NOTICE: Approved for public release. Distribution is unlimited.

ORNL/TM-2019/1233

Investigating the Impact of Connected Vehicle Market Share on the Performance of Reinforcement-Learning Based Traffic Signal Control

SMA Bin Al Islam
H M Abdul Aziz
Hong Wang
Stanley Young

30 June 2019

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE US DEPARTMENT OF ENERGY

ORNL/TM-2019/1233

Computational Sciences and Engineering Division

INVESTIGATING THE IMPACT OF CONNECTED VEHICLE MARKET SHARE ON THE PERFORMANCE OF REINFORCEMENT-LEARNING BASED TRAFFIC SIGNAL CONTROL

SMA Bin Al Islam

H M Abdul Aziz

Hong Wang

Stanley Young

Date Published: 30 June 2019

Prepared by

OAK RIDGE NATIONAL LABORATORY

Oak Ridge, TN 37831-6283

Managed by

UT-BATTELLE, LLC

for the

US DEPARTMENT OF ENERGY

Under contract DE-AC05-00OR22725

LIST OF ACRONYMS

CAV	Connected and Automated Vehicle
CV	Connected Vehicle
V2V	Vehicle-to-Vehicle
V2I	Vehicle-to-Infrastructure
I2I	Infrastructure-to-Infrastructure
SPaT	Signal Phase and Timing
RL	Reinforcement Learning
RSU	Road Side Unit
BSM	Basic Safety Message
DSRC	Dedicated Short Range Communications

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

We aim to understand and explore the performance of reinforcement learning based signal control algorithms in a mixed environment with less than 100% market share of connected and automated vehicles (CAVs). Within a simulation environment, we have considered partial connectivity—less than 100% market share of CAVs—in the network and investigated the impact on the performance of the signal control algorithm. Two test networks including a four-intersection arterial in Lankershim Boulevard, California and a portion of downtown Springfield, Illinois with 20 intersections. The first network is calibrated in the micro-simulator PTV Vissim with the US DOT provided NGSIM datasets. The results provide insights regarding the impact of the connectivity and sensing technologies on the practical implementation of traffic signal control algorithms that leverage the data sharing capability of a connected environment. For scenarios with 40% or more market share of CAVs, we observed improvement in the performance metrics—travel time, queue time, and energy consumption—with the increase in market share. Results from our experiments do not indicate any clear trend when the networks have low (less than 40%) market share of CAVs. The higher standard deviations as obtained from the statistical analyses of the performance metrics at low market share may indicate the instability of the RL controller arising from the partial (if not zero) observability of the traffic states. Further, we have conducted simplified scenario analyses to explore the impact of the market share of battery electric vehicles (BEVs) on energy consumption due to the regenerative braking feature. Our results and findings will be the foundation for the future reinforcement learning based control algorithm development that accounts for partial connectivity—less than 100% CAV market share, and the presence of BEVs in a network of connected and automated signalized intersections.

Acknowledgment: This work is funded by the Vehicle Technology Office (VTO) of the Department of Energy (DOE) under the Energy Efficient Mobility Systems (EEMS) program. The research was highly benefitted from the suggestions of the DOE program managers.

1. MOTIVATION AND BACKGROUND

The connected and automated Vehicle (CAV) technology is a promising platform that enables a new dimension of real-time data sharing among vehicles and infrastructure components in transportation networks. Such real-time information set that includes high-resolution position updates, speed, and acceleration profiles, has been recently being used to optimize and design traffic signal control algorithms. Many recent studies (Christian Priemer and Friedrich, 2009; He, Head and Ding, 2012; Lee, Park and Yun, 2013; Feng *et al.*, 2015; Islam and Hajbabaie, 2017; Zheng and Liu, 2017) have demonstrated the success of using high-resolution CAV data. Nevertheless, the effectiveness of traffic signal control using CAV technology highly depends on the level of connectivity and observability of the system as a function of the market share of CAVs along with the developed signal control strategy and the underlying traffic conditions (Argote-Cabañero, Christofa and Skabardonis, 2015). A few studies suggest that the current CAV-based signal control strategies require a range of 20% to 50% CAV market shares to outperform the existing signal control systems those do not leverage connectivity in a CAV environment (Christian Priemer and Friedrich, 2009; Lee, Park and Yun, 2013; Feng *et al.*, 2015). Also, the less than 100% CAV market share condition may significantly impact the performance of data-driven signal control strategy—for instance, a reinforcement learning (RL) based signal control. Though many studies have made great efforts in developing traffic signal control using RL, most of them rely on the assumption of full observability of the vehicles in the network—a signal controller can fully observe the traffic states with 100% connectivity. However, we do not anticipate a 100% connected vehicle environment soon, and the signal controllers may have only partial observability where traffic state measurements are often imperfect due to a low CAV market share. As such, RL-based signal control algorithm can perform in a sub-optimal manner due to a partial information sharing among vehicles and infrastructure components.

Recently we have developed RL-based control algorithm (Islam, H. M. A. Aziz, *et al.*, 2018) where the signal controllers learn to optimize over time through observing the transition of traffic states resulting from exploring and exploiting controller settings such as adjusting phase sequences and green durations. The developed RL algorithm exclusively utilizes the vehicle to infrastructure (V2I) communication capability. A connected signalized intersection equipped with a roadside unit (RSU) collects all the basic safety messages (BSM) from approaching equipped vehicles and the traffic states—such as queue lengths of neighboring intersections using infrastructure to infrastructure (I2I) communication. The control algorithm assumes a 100% market share of equipped vehicles that can exchange data with the controller.

As discussed, it is critical to investigate and understand the impact of having partial connectivity on the algorithm performance when the controller can access data from only a specific portion of the vehicles near the intersection. To understand the impact of partial connectivity, we have designed experiments that represent different market share compositions of CAVs and legacy vehicles in the network. The key goal is to conduct a simulation-based statistical assessment of the market share impact. Additionally, we have conducted a simplified sensitivity analysis that involves the market share of electric vehicles (EVs) to explore the impact of energy consumption when EVs are present in the road network.

2. FINDINGS FROM EXISTING LITERATURE

Recent studies in signal control mostly focus on designing algorithms by integrating high resolution CAV data. Acknowledging the vast literature on signal control algorithm, we focused only on the penetration rate that requires for successful implementation of the proposed algorithm without any adjustment for low penetration rate. (C. Priemer and Friedrich, 2009) applied dynamic programming with completer enumeration over the states to control signals in a CV environment. The proposed algorithm required 33% CAV penetration rate to outperform actuated-coordinated signal. Furthermore, the signal control heuristic proposed in (Smith *et al.*, 2011) improves the mobility performance in the network with more than 25%

CAV market penetration compared to coordinated-actuated signal control. (He, Head and Ding, 2012) integrated CAV data to estimate the size of approaching platoons to control a platoon-based signal by minimizing platoon splitting. The proposed method requires at least 40% CAV data to correctly estimate platoon size. (Goodall, Smith and Park, 2013) evaluated best signal control strategy by importing CAV data in a simulated network. The proposed algorithm did not work well in oversaturated traffic condition even with 100% CAV market share. However, the algorithm outperformed the actuated-coordinated signal with 50% CAV in undersaturated condition. (J, no date) required 70% CAV data to estimate cumulative travel time of different to control a travel-time responsive intersection. (Day and Bullock, 2016) required that 5% CAV data to generate vehicle arrival pattern in an intersection for 15 minutes. Similarly, (Islam and Hajbabaie, 2017) required 40% CAVs to outperform actuated-coordinated signal system in their proposed distributed-coordinated algorithm. (Mohebifard and Hajbabaie, 2018) achieved a significant reduction in system-level delay of network traffic at 40% CAV market penetration rate.

