Eco-Driving With Partial Wireless Charging Lane at Signalized Intersection: A Reinforcement Learning Approach

Xinxing Ren, *Student Member, IEEE*, Chun Sing Lai, *Senior Member, IEEE*, Zekun Guo, *Member, IEEE*, and Gareth Taylor, *Senior Member, IEEE*

***Abstract*—Consumer electronics such as advanced GPS, vehicular sensors, inertial measurement units (IMUs), and wireless modules integrate vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) within Internet of Things (IoT), enabling connected autonomous electric vehicles (CAEVs) to optimize energy optimization through eco-driving. In scenarios with traffic light intersections and partial wireless charging lanes (WCL), an eco-driving algorithm must consider net and gross energy consumption, safety, and traffic efficiency. We introduced a deep reinforcement learning (DRL) based eco-driving control approach, employing a twin-delayed deep deterministic policy gradient (TD3) agent for real-time acceleration planning. This approach uses reward functions for acceleration, velocity, safety, and efficiency, incorporating a dynamic velocity range model which not only enables the vehicle to smoothly pass the signalized intersections but also uses partial WCL efficiently and time-adaptively while ensuring traffic efficiency in diverse traffic scenarios. Tested in Simulation of Urban Mobility (SUMO) across various intersections with partial WCL, our method significantly lowered net and gross energy consumption by up to 44.01% and 17.19%, respectively, compared to conventional driving, while adhering to traffic and safety norms.**

***Index Terms*—Consumer electronics, vehicle-to-vehicle communications, vehicle-to-infrastructure communication, connected autonomous electric vehicles, autonomous electric vehicles, eco-driving, wireless charging lane, deep reinforcement learning.**

NOMENCLATURE

Symbols

a_{ego}	Acceleration of the ego vehicle.
a_{IDM}	Acceleration generated by the Intelligent Driving Model.
a_{min}	Minimum acceleration.
a_{max}	Maximum acceleration.
a_{safe}	Safe acceleration derived from the rule-based car-following model.
a_{TD3}	Acceleration determined by the Twin Delayed Deep Deterministic (TD3) policy gradient agent.
$a^{(a)}$	Maximum possible acceleration in the IDM.
b	Action of agent in current state.
$b^{(a)}$	Maximum possible deceleration in the IDM.
c_1, c_2, c_3, c_4	Weighting factors in the reward function.
D_{leader}	Distance to the leading vehicle.
$D_{traffic\ light}$	Distance to the upcoming traffic light.
D_{WCL_remain}	Remaining distance to traverse within the Wireless Charging Lane.
E_{Bat}	Energy stored in the vehicle's battery.
ΔE_{Bat}	Net energy consumption of the battery pack at each time step.
I_1, I_2, I_3, I_4	Indicator variables for phases or conditions in the dynamic velocity range model.
k	Delay cycle number.
P_{chrg}	Wireless charging power.
$r_{acceleration}$	Reward functions for acceleration.
$r_{velocity}$	Reward functions for velocity.
r_{safety}	Reward functions for safety.
$r_{efficiency}$	Reward functions for traffic efficiency.
s	Current state in the Markov Decision Process.
$s^*, s_0^{(a)}, s_a$	Desired distance, minimum distance, and actual distance with leading vehicle in the IDM.
t_{green_cycle}	Duration of the green phase of a traffic signal.
t_{red_cycle}	Duration of the red phase of a traffic signal.
t_{green_remain}	Remaining time of the green phase.
t_{red_remain}	Remaining time of the red phase.
t	Current time step.
T_{WCL}	Time spent within the WCL by the vehicle.
T^a	Constant time headway in the IDM.
V_{ego}	Speed of the ego vehicle.
V_{leader}	Speed of the leading vehicle.
V_{limit}	Speed limit on the road.
V_{max}, V_{min}	Maximum and minimum speeds determined by the dynamic velocity range model for standard driving scenarios.

Received 30 June 2024; revised 26 September 2024; accepted 11 October 2024. Date of publication 16 October 2024; date of current version 31 December 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62206062, and in part by the Brunel University of London-Chongqing University of Posts and Telecommunications Doctoral Studentship. *(Corresponding author: Chun Sing Lai.)*

Xinxing Ren and Gareth Taylor are with the Department of Electronic and Electrical Engineering, Brunel University of London, UB8 3PH London, U.K. (e-mail: xinxing.ren@brunel.ac.uk; gareth.taylor@brunel.ac.uk).

Chun Sing Lai is with the Department of Electronic and Electrical Engineering, Brunel University of London, UB8 3PH London, U.K., and also with the School of Automation, Guangdong University of Technology, Guangzhou 510006, China (e-mail: chunsing.lai@brunel.ac.uk).

Zekun Guo is with the Department of Engineering, University of Hull, HU6 7RX Hull, U.K. (e-mail: z.guo2@hull.ac.uk).

Digital Object Identifier 10.1109/TCE.2024.3482101

V_{max_w}, V_{min_w}	Maximum and minimum speeds determined by the dynamic velocity range model for driving scenarios with wireless charging lane.
v_a, $v_0^{(a)}$, Δv	Current speed, maximum possible speed, and speed variation in the IDM.
δ	Acceleration exponent in the IDM model.
η_{chrg}	Efficiency of the wireless charging process.
γ	Discount factor in the reinforcement learning framework.
w_1, w_2	Weighting coefficients in the multi-objective function of eco-driving.

I. Introduction

THE ADOPTION of the electric vehicles (EVs) industry is impeded primarily by the constrained driving range [1]. According to a study conducted by [2], it has been shown that the typical customer expectation for the driving range of passenger vehicles falls within the range of 300-400 miles. However, in the United Kingdom, the average driving range of EVs sold in the year 2023 was recorded to be just 212 miles [3]. Both industry and academics have devised many ways to address the issue of driving range anxiety. These options include enhancements in battery capacity [4], battery swapping [5], and plug-in fast charging technologies [6]. Although increasing battery capacity is the most direct solution to range anxiety, a typical new EV costs between £23,000 and £43,000 more than an equivalent petrol or diesel model [7], and batteries account for between 25% and 40% of the cost of an EV [8], making it less likely that consumers will be able to afford battery upgrades. Although the primary advantage of battery swapping is speed, brand compatibility and cross-platform features may hinder the success of this technology, and battery degradation and the issue of battery ownership will be obstacles to its widespread adoption [5]. Although the plug-in fast charging technology can significantly shorten the charging time, it has the following disadvantages: first, the plug-in fast charging technology will significantly reduce the battery life [9], and second, a short circuit or breakdown of the insulation of the charging wire due to factors such as high temperature, friction with the ground, or the charging device itself can result in a fatal electric shock [10].

By integrating consumer electronics such as advanced GPS systems, vehicular sensors, IMUs, and wireless communication modules, CAEVs are potential to save energy consumption up to 47.5% while ensuring traffic efficiency and safety by planning its speed trajectory [11], which is known as eco-driving. These consumer electronics facilitate precise real-time data exchange, enabling vehicles to interact seamlessly with their environment. IMUs and vehicular sensors, for instance, provide accurate positional data and vehicle dynamics information, which are essential for maintaining safety and improving driving precision [12]. Wireless communication modules power V2V and V2I communications and ensure reliable connectivity between vehicles and infrastructure, allowing CEVs to obtain signal phase and timing (SPaT) information [13], which is crucial for planning speed trajectory and allows CAEVs to smoothly pass the signalized intersections so that avoid unnecessary energy consumption. Furthermore, these components support advanced vehicular functions such as cooperative sensing and secure data sharing [14], enhancing the overall functionality of eco-driving.

At the same time, the wireless charging system (WCS) is now being extensively investigated as a possible solution to address range anxiety. Notably, [15] highlighted an adaptive technique poised to significantly improve the efficiency of magnetic resonance-based wireless power transfer systems. This innovation is particularly geared towards portable consumer electronics but holds profound implications for EV charging solutions, potentially transforming the existing landscape of electric mobility by optimizing energy transfer over distances. Compared to other solutions for addressing driving range, WCS has several benefits [16]. WCS is more user friendly as EVs can be charged while in motion, so they do not have to deal with issues associated with plug-in charging station, such as charging time, location within an EV's range, potential traffic congestion, and waiting time at charging station [1]. Second, WCS charges the battery frequently but slowly, which is able to reduce battery size and increase battery life [17]. At the same time, WCS equipment can be concealed beneath the road, eliminating the danger of electric shock for road users. In addition, WCS emits even less radiation than mobile phones [18]. Moreover, compared to plug-in charging, WCS is more economical and more environmentally friendly. A survey was done by Brown [19] at the University of Michigan, using a 12-year framework study, in order to assess the efficacy of plug-in and WCS. Two conclusions were reached. 1) WCS systems minimize battery use, reducing greenhouse gas (GHG) emissions and wireless charging energy needs. 2) Reducing battery size and weight may offset the additional costs of wireless system installation.

