

Integrated velocity optimization and energy management for FCHEV: An eco-driving approach based on deep reinforcement learning

Weiqi Chen^a, Jiankun Peng^{a,*}, Tinghui Ren^a, Hailong Zhang^b, Hongwen He^c, Chunye Ma^a

^a School of Transportation, Southeast University, Nanjing 211102, China

^b College of Mechanics Engineering, North University of China, Taiyuan 030051, China

^c School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China



ARTICLE INFO

Keywords:

Ecological driving
Fuel cell hybrid electric vehicle
Multiple-objective optimization
Deep deterministic policy gradient

ABSTRACT

Ecological driving (eco-driving) is a promising technology for transportation sector to save energy and reduce emission, which works by improving vehicle behaviors in traffic scenarios. Fuel cell hybrid electric vehicles (FCHEV) are receiving extensive attentions due to global fossil energy crisis, but whose implementations for eco-driving result in multiple objective collaborative optimization problems. In this paper, an eco-driving framework for FCHEV is proposed based on deep deterministic policy gradient (DDPG) algorithm. And it combines adaptive cruise control (ACC) and energy management strategy (EMS) into an integrated architecture. Firstly, in order to achieve excellent balance between driving behaviors and fuel economy, an appropriate weight coefficient value is determined after adequate explorations. Secondly, power-varying equivalent hydrogen conversion coefficient function is constructed to save fuel consumption by 8.97%. Thirdly, ablation experiments for health state of fuel cell system present 19.95% decrease in terms of health degradation. Then, comparison experiments indicate that the DDPG-based eco-driving strategy can reach 94.16% of that of dynamic programming with respect to equivalent hydrogen consumption, meanwhile with best ride comfortability. Moreover, simulation results under validation driving cycle manifest its excellent adaptability.

1. Introduction

Ecological driving is an important automotive application technology which reduces fuel consumption as much as possible by improving driving behaviors under the premise of satisfying traffic needs. Thus, there are two basic components of ecological driving, and one is energy management strategy (EMS) to save fuel, reduce health degradation of power sources, and keep state of charge (SOC) margin. And the other is optimization and control of vehicle dynamics in specific traffic scenarios, such as approach-departure strategy at intersection [1] and longitudinal control during car-following [2].

FCHEV are getting more and more attention due to simple utilization, quiet operation, little emissions and high efficiency [3,4]. However, the low power density and low dynamic characteristics of fuel cells require power battery pack as auxiliary energy storage devices to provide peak power and recover brake energy [5]. Given the complex and changeable road environment and traffic needs, it is of great significance to develop ecological driving strategy for FCHEV to maximize its economic potential.

1.1. Energy management for FCHEV

The cornerstone of eco-driving is energy management, which distributes power flow among multiple energy sources to maximize fuel economy, reduce health degradation, and keep SOC. The current research on EMS for FCHEV can be divided into three categories [6]: (1) rule-based, (2) optimization-based, (3) learning-based.

Rule-based EMS mainly includes fuzzy logic controller [7], state machine strategy [8], wavelet transform [9], on-off switching control [10], stiffness coefficient model [11] and charge-depleting and charge sustaining [12], etc. These strategies are primarily established on engineering experience and operation rules, and are practical for single-objective optimization of FCHEV. However, they cannot guarantee optimal performances, and are incompetent to deal with complex and changeable driving environment due to the lack of self-adaptability [13].

Optimization-based EMS can achieve better performance based on numerical optimization methods. Dynamic programming (DP) is utilized to develop off-line EMS under the premise of knowing whole driving

* Corresponding author.

E-mail address: jkpeng@seu.edu.cn (J. Peng).

cycle in advance [14], but it is mainly used as benchmark due to heavy computation burden and global optimality [15]. Although the Pontryagin's minimum principle (PMP) based EMS can be implemented on real-time vehicle control unit [16], it suffers from high sensitivity to operating conditions due to the iteration optimization way of co-state matrix [17]. Another important method is model predictive control (MPC). It is usually employed to establish hierarchical control architecture, where the upper layer deals with traffic information, such as speed planning [18,19] and traffic lights [20], while the lower layer decides energy distribution through MPC [21]. Although MPC-based approaches satisfy the requirements of on-line application, the incompatibility and incoordination of models in different levels hinders its further improvement of optimization performance.

EMS based on reinforcement learning (RL) and deep reinforcement learning (DRL) have become research focus in recent years. Tang et al. [22] developed DRL-based EMS for FCHEV for the first time, they utilized deep Q-network (DQN) algorithm to control the output power of fuel cell system (FCS). However, discrete action results in loss of control accuracy, making it difficult to achieve optimal performances [23]. Wu et al. [24] designed EMS with hybrid action space based on deep twin delayed deep deterministic policy gradient (TD3). And performance gap to the DP benchmark is only 2.55% in terms of fuel consumption. Wu et al. [25] proposed soft actor critic (SAC)-based EMS with continuous control actions to optimize fuel economy, as well as enhance battery thermal safety and health performance. The advantages of DRL-based EMS lie in the independence from global driving conditions [26] and the approximation to the globally optimal performance [27], as well as the application capability in real-time control systems [28].

1.2. Ecological driving

The core concept of ecological driving is the collaborative optimization of speed control and energy management. The EMS module provides optimization suggestions for speed curves, while smoother speed curves improve the performance limit of EMS [29]. Zhao et al. [30] designed a cooperative ecological driving strategy for a group of vehicles at a signalized intersection based on MPC, ensuring pass ability while reducing fuel consumption. The prediction-based ecological approach and departure strategy at urban intersection proposed by Ye et al. [31] achieved 4.0% energy saving. A hybrid DRL framework which combines rule-based strategy and Dueling DQN is proposed in [32], to support connected ecological driving at signalized intersections. Their method can reduce energy consumption by 12.70% when compared with model-based approach. To sum up, these studies mainly achieve energy saving by avoiding frequent stop and start of vehicles in signalized intersection scenario.

Constant velocity is an optimal speed profile for fuel consumption, thus utilizing cruise control when possible is a good recommendation for ecological driving [33]. The primary scenarios for cruise control are free flow and car-following [34]. Sohn C et al. [35] proposed pulse-and-gliding strategy in car-following scenario, which can achieve pretty well balance between fuel consumption and ride comfort. Li et al. [36] combined ACC with EMS into a multiple-objective optimization problem based on DP, which can lead to near-optimal fuel economy and comfortable driving process. Nie Z et al. [2] proposed a real-time dynamic predictive cruise control system, working based on a bi-level MPC algorithm. The car-following-oriented MPC system can realize energy-saving rate by 8.5%~15.6%, while work well and robustly under high-speed driving situation. The above eco-driving strategies are all constructed on numerical optimization methods, and there is still room for performance improvement especially facing with changeable driving conditions.

Wang et al. [37] designed DRL-based eco-driving strategy for plug-in hybrid electric vehicle (HEV) in car-following scenario. Compared with MPC-based strategy, their method can reduce fuel consumption by 15.8% while ensuring safety and comfortability. While in our previous

work, ACC and EMS were collaboratively optimized by an integrated eco-driving framework based on DDPG [38,39] and multi-agent DDPG [40] respectively. Note that their research objectives were hybrid electric vehicle with internal combustion engine. Their simulation results manifested that, both of the approaches can achieve near-optimal fuel economy and comfortable car-following performances.

1.3. Motivation and contribution

Eco-driving includes two essential objectives: (1) improve driving behaviors and (2) improve fuel economy, and the former is aiming to serve the latter. The mutual coupling of the two problems leads to a difficult multi-objective collaborative optimization task, especially when traffic scenarios such as intersections or car-following are taking into consideration. Conventional optimization algorithms such as DP [36], MPC [2,41] were attempting to solve this tricky task. Despite the achievements of these methods, complex models, heavy computation burden, and unsatisfactory performance still hinder their further applications.

DRL-based methods have been developed for eco-driving strategy in intersection scenarios, and achieved excellent mobility, driving comfort, fuel economy and ability for real-time application [32,42–44]. However, compared with the more common cruising scenarios in long continuous road sections, the energy saving potential of intersection scenarios is inherently limited. Therefore, it is necessary to combine EMS and ACC together to realize ecological driving during cruise.