It is evident that the performance of CAV data-based signal control depends on the connectivity and data quality. Although several studies have applied RL-based techniques to solve signal control problems (Zhu *et al.*, 2015; Mannion, Duggan and Howley, 2016; Zou *et al.*, 2016; Islam, H M A Aziz, *et al.*, 2018; Zhao *et al.*, 2018); the impacts of CAV penetration rate were not explored in a detail manner. Our efforts fill this gap by evaluating the performance of RL-based signal control algorithm in different CAV market penetration rate and establish the basis of developing RL-based controls that account for mixed traffic with less than 100% market share of CAVs.

2.1 RESEARCH APPROACH

The sensitivity analyses reported here are based on our previous work (Islam, H. M. A. Aziz, *et al.*, 2018) where we have developed RL-based signal control strategies. For the details of the algorithm readers are referred to our published work (Islam, H. M. A. Aziz, *et al.*, 2018). RL-based techniques are suitable in a dynamic environment like traffic in a network of signalized intersections. Furthermore, the successful implementation of RL-based algorithm largely depends on the availability of the reliable and real-time data. Recent advances in CV environment offer useful and reliable technologies in acquiring high resolution data from transportation network. The next few subsections will define and describe these components in detail.

2.2 DATA COLLECTION ARCHITECTURE IN CV ENVIRONMENT

CV environment facilitates communication platform between vehicles and intersection controllers through vehicle to infrastructure (V2I) and infrastructure to infrastructure (I2I) communication. In a multi-agent system architecture, because of the decentralized control strategies, agents have a limited view of their surroundings. As a result, all the adjacent agents communicate back and forth through I2I communication to ensure the improvement in network-wide performance. The communication between agents increases the global view of an agent and so do the coordination (Balaji, German, & Srinivasan, 2010). Using dedicated short-range communication (DSRC), vehicles continuously broadcast their speed, acceleration, position to through Basic Safety Message (BSM) to the roadside unit. Using the shared information from CVs, an intersection controller equipped with RSU determines the condition of the network.

2.3 MECHANISM OF REINFORCEMENT LEARNING

In the context of RL techniques to control traffic signals, traffic network and controllers acted as an environment and agents, respectively. An intersection controller allocates right-of-way to a set of nonconflicting movements at each decision period; thereby, directly influences the transition of the states in a traffic network. The optimal policy is to find an action that is beneficial for the overall environment. This research utilizes a decentralized multi-agent architecture where each signal controller capable of estimating traffic condition and controlling the traffic signal without any central supervising agent. An RL-based algorithm requires essential components–state, action, and reward–to be defined specific to the problem at hand. At each time-step, a signal controller agent takes an action on termination or continuation of existing signal phase for the next time step. The proposed signal control system imposed minimum green constraints; however, do not follow any restrict phase sequence. The action sets could be different for each intersection and they were configured based on the signal phasing used in the study area.

During learning phase, an agent corresponding to an intersection randomly selected a phase. However, during the implementation, in the context of RL in traffic signal control, phase selection strategy involves an art of balancing the exploration and exploitation tradeoff meaningfully. The selecting action by the signal controller can be either entirely greedy i.e. selecting the action with the maximal benefits or exploratory by selecting random phases with pre-assigned probabilities, ε. During exploration-exploitation suggested by Sutton and Barto (Sutton & Barto, 1998), the authors used ε – *greedy method* for selecting a phase. This method involves in choosing an action that maximizes reward by each agent. However, at some cases it chooses a random action with a probability, ε (Aziz, Zhu, & Ukkusuri, 2017). Finally, during the evaluation of the proposed method, the selected actions were entirely greedy that involved in choosing a phase that maximizes reward by each agent.

2.3.1 Signal control strategies

Strategy-I: minimizing control delay

The definition of reward (a penalty in this case) is relatively straightforward in the single-agent case for minimizing control delay. The control delay of a vehicle was determined according to (1**Error! Reference source not found.Error! Reference source not found.Error! Reference source not found.**). The reward was calculated by the total control delay incurred between successive decision points by vehicles in the queues of all the approaches in an intersection. As such, the reward (penalty) function for an intersection-level agent can be written as

$$R_D = \sum_{\forall i \in I^t} d_i^t \tag{1}$$

d_i^t = control delay at time t in an intersection i.

I^t = set of all CAVs at time t.

Strategy-II: minimizing energy consumption

The reward was calculated by total energy consumed by all the vehicles approaching to an intersection. VSP and fuel consumption during idling phase were converted to gallons per unit time. The reward function for an intersection-agent is expressed as:

$$R_E = \sum_{i \in I_c^t} \phi(VSP_i^t) + \sum_{j \in I_s^t} \overline{E}_J \tag{2}$$

$\phi(VSP_i^t)$ = fuel consumption as a function VSP of vehicle $i \in I_c^t$ at time t,

I_c^t = set of all CAVs that are in cruising state at time t,

I_s^t = set of all CAVs that are in idling state at time t,

$\overline{E}_J$ = Energy consumption of idle-state vehicle j .

Strategy-III: minimizing energy consumption with penalty for stops

In order to reduce the unnecessary stops in a signalized intersection, queue stops are penalized in the strategy-III. We used an exponential function Π to penalize stops in an intersection. The reward function is calculated as (3).

$$R_{E-D} = R_E - \Pi(k) \tag{3}$$

$\Pi(k)$ = exponential function as a function of number of stops k.

= $\delta \times \rho \times \exp(\mu \times k)$

δ, μ = constant, 23.5 and 0.05 respectively

ρ = penalty factor

Our objective is to use signal timing at intersections to realize the above three strategies. This indicates that the decision variables for the above optimization are the set of signal controls (agent in RL scheme) at the concerned intersections.

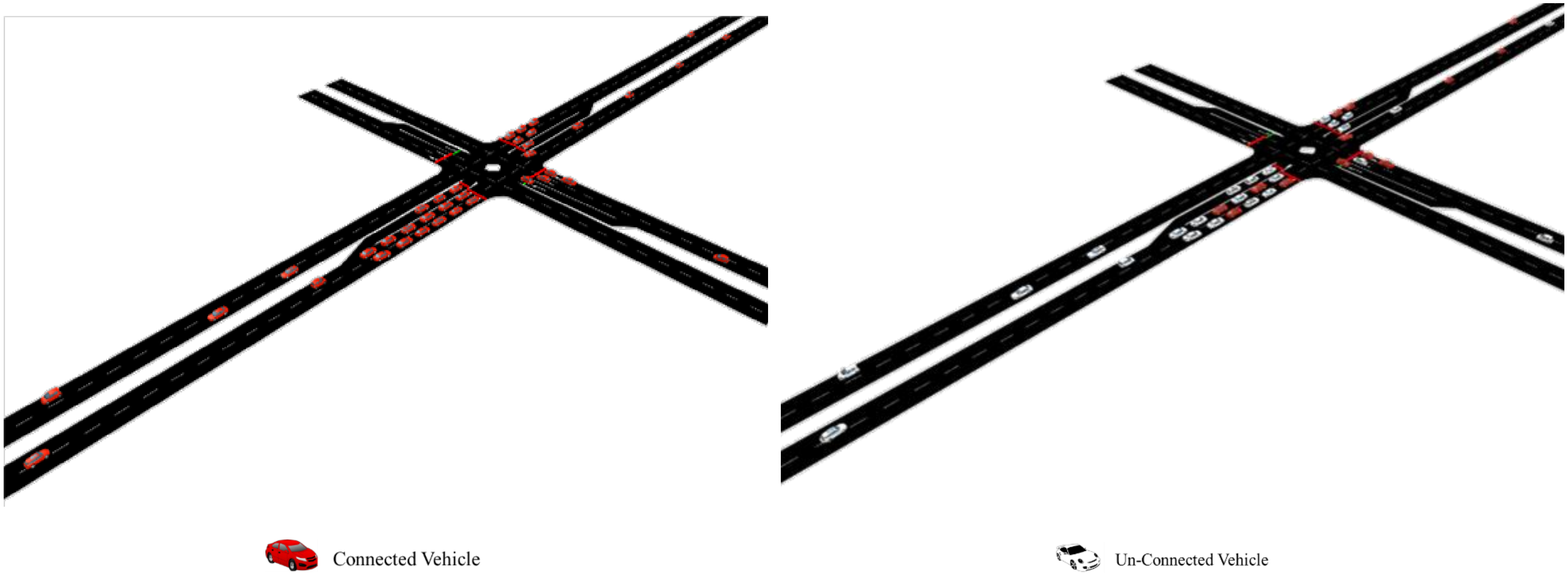
2.3.2 Conceptual framework for partial connectivity at low CAV market share