The urban transport system is a potential application for wireless charging. Since many EVs will be halted or travelling slowly in front of the intersection, this will be an ideal time to charge these EVs. Mohrehkesh and Nadeem [20] have demonstrated that wireless charging at traffic intersections extends the range of electric vehicles. In their work, stationary vehicles at red traffic signals were considered. However, to enhance traffic efficiency, EVs should avoid stopping at traffic lights and aim to pass through them as swiftly as possible. Given cost considerations, a more practical approach is to install partial wireless charging lane (WCL) in specific sections, such as just before traffic lights, rather than laying them continuously throughout the area. Consequently, EVs face a balance between fully utilizing partial WCL and maintaining traffic efficiency. Developing a speed planning strategy for CAEVs that navigates this balance effectively represents a significant challenge.

A summary of eco-driving algorithms is shown in Table I. Many eco-driving algorithms have been proposed for CAEVs to reduce gross energy consumption while guarantee safety and traffic efficiency, although the impact of WCL was not

TABLE I
A SUMMARY OF ECO-DRIVING ALGORITHMS

Algorithm	Methodology	SPaT information	Car-following	Wireless charging lane/ Timing-adaptive	Real-time
[21]	Mathematical Optimization	√	–	–	–
[22]	Convex Optimization	√	–	–	√
[23]	MPC	√	√	–	√
[24]	DRL	√	√	–	√
[25]	Mathematical Optimization	√	–	√/–	–
Proposed	DRL	√	√	√/√	√

considered in these methods. (In this paper, Gross energy consumption refers to the total energy drawn from the battery. In contrast, net energy consumption is the difference between the energy used from the battery and the energy that has been recharged.) These algorithms can be evaluated from the considered scenarios and methodology. The basic eco-driving algorithm considered SPaT information based on the scenario [21], [22]. By calculating a velocity range based on the phase duration and the distance to the traffic light, EVs can past the traffic light without halting. Subsequent studies [23], [24] progressively included SPaT information and integrated the car-following capability into the algorithm. This enhancement enabled CAEVs to optimize energy usage while driving in complex traffic scenarios. The implementation of eco-driving has gone through a series of methodological iterations, beginning with global mathematical optimization methods [21] and progressing to convex optimization [22], model predictive control (MPC) [23] and DRL [24]. Global optimization mathematical methodologies are capable of determining the global optimal solution, but their numerous calculation times prevent them from supporting autonomous driving decisions in real time. Conversely, convex optimization is able to significantly reduce computation time and MPC can achieve a near-optimal solution while further reducing the computational time for each eco-driving decision to less than 0.1 seconds. However, a computational time of 0.1 seconds still presents a significant safety risk for CAEVs. By incorporating neural networks to estimate the value function, DRL can reduce computational time by an order of magnitude compared to MPC for the same task, while maintaining similar performance levels [26]. Moreover, the advanced generalization capability of DRL enables it to handle complex scenarios effectively.

As WCS becomes a hot research topic, Zhang et al. [25] proposed an eco-driving algorithm considering WCL while assuming the road is empty, which is not suitable to be used in realistic traffic scenarios. At the same time, this method lacks the capability of timing-adaptive. This may lead to controlled

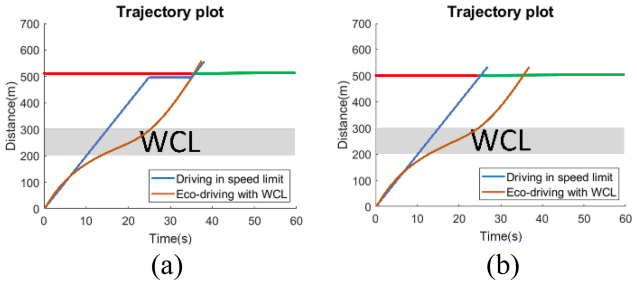


Fig. 1. (a) Extra time is given to allow EVs to spend more time within WCL. (b) No extra time is given to allow EVs to spend more time within WCL.

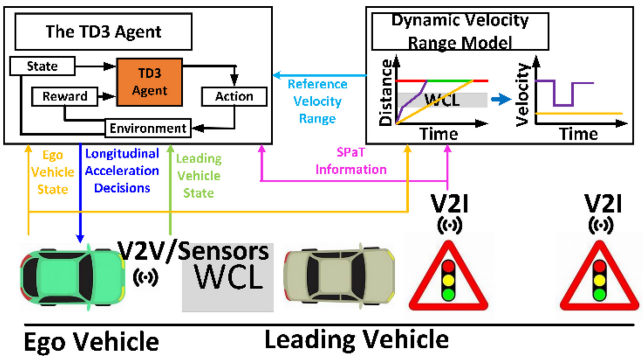


Fig. 2. Illustration of proposed eco-driving algorithm at signalized intersections with partial WCL.

vehicles seeking excessive charging even when time is limited, resulting in compromised traffic efficiency. For example, as shown in Fig. 1 (a), assume that a road is 500m long, there is a WCL located between 200-300m, the road speed limit is 20m/s, and there are 35 seconds left at the red phase. If an EV drives at a speed of 20m/s, it will eventually stop in front of the traffic light. This means giving an EV extra time to stay in WCL for more time to increase its range without affecting traffic efficiency. However, if in the second case, as shown in Fig. 1 (b), the red phase only has 25 seconds left, which means that an EV can directly pass the traffic light at a speed of 20m/s. In this case, the EV should not stay in the WCL for any additional time, while [25] does not balance this trade-off.

Apart from accounting for Timing-adaptive, balancing the simultaneous consideration of SPaT information, car-following behaviors, wireless charging lane integration, and real-time responsiveness presents a notable challenge. On one hand, the extensive computational requirements preclude the application of global mathematical optimization techniques. Conversely, attempts to merge prior research efforts, which have explored diverse scenarios and factors, encounter difficulties due to fundamental differences in methodology. The varying processes of problem formulation and solution strategies employed in these methodology present significant disparities.

Considering the above-mentioned factors, this paper proposes a DRL-based real-time eco-driving algorithm, to allow the controlled vehicle to utilize WCL efficiently and timing-adaptively while ensuring safety and maintaining traffic efficiency. This algorithm is illustrated in Fig. 2. The dynamic velocity range model generates a reference speed to allow the

controlled vehicle to efficiently and adaptively utilize WCL while maintaining traffic efficiency. The generated reference speed is then given to the TD3 agent as an input, which determines the acceleration for the controlled vehicle in real-time to follow the reference speed while ensuring traffic safety. The contribution of this study is to highlight: 1) A dynamic velocity range model that enables the controlled vehicle to efficiently and adaptively utilize WCL while maintaining traffic efficiency, 2) A DRL-based eco-driving controller that determines real-time acceleration for the controlled vehicle to adhere to the reference speed and ensure traffic safety, 3) A novel eco-driving method that allows CAEVs to maximize the usage of partial WCL while ensure safety and traffic efficiency at signalized intersections with partial WCL. To the best of our knowledge, this is the first method to creatively address the challenge of integrating SPaT information, car-following, WCL, and real-time operation. In general, this study introduces a potentially effective eco-driving approach at signalized intersections with partial WCL, hence offering potential contributions towards the establishment of a transport system that is both sustainable and efficient.

The remaining sections of this paper are organized as follows: Section II provides energy consumption models of the vehicle and wireless charging model and formulates the optimal control problem (OCP) for eco-driving. In Section III, the eco-driving control is elaborated by considering the partial WCL. In Section IV, the simulation setup and benchmarks are introduced. In Section V, the simulation results used to validate the proposed method are presented. In Section VI, the major conclusions and future work are discussed.

II. Modelling and Problem Statement

In this section, the energy consumption models for the electric vehicle and wireless charging system are presented. In addition, this section formulates the OCP for eco-driving.

A. Energy Consumption Model

To enhance the precision of performance evaluation for the proposed model, it is crucial to select a highly accurate EV energy consumption model. In this work, the Electric Vehicle Emission Model (MMPEVEM) from the Mechatronics in Mobile Propulsion of RWTH Aachen University is employed, as it accounts for every component of the powertrain to provide an accurate estimation of power consumption. The model achieves a root mean square error (RMSE) of only 4.99 kW when compared to chassis dynamometer measurements over the world vehicles test cycle (WLTC), where the tested battery power ranges from −40 kW to 50 kW. We kindly direct readers to refer to [27] for further details.