Many works have reported DRL-based eco-driving in car-following scenarios [37–40], however, they were all implemented on hybrid electric vehicle (HEV) with internal combustion engine. Given the significant differences in topology and energy system operation between the two kinds of vehicle, it is of great importance to design specific eco-driving strategy based on DRL method for FCHEV.

To bridge the afore mentioned research gaps, this paper proposes an eco-driving framework for FCHEV based on DDPG, which has continuous control capability. The major contributions are as follows:

- (1) An integrated eco-driving strategy in car-following scenario, which combines EMS and ACC into a collaborative multiple-objective optimization problem, is proposed for FCHEV. And the DDPG method is applied for optimization and control.
- (2) The power-varying equivalent hydrogen conversion coefficient function is established by calculating the final power released by per unit of hydrogen. And it is utilized to guide FCS work in high efficiency ranges.
- (3) The weight coefficient with near-optimal tradeoff between ACC and EMS, is selected after adequate exploration. And ablation experiments for FCS health state is conducted to present the validity of health management. Otherwise, simulation results on validation cycle manifest that the proposed method has excellent adaptability.
- (4) The proposed DDPG-based eco-driving strategy is better than other DRL-based method and achieves very close performance to DP benchmark in terms of energy-saving. Meanwhile, it can achieve much better car-following performance than conventional algorithms.

The remainder of this paper is organized as follows. In Section 2, car-following model, vehicle powertrain model, fuel cell model and power battery model are described respectively. In Section 3, the principle and application of DDPG are depicted briefly, and reward function are designed. In Section 4, simulation results are analyzed. And Section 5 concludes this paper.

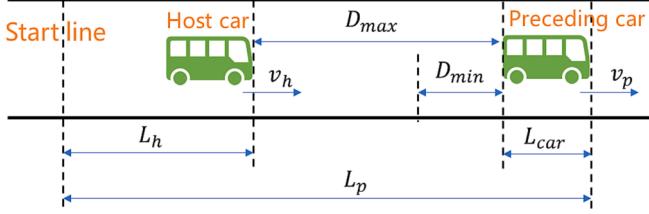


Fig. 1. The car-following scenario.

Table 1
Parameters configuration of the FCHEV.

Section	Parameter Description & Symbol	Value
Vehicle	Mass, m	8400kg
	Front window area, A	6.56m ²
	Wheel radius, r_w	0.467m
	Coefficient of air resistance, C_D	0.55
	Coefficient of rolling resistance, f	0.012
	Final drive ratio, R_{fd}	6.2
Motor	Rated / peak power, P_m^r / P_m^p	100/200kW
	Rated / peak torque, T_m^r / T_m^p	1200/2400Nm
	Rated / peak speed, W_m^r / W_m^p	800/3000rpm
FCS	Rated power, P_{FCS}^r	60kW
DC/DC converter	Rated power, $P_{DC/DC}^r$	60kW
Power battery	Efficiency, $\eta_{DC/DC}$	[0.90, 0.95]
	Capacity, Q_0	108.14kWh
	Rated voltage, V_{OC}	633V

2. Model description

2.1. Car-following model

The traffic environment in this study is a car-following scenario on a single lane without traffic lights, where the host car must keep a safe and reasonable distance from the preceding car, as Fig. 1 shows. Note that the host bus cannot overtake. Symbols L , v , a denote: distance traveled, velocity, acceleration or deceleration, respectively. And subscript h , p represent the host bus and the preceding bus respectively. L_{car} is the length of vehicle, which is 10 m. The distance between two cars is denoted by D .

$$\begin{cases} L_h = \int v_h dt \\ v_h = \int a_h dt \\ D = L_p - L_h - L_{car} \end{cases} \quad (1)$$

To ensure security and effectiveness of car-following function of the adaptive cruise control model, the minimum distance D_{min} [45] and the ideal maximum distance D_{max} [46] are defined. The distance D need be kept between D_{min} and D_{max} .

$$\begin{aligned} D_{min} &= v_h \bullet t_d + v_h^2 / a_{max} + 3 \\ D_{max} &= 0.0825 \bullet v_h^2 + v_h + 10 \end{aligned} \quad (2)$$

where t_d is the sum of brake delay and reaction time, which equals to 1.5 s. And a_{max} denotes the emergency deceleration, which is $7.5m/s^2$. The two values are both from [45].

2.2. Vehicle powertrain model

The research object in this paper is a FCHEV whose main parameters are listed in Table 1 [4], and the structure of powertrain system is shown in Fig. 2(a). The requested traction force of the car is calculated according to the longitudinal dynamics as follows:

$$F_t = mgf\cos\theta + mgsin\theta + \frac{AC_Dv_h^2}{21.15} + \delta ma_h \quad (3)$$

where m is total mass of the vehicle, f is the coefficient of rolling resistance, θ is the road slope, A is the front window area, C_D is the coefficient of air resistance, δ is the rotation mass convention coefficient, and g denotes the acceleration of gravity. Then, rotation speed of wheel W_w and torque of drive shaft T_w are derived:

$$\begin{cases} W_w = v_h / r_w \\ T_w = r_w \bullet F_t \end{cases} \quad (4)$$

where r_w is wheel radius. The rotation speed W_m and torque T_m of motor are then calculated as follows:

$$\begin{cases} W_m = W_w \bullet R_{fd} \\ T_m = \begin{cases} T_w / (\eta_{fd} \bullet R_{fd}), T_w \geq 0 \\ T_w \bullet \eta_{fd} / R_{fd}, T_w < 0 \end{cases} \end{cases} \quad (5)$$

where R_{fd} is the final drive gear ration, and η_{fd} is the efficiency of drive shaft. The power requested by vehicle is calculated as follows:

$$P_{req} = \begin{cases} T_m \bullet W_m / \eta_m, T_m \geq 0 \\ T_m \bullet W_m \bullet \eta_m, T_m < 0 \end{cases} \quad (6)$$

where η_m is motor efficiency, interpolated from the quasi-steady motor model. And Fig. 2(b) shows the motor efficiency map [4]. Since the study object is a FCHEV, which is equipped with fuel cell system and Li-ion power battery pack, the P_{req} is provided by the two energy sources:

$$P_{req} = P_{DC/DC} + P_{bat} \quad (7)$$

where $P_{DC/DC}$ is output power of DC/DC converter, and it is

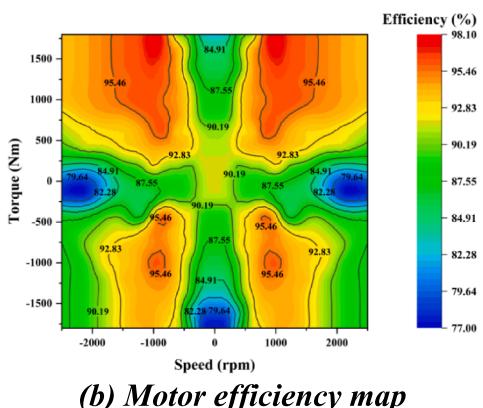
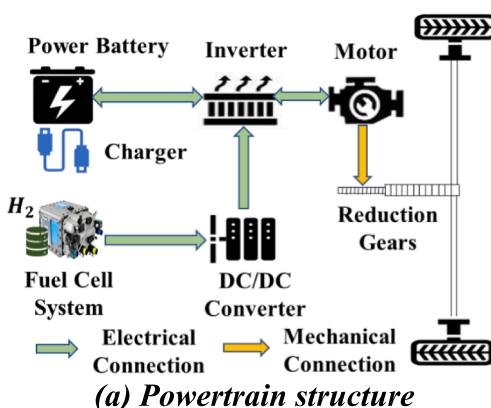


Fig. 2. Powertrain structure and motor efficiency map of the FCHEV.

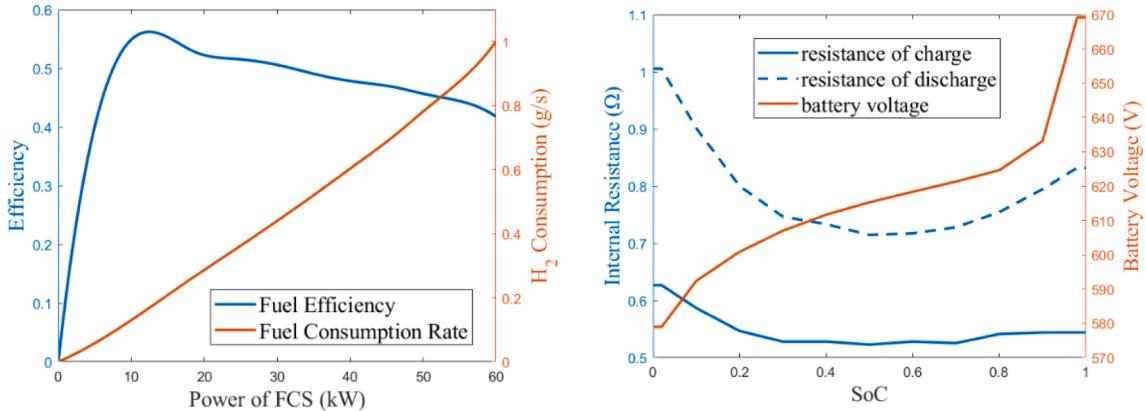


Fig. 3. Characteristic curves of fuel cell and power battery.

determined by the output power of fuel cell system (P_{FCS}). While P_{bat} is power of Li-ion power battery pack, including two modes: charge and discharge.

2.3. Fuel cell system model

The fuel cell system (FCS) converts chemical energy released by hydrogen and oxygen reactions into electricity, whose power is used to drive the motor or stored into power battery pack through the inverter. A physically empirical model considering physical laws and operating conditions of FCS is employed in this paper. The hydrogen consumption rate depending on the output power of FCS is calculated as follows [47]:

$$\dot{m} = \frac{P_{FCS}}{\eta_{FCS} \cdot L_v} \quad (8)$$

where L_v is the chemical energy density of hydrogen, equaling to 120MJ/kg . And η_{FCS} denotes the efficiency of FCS. The relationships between the output power P_{FCS} and hydrogen consumption rate \dot{m} and efficiency η_{FCS} are respectively illustrated in Fig. 3(a), and the data is from [4]. The power outputted by FCS needs to be adjusted by DC/DC converter to match the voltage level with power battery pack. The efficiency of DC/DC converter is obtained by interpolation according to laboratory experiment data [4], in order to reduce computation cost on the premise of ensuring accuracy. And the corresponding figure is not

Simulation Environment

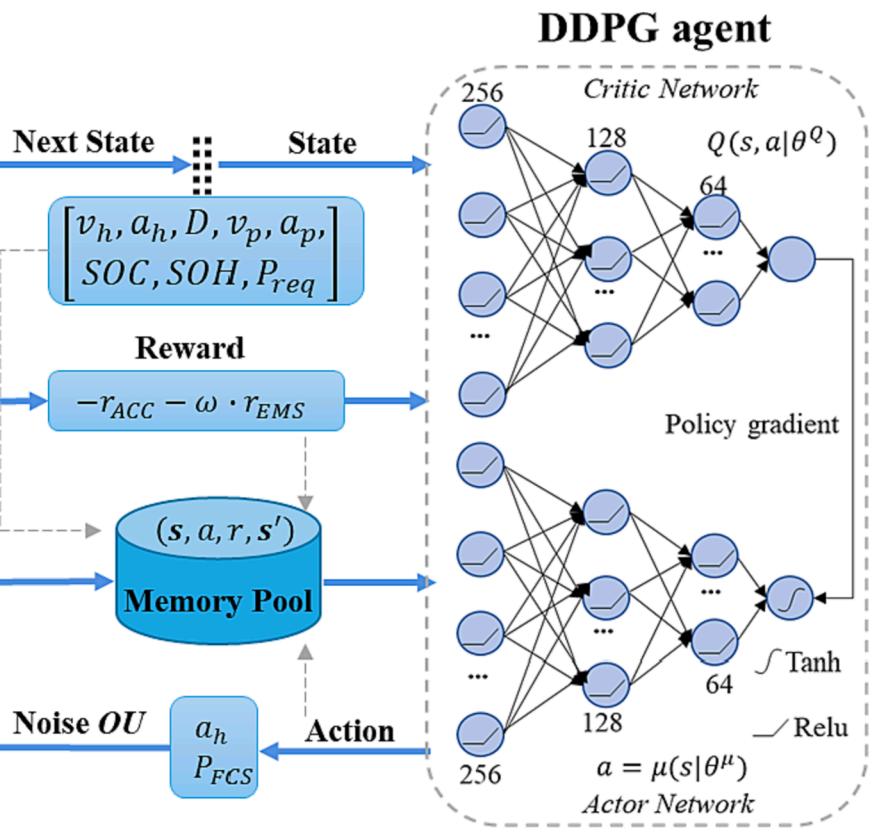
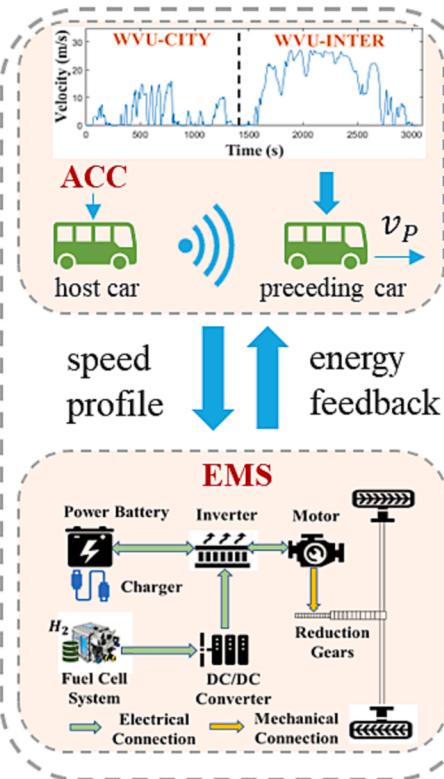


Fig. 4. The architecture of the DDPG-based eco-driving framework.

presented for simplicity. The output power of DC/DC converter $P_{DC/DC}$ is calculated as follows:

$$P_{DC/DC} = \eta_{DC/DC} \bullet P_{FCS} \quad (9)$$

Four different types of disadvantageous driving conditions: (1) load changing cycle, (2) start-stop cycle, (3) low-power load, (4) high-power load, are the primary contributors of FCS health degradation. According to the findings of Raeesi M et.al [48], the total health degradation of FCS can be formulated by the discrete expression:

$$SOH = 1 - \sum_{t=0}^n [d_{ss}(t) + d_{low}(t) + d_{high}(t) + d_{cha}(t)] \quad (10)$$

where SOH is the state of health, and its value is between $[0,1]$. While n is the total number of time steps t . $d_{ss}(t)$, $d_{low}(t)$, $d_{high}(t)$, $d_{cha}(t)$ are the SOH degradation caused by start-stop cycle, low-power load, high-power load, and load changing cycle at time step t respectively.

2.4. Power battery model

As the other energy source of FCHEV, the Li-ion power battery pack is primarily employed to provide peak power, thus flattening the output power of FCS for protection. Another important role is to store excess energy, which is from FCS and brake recovery. The power battery pack is established as an equivalent circuit model [38] as follows:

$$\begin{cases} P_{bat} = V_{OC}I - R_{bat}I^2 \\ I = \frac{V_{OC} - \sqrt{V_{OC}^2 - 4R_{bat}P_{bat}}}{2R_{bat}} \\ SOC = SOC_0 - \frac{\int I dt}{Q_0} \end{cases} \quad (11)$$

where V_{OC} denotes open circuit voltage, and I is the load current of battery pack. R_{bat} represents internal resistance, which is different when charging and discharging. Q_0 denotes the nominal capacity of power battery pack. SOC is the state of charge, and SOC_0 is its initial value. The relationships between SOC and internal resistance and open circuit voltage are as shown in Fig. 3(b), and the data can be found in [39].

3. Ddpg-based eco-driving strategy

3.1. DDPG algorithm

The DDPG algorithm is a classical and effective method in solving multi-objective optimization with continuous control actions, which is based on actor-critic framework [49]. The DDPG agent interacts with simulation environment through action $a_t = \mu(s_t|\theta^\mu) + N$ outputted from Actor network $\mu(s_t|\theta^\mu)$. And N is the action noise sampled from the Ornstein-Uhlenbeck process for exploration of better policy. Then, the agent obtains reward r_t from environment, and the environment steps into next state s_{t+1} . The return from a state is defined as the sum of discounted future reward $R_t = \sum_{i=t}^T \gamma^{i-t} r_i(s_i, a_i)$ with discounting rate γ . The goal of DDPG agent is to learn a policy with maximum expected return from the start distribution $J = E_{r_t, s_t, SE, a_t, \mu}[R_1]$ [50]. The Critic network $Q(s_t, a_t|\theta^Q)$ evaluates the performance of Actor network and guides the DDPG agent to behave better. To improve stability of training process, target Actor network μ' and target Critic network Q' are employed. The architecture of the proposed eco-driving framework is illustrated in Fig. 4.

The parameters of Critic network θ^Q are optimized by minimizing the temporal difference error, in which (s_t, a_t, r_t, s_{t+1}) is mini-batch sampled stochastically from the experience replay buffer M .

$$\begin{cases} L(\theta^Q) = E[(Q(s_i, a_i|\theta^Q) - y_i)^2] \\ y_i = r_i + \gamma Q'[s_{i+1}, \mu'(s_{i+1}|\theta^{Q'})|\theta^Q] \end{cases} \quad (12)$$

The parameters of Actor network θ^μ are updated by applying the chain rule to the expectation return from the start distribution J :

$$\nabla_{\theta^\mu} J = E\left[\nabla_a Q(s, a|\theta^Q)|_{s=s_t, a=\mu(s_t)} \nabla_{\theta^\mu} \mu(s|\theta^\mu)|_{s=s_t}\right] \quad (13)$$

After the update of Actor and Critic network parameters, the parameters of two target networks are optimized by soft updating with soft factor $\tau = 0.005$:

$$\begin{cases} \theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'} \end{cases} \quad (14)$$

3.2. State and action space

The state space consists of what the DDPG agent observes from the interactive environment, and is inputted into the Actor and Critic network after normalization. Therefore, it is ought to be composed of critical information for decision-making. The velocity, acceleration and distance are three essential features for ACC, and SOC , SOH , P_{req} are the most important information for EMS, so that the state space is defined as follows:

$$s = [v_h, a_h, D, v_p, a_p, SOC, SOH, P_{req}] \quad (15)$$

The action space decides what the DDPG agent can do under specific situation, which is defined as follows:

$$\begin{cases} a = [a_h, P_{FCS}] \\ a_h \in [-2.5, 2.5] m/s^2 \\ P_{FCS} \in [0, 60] kW \end{cases} \quad (16)$$

3.3. Reward function

The core idea of reinforcement learning is that the agent learns the policy which has the maximum expectation of reward during its interaction with environment. Thus, a well-designed reward function is of great importance for the training of the DDPG-based strategy. The reward function consists of two parts in the proposed eco-driving framework:

$$r(t) = -[r_{ACC}(t) + \omega \bullet r_{EMS}(t)] \quad (17)$$

where r_{ACC} and r_{EMS} represents reward from ACC and EMS respectively, and weight coefficient ω adjusts their relative importance. Security and comfortability are two main factors that need to be considered during car-following, thus deriving:

$$\begin{cases} r_{ACC}(t) = \sigma \bullet r_c(t) + r_s(t) \\ r_c(t) = \frac{|a_h(t-1) - a_h(t)|}{(a_h^{max} - a_h^{min})} \\ r_s(t) = \begin{cases} v_h^{max}, if D(t) \leq 0 \\ v_h(t), if 0 < D(t) \leq D_{min}(t) \\ 1, if D_{min}(t) < D(t) \leq D_{max}(t) \\ D(t) - D_{max}(t), if D(t) > D_{max}(t) \end{cases} \end{cases} \quad (18)$$

where r_s and r_c denote reward about security and comfortability respectively, σ is a coefficient to make them compatible, and selected as 10 after numerous attempts.

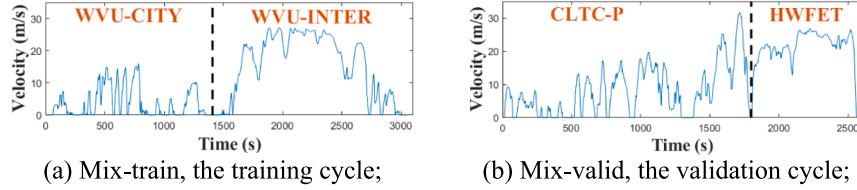


Fig. 5. The mixture driving cycles.

Table 2
Key hyper parameters for training DDPG algorithm.

Parameter Description	Value
Number of hidden layers of Actor and Critic (target) network	3
Neurons distribution of Actor and Critic (target) network	256, 128, 64
Neural network connection	Fully connected
Optimizer	Adam
Learning rate scheduler	Cyclical [51] [5e-5, 1e-4]
Learning rate range of Actor network	[5e-5, 1e-3]
Learning rate range of Critic network	[5e-5, 1e-3]
Size of experience replay buffer	3e4
Minibatch size	64
Number of training episodes	500

The objectives of EMS involve three aspects: (1) minimize hydrogen consumption, (2) reduce health degradation of FCS, (3) and keep SOC within reasonable range.

$$r_{EMS}(t) = \rho_1 [\dot{m}(t) + \delta \bullet P_{bat}(t)] + \rho_2 \Delta SOH(t) + \rho_3 |SOC(t) - SOC_{ref}| \quad (19)$$

where δ is the conversion coefficient which measures the equivalence between the power of power battery pack P_{bat} and hydrogen consumption. Then, the hydrogen directly consumed by FCS and the hydrogen equivalently converted from power battery pack are collectively referred to as the equivalent hydrogen consumption. The impact of the conversion coefficient on EMS performance is described detailedly in Section 4.2.

And ρ_1, ρ_2 are hydrogen price, replacement price of FCS, respectively. The former two objectives are uniformed by money cost. And ρ_3 is the weight coefficient which determines the relative importance of the money cost versus battery SOC value. It is set as 20 after numerous experiments. SOC_{ref} is the reference value of SOC, set as 0.5.

3.4. Training configuration

The kinematics of host vehicle is controlled by acceleration a_h outputted from DDPG agent, while the preceding car follows driving cycles. Two driving cycles are constructed in this work, named Mix-train and Mix-valid respectively. The Mix-train cycle is used for training the DDPG network parameters, and it covers low-speed to medium-speed

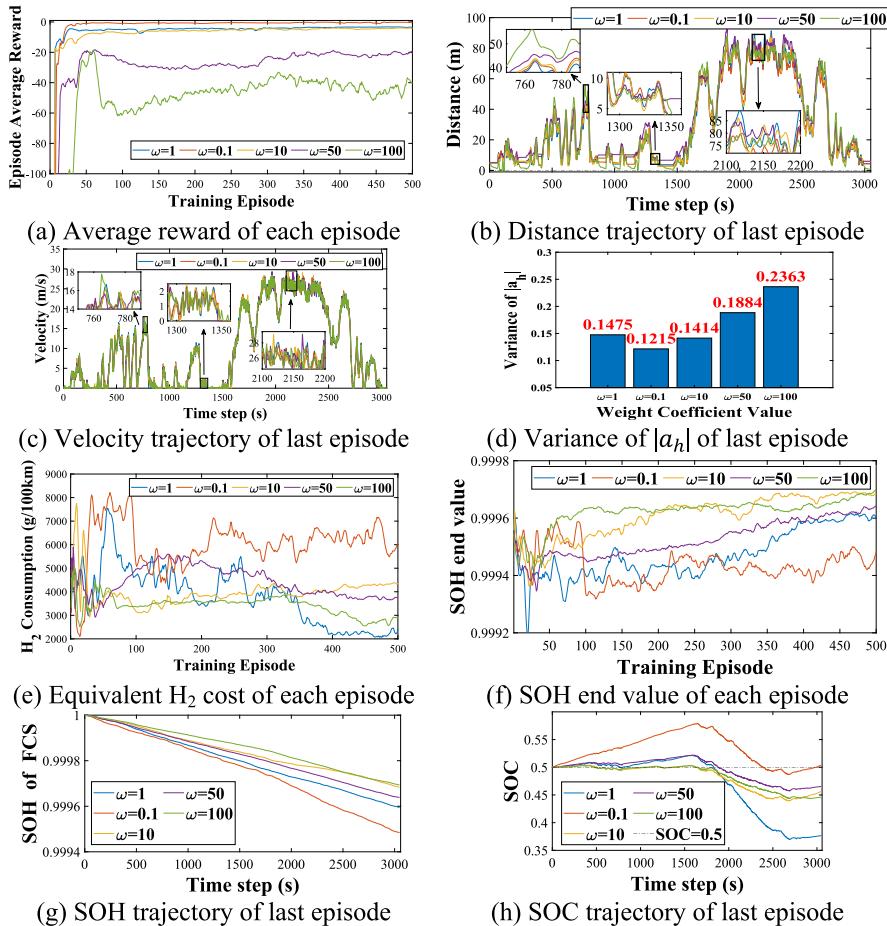
Fig. 6. Performances comparison under different ω .

Table 3Comparison of EMS performances under different ω

Value of ω	Equivalent H ₂ cost (g/100 km)	Comparison	Degradation of SOH	Comparison	Money cost (CNY)	Comparison
0.1	9391.0	100.00%	0.052%	100.00%	321.28	100.00%
1	7103.3	75.64%	0.040%	76.92%	245.02	76.26%
10	7551.7	80.41%	0.031%	59.62%	225.91	70.32%
50	6914.7	73.63%	0.036%	69.23%	229.70	71.49%
100	6305.8	67.14%	0.030%	57.69%	200.98	62.56%

and high-speed, in order to enable the learned strategy can cope with different traffic flow conditions. Meanwhile, the Mix-valid cycle consists of totally different driving cycles, which is used for checking whether overfitting has occurred during training process. As shown in Fig. 5, the Mix-train cycle consists of two standard driving cycles: West Virginia University City (WVU-CITY) and West Virginia University Interstate (WVU-INTER). While the Mix-valid consists of another two: the China light-duty vehicle test cycle-passenger car (CLTC-P) and the Highway Fuel Economy Test (HWFET). The key hyper parameters of the proposed DDPG-based eco-driving strategy are listed in Table 2.

4. Simulation results and discussions

In this section, the DDPG-based eco-driving strategy is learned and analyzed. Firstly, to obtain near-optimal tradeoff between ACC and EMS, numerous simulation experiments are implemented to explore an appropriate value of weight coefficient ω . Secondly, the influences of the power-varying equivalent hydrogen conversion coefficient on EMS are analyzed. Thirdly, ablation experiments for SOH are conducted to manifest the importance and validity of health management of EMS. Then, car-following models and EMS based on different algorithms are respectively compared with the DDPG-based method to demonstrate its optimality. Last but not least, to evaluate the adaptability of the proposed method, optimization performances when the driving cycle of the preceding car is set Mix-valid are analyzed.

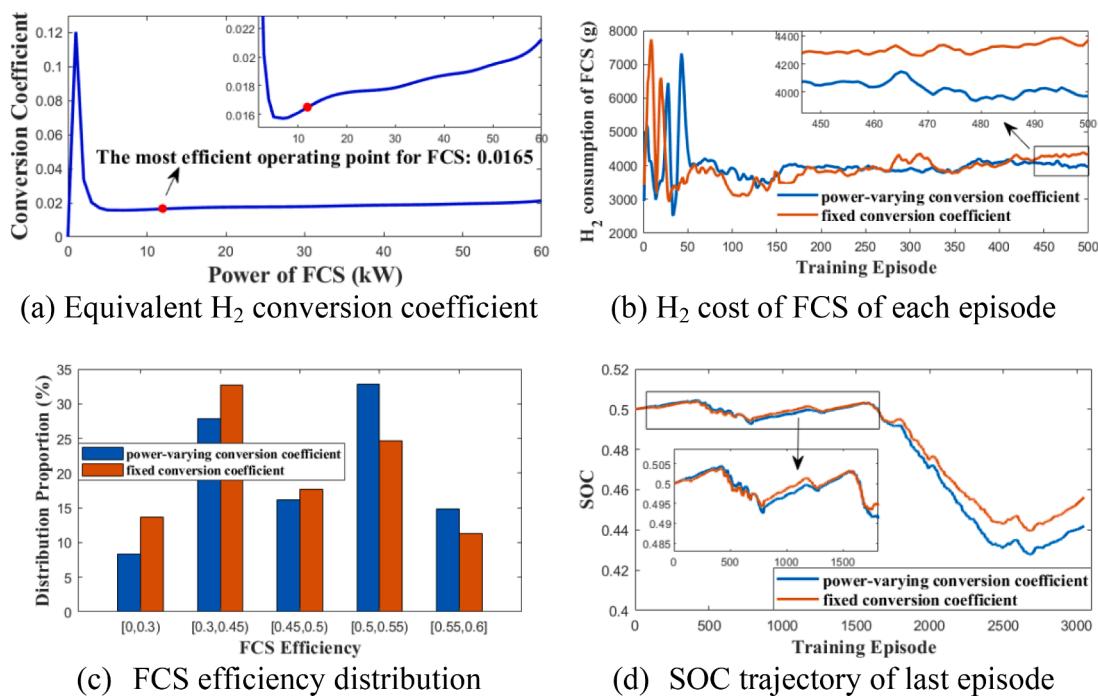
4.1. Tradeoff between multiple objectives

The eco-driving strategy proposed in this study combines ACC and

EMS into an integrated framework. The ACC module focuses on traffic-related concerns, whereas the EMS module tackles power distribution within the FCHEV. Given the distinct interpretations and scales of the reward functions associated with the two modules, it becomes imperative to explore the most appropriate weight coefficient for reward allocation. In this section, the ACC module keeps a fixed weight coefficient, while the reward weight coefficient of the EMS module, ω , is investigated. By comparing the optimization performances with various ω , the desired value of weight coefficient is identified.

Firstly, Fig. 6(a) presents the episode average reward curves for different weight coefficients. Among the selected five values, $\omega = \{0.1, 1, 10\}$ demonstrate high rewards and very good convergence, indicating that the chosen agents achieve favorable learning outcomes under the three weight coefficients.

Secondly, the car-following performances of ACC module are analyzed in terms of safety and comfortability. Fig. 6(b) illustrates the tracking distance between the host car and the preceding car during training, while Fig. 6(c) displays the speed trajectory of the host car. It is verified that when coefficients $\omega = \{0.1, 1, 10\}$ are utilized, the host car can maintain effective and safe distances from the preceding vehicle across different speed ranges—low, medium, and high. Noting that when $\omega = 10$ is applied, the host car exhibits smaller distances at low-speed range, appropriate distances at medium-speed range, and larger distances at high-speed range, demonstrating good considerations for safety and effectiveness. However, when coefficient $\omega = 100$ is utilized, the rewards proportion of ACC module becomes excessively low, resulting in insufficient considerations for car-following objectives, and consequently leading to inappropriate distances between vehicles and even collisions. Fig. 6(d) illustrates the variance of absolute value of

Fig. 7. Performance comparison of equivalent H₂ conversion coefficient.

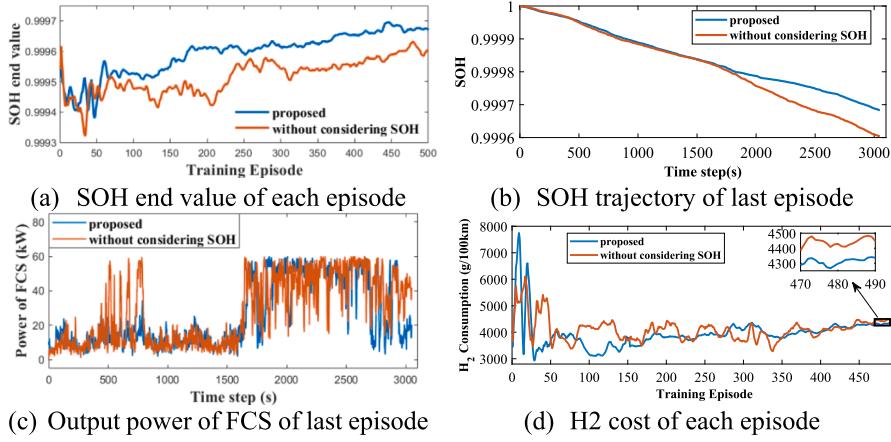


Fig. 8. SOH ablation experiment results.

acceleration with different ω , which serve as indirect indicator of ride comfortability. Notably, smaller variances occur with $\omega = \{0.1, 1, 10\}$. In short, it means that when $\omega = \{0.1, 1, 10\}$ are applied, the proposed strategy can realize effective car-following with pretty good safety and comfortability.

Thirdly, the performances of EMS module are discussed. The equivalent hydrogen consumption serves as an indicator of overall energy consumption level of the FCHEV. As shown in Fig. 6(e), a clear trend can be observed, that the agents trained with $\omega = \{1, 10\}$ demonstrate significantly lower equivalent hydrogen consumption compared to the agent trained with $\omega = 0.1$. As shown in Table 3, in the last training episode, the agents with $\omega = \{1, 10\}$ generated equivalent hydrogen consumption of 75.64% and 80.41% respectively, compared to the agent with $\omega = 0.1$. In short, significant energy-saving advantages can be observed with $\omega = \{1, 10\}$.

Then, the SOH degradation of FCS is compared. Fig. 6(f) demonstrates the evolving SOH value at the end of each training episode for agents trained with various weight coefficients. It is worth noting that agents with $\omega = \{10, 100\}$ exhibit the most pronounced improvements in SOH, indicating substantial learning outcomes in terms of preserving and enhancing FCS health. As shown in Fig. 6(g), the agent with $\omega = 10$ yields the second-least degradation of FCS SOH. According to Table 3, the SOH degradation percentages for agents with $\omega = \{1, 10, 50, 100\}$ are 76.92%, 59.62%, 69.23%, and 57.69% respectively, compared to the agent with $\omega = 0.1\omega = 0.1$. Although it exhibits the best health performance at $\omega = 100$, this is achieved by significantly sacrificing ACC performance and is not desirable. Taking money cost to normalize hydrogen consumption and SOH degradation, the agent with $\omega = 10$ outperforms others. Furthermore, the agent with $\omega = 10\omega = 10$ demonstrates a reasonable and smooth SOC trajectory, as illustrated in Fig. 6(h).

In summary, the agent with $\omega = 10$ exhibits excellent performances in terms of training convergence, safety and comfortability during car-following, and relatively better performances with respect to fuel-saving, SOH and SOC maintenance. Thus, $\omega = 10$ is the preferred weight coefficient in subsequent experiments.

4.2. Analysis of equivalent hydrogen conversion coefficient

Consider the situation that the power of power battery pack is solely provided by FCS. According to the FCS model described in Section 2.3, the function relationship between the equivalent hydrogen conversion coefficient δ and the FCS output power P_{FCS} , is established by calculating the final power released by per unit of hydrogen, as shown in Fig. 7(a). It is referred as the power-varying conversion coefficient. The red dot marks the operating point with the highest FCS efficiency, which is near the local minimum of the power-varying conversion coefficient function

Table 4
Comparison of H₂ consumption and SOH.

Method	FCS SOH degradation	Comparison	H ₂ cost (g/100 km)	Comparison
Proposed	0.0310%	80.05%	4336.9	97.52%
Without SOH	0.0396%	100%	4447.8	100%

with the value of 0.0165. This value is considered to represent to optimal FCS operating point, and is widely used to calculate equivalent hydrogen consumption [20].

Fig. 7(b) illustrates the hydrogen consumption trajectory directly from FCS during training process for fixed and power-varying coefficient respectively. From the 413rd episode, the power-varying coefficient method consistently exhibits lower FCS hydrogen consumption compared to the fixed coefficient method, continuing until to the end of training. When training finished, the FCS hydrogen consumption for the fixed coefficient method is 4371.6 g/100 km, whereas for the power-varying coefficient method, it is 3979.2 g/100 km, representing an 8.98% decrease. Fig. 7(c) shows the distribution of FCS efficiency. It can be obviously observed that the operating points of power-varying coefficient method are much more distributed in high efficiency range. This indicates that the power-varying conversion coefficient function that reflects real operation situations can guide the DDPG agent to output more efficient actions, thereby improving energy efficiency and reducing hydrogen consumption. Fig. 7(d) illustrates the SOC trajectory for the fixed and varying coefficient methods in the last episode. Under both methods, the DDPG agent can stabilize SOC within [0.4, 0.5].

In summary, thanks to more actual reflection of FCS operation, the power-varying conversion coefficient method can achieve better performance in terms of energy efficiency and fuel-saving, presenting an 8.98% decrease of hydrogen consumption.

4.3. Ablation experiment for SOH

In this section, ablation experiments for health management are conducted. The proposed method considers the SOH of FCS, while the baseline method does not consider it. The trajectories of SOH end values for different methods are shown in Fig. 8(a). The proposed method exhibits an upward trend as the training progresses and quickly surpasses the baseline method in the early stages of training. This advantage maintains until the end of the training. Fig. 8(b) illustrates the SOH trajectory in the last training episode. The proposed method consistently maintains a stable rate of SOH degradation, while the baseline method experiences a faster decline from 1700th time-step until the end. The statistical data in Table 4 shows that the health performance of the

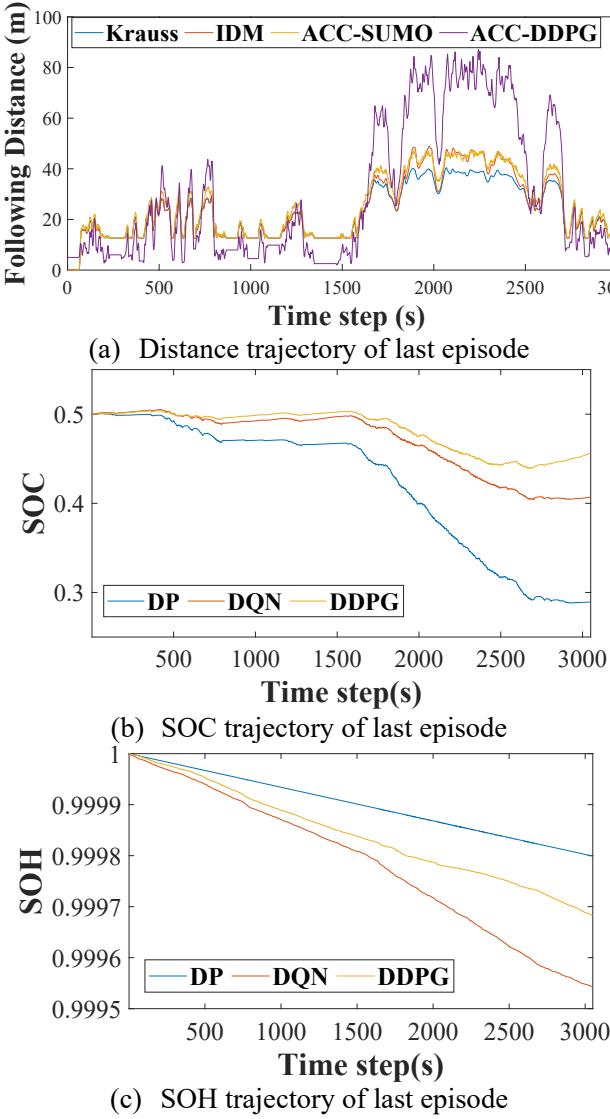


Fig. 9. Performances comparison of different methods.

proposed method has improved by 19.95% compared to the baseline method.

Fig. 8(c) presents FCS output power curves for the two methods. Compared to the baseline method, the FCS operating points of the proposed method are noticeably reduced under high-load conditions. The variances of output power P_{FCS} for the proposed and baseline method are 473.30 and 494.99 respectively. The 4.38% reduction of P_{FCS} variance indicates more smoother output characteristics of FCS, and it is beneficial for health maintenance [4]. Fig. 8(d) illustrates FCS hydrogen consumption curves during training. Excellent health performance often corresponds to more efficient and reasonable operating points, resulting in reduced fuel consumption. The proposed method achieves significantly lower hydrogen consumption compared to the baseline method from the early stages of training until the end. As the quantitative data shown in Table 4, the proposed method achieves a lower hydrogen consumption of 4336.9 g/100 km, which has 2.48% reduction compared to the baseline method.

In summary, the results highlight the superiority and importance of considering health degradation of fuel cell system in eco-driving, for that it can reduce SOH degradation by 19.95% and hydrogen consumption by 2.48% respectively.

Table 5

Variances of $|a_h|$ of different car-following methods.

Method	Variance of $ a_h $	Comparison	Method	Variance of $ a_h $	Comparison
Krauss	0.4807	100%	ACC-SUMO	0.2502	52.04%
IDM	0.2483	51.65%	ACC-DDPG	0.1414	29.42%

4.4. Analysis of optimality

This section discusses the optimality of the proposed method from two aspects, car-following control and EMS. In terms of car-following performance, three traditional methods, i.e., Krauss, Intelligent Driver Model (IDM) and ACC, form the Simulation of Urban Mobility (SUMO), are utilized. As shown in Fig. 9(a), compared to other car-following models, the proposed ACC module achieves better balance between tracking efficiency and safety. The proposed method maintains shorter following distances in the short gap range, indicating improved efficiency under the premise of non-collision during low-speed driving. Conversely, in the long gap range, it maintains longer following distances, enhancing safety in high-speed situation. The variance of absolute value of acceleration are listed in Table 5, in which the proposed method shows nearly 70% improvement with respect to comfortability performance. In short, the DDPG-based method outperforms other conventional car-following models.

In order to demonstrate superior performances of EMS, besides DDPG, DQN and DP are implemented in the same driving cycle. The SoC trajectories in Fig. 9(b) illustrate very effective energy management process. Under the consideration of FCS SOH, the DP method consistently attempts to utilize power battery pack to drive vehicle, resulting in rapid decrease in SOC until it eventually achieves at 0.2893. In contrast, the SOC trajectories of DDPG and DQN are more similar, and shows smoother variations, ultimately remaining close to the predefined value. This discrepancy could be attributed to the heightened sensitivity to reward changes of DRL-based approaches. The DRL agents are unwilling to make SOC decrease significantly, as it would lead to reward loss associated with maintaining SOC.

Fig. 9(c) illustrates the SOH trajectories for different methods. Due to the DP method consistently attempting to use power battery pack to drive the vehicle as possible, it exhibits the smallest decrease in SOH over time. Meanwhile, the DRL-based methods show promising health performance while also striving to maintain SOC. As shown in Table 6, the SOH end value for DQN and DDPG is 99.95% and 99.99% of that observed in the DP method, respectively. Taking DP results as the global optimal benchmark, as the data listed in Table 6, in terms of fuel consumption, DQN and DDPG can reach 92.07% and 94.16% respectively of that of DP. Additionally, the DDPG exhibits money cost that is only 96.56% of that observed in the DP method, while DQN can only reach 81.02%. This further demonstrates pretty well fuel economy of the proposed strategy.

In summary, compared with conventional car-following methods, the proposed DDPG-based eco-driving can achieve excellent ride comfortability on the premise of safety and effectiveness. On the other hand, it can reach 94.16% and 96.56% of that of the DP benchmark, in terms of equivalent hydrogen consumption and money cost respectively. Moreover, it is an integrated framework where ACC and EMS are controlled synchronously.

4.5. Analysis of adaptability

To evaluate the adaptability of the proposed method, the saved neural network parameters are loaded to execute the pre-trained policy under the Mix-valid cycle. Fig. 10(a) shows that the proposed strategy can achieve car-following performance with security and effectiveness

Table 6

EMS performances of different methods.

Method	Equivalent H ₂ cost (g/100 km)	Comparison	FCS SOH end value	Comparison	Money cost (CNY)	Comparison
DP	7110.3	100%	99.98%	100%	218.14	100%
DQN	7722.4	92.07%	99.95%	99.97%	270.91	81.02%
DDPG	7551.7	94.16%	99.97%	99.99%	225.91	96.56%

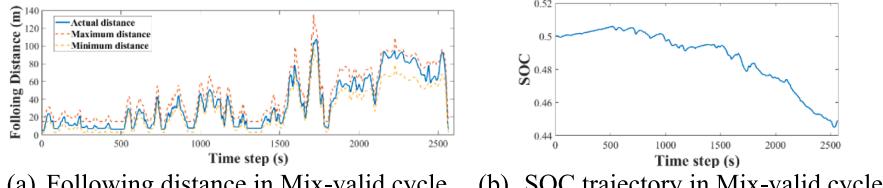


Fig. 10. Car-following performance and SOC trajectory in Mix-valid cycle.

Table 7

Performances comparison between Mix-valid and Mix-train cycle.

Driving cycle	Variance of a _h	Comparison	H ₂ cost (g/100 km)	Comparison	FCS SOH degradation	Comparison
Mix-train	0.1414	100%	4371.6	100%	0.0310%	100%
Mix-valid	0.1206	85.29%	3314.1	75.81%	0.0049%	15.81%

even in new driving cycles which never been known, the actual distance can be controlled between expected maximum and minimum distance. Fig. 10(b) shows a stable and reasonable SOC trajectory within [0.4, 0.5]. The comparison data in Table 7 shows that the ride comfortability of the proposed method on Mix-valid cycle is also guaranteed, the variance of absolute value of acceleration is 85.29% of that on the Mix-train cycle. More importantly, the hydrogen consumption of the proposed method is also lower on the Mix-valid cycle, 75.81% of that on the training cycle. Meanwhile, the degradation of SOH under Mix-valid cycle exhibits remarkable performance, amounting to only 15.81% of that on the Mix-train cycle, which means the proposed method can maintain favorable performance in terms of health. In summary, the above results indicate that the proposed method has good adaptability in urban and highway car-following conditions.

5. Conclusion

This paper proposes an integrated ecological driving framework for FCHEV, where longitudinal ACC and EMS are optimized collaboratively and controlled synchronously. The DDPG algorithm with real-time continuous control capacity is employed for training. Firstly, the weight coefficient value of ACC reward versus EMS reward is determined after numerous experiments. Secondly, the equivalent hydrogen conversion coefficient function which is varying with FCS output power is constructed, and it is indeed helpful to reduce hydrogen consumption. Then, ablation experiments for SOH are conducted to emphasize the significance of health management. Last, comparative experiments with different methods and driving cycle manifest optimality and adaptability of the proposed eco-driving strategy. The primary conclusions of this work are as follows:

- (1) After adequate explorative experiments, the proposed framework achieves an excellent balance between the two optimization tasks, ACC and EMS. And a stable strategy which performs well can be learned in such configurations. An appropriate value of weigh coefficient is determined, $\omega = 10$.
- (2) The EMS module takes the power-varying equivalent hydrogen conversion coefficient into consideration, which shows 8.97% decrease in terms of hydrogen consumption when compared with fixed equivalent conversion coefficient method.

- (3) The proposed eco-driving strategy takes FCS health degradation into consideration, and the SOH ablation experiments indicate that it can improve health performance by 19.95%.
- (4) The car-following performance of the DDPG-based eco-driving strategy is compared with three car-following models in SUMO. Results show that it not only achieves more effective trace, but also the best ride comfortability with security.
- (5) In terms of EMS, the proposed eco-driving strategy is compared with that based on DP and DQN. Results show that it performs much better than DQN, and reaches 94.16% and 96.56% of that of the DP benchmark, with respect to equivalent hydrogen consumption and money cost respectively.
- (6) Simulation results under the pre-constructed validation cycle present better performance than that on the training cycle, which manifests excellent adaptability of the proposed strategy.

CRediT authorship contribution statement

Weiqi Chen: Methodology, Software, Investigation, Writing – original draft, Writing – review & editing. **Jiankun Peng:** Supervision, Funding acquisition. **Tinghui Ren:** Data curation, Formal analysis, Software, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported in part by the National Key R&D Program of China (Grant No.2022YFB4300300), the National Natural Science Foundation of China (Grant No.52072074 & No.52372380), the Fundamental Research Funds for the Central Universities (Grant No.2242021R40007), Emission Peak & Carbon Neutrality Innovation

S&T Project of Nanjing (No.202211018), and “Zhishan” Scholars Programs of Southeast University.

References

- [1] Wang Z, Wu G, Barth MJ. Cooperative eco-driving at signalized intersections in a partially connected and automated vehicle environment[J]. *IEEE Trans Intell Transp Syst* 2019;21(5):2029–38.
- [2] Nie Z, Farzaneh H. 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[J]. *Energy* 2022;241:122888.
- [3] Inci M, Büyükk M, Demir MH, et al. A review and research on fuel cell electric vehicles: Topologies, power electronic convert Cooperative Eco-Driving at Signalized Intersections in a Partially Connected and Automated Vehicle Environmenters, energy management methods, technical challenges, marketing and future aspects[J]. *Renew Sustain Energy Rev* 2021;137:110648.
- [4] Chen W, Peng J, Chen J, et al. Health-considered energy management strategy for fuel cell hybrid electric vehicle based on improved soft actor critic algorithm adopted with Beta policy[J]. *Energ Conver Manage* 2023;292:117362.
- [5] Kandidayeni M, Trovão JP, Soleymani M, et al. Towards health-aware energy management strategies in fuel cell hybrid electric vehicles: A review[J]. *Int J Hydrogen Energy* 2022;47(17):10021–43.
- [6] Zhao X, Wang L, Zhou Y, et al. Energy management strategies for fuel cell hybrid electric vehicles: Classification, comparison, and outlook[J]. *Energ Conver Manage* 2022;270:116179.
- [7] Ibrahim O, Bakare MS, Amosa TI, et al. Development of fuzzy logic-based demand-side energy management system for hybrid energy sources[J]. *Energy Conversion and Management*: X 2023;18:100354.
- [8] Fernandez AM, Kandidayeni M, Boulon L, et al. An adaptive state machine based energy management strategy for a multi-stack fuel cell hybrid electric vehicle[J]. *IEEE Trans Veh Technol* 2019;69(1):220–34.
- [9] Robayo M, Mueller M, Shark S, et al. Assessment of supercapacitor performance in a hybrid energy storage system with an EMS based on the discrete wavelet transform[J]. *J Storage Mater* 2023;57:106200.
- [10] Vadiee A, Yaghoubi M, Martin V, et al. Energy analysis of solar blind system concept using energy system modelling[J]. *Sol Energy* 2016;139:297–308.
- [11] Yun H, Liu S, Zhao Y, et al. Energy management for fuel cell hybrid vehicles based on a stiffness coefficient model[J]. *Int J Hydrogen Energy* 2015;40(1):633–41.
- [12] Jungem M, Kimmig N, Langwiesner M, et al. Analysis of the Optimal Operating Strategy of a P24-Hybrid for Different Electric Power Distributions in Charge-Depleting and Charge-Sustaining Operation[R]. SAE Technical Paper, 2021.
- [13] Inci M, Büyükk M, Demir MH, et al. A review and research on fuel cell electric vehicles: Topologies, power electronic converters, energy management methods, technical challenges, marketing and future aspects[J]. *Renew Sustain Energy Rev* 2021;137:110648.
- [14] Peng J, He H, Xiong R. Rule based energy management strategy for a series-parallel plug-in hybrid electric bus optimized by dynamic programming[J]. *Appl Energy* 2017;185:1633–43.
- [15] Ali AM, Ghanbar A, Söfker D. Optimal control of multi-source electric vehicles in real time using advisory dynamic programming[J]. *IEEE Trans Veh Technol* 2019; 68(11):10394–405.
- [16] Song K, Wang X, Li F, et al. Pontryagin’s minimum principle-based real-time energy management strategy for fuel cell hybrid electric vehicle considering both fuel economy and power source durability[J]. *Energy* 2020;205:118064.
- [17] Sun X, Zhou Y, Huang L, et al. A real-time PMP energy management strategy for fuel cell hybrid buses based on driving segment feature recognition[J]. *Int J Hydrogen Energy* 2021;46(80):39983–40000.
- [18] Nie Z, Jia Y, Wang W, et al. Co-optimization of speed planning and energy management for intelligent fuel cell hybrid vehicle considering complex traffic conditions[J]. *Energy* 2022;247:123476.
- [19] Guo J, He H, Wei Z, et al. An Economic Driving Energy Management Strategy for the Fuel Cell Bus[J]. *IEEE Trans Transp Electrif* 2022.
- [20] Yan M, Li G, Li M, et al. Hierarchical predictive energy management of fuel cell buses with launch control integrating traffic information[J]. *Energ Conver Manage* 2022;256:115397.
- [21] Hu X, Zou C, Tang X, et al. Cost-optimal energy management of hybrid electric vehicles using fuel cell/battery health-aware predictive control[J]. *IEEE Trans Power Electron* 2019;35(1):382–92.
- [22] Tang X, Zhou H, Wang F, et al. Longevity-conscious energy management strategy of fuel cell hybrid electric Vehicle Based on deep reinforcement learning[J]. *Energy* 2022;238:121593.
- [23] Tan H, Zhang H, Peng J, et al. Energy management of hybrid electric bus based on deep reinforcement learning in continuous state and action space[J]. *Energ Conver Manage* 2019;195:548–60.
- [24] Wu C, Ruan J, Cui H, et al. The application of machine learning based energy management strategy in multi-mode plug-in hybrid electric vehicle, part I: Twin Delayed Deep Deterministic Policy Gradient algorithm design for hybrid mode[J]. *Energy* 2023;262:125084.
- [25] Wu J, Wei Z, Li W, et al. Battery thermal-and health-constrained energy management for hybrid electric bus based on soft actor-critic DRL algorithm[J]. *IEEE Trans Ind Inf* 2020;17(6):3751–61.
- [26] Wu Y, Tan H, Peng J, et al. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus[J]. *Appl Energy* 2019;247:454–66.
- [27] Hu X, Liu T, Qi X, et al. Reinforcement learning for hybrid and plug-in hybrid electric vehicle energy management: Recent advances and prospects[J]. *IEEE Ind Electron Mag* 2019;13(3):16–25.
- [28] Lian R, Peng J, Wu Y, et al. Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle[J]. *Energy* 2020;197:117297.
- [29] Huang Y, Ng ECY, Zhou JL, et al. Eco-driving technology for sustainable road transport: A review[J]. *Renew Sustain Energy Rev* 2018;93:596–609.
- [30] Zhao W, Ngoduy D, Shepherd S, et al. A platoon based cooperative eco-driving model for mixed automated and human-driven vehicles at a signalised intersection [J]. *Transportation Research Part C: Emerging Technologies* 2018;95:802–21.
- [31] Ye F, Hao P, Qi X, et al. Prediction-based eco-approach and departure at signalized intersections with speed forecasting on preceding vehicles[J]. *IEEE Trans Intell Transp Syst* 2018;20(4):1378–89.
- [32] Bai Z, Hao P, Shangguan W, et al. Hybrid reinforcement learning-based eco-driving strategy for connected and automated vehicles at signalized intersections[J]. *IEEE Trans Intell Transp Syst* 2022;23(9):15850–63.
- [33] Singh H, Kathuria A. Profiling drivers to assess safe and eