh
$Dj,i/Dj,c$	Driving distance between vehicle j and request i/ CS c	km
$w_s/w_c/w_d/w_w/w_p$	Weight for reward of serving/charging/driving/waiting/penalty	float
$u_i/d_i/p_i/b_i$	Information of trip request i (origin/destination/price/patience-time)	mixed
$Z_{j,j'}$	A binary indicator if vehicle j has lower SoC of vehicle j'	boolean
Variables		
$SoC/m/l$	Energy/mode/location of the corresponding vehicle	%/integer/integer
s	General state of the system regarding the corresponding vehicle	mixed
$s_t/s_v/s_l/s_c$	State of time/vehicle/local supply & demand/CSs	mixed
n_s/n_d	Number of local supply/demand around the corresponding vehicle	integer
q_c	Number of waiting vehicles in CS c	integer
c_c	Decision variable of charging the corresponding vehicle in CS c	boolean
$r_s/r_c/r_d/r_w/r_p$	Reward of serving/charging/driving/waiting/penalty of a vehicle	USD
r_{profit}/r_{missed}	Reward of profits and missed trips penalty of the fleet	USD
τ	Termination duration of an operational action	period
τ_{e_h}	Period between triggering and starting task h	period
τ_{o_h}	Period between starting and finishing task h	period
a	General operational action (charging and allocating to CS) of vehicles	integer
a_{meta}	Action of the meta controller (charging or not)	integer
a_{sub}	Action of the sub controller (which CS)	integer
$t_{j,c}^{Driving}$	Driving time between CS c and vehicle j	period
$t_{j,c}^{Charging}$	Charging time of vehicle j at CS c	period
$t_{j,c}^{Queue}$	Waiting time of vehicle j at CS c	period
$y_{j,c}$	Indicator whether vehicle j is assigned to CS c	boolean
$x_{i,j}$	Indicator whether vehicle j is assigned to trip request i	boolean

Table B.4 A Summary of Notations