B. Wireless Charging Model

Several dynamic wireless charging systems have been developed in recent years [28], [29], [30]. The power electronics team at Oak Ridge National Laboratory (ORNL) developed the world's first 22 kW dynamic wireless charging system for passenger cars with 90% efficiency [31], which is used in this study. In addition, ORNL aims to develop 50kW dynamic wireless charging system as well [32]. When WCL is widely used, its charging power may be between 22kW and 50kW, so this study will also explore the impact of different charging powers on the proposed model. If a vehicle moves or stops above or within a WCL region, its battery is charged according to (1), which contains charging power P_{chrg}, charging efficiency η_{chrg}, and the duration Δt between two discrete time steps t.

$$E_{Bat}(t+1) = E_{Bat}(t) + P_{chrg} \cdot \eta_{chrg} \cdot \Delta \mathrm{t} \tag{1}$$

C. Quantifying the Optimal Control Problem for Eco-Driving

This algorithm aims to minimize the net energy consumption of the vehicle while guarantee safety and traffic efficiency. Therefore, the multi-objective function could be quantified as follow:

$$\begin{aligned} min_{a_{ego}} J = w_1 \sum_{i=0}^{N} \Delta E_{Bat}\left(V_{ego}(a_{ego}(t))\right) \\ -w_2 \sum_{i=0}^{N} V_{ego}\left(a_{ego}(t)\right)^2 \\ \text{s.t.} \quad a_{min} < a_{ego}(t) < a_{max} \\ a_{ego}(t) < a_{safe}(t) \\ V_{min}(t) < V_{ego}(t) < min(V_{max}(t), V_{limit}) \end{aligned} \tag{2}$$

where ΔE_{Bat} is the net energy consumption of battery pack in each time step, a_{ego} is the acceleration of ego vehicle and V_{ego} is the speed of ego vehicle. Meanwhile, w_1 and w_2 are weighting coefficient. In addition, $a_{ego}(t)$ must be within the range of maximum acceleration and deceleration, and it should be lower than $a_{safe}(t)$ to guarantee safety, where $a_{safe}(t)$ is the safe acceleration generated by rule-based car-following model, which will be explained in Section III. Meanwhile, to allow for the ego vehicle to smoothly pass through the signalized intersections during green phase, $V_{ego}(t)$ should be within the velocity range $[V_{min}(t), V_{max}(t)]$ calculated by SPaT information. The eco-driving aims to minimize the multi-objectives function $f^* = min(J)$. Based on the established multi-objectives function of the eco-driving problem, the proposed DRL based strategy will be discussed in the next step.

III. Eco-Driving Control Considering Partial WCL at Signalized Intersections

For the purpose of solving the multi-objectives OCP of eco-driving control at signalized intersections with partial WCL, DRL is used to plan the controlled vehicle's velocity profiles. Fig. 2 depicts the proposed scheme, in which a hierarchical framework is implemented. In this study, the ego vehicle is assumed to be communicated with signalized intersections via V2I communication. Furthermore, if the leading vehicle is also a CV, the ego vehicle can establish V2V communication with it. If not, the ego vehicle can still perceive the state of the leading vehicle through onboard sensors such as cameras and LiDAR. Consequently, SPaT information, the state of

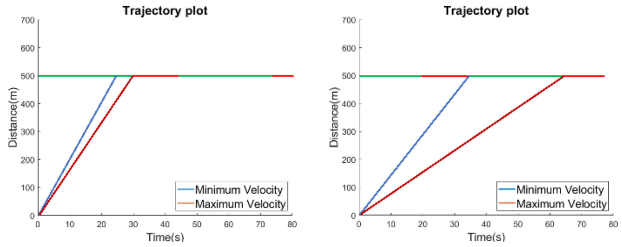


Fig. 3. (a) The vehicle can pass the current green phase at speed limit. (b) The vehicle cannot pass the current green phase at speed limit.

the preceding vehicle, and the state of the ego vehicle itself can all be acquired in real time. Given the SPaT information and ego vehicle state, a dynamic velocity range model is devised to generate reference speed to direct an CAEV fully utilize WCL without impeding traffic. Moreover, a Twin Delayed Deep Deterministic policy gradient (TD3) agent was developed to derive control decisions for acceleration at multiple intersections with partial WCL by taken SPaT information, ego vehicle state and leading vehicle state as input.

A. Dynamic Velocity Range Model

For general eco-driving, e.g., eco-driving that does not take WCL into account, the purpose of calculating the velocity range is to enable the vehicle to pass signalized intersections without halting while maintaining traffic efficiency.

If the vehicle can pass through the green phase at the speed limit when the phase is green, as shown in Fig. 3 (a), the maximum speed should equal to the speed upper limit, and the minimum speed should exactly allow the vehicle to pass by the end of the green phase. If the vehicle cannot pass through the green phase at the speed limit, as shown in Fig. 3 (b), the maximum speed should exactly allow it to pass at the beginning of the next green phase, and the minimum speed should allow it to pass at the end of the next green phase. Therefore, the reference velocity range when phase is green could be formulated as Eq. (3), shown at the bottom of the page, where V_{limit} is the road speed limit, $D_{traffic\ light}$ is the distance to traffic light, t_{green_cycle} is the duration of green phase, t_{red_cycle} is the duration of red phase, t_{green_remain} is the remaining green time, t_{red_remain} is the remaining red time and $k = 1, 2, \ldots, \infty$ is the delay cycle. $k = 0$ if the vehicle can pass the current green phase, otherwise, the appropriate green phase could be chosen by adjusting k.

Similarly, at a red phase, the reference velocity range can be defined as follows:

$$V_{max} = min\Big\{V_{limit}, D_{traffic\ light} \div \big[t_{red\_remain} + k\big(t_{green\_cycle} + t_{red\_cycle}\big)\big]\Big\}$$

$$V_{min} = D_{traffic\ light} \div \big[t_{red\_remain} + k\big(t_{green\_cycle} + t_{red\_cycle}\big)\big] \tag{4}$$

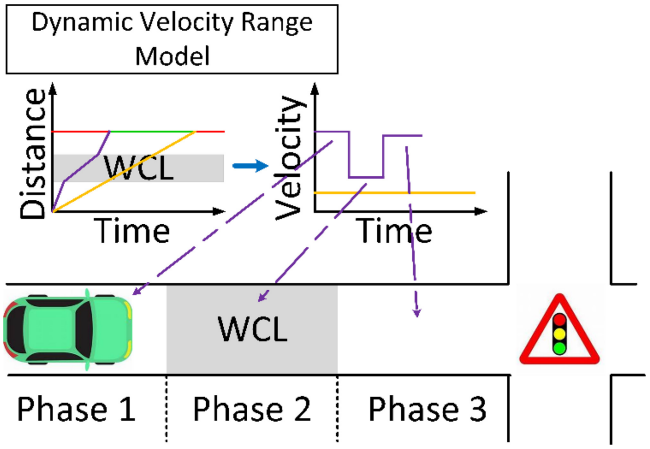


Fig. 4. Three phases at signalised intersection with partial WCL.

However, for eco-driving that takes partial WCL into account, the purpose of calculating the reference velocity range is not only to enable the vehicle to pass the traffic light efficiently while maximizing traffic efficiency, but also to maximize the vehicle's utilization of WCL. There are 2 cases for eco-driving accounts for partial WCL. 1) When $V_{max} = V_{limit}$, to maintain traffic efficiency the reference velocity range should follow the above mentioned one's. 2) When $V_{max} < V_{limit}$, it means that there is extra time for a vehicle to spend within the WCL. In this case, as shown in Fig. 4, the maximum speed of reference velocity range V_{max_w} could be divided into 3 phases, e.g., $V_{max\_w} = [V_{max\_1}, V_{max\_2}, V_{max\_3}]$. For phase 1 and phase 3, the vehicle is within non-WCL region, it should be drive as fast as possible to maximize the time it spends within WCL region, therefore, V_{max_1} and V_{max_3} could be formulated as follow:

$$V_{max\_1} = min\{1.5 * V_{max}, V_{limit}\}$$
$$V_{max\_3} = V_{max} \tag{5}$$

There is a certain speed difference between WCL area and non-WCL area. To avoid the discomfort caused by excessive jerk to passengers, V_{max_1} is not directly set to V_{limit}, but is set to 1.5 times V_{max}.

To calculate V_{max_2}, it is necessary to estimate the remaining time the vehicle travels in the WCL area T_{WCL} first, which is equal to the difference between the time the vehicle travels to the signalized intersection and the time the vehicle travels in the non-WCL area. To ensure traffic efficiency, the former term is set as the distance from the vehicle to the signalized

$$V_{max} = \begin{cases} V_{limit}, & k = 0 \\ min\begin{cases} V_{limit}, \\ D_{traffic\ light} = t_{green\_remain} + (k-1)\big(t_{green\_cycle} + t_{red\_cycle}\big) + t_{red\_cycle} \end{cases}, & k = 1, 2, \ldots, \infty \end{cases}$$
$$V_{min} = D_{traffic\ light} \div \big[t_{green\_remain} + k\big(t_{green\_cycle} + t_{red\_cycle}\big)\big] \tag{3}$$

intersection divided by the V_{max}. To allow the charging time as long as possible, it is therefore assumed that the vehicle is traveling at the V_{limit} within the non-WCL area. Therefore, T_{WCL} can be formulated as follow:

$$T_{WCL} = \frac{D_{traffic\ light}}{V_{max}} - \frac{D_{traffic\ light} - D_{WCL\_remain}}{V_{limit}} \tag{6}$$

where D_{WCL_remain} is the remaining length of WCL, therefore, V_{max_2} could be computed as follow:

$$V_{max\_2} = \frac{D_{WCL\_remain}}{T_{WCL}} \tag{7}$$

Overall, the reference speed range $[V_{max\_w}, V_{min\_w}]$ for a vehicle in partial WCL scenario could be summarised as:

$$V_{max\_w} = \begin{cases} V_{max}, & I_1 = 0 \\ V_{max\_1}, & I_1 = 1 \ and \ I_2 = 1 \\ V_{max\_2}, & I_1 = 1 \ and \ I_3 = 1 \\ V_{max\_3}, & I_1 = 1 \ and \ I_4 = 1 \end{cases} \tag{8}$$

$$V_{min\_w} = V_{min} \tag{9}$$

where I_1 is the indicator to indicate if the vehicle can pass the signalised intersection if it is drive in speed limit and I_2, I_3 and I_4 are the indicators to indicate if the vehicle is within phase 1 or phase 2 or phase 3 respectively.

B. Implementation of TD3 for Eco-Driving Control

Reinforcement learning techniques often use a simplification of the real-world environment known as the Markov decision process (MDP) [33]. In this framework, the probabilities of transitioning to distinct future states are only determined by the present state, s. The determination of action b may be achieved by the use of a policy π, which represents a probability distribution encompassing the many actions available at the present state. The state transition probability after the execution of an action b_t in state s_t can be formally defined as:

$$p(s_{t+1}|s_t, b_t) = \ Pr\{s\_\{t+1\} = s'|s\_t = s, b\_t = b\} \tag{10}$$

where the variable s' represents the subsequent state. The current state and action are denoted by the variables s and a, respectively. When the policy π is deterministic, the action-state value (also known as the Q-value) can be inferred as the expected cumulative reward:

$$Q^{\pi}(s_t, b_t) = E_{r_{i\geq t}, s_{i>t} \sim E}\left[\sum_{i=t}^{T} \gamma^{i-t} r_i(s_i, b_i)\right] \tag{11}$$

where the discounting factor $\gamma \in [0, 1]$. The variable T represents the duration of the finite MDP. The notation $r_i(s_i, b_i)$ denotes the reward obtained after executing action b_i in state s_i. The set E represents a collection of states. The MDP can be represented by a recursion relation, which can be expressed as follows:

$$Q^{\pi}(s_t, b_t) = r(s_t, b_t) + \gamma E_{r_t, s_{t+1} \sim E}\left[Q^{\pi}(s_{t+1}, b_{t+1})\right] \tag{12}$$

Equation (15) is often referred to as the Bellman equation [34]. The objective of the reinforcement learning algorithm is to identify a policy π^* that can maximise the expected cumulative reward. The TD3 method is a cutting-edge DRL technique that is capable of functioning inside continuous action spaces and mitigating the overestimation bias associated with action value estimation [35]. Hence, the TD3 algorithm is chosen as the solution for addressing the eco-driving control problem for ACEVs at signalized intersections with partial WCL.

C. Transforming the Multi-Objectives Function to Reward Function

It is hard to directly set net energy consumption as reward function because the vehicle controlled by the agent tends to stop within WCL area and causes traffic congestion. Instead, the reward function of velocity range $r_{velocity}$ is set to allow the vehicle to maximise the time spent in WCL by following the speed produced by dynamic velocity range model. Meanwhile, the reward function of acceleration $r_{acceleration}$ is set to reduce the vehicle's acceleration in each time step so that minimizing the gross energy consumption indirectly. The second term of the multi-objectives function and the constraint of safety are separately converted into reward function of traffic efficiency $r_{efficiency}$ and reward function of safety r_{safety}. Therefore, the multi-objectives function is transformed to reward function directly as instantaneous cost as follow:

$$\begin{aligned} r(s_t, b_t) = c_1 r_{acceleration} + c_2 r_{velocity} \\ + c_3 r_{safety} + c_4 r_{efficiency} \end{aligned} \tag{13}$$

where c_1, c_2, c_3 and c_4 are weighting factors and $r_{acceleration}$ is defined as:

$$r_{acceleration} = -{a_{TD3}}^2 \tag{14}$$

To guarantee the vehicle to use the WCL efficiently and time-adaptively while avoid unnecessary stopping at signalized intersections, it is important to encourage it drives within reference velocity range, therefore, $r_{velocity}$ is formulated as:

$$r_{velocity} = \begin{cases} -\left(V_{ego} - V_{max\_w}\right)^2, & V_{ego} > V_{max\_w} \\ -\left(V_{ego} - V_{min\_w}\right)^2, & V_{ego} < V_{min\_w} \\ 0, & V_{min\_w} < V_{ego} < V_{max\_w} \end{cases} \tag{15}$$

It is necessary to ensure that the acceleration generated by TD3 agent is secure in order to assure security. Intelligent Driving Model (IDM) is a rule-based car-following algorithm that simulates human drivers. The acceleration it produces is strictly secure, which is a nonlinear differential equation to describe vehicular acceleration:

$$a_{IDM} = a^{(a)}\left[1 - \left(\frac{v_a}{v_0^{(a)}}\right)^{\delta} - \left(\frac{s^*(v_a, \Delta v_a)}{s_a}\right)^2\right]$$

$$\text{with } s^*(v_a, \Delta v_a) = s_0^{(a)} + T^a v_a + \frac{v_a \Delta v_a}{2\sqrt{a^{(a)} b^{(a)}}} \tag{16}$$

where $a^{(a)}$ and $v_0^{(a)}$ represent the maximum possible acceleration and speed; The acceleration exponent is δ; The desired distance and minimum safety distance are represented by $s^*(v_a, \Delta v_a)$ and $s_0^{(a)}$; The actual gap and speed variation are s_a and Δv_a, respectively; The reaction time is denoted by T^a, and the desired deceleration is given by $b^{(a)}$. In particular, the

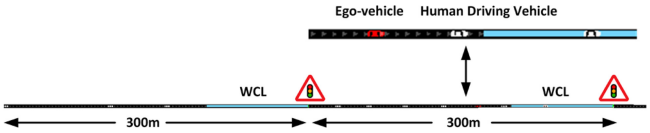


Fig. 5. The illustration of simulation traffic scenario in SUMO.

minimum safety distance refers to the baseline distance that should be maintained even at zero speed, while the reaction time accounts for the period it takes for the driver or the automated driving system to react to the behavior changes of the preceding vehicle. During this time, the vehicle continues to travel at its current speed, which necessitates considering the distance traveled during this reaction period in the safety distance calculation. By incorporating minimum safe distance and reaction time, the IDM ensures safe acceleration. Therefore, setting the TD3-based eco-driving algorithm's acceleration lower than the IDM's enhances safety, providing a greater safety buffer. Therefore, r_{safety} is given as:

$$r_{safety} = \begin{cases} -(a_{TD3} - a_{IDM})^2, & a_{TD3} > a_{IDM} \\ 0, & a_{TD3} < a_{IDM} \end{cases} \tag{17}$$

It is worth to mention that, setting the TD3 algorithm's acceleration to be lower than that of the IDM may seem to limit the optimization space, but it is crucial for maintaining driving safety in different conditions. This restriction helps prevent the vehicle from accelerating beyond safe limits, thus avoiding potential collision risks and ensuring driving stability. While this might slightly reduce the room for optimization, it provides a reliable safety baseline, allowing the TD3 algorithm to optimize in a safe and robust manner. Generally, the speed of the vehicle controlled by the eco-driving algorithm is lower than that of the leading vehicle to ensure a smooth passage through the signalized intersection. In this scenario, the actual gap between the ego-vehicle and the leading vehicle will be relatively large, allowing ample optimization space. Conversely, when the ego-vehicle approaches the leading vehicle at signalized intersections, its optimization space becomes limited to ensure safety.

Although $r_{velocity}$ ensures that vehicles pass signalized intersections without incident, vehicles may choose to pass signalized intersections when the green phase ends to save energy, which can easily lead to congestion. To maximize traffic efficiency, vehicles should travel at the maximum speed of the reference velocity range as possible. Thus, $r_{efficiency}$ can be described as follows:

$$r_{efficiency} = {V_{ego}}^2 \tag{18}$$

D. TD3 Agent State and Action Space Selection

The selection of state should include the situation of ego-vehicle and traffic that affects the decision of agent. Therefore, the state variables of TD3 agent includes:

$$S_t = \begin{bmatrix} V_{ego}, a_{ego}, D_{leader}, V_{leader}, a_{leader}, D_{traffic\ light}, \\ V_{max\_w}, V_{min\_w}, I_1, I_2, I_3, I_4 \end{bmatrix} \tag{19}$$

In this paper, ego-vehicle's longitudinal acceleration a_{ego} is selected as action variable.

IV. Simulation Setup

A. Simulation Environment and Benchmarks

The simulation environment used for both training and case studies was implemented in the SUMO [36]. The simulation platform includes one lane, ego-vehicle, leading vehicles, 2 WCLs and 2 traffic lights, as shown in Fig. 5. The length of the road is set to 600 meters, with two traffic lights positioned at 300 meters and 600 meters, respectively. Each signal phase has a duration of 30 seconds for green and 15 seconds for red. In this study, the speed limit of the road is 20m/s. All surrounding vehicles in the case studies were modelled by IDM to simulate human driving vehicles without intelligent control.

To rigorously assess the efficiency and time-adaptiveness of the proposed method for utilizing WCL, as well as its performance in terms of energy saving, safety, and real-time capability, three case studies were conducted alongside three benchmark comparisons. The three case studies will be presented in the following subsections, and the three benchmarks include: 1) An IDM was employed to simulate human driving behaviors at signalized intersections with partial WCL, 2) Li's DRL-based model adapted from [28], was used to simulate standard eco-driving method without focusing on maximizing WCL utilization, 3) Zhang's optimization-based model from [30], which aims to efficiently utilize the WCL. The basic parameters of the IDM that apply to both benchmark and surrounding vehicles are as follows: The maximum possible speed is 20 m/s. The maximum possible acceleration is 3 m/s^2. The constant time headway is 3 seconds, referring to the consistent time gap maintained from the vehicle ahead for safety. The acceleration exponent is set at 4, a unitless factor that contributes to the calculation of the vehicle's acceleration. The desired deceleration is 1.6 m/s^2, representing the preferred rate at which the vehicle slows down. Lastly, the minimum distance kept from the vehicle in front is 3 meters, ensuring an adequate space cushion to react and stop if necessary. These parameters collectively define the behavior of vehicles within the IDM framework.

The state, action, and reward function of Li's model is basically similar as the proposed model. However, because this model does not consider the effect of WCL, the state of this model is constructed as follow:

$$S_{t_{benchmark}} = \begin{bmatrix} V_{ego}, a_{ego}, D_{leader}, V_{leader}, a_{leader}, D_{traffic\ light}, \\ V_{max}, V_{min} \end{bmatrix} \tag{20}$$

where V_{max} and V_{min} are the maximum and minimum speed computed from normal eco-driving reference range. At the same time, the decision action of Li's model is longitudinal acceleration as well. In terms of the reward function, it is the same as that of proposed model. Apart from that, the weight of reward function and the construction of neural network of the Li's model are same as that of the proposed model.

In terms of Zhang's model, it aims to minimize the net energy consumption while ensure traffic efficiency, thus its objective function is defined as follow:

$$\min_{a(t)} L = \alpha_1 (t_f - t_0) + \alpha_2 (E_{NC} - E_{Charge}) \tag{21}$$

TABLE II
THE BASIC PARAMETERS OF THE EGO-VEHICLE

Parameter	Value
Actor network layer	4
Critic network layer	4
Hidden layer width	48
Actor network learning rate	0.0001
Critic network learning rate	0.005
Hidden layer activation function	ReLU [1]
Output layer activation function	Tanh [2]
Batch size	512
Discount factor	0.999
Gaussian exploration noise	0.1
Target network update rate	0.005
Range to clip target policy noise	0.5
Critic update noise	0.5
Frequency of delayed policy updates	2

where t_f and t_0 are the initial time and the finish time respectively, while E_{NC} and E_{Charge} are the total consumed and charged energy respectively. α_1 is the weight coefficient of time and α_2 is the electricity price. More information about these parameters and the constraints of the above objective function can be found in [31]. The basic parameters of the ego-vehicle are as follows: The mass is 1830 kg, the front area is 2.6 m^2, and the air drag coefficient is 0.35. The rolling resistance coefficient is listed as 0.01, with an internal moment of inertia at 0.01 as well. The propulsion efficiency reaches 0.98, while the recuperation efficiency is slightly lower at 0.96. Finally, the maximum battery capacity is indicated as 64 kWh.

B. DRL Parameter Prescription and Training

The principal DRL algorithm parameters are presented in Table II. The actor network and the critic network, each consisting of four hidden layers, are respectively designed to be constructed with one deep network. The comprehensive network is trained on an actor-critic structure, with the reward of each episode calculated using the (13). To train the TD3 agent, a route with two traffic lights is chosen, and the total training episode is set to 245, while the simulation sample duration is set to 0.1 second.

Three proposed models were trained separately with WCL located at 200-300m/500-600m (region A), 100-200m/400-500m (region B) and 0-100m/300-400m (region C). This is because the model was hard to converged if the training environment included three locations of WCL at the same time. Solving this problem and training a general model will be the future work. For each training environment, several system parameters, such as the initial state of the ego-vehicle, leading vehicles and SPaT information, are randomized for each episode to train the agent to adapt for various driving scenarios. Meanwhile, the DRL-agent of Li's model is trained in the same environment without WCL.

The training outcome is depicted in Fig. 6. Overall, the converged accumulated reward of Li's model is higher than that of the proposed models. This is because the reward

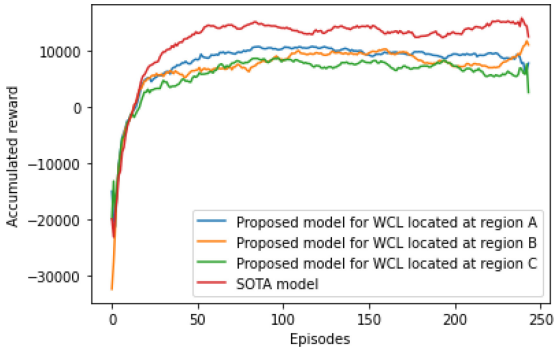


Fig. 6. Convergence of training of proposed model and SOTA model.

function is highly related to the speed of the vehicle in each training time step, while the reference velocity range models followed by proposed model and Li's model are different. The average accumulative reward of three proposed models rises abruptly at the start of training and stabilizes primarily after approximately 90 training episodes, however, the convergence of training of Li's model is relatively earlier, at around 60 episodes because the state space of Li's model is simpler than that of proposed model. It is important to note that each episode's cumulative reward differs marginally after convergence due to the sporadic nature of initial conditions, resulting in various reference speeds. The cumulative reward varies depending on the speed followed by the ego-vehicle. With such conditions considered, four DRL-based controllers for online execution were realized, in which the controllers can derive acceleration decisions directly based on the system's state.

V. CASE STUDIES AND DISCUSSION

A. Analysis of Energy Savings and Traffic Efficiency of the Proposed Method in Zero Traffic Scenario

To evaluate the energy savings and traffic efficiency of the proposed method, a 600-meter road with two signalized intersections and two WCLs located between 100-200 meters and 400-500 meters was selected. The initial timings for the traffic lights were set to green for 21 seconds and red for 7 seconds, respectively. Under these settings, a vehicle adhering to the speed limit would stop at the first intersection but pass through the second intersection without stopping. In this scenario, traffic was set to zero because Zhang's model does not handle car-following scenarios effectively. Key metrics such as net energy consumption, gross energy consumption, charging time, travel time, and algorithm execution time were recorded.

Fig. 7 (a) and (b) show that the vehicle controlled by the IDM model stopped at the first signalized intersection and passed through the second without stopping, due to its strict adherence to the speed limit. The vehicle controlled by Li's model smoothly passed both intersections without stopping, although it did not optimize for WCL usage and did not spend additional time in the WCL area.

In contrast, Zhang's model allowed the vehicle to accelerate in non-WCL areas and decelerate in the WCL area,

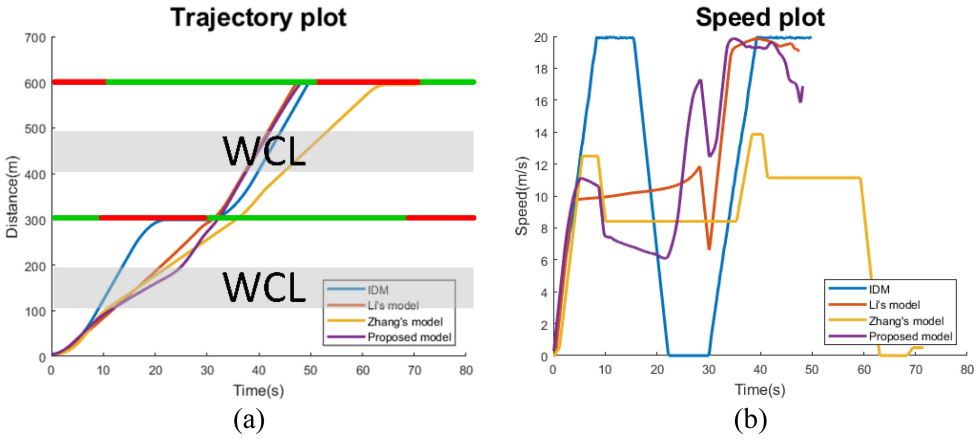


Fig. 7. (a) Distance profile of 4 models and (b) speed profile of 4 models.

TABLE III
SIMULATION RESULTS UNDER DIFFERENT APPROACHES IN THE ZERO TRAFFIC SCENARIO

Model	Net energy consumption (First signalized intersection) (kWh)	Gross energy consumption (First signalized intersection) (kWh)	Charging time (First signalized intersection) (s)	Travel time (Total/First signalized intersection) (s)	Algorithm execution time (s)
IDM	$74.63*10^{-3}$	$101.32*10^{-3}$	4.99	49.9/32.0	0.0001
Li's model	$9.85*10^{-3}$	$61.65*10^{-3}$	9.69	47.5/31.5	0.0023
Zhang's model	$-9.87* 10^{-3}$[1]	$63.75*10^{-3}$	12.89	59.03/35.4	8.7348
Proposed model	$-5.64* 10^{-3}$[1]	$69.75*10^{-3}$	14.09	48.3/31.6	0.0028

[1] Negative energy consumption means that the charging electricity for the vehicle is greater than the consuming electricity

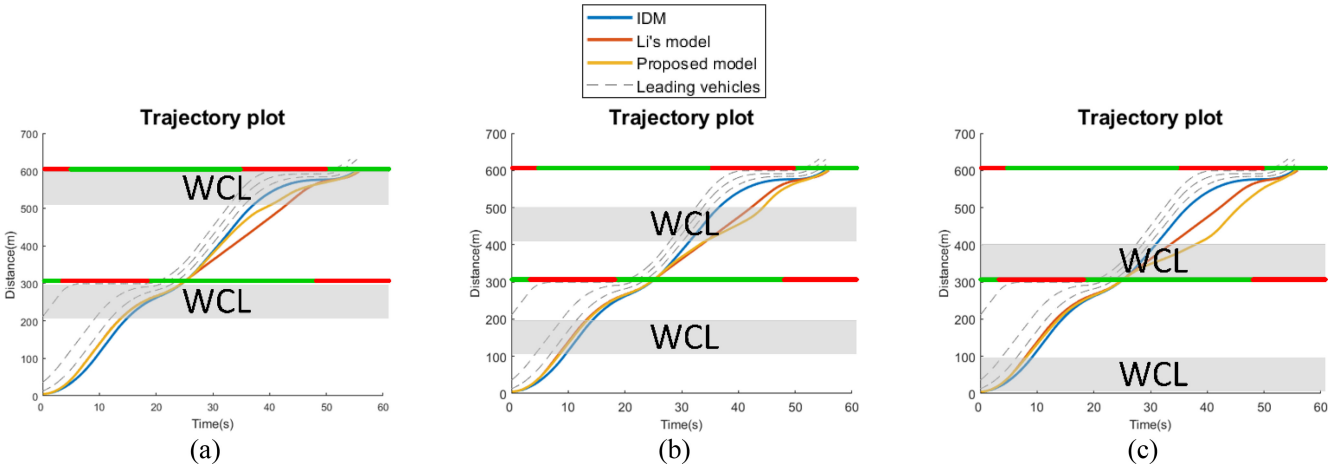


Fig. 8. (a) Distance profile of 3 models when WCL is located at 200-300m and 500-600m, (b) 100-200m and 400-500m, (c) 0-100m and 300-400m.

passing the first intersection right after the light turned green. However, it slowed down in the WCL area near the second intersection and stopped before the light turned red, indicating a lack of time-adaptiveness despite efficient WCL usage.

The proposed method showed similar behavior to Zhang's model at the first intersection but optimized for traffic efficiency at the second intersection by passing the light before it turned red, thanks to the dynamic velocity model, which adjusts the reference speed to ensure efficient WCL usage only when there is extra time.

Quantitative results in Table III show that the travel times at the first signalized intersection for all four methods are around 32 seconds, but the vehicle controlled by Zhang's method had an overall travel time about 10 seconds longer than the others. Therefore, the energy and charging data presented focus on net energy consumption, gross energy consumption, and charging time at the first intersection to ensure meaningful comparison. The proposed method achieved net energy consumption similar to Zhang's model, but with slightly higher gross energy consumption compared to Li's model. Additionally, the proposed model's execution time was

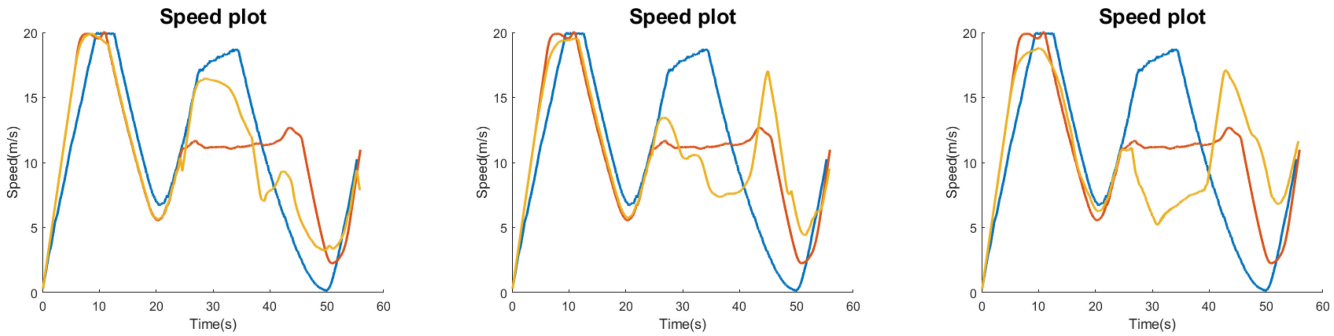


Fig. 9. (a) Speed profile of 3 models when WCL is located at 200-300m and 500-600m, (b) 100-200m and 400-500m, (c) 0-100m and 300-400m.

significantly faster, at 0.0028s (The simulation was conducted on a PC equipped with an 11th Gen Intel Core i7-11700 processor), highlighting its real-time capabilities and efficient, time-adaptive use of WCLs, while maintaining traffic efficiency.

B. Analysis of Energy Savings, Traffic Efficiency, and Safety of the Proposed Method in Scenarios With Traffic

To assess the safety of the proposed model, in addition to its energy-saving capabilities and traffic efficiency, three driving scenarios with two signalized intersections and a 600-vehicle-per-hour traffic volume were constructed to simulate and compare various eco-driving approaches. Meanwhile, the respective positions of WCL are 200-300m/500-600m, 100-200m/400-500m, and 0-100m/300-400m. In this case study, the initial timings of the first and second traffic light were set as green at 8s and red at 7s respectively. Under this timing configuration, a vehicle following the speed limit will be able to pass the first signalized intersection but will stop at the second intersection if it adheres to the speed limits. Notably, Zhang's method was excluded from this case study due to its inability to handle scenarios involving a leading vehicle.

It can be observed from Fig. 8 that the vehicle controlled by the IDM model adhered to the speed limit, passing through the first signalized intersection but stopping at the second one. In contrast, the vehicle controlled by Li's method smoothly passed both traffic lights without any unnecessary idling. It is noteworthy that since neither of these methods takes the WCL into account, their distance profiles remain the same regardless of the WCL location. This consistency is also evident in their speed profiles, as shown in Fig. 9. In contrast, the vehicle controlled by the proposed method smoothly passed the first signalized intersection but slowed down in the WCL area before proceeding through the second signalized intersection. Notably, the dynamic velocity range model enabled the proposed method to dynamically adjust the vehicle's speed according to the different locations of the WCLs, as evidenced in Fig. 9.

It can also be observed from both Figs. 8 and 9 that, unlike in Case Study A, in this case study involving leading vehicles, vehicles controlled by the three methods slowed down when approaching leading vehicles to maintain a minimum safety distance and avoid collisions. The IDM model achieved this by incorporating the difference between the actual gap and the minimum gap into its differential equation. In contrast, the proposed method and Li's method accomplished this by including the acceleration generated by the IDM model in their reward functions.

TABLE IV
AVERAGE SIMULATION RESULTS UNDER DIFFERENT APPROACHES IN THREE WCL LOCATIONS

Model	Net energy consumption (kWh)	Gross energy consumption (kWh)	Charging time (s)	Travel time (s)
IDM	$48.29*10^{-3}$	$144.54*10^{-3}$	17.4	56.4
Li's model	$24.35*10^{-3}$ (-49.57%)	**$123.35*10^{-3}$(-14.46%)**	17.92	57
Proposed model	**$8.61*10^{-3}$(-82.17%)**	$127.04*10^{-3}$(-12.10%)	21.39	56.86

The average simulation results for different approaches in WCL locations are presented in Table III. The travel times for vehicles controlled by the three models are similar, at around 57 seconds. However, the charging time for the vehicle under the proposed method is approximately 3.5 seconds longer than those of the two benchmark models. The proposed method's net energy consumption is 82.17% lower than that of the IDM model and significantly lower than that of Li's model. Regarding gross energy consumption, Li's model performed the best, with a 14.46% reduction compared to the IDM model, while the proposed method was close, showing a 12.10% reduction. The extra acceleration when entering and leaving the WCL area contributed to the slightly higher gross energy consumption for the proposed model. Overall, the study concludes that the proposed model can substantially reduce net energy consumption in WCL areas while maintaining traffic efficiency and safety, despite a modest increase in gross energy consumption.

C. Analysis of Adaptability of the Proposed Method

To assess the proposed eco-driving method's flexibility, three simulation sets based on varying traffic volumes (300, 600, and 1200 vehicles per hour) and WCL positions (200-300m/500-600m, 100-200m/400-500m, and 0-100m/300-400m) were conducted, each comprising 36 runs with different traffic signal patterns. The comprehensive simulations evaluated net and gross energy consumption, charging and travel times, as detailed in Table IV.

TABLE V
AVERAGE SIMULATION RESULTS OF DIFFERENT APPROACHES AND TRAFFIC VOLUME IN THREE WCL LOCATIONS

Model	Net energy consumption (kWh)	Gross energy consumption (kWh)	Charging time (s)	Travel time (s)	Traffic volume (vehicles/hour)
IDM	$108.24*10^{-3}$	$193.01*10^{-3}$	15.38	50.08	1200
Li's model	$68.92*10^{-3}$ (-36.32%)	$156.34*10^{-3}$(-18.99%)	15.83	50.56	1200
Proposed model	$60.60*10^{-3}$ (-44.01%)	$159.82*10^{-3}$(-17.19%)	17.98	50.65	1200
IDM	$108.66*10^{-3}$	$187.46*10^{-3}$	14.20	46.61	600
Li's model	$76.66*10^{-3}$ (-29.44%)	$155.44*10^{-3}$(-17.08%)	14.19	45.79	600
Proposed model	$71.50*10^{-3}$ (-34.19%)	$160.67*10^{-3}$(-14.29%)	16.16	45.93	600
IDM	$113.95*10^{-3}$	$189.83*10^{-3}$	13.78	45.30	300
Li's model	$81.05*10^{-3}$ (-28.87%)	$155.92*10^{-3}$(-17.86%)	13.55	43.87	300
Proposed model	$76.95*10^{-3}$ (-32.47%)	$161.84*10^{-3}$(-14.74%)	15.38	44.22	300

The simulations revealed no collisions across scenarios. Higher traffic volumes resulted in longer travel and charging times. The IDM vehicle's gross energy consumption rose with traffic density due to its speed-oriented programming, which led to increased energy expenditure, especially at 1200 vehicles/hour. Despite more time in the WCL at high volumes, the IDM's net energy use remained high. The Li's model prioritized traffic efficiency and minimized gross energy use, outperforming the IDM in energy conservation but not optimizing WCL use, resulting in higher net energy consumption than the proposed method. The proposed model balanced charging time and travel efficiency, achieving the lowest net energy use despite slightly higher gross consumption due to acceleration changes. This was most notable at 1200 vehicles/hour, where the proposed model's net energy use was 44.01% less than the IDM's, confirming its superior WCL utilization and efficiency across varying traffic volumes.

VI. CONCLUSION

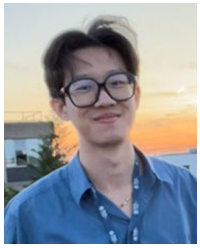
This article proposes a DRL-based eco-driving strategy for CAEVs in a connected traffic environment featuring signalized intersections and a partial WCL. This innovative method addresses the challenge of simultaneously ensuring efficient and time-adaptive use of the WCL, maintaining traffic efficiency and safety, and achieving real-time algorithm performance. This method accomplishes this by integrating a novel dynamic velocity range model with a TD3 controller. The dynamic velocity range model dynamically computes reference speeds for the TD3 agent based on the varying locations of the WCL and different SPaT information. Meanwhile, the TD3 agent continuously generates acceleration by considering the reference speed and safety information. Based on the simulation results, the proposed strategy exhibits a notably lower net energy consumption in comparison to both the manually driven vehicle and the benchmarks models at the premise of ensuring traffic efficiency and safety. Specifically, the energy consumption of the proposed model could reduce up to 44.01% compared with the manually driven vehicle when traffic volume reaches 1200. With three different traffic volume, the proposed method's gross energy efficiency still holds compared to the other two benchmark models.

For future work, obtaining a more general model, improving the reward function and state space of the model so that the training of environments with different WCL positions can be completed at one time will be a meaningful focus. Furthermore, WCL is a relatively new concept, and although many research institutions are actively studying this technology, the number of roads globally equipped with WCL is still quite limited. As a result, we are currently unable to access real-world data regarding WCL lengths. Additionally, since traffic light timings are controlled by local government agencies, relevant data is also unavailable at this stage. Nevertheless, we are keen to obtain real-world data on the length ratio between the road and WCL, and traffic light settings in the future, in order to further enhance our approach.

REFERENCES

[1] L. Pan, E. Yao, Y. Yang, and R. Zhang, "A location model for electric vehicle (EV) public charging stations based on drivers' existing activities," *Sustain. Cities Soc.*, vol. 59, Aug. 2020, Art. no. 102192.
[2] M. Fuller, "Wireless charging in California: Range, recharge, and vehicle electrification," *Transp. Res. Part C, Emerg. Technol.*, vol. 67, pp. 343–356, Jun. 2016.
[3] "Average electric car range in the U.K. 2023." 2023. [Online]. Available: https://www.nimblefins.co.uk/average-electric-car-range#nogo
[4] X.-G. Yang, T. Liu, and C.-Y. Wang, "Thermally modulated lithium iron phosphate batteries for mass-market electric vehicles," *Nat. Energy*, vol. 6, no. 2, pp. 176–185, 2021.

[5] H. R. Sayarshad and V. Mahmoodian, "An intelligent method for dynamic distribution of electric taxi batteries between charging and swapping stations," *Sustain. Cities Soc.*, vol. 65, Feb. 2021, Art. no. 102605.
[6] D. Jaraniya and S. Kumar, "Power quality signal conditioning and mitigation by using LMSEXF based control strategy for grid-tied PV assisted EV charging station," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 564–573, Feb. 2024.
[7] "Fuel prices: Could I save money driving an electric car?" 2022. [Online]. Available: https://www.bbc.co.uk/news/ explainers-60943322
[8] "Q-series: Tearing down the heart of an electric car lap 2: Cost parity a closer reality?" 2022. [Online]. Available: https://www.ubs.com/global/en/investment-bank/in-focus/2020/heart-of-electric-car.html
[9] A. Y. S. Lam, Y.-W. Leung, and X. Chu, "Electric vehicle charging station placement: Formulation, complexity, and solutions," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2846–2856, Nov. 2014.
[10] D. M. Vilathgamuwa and J. P. K. Sampath, "Wireless power transfer (WPT) for electric vehicles (EVS)—Present and future trends," in *Plug in Electric Vehicles in Smart Grids: Integration Techniques*, S. Rajakaruna, A. Garcia-Cerrada, and A. Ghosh, Eds., Singapore: Springer, 2015, pp. 33–60.
[11] W. Li, H. Ding, N. Xu, and J. Zhang, "Toward carbon–neutral transportation electrification: A comprehensive and systematic review of eco-driving for electric vehicles," *IEEE Trans. Transport. Electrific.*, vol. 10, no. 3, pp. 6340–6360, Sep. 2024.
[12] J. Georgy, A. Noureldin, and C. Goodall, "Vehicle navigator using a mixture particle filter for inertial sensors/odometer/map data/gps integration," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 544–552, May 2012.
[13] M. A. Saleem et al., "Provably secure conditional-privacy access control protocol for intelligent customers-centric communication in VANET," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 1747–1756, Feb. 2024.
[14] K. Liu, X. Xu, P. Dai, and B. Chen, "Cooperative sensing and uploading for quality-cost tradeoff of digital twins in VEC," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 3614–3625, Feb. 2024.
[15] H. Hoang, S. Lee, Y. Kim, Y. Choi, and F. Bien, "An adaptive technique to improve wireless power transfer for consumer electronics," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 327–332, May 2012.
[16] G. Duarte, A. Silva, and P. Baptista, "Assessment of wireless charging impacts based on real-world driving patterns: Case study in Lisbon, Portugal," *Sustain. Cities Soc.*, vol. 71, Aug. 2021, Art. no. 102952.
[17] S. Jeong, Y. J. Jang, and D. Kum, "Economic analysis of the dynamic charging electric vehicle," *IEEE Trans. Power Electron.*, vol. 30, no. 11, pp. 6368–6377, Nov. 2015.
[18] "A demo of wireless electricity." 2009. [Online]. Available: https://www.ted.com/talks/eric_giler_a_demo_of_wireless_electricity
[19] W. C. Brown, "The history of power transmission by radio waves," *IEEE Trans. Microw. Theory Techn.*, vol. 32, no. 9, pp. 1230–1242, Sep. 1984.
[20] S. Mohrehkesh and T. Nadeem, "Toward a wireless charging for battery electric vehicles at traffic intersections," in *Proc. 14th Int. IEEE Conf. Intell. Transport. Syst. (ITSC)*, 2011, pp. 113–118.
[21] S. Mandava, K. Boriboonsomsin, and M. Barth, "Arterial velocity planning based on traffic signal information under light traffic conditions," in *Proc. 12th Int. IEEE Conf. Intell. Transport. Syst.*, 2009, pp. 1–6.
[22] B. Liu et al., "Bi-level convex optimization of eco-driving for connected fuel cell hybrid electric vehicles through signalized intersections," *Energy*, vol. 252, Aug. 2022, Art. no. 123956.
[23] Z. Nie and H. Farzaneh, "Real-time dynamic predictive cruise control for enhancing eco-driving of electric vehicles, considering traffic constraints and signal phase and timing (SPaT) information, using artificial-neural-network-based energy consumption model," *Energy*, vol. 241, Feb. 2022, Art.no. 122888.
[24] J. Li, X. Wu, M. Xu, and Y. Liu, "Deep reinforcement learning and reward shaping based eco-driving control for automated HEVs among signalized intersections," *Energy*, vol. 251, Jul. 2022, Art. no. 123924.
[25] J. Zhang, T.-Q. Tang, Y. Yan, and X. Qu, "Eco-driving control for connected and automated electric vehicles at signalized intersections with wireless charging," *Appl. Energy*, vol. 282, Jan. 2021, Art. no. 116215.
[26] N. Xu et al., "Towards a smarter energy management system for hybrid vehicles: A comprehensive review of control strategies," *Appl. Sci.*, vol. 9, no. 10, p. 2026, 2019.
[27] L. Koch et al., "Accurate physics-based modeling of electric vehicle energy consumption in the SUMO traffic microsimulator," in *Proc. IEEE Int. Intell. Transport. Syst. Conf. (ITSC)*, 2021, pp. 1650–1657.
[28] S. Li and C. C. Mi, "Wireless power transfer for electric vehicle applications," *IEEE J. Emerg. Select. Topics Power Electron.*, vol. 3, no. 1, pp. 4–17, Mar. 2015.
[29] H. Wang and K. W. E. Cheng, "An improved and integrated design of segmented dynamic wireless power transfer for electric vehicles," *Energies*, vol. 14, no. 7, p. 1975, 2021.
[30] G. A. Covic and J. T. Boys, "Inductive power transfer," *Proc. IEEE*, vol. 101, no. 6, pp. 1276–1289, Jun. 2013.
[31] "Ornl surges ahead with 20-kilowatt wireless charging system for vehicles." 2016. [Online]. Available: https://web.ornl.gov/info/reporter/no158/spring_2016.pdf
[32] "Ornl achieves 90% efficiency with 20 kW wireless charging for vehicles, looking ahead to 50 kW." 2016. [Online]. Available: https://www.greencarcongress.com/2016/03/20160331-ornl.html
[33] R. S. Sutton and A. G. Barto, *Reinforcement learning: An Introduction*. Cambridge, MA, USA: MIT, 2018.
[34] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.
[35] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 1587–1596.
[36] P. A. Lopez et al., "Microscopic traffic simulation using SUMO," in *Proc. 21st Int. Conf. Intell. Transport. Syst. (ITSC)*, 2018, pp. 2575–2582.



Xinxing Ren (Student Member, IEEE) received the B.Eng. degree (with First-Class Hons.) in electronic and information engineering from the Brunel University of London, Uxbridge, U.K., in 2022, where he is currently pursuing the Ph.D. degree.

He is a member of the Brunel Interdisciplinary Power Systems Research Centre, Brunel University of London. His research interests include connected vehicles, reinforcement learning, imitation learning, end-to-end autonomous driving, and the integration of large language models into transportation systems.

Chun Sing Lai (Senior Member, IEEE) received the B.Eng. degree (First-Class Hons.) in electronic and electrical engineering from the Brunel University of London, U.K., in 2013, and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 2019.

He is currently a Senior Lecturer with the Department of Electronic and Electrical Engineering and the Course Director of M.Sc. Electric Vehicle Systems with the Brunel University of London. Dr. Lai is an Academic Visitor with the School of Automation, Guangdong University of Technology, China. His research interests are in power system optimization and electric vehicle systems. He is a recipient of the 2022 Meritorious Service Award from the IEEE SMC Society for "meritorious and significant service to IEEE SMC Society technical activities and standards development." He was a Technical Program Co-Chair for 2022 IEEE International Smart Cities Conference. He is the Vice-Chair of the IEEE Smart Cities Publications Committee. He is an Associate Editor of IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, and *IET Energy Conversion and Economics*. He is the Working Group Chair for IEEE P2814 and P3166 Standards, an Associate Vice President of Systems Science and Engineering of the IEEE Systems, Man, and Cybernetics Society (IEEE/SMCS), and the Co-Chair of the IEEE SMC Intelligent Power and Energy Systems Technical Committee. He is an IET Member, a Chartered Engineer, and a Fellow of the Higher Education Academy.

Zekun Guo (Member, IEEE) received the B.Eng. degree in energy and environment system engineering from Shandong University, Shandong, China, in 2018, the M.Sc. by research degree in energy systems from the University of Edinburgh, Edinburgh, U.K., in 2019, and the Ph.D. degree in electronic and electrical engineering from the Brunel University of London, Uxbridge, U.K., in 2023. He is currently a Lecturer and a Postgraduate Research Director of the Data Science, Artificial Intelligence, and Modelling Centre, University of Hull, where he leads research in artificial intelligence applications for engineering. His research interests include power system planning, transport electrification, and the integration of AI techniques in energy systems.

Gareth Taylor (Senior Member, IEEE) is the Director of the Brunel Interdisciplinary Power Systems Research Centre and was the Head of the Department of Electronic and Electrical Engineering from 2019 to 2023 with the Brunel University of London. He was also a Project Coordinator for a 4.2M euro H2020 Energy Project (2017–2020/774500) titled "Coordination of Transmission and Distribution Data Exchanges for Renewables Integration in the European Marketplace Through Advanced, Scalable and Secure ICT Systems and Tools" TDX-ASSIST. He has contributed to over 300 research publications and his current research interests include power system operation, smart grids, renewable energy systems, and energy and power system information systems and communications.