The goal of this research is to investigate the impact of partial information in low CAV penetration rate on the performance of a trained RL-based signal controller. After initial learning and exploration-exploitation phase, we implemented greedy action from Q table in different CAV market penetration rates. While testing the performance in low penetration rates, we did not update Q-values in the table. As such, recording reward function was not required. However, the signal controller needs to estimate state for finding an appropriate action from Q-table. In low CAV penetration rates, the signal controller receives BSMs from partial vehicles which leads to imperfect estimation of states. Figure 1(a) shows actual number of vehicles in queue at different movements. At low CAV penetration rates as shown in Figure 1(b), the estimated number of vehicles in queue is smaller than actual number. Therefore, in low penetration rate, queue state is estimated based on the position of the last CAV that joins in the queue assuming all the vehicle in front of the last CAV are in queue. If the average vehicle length is C, queue-state at a movement m can be calculated as (4). Note that, we did not change the logic for estimating inflow state in low penetration rate. Entry of a CAV is detected as the signal control receive BSMs at the entrance of a link.

$$\psi_m^{t,n} = \frac{\mathcal{L}_m^{t,n}}{C} \tag{4}$$

$\mathcal{L}_m^{t,n}$ = position of last CAV in the queue of movement m in an intersection $n \in N$

C = average length of a car.



(a) Actual queue state in 100% CAV penetration rate (b) Queue state in CAV penetration rate less than 100%

Figure 1. Queue-state in different CAV penetration rates.

3. EXPERIMENTAL SETUP

We developed and evaluated the performance of the signal control strategies in two study networks using a traffic microsimulation tool—PTV VISSIM (Vissim, 2013). The first network is a four-intersection arterial in Lakershim Boulevard, California. The second case study network is a portion of downtown Springfield, Illinois with 20 intersections. Next section provides details on the case study networks.

3.1 NETWORK-I: ARTERIAL (LANKERSHIM BLV, CA)

This case study network was simulated using microscopic traffic data that were collected under the Next Generation Simulation (NG-SIM) program in 2005. This high resolution dataset includes vehicle trajectories at a frequency of 0.1 seconds (Alexiadis *et al.*, 2004). The study area is approximately 1600ft long corridor consisting of bidirectional movements with three- to four-lane arterial segments and 35 mph speed limit. The arterial connects U.S. Highway 101 to the Universal City neighborhood as shown in Figure 2. This research utilized 15 minutes data sample, recorded from 8:30 am to 8:45 am on June 16, 2005 (Alexiadis *et al.*, 2004). We simulated the network based on the demands at all origins, predefined turning percentage at intersections defined in NG-SIM data. The network is calibrated before developing RL-techniques for signal control.

Figure 2. Case study network-I in Lankershim Blv, CA

This research follows the methodology proposed in Park et al., (2006) for the calibration process. This method involved establishing a mathematical model that estimates the measure of effectiveness (MOE) of simulated response based on the calibration parameters and finding optimal parameters so that performance measures of simulated response matches with field condition. A random sample considering all the controllable parameters in VISSIM was created based on Latin Hypercube sampling technique (Iman, 2008).

3.2 NETWORK-II: GRID NETWORK (SPRINGFIELD, IL)

The network consists a mix of one-way and two-way streets with different number of lanes and turning configurations at signalized intersections. The speed limit in the network is 25 mph. The base demand pattern for north and south direction is shown in Figure 3. Case study network-II in Springfield, IL. We used default parameters of VISSIM to simulate Springfield network as the actual vehicle trajectory is not available.

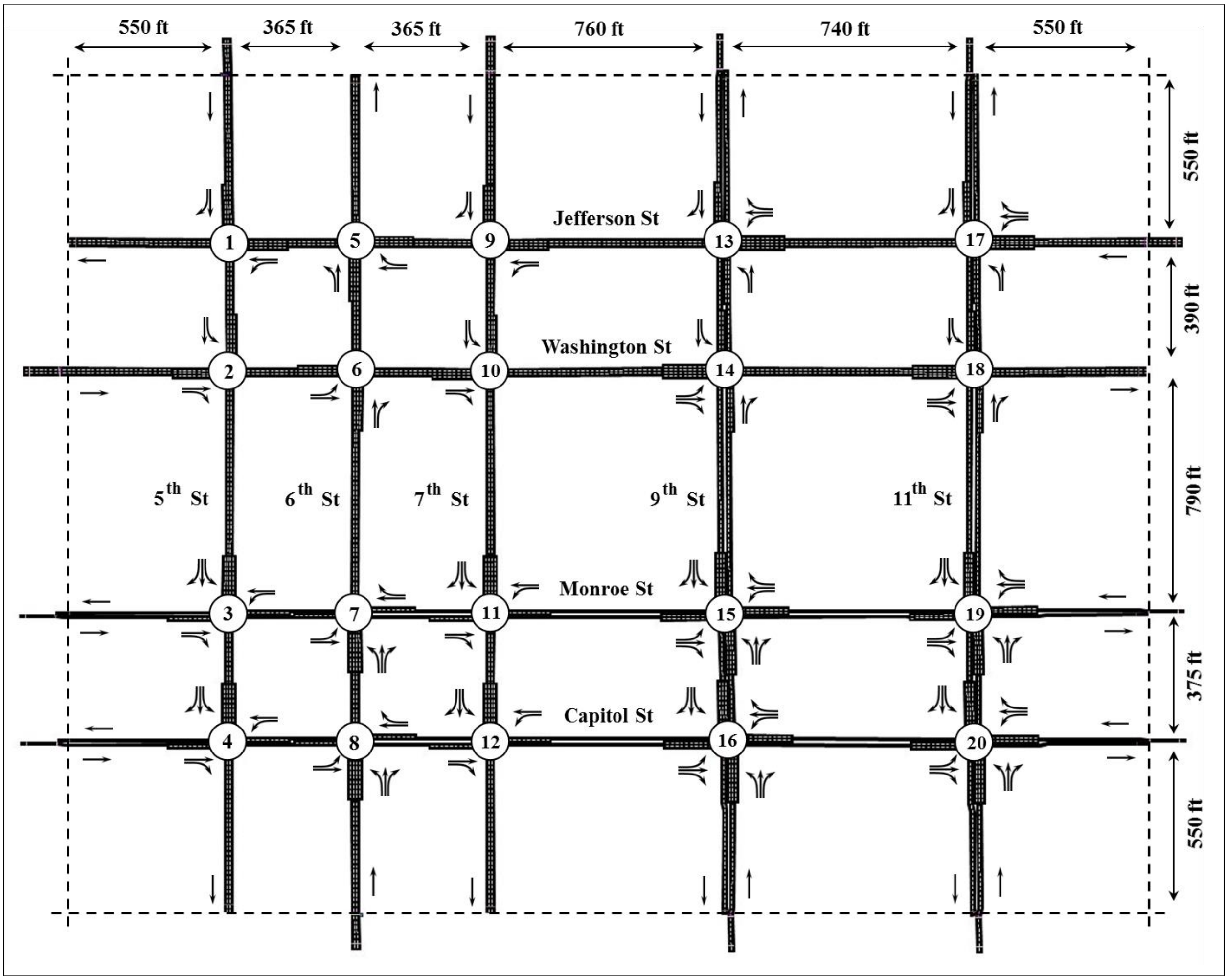


Figure 3. Case study network-II in Springfield, IL (Mohebifard and Hajbabaie, 2019; Mohebifard, Al Islam and Hajbabaie, 2019)

3.3 SIMULATION SETUP

This research uses traffic simulation tool VISSIM to simulate the environment. The signal control system based on RL was coded in Java inter-acting with VISSIM through the Component Object Model (COM) interface.

3.3.1 Learning and exploration-exploitation

At first, both of the networks were trained using random action for 2000 times with different random seeds. Each of the runs in Network-I and II were simulated for 15 minutes and 1 hours respectively. Netowrk-II is comparatively larger than network-I. As such, to capture the flow dynamics completely, the simulation time for each run in Network-II is higher compared to Network-I. The rewards and other performance metrics were obtained directly from VISSIM. After training the initial Q-table, we explored both networks for 100 times and prepared the final table for implementation.

3.3.2 Market share sensitivity

Finally, in order to evaluate the impact of low CAV penetration rate on RL-based signal control strategy, the proposed signal control strategies were tested for nine CAV market penetration rates (from 10% to 90% CAVs at 10% increments). During this evaluation, we implemented greedy actions from the Q-table that is learned and explored in 100% CAV penetration rate. The performances during implementation phase, a sample of 33 VISSIM simulation instances each with a different random seed were evaluated and their average was reported. Average travel time, average delay, number of completed trips, number of stops, network wide queue time, travel time and fuel consumptions were chosen as the measures of effectiveness in this research. Average travel time and delay were calculated for the completed trips only. In order to compare the performance for all the vehicles in the network, the authors reported number of stops, system travel time, total queue, number of stops for all the vehicles that were in the network. The detail on simulation setup in both of the case studies are given in Table 1.

Table 1. Description of the parameters during learning and exploration

Case study	Network-I (Lankershim Blv, CA)			Network-II (Springfield, IL)		
	Learning	**Exploration-exploitation**	**Implementation**	**Learning**	**Exploration-exploitation**	**Implementation**
Number of demand levels tested	-	3	3	-	3	3
CAV penetration rates	100%	100%	0~100%	100%	100%	0~100%
Number of runs	2000	100	33	2000	100	33
Simulation time (s)	900	900	900	3600	3600	3600
Action taking interval (s)	6	6	6	6	6	6
Action taking type	Random	ε – greedy	greedy	Random	ε – greedy	greedy

The proposed signal control methodology was tested for three demand patterns for both of the case study networks. All the demand patterns tested are detailed below:

- Network-I (Lankershim Blv., CA): 0%, 25% and 50% increase in the existing demand in the network. Current demand in the case study is taken as the base demand (v).
- Network-II (Springfield, IL): 50%, 75% and 100% of the saturated demand pattern in Springfield network. In saturated condition has 900 veh/hr/ln demand on all entry points.

4. RESULTS: IMPACT OF CAV MARKET SHARE

4.1 RESULTS FOR NETWORK-I (LANKERSHIM BLV, CA)

We implemented greedy action-based control using the final Q-table constructed in the strategy-I (minimizing delay) and strategy-III (minimizing energy with penalty for stops) to evaluate their performances under various CAV market penetration rates. We left strategy-II from further analysis due to its undesirable mobility performance. Figure 4 shows the mobility (completed trips, queue time, total travel time) and energy (energy and fuel consumption) performances of strategy-I and III for three different demand patterns in Lankershim Blvd under various penetration rates. The trends in Figure 4 suggests an improved performance for both strategies with increasing penetration rate from 20%. With increasing CAV market share, the accuracy of estimated states from CAV data is improved resulting into better trends in network performance metrics.

As shown in Figure 4, total travel time in the network for both strategy-I and III outperforms the performance of current signal control system at 70% CAV penetration rate under current demand pattern in the study area. However, the penetration rate reduced to 60% to outperform the current signal while the demand was increased by 50%. Increasing traffic volume increases the number of CAVs in the network; therefore, increases the probability of last vehicle in the queue to be a CAV. As such, the estimated queue state in low CAV market share became closure to actual one with increasing traffic volume. The authors acknowledged that the CAV penetration rate to outperform the performance of state-of-the-art signal control system cannot be generalized and highly depends on the design on the scenario *per se*.

Finally, the reduction of average trip delay in strategy-I is better compared to strategy-III as shown in Figure 4 (total travel time) as strategy-I directly emphasized reducing control delay in an intersection. The opposite is true for Figure 4 (Energy consumption) where strategy-III performs better than strategy-I in terms of total energy consumption the network. However, overall fuel consumption in strategy-III is higher than strategy-I as fuel consumption considers both queuing and cruising state, while energy consumption considers only cruising state.

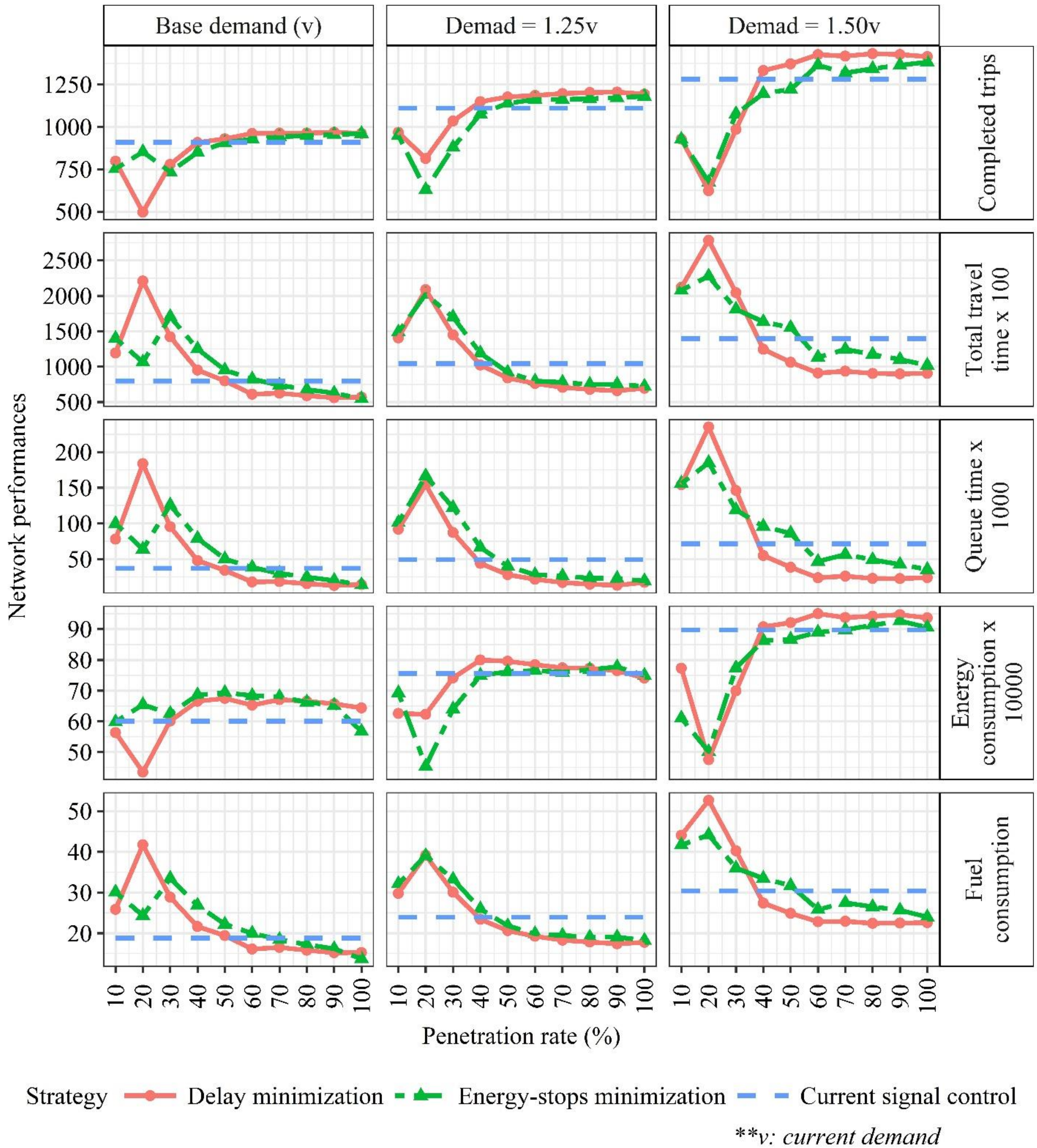


Figure 4. Network performances (Completed trips, total travel time (s), queue time (s), energy consumption (kW/ton) and fuel consumption (gal)) for different demand patterns in the case study network-I under various CAV market shares.

We observed that the performance of the algorithms at 10% CAV penetration rate did not align with the general trend in Figure 4. This off-the-trend performance at 10% CAV penetration rate does not mean the better performance in general. The underlying reason may be referred to the inherent stochastic nature of the transportation networks and the traffic states become unstable at lower penetration rate. It is possible that, in some cases the environment favors the state estimation at 10% CAV penetration rate which helps the algorithm to have a better performance. However, the better performance is not always guaranteed which is shown by the standard error of the performance matrices under several penetration rates in figure 5. We showed the network performances i.e. total travel time and energy consumption for strategy-III only as strategy-I also followed similar trend. Figure 5 also shows that the value of standard error reduced with CAV penetration rate. This implies that as the CAV market share increases the system becomes more stable in network-I.

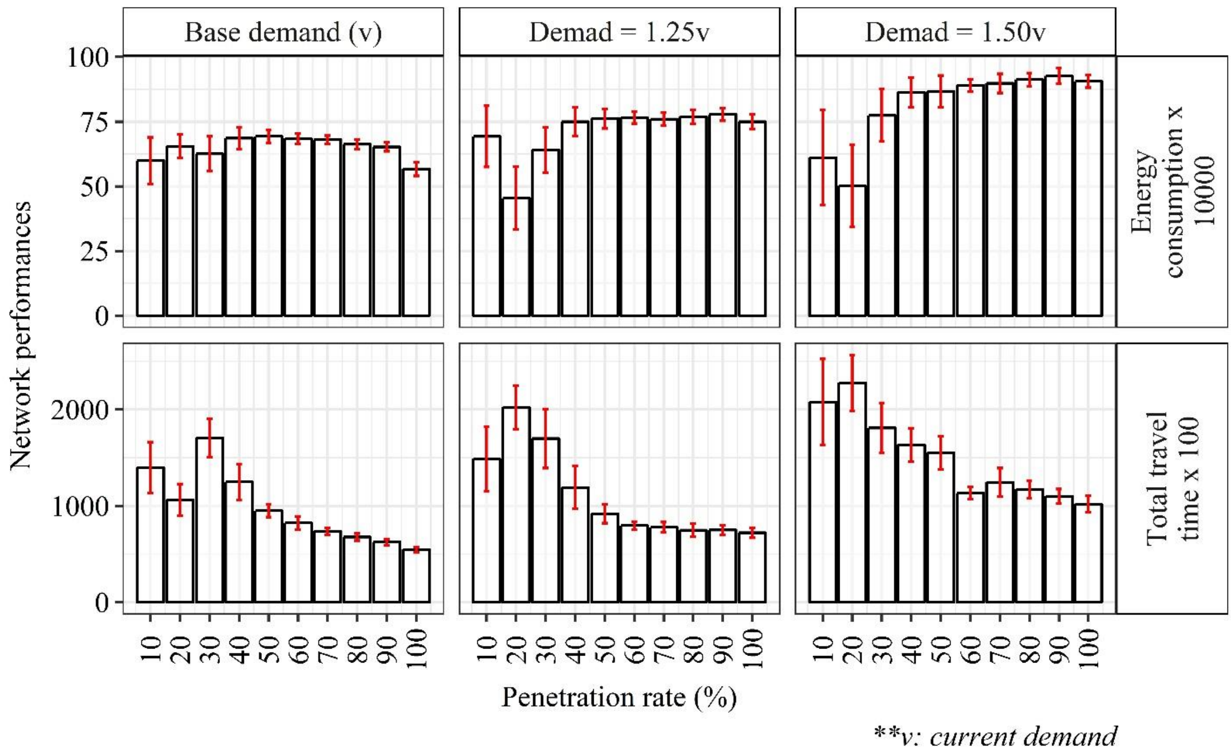


Figure 5. Mean value with standard errors of 33-simulation runs for network performances (total travel time (s) and energy consumption (kW/ton)) in Strategy-III under different demand pattern in case study network-I

4.1.1 Network-II (Springfield, IL)

Similar to network-I, we analyzed the mobility and energy performances of signal control based on strategy-III in various CAV market penetration rates for network-II. Figure 6 shows an increasing trend in both mobility and energy performances with CAV penetration rates. At 100% CAV penetration rate, as the traffic volume increases, number of completed trips increases. Therefore, other network performance measures i.e. total travel time, queue time, fuel and energy consumption would increase. However, these performances are not consistent at CAV penetration rate less than 100%. For instance, at 90% CAV market share, case study with under-saturated traffic demand completed more trips than conditions with higher traffic demand. This implies that, at high traffic volume, the underestimation of traffic state leads the system to be instable that seriously affects the performance of the signal control. In low traffic demand, the signal controller gets enough room to bring the network back to stable condition. However, increasing number of approaching vehicles in the case studies with higher demand do not left enough room for the signal controller to handle instability which causes the queue to overflow the link capacity. As such, queue spillback and grid-lock become common phenomena in high traffic demand. Figure 6 implies that higher demand may cause greater order of inaccurate estimation of the traffic state at low penetration rate. We also observe that, saturated demand would reach the grid-lock condition quicker than other demand patterns and accordingly, we have fewer completed trips at CAV penetration rate less than 100%.

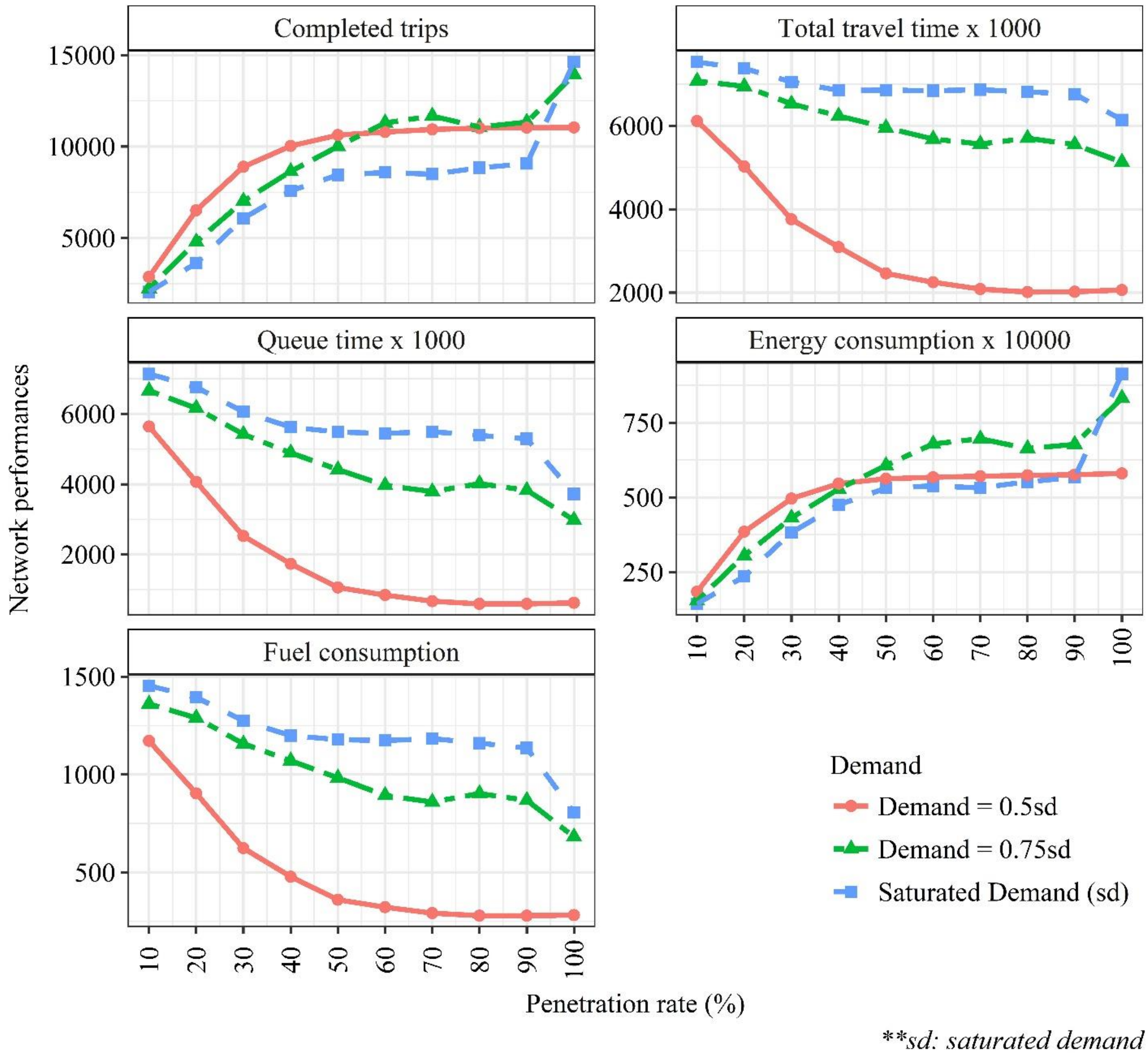


Figure 6. Network performances (Completed trips, total travel time (s), queue time (s), energy consumption (kW/ton) and fuel consumption (gal)) of the signal control strategy-III for different demand patterns under various CAV penetration rate in case study network-II.

4.2 SENSITIVITY TO MARKET SHARE OF BATTERY ELECTRIC VEHICLES

We analyzed the performance of the signal control strategy-III under various market penetration rate of battery electric vehicles (BEV). We implemented greedy action from final Q-table constructed in the strategy-III assuming 100% gasoline vehicle. In this analysis, our goal is to find out the performance of RL-based signal control that is completely trained for gasoline vehicle in a transportation network with various market penetration of BEVs.

It is expected that increasing percentage of BEVs reduces energy consumption in the network accounting for the energy-gains from the regenerative braking phenomenon. BEVs solely use one or more electric motor(s) for traction with batteries as energy resources (Ehsani *et al.*, 2018). It has the bidirectional power flows and power recovery (battery charging) is enabled from regenerative braking. The efficiency of a traction motor varies with its operating points on the speed–torque (speed–power) plane. For Nissan Leaf, the motor efficiency is normally around 80-95% (Brooker *et al.*, 2015). The regenerative braking efficiency can be as high as 90%, depending on vehicle speed. However, the efficiency for barking to chemical energy (Battery) is quite stochastic and function of many different parameters (Rask, Santini and Lohse-Busch, 2013). Note that our analysis is rather a simpler version of this complex dynamics to qualitatively find the order of gains due to regenerative braking. In this analysis, we computed the energy consumption of conventional gasoline vehicle using (**Error! Reference source not found.**) considering no energy gain due to braking. In this study we used the following equation to compute energy consumption of BEV:

$$VSP_i^t = \frac{a_i^t w_i + 0.5\rho A_i C_d {v_i^t}^2 + R + w_i g \sin\theta}{w_i 1000} \tag{1}$$

VSP_i^t	=	Vehicle specific power in watt per ton of vehicle i ,
w_i	=	Mass of the vehicle,
g	=	Acceleration due to gravity,
ρ	=	Air density ($=\ 1.1985\ kg/m^3$),
A_i	=	Frontal area (m^2),
C_d	=	Drag coefficient,
R	=	Rolling resistance ($= (\mu_{r_1} + v_i^t \mu_{r_2}) w_i g \cos\theta$),
μ_{r_1}, μ_{r_2}	=	Rolling resistance coefficients.

We assumed that 70% of the energy could be regained when the calculated VSP is negative. In our analysis, we assumed all BEVs as 2016 Nissan Leaf and all the vehicle specific parameters are taken from Autonomie (Karbowski, Pagerit and Calkins, 2012).

4.2.1 Results

Figure 7 shows network-level energy consumption (reported with quartile values) in case study network-I for all the demand patterns. As the percentage of BEVs increases, the energy consumption in the network decreases as expected. Higher number of BEVs enable the system to regain some energy from braking process. As shown in the figure, with 100% BEVs in Lankershim Blvd, we could save up to 20% energy consumption compared to the base case of 0% BEV. However, we acknowledge that the amount of energy consumption reported in this study may not reflect the actual value as we used a very simpler equation for energy gain in regenerative braking process. However, we expected a similar downward trend of reduction in energy consumption in real condition. Similar to Figure 7, Figure 8 shows the network-level energy consumption (reported with quartile values) for network-II. Furthermore, a general decreasing trend in energy consumption is also found in Figure 8. With 100% BEVs, we can regain up to 40% of energy consumption compared to the base case of 0% BEVs for network-II.

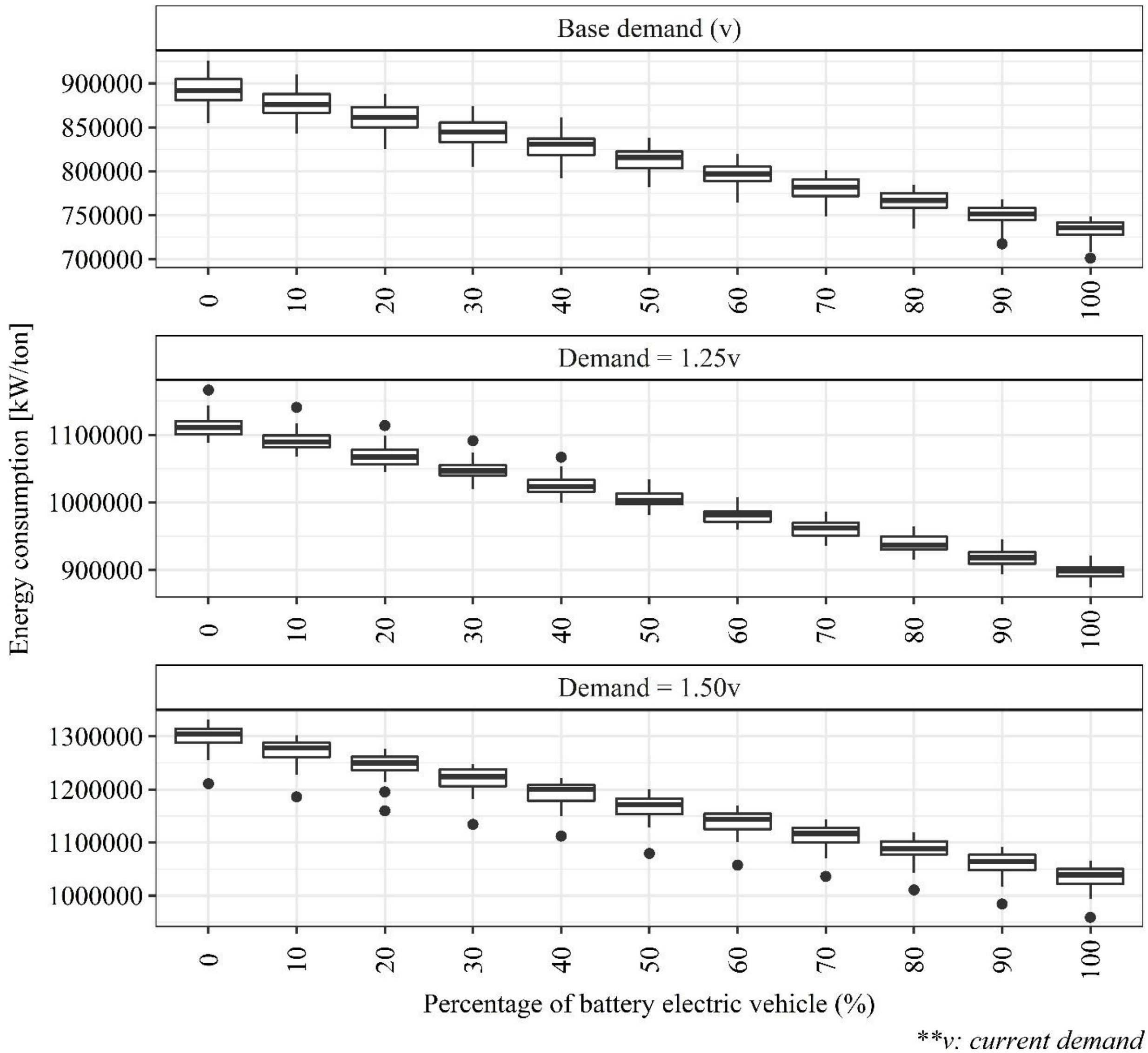


Figure 7. Energy consumption (kW/ton) in signal control strategy-III under various percentage of BEVs in case study network-I for different demand patterns.

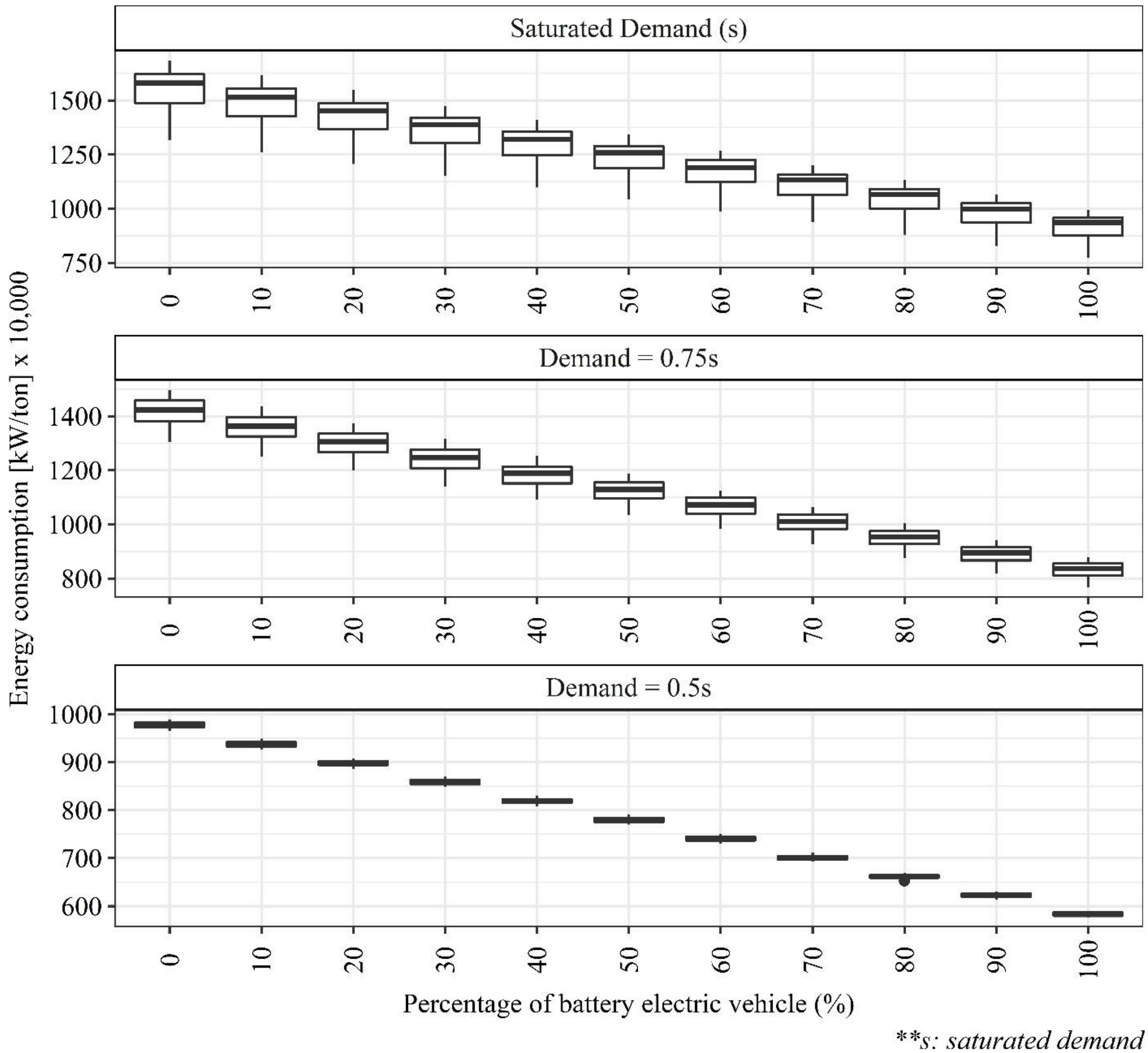


Figure 8. Energy consumption (kW/ton) in signal control strategy-III under the various percentage of BEVs in case study network-II for different demand patterns

5. CONCLUDING REMARKS

Our primary objective in this research is to explore and understand the impact of partial-connectivity—leading to an inaccurate observation of the traffic state—on the performance of a reinforcement learning-based signal control that highly depends on the availability and the quality of the data exchange in a connected-automated traffic environment. Previously, we have developed reinforcement learning (RL) based algorithms where the signal leverages data connectivity and exchange capability in a connected-automated environment. This control algorithm assumes a 100% market share of equipped vehicles (CAVs) enabling perfect data exchange with the controller. Now, we have simulated the market share of CAVs through the representation of partial connectivity. Following a defined statistical distribution of vehicles arriving at the intersection, we assigned a certain percentage of the vehicles to be incapable of exchanging data with the controller or the road-side unit. Compared to a point estimate, we produced a distribution of change in the signal control performance metrics (e.g., delay, energy, stops, and so on) for a partial market share of CAVs—see figure 5. Our analyses provide an answer to the question—if we have only 30% of CAVs that exchange data with the controller, what would be the performance of the RL-based control compared to the case when we have 100% market share.

Except for low market share (below 40% in most cases) values, the trend is intuitive. The greater availability of data through connectivity improves the performance of the control algorithms. The gains are marginal after a 90% market share on an average. We must note that the results cannot be generalized, instead, the findings are only valid for the two networks with our experimental designs. Our findings also indicate the instability of the performance of reinforcement learning based control arising from the low market share of CAVs. Any decision made by the signal controller with inaccurate traffic state estimation has a long-term effect and the propagation is stochastic in nature—can lead to excessive queue building up or being dampened by the dominating flow from conflicting approaches. Extensive analyses will be required to understand and identify the instability pattern at a low market share of CAVs.

Our results and findings will be the base to enhance the existing RL algorithm to incorporate the uncertainty in state estimation that impacts the performance metrics. As future work, we plan to adapt Partially-Observed-Markov-Decision-Process (POMDP) theory to reformulate the RL-based control which will allow finding optimal control settings when the traffic states are not accurate all the time—partially observable.

6. REFERENCES

Alexiadis, V. *et al.* (2004) 'The next generation simulation program', *Institute of Transportation Engineers. ITE Journal*. Institute of Transportation Engineers, 74(8), p. 22.

Argote-Cabañero, J., Christofa, E. and Skabardonis, A. (2015) 'Connected vehicle penetration rate for estimation of arterial measures of effectiveness', *Transportation Research Part C: Emerging Technologies*, 60, pp. 298–312. doi: 10.1016/j.trc.2015.08.013.

Brooker, A. *et al.* (2015) 'FASTSim: A Model to Estimate Vehicle Efficiency , Cost and Performance', *SAE Technical Paper*, (April), pp. 21–23. doi: 10.4271/2015-01-0973.Copyright.

Day, C. M. and Bullock, D. M. (2016) 'Detector-Free Signal Offset Optimization with Limited Connected Vehicle Market Penetration', *Transportation Research Record: Journal of the Transportation Research Board*, 2558, pp. 54–65. doi: 10.3141/2558-06.

Ehsani, M. *et al.* (2018) *Modern electric, hybrid electric, and fuel cell vehicles*. CRC press.

Feng, Y. *et al.* (2015) 'A real-time adaptive signal control in a connected vehicle environment', *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 55, pp. 460–473. doi: 10.1016/j.trc.2015.01.007.

Goodall, N., Smith, B. and Park, B. (2013) 'Traffic Signal Control with Connected Vehicles', *Transportation Research Record: Journal of the Transportation Research Board*, 2381(2381), pp. 65–72. doi: 10.3141/2381-08.

He, Q., Head, K. L. and Ding, J. (2012) 'PAMSCOD: Platoon-based arterial multi-modal signal control with online data', *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 20(1), pp. 164–184. doi: 10.1016/j.trc.2011.05.007.

Iman, R. L. (2008) 'Latin hypercube sampling', *Encyclopedia of quantitative risk analysis and assessment*. Wiley Online Library.

Islam, S. M. A. B. Al, Aziz, H M A, *et al.* (2018) 'Minimizing energy consumption from connected signalized intersections by reinforcement learning', in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1870–1875. doi: 10.1109/ITSC.2018.8569891.

Islam, S. M. A. B. Al and Hajbabaie, A. (2017) 'Distributed coordinated signal timing optimization in connected transportation networks', *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 80, pp. 272–285. doi: 10.1016/j.trc.2017.04.017.

J, L. (no date) 'Cumulative Travel-Time Responsive Real-Time Intersection Control Algorithm in the Connected Vehicle Environment.', *Journal of Transportation Engineering*.

Karbowski, D., Pagerit, S. and Calkins, A. (2012) 'Energy consumption prediction of a vehicle along a user-specified real-world trip', *World Electric Vehicle Journal*, 5(4), pp. 1109–1120.

Lee, J., Park, B. (Brian) and Yun, I. (2013) 'Cumulative Travel-Time Responsive Real-Time Intersection Control Algorithm in the Connected Vehicle Environment', *Journal of Transportation Engineering*, 139(10), pp. 1020–1029. doi: 10.1061/(ASCE)TE.1943-5436.0000587.

Mannion, P., Duggan, J. and Howley, E. (2016) 'An Experimental Review of Reinforcement Learning Algorithms for Adaptive Traffic Signal Control', *Autonomic Road Transport Support Systems*, pp. 47–66. doi: 10.1007/978-3-319-25808-9_4.

Mohebifard, R. and Hajbabaie, A. (2018) 'Distributed optimization and coordination algorithms for real-time traffic metering in connected urban street networks', *IEEE Transactions on Intelligent transportation systems*. doi: 10.1109/TITS.2018.2848246.

Mohebifard, R. and Hajbabaie, A. (2019) 'Optimal network-level traffic signal control: A benders decomposition-based solution algorithm', *Transportation Research Part B: Methodological*, 121, pp. 252–274. doi: https://doi.org/10.1016/j.trb.2019.01.012.

Mohebifard, R., Al Islam, S. M. A. Bin and Hajbabaie, A. (2019) 'Cooperative traffic signal and perimeter control in semi-connected urban-street networks', *Transportation Research Part C: Emerging Technologies*. Elsevier, 104, pp. 408–427.

Park, B. *et al.* (2006) *Microscopic simulation model calibration and validation handbook.*

Priemer, Christian and Friedrich, B. (2009) 'A decentralized adaptive traffic signal control using V2I communication data', in *Intelligent Transportation Systems, 2009. ITSC'09. 12th International IEEE Conference on*, pp. 1–6.

Priemer, C. and Friedrich, B. (2009) 'A decentralized adaptive traffic signal control using V2I communication data', *2009 12th International IEEE Conference on Intelligent Transportation Systems*, pp. 765–770. doi: 10.1109/ITSC.2009.5309870.

Rask, E., Santini, D. and Lohse-Busch, H. (2013) ‘Analysis of input power, energy availability, and efficiency during deceleration for X-EV vehicles’, *SAE International Journal of Alternative Powertrains*. JSTOR, 2(2), pp. 350–361.

Smith, B. L. *et al.* (2011) *IntelliDrive Traffic Signal Control Algorithms, University of Virginia*.

Vissim, P. (2013) ‘PTV Vissim 7 User Manual’, *PTV AG*.

Zhao, P. *et al.* (2018) ‘A deep reinforcement learning framework for optimizing fuel economy of hybrid electric vehicles’, in *Design Automation Conference (ASP-DAC), 2018 23rd Asia and South Pacific*, pp. 196–202.

Zheng, J. and Liu, H. X. (2017) ‘Estimating traffic volumes for signalized intersections using connected vehicle data’, *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 79, pp. 347–362. doi: 10.1016/j.trc.2017.03.007.

Zhu, F. *et al.* (2015) ‘A junction-tree based learning algorithm to optimize network wide traffic control: A coordinated multi-agent framework’, *Transportation Research Part C: Emerging Technologies*. Elsevier.

Zou, Y. *et al.* (2016) ‘Reinforcement learning-based real-time energy management for a hybrid tracked vehicle’, *Applied energy*. Elsevier, 171, pp. 372–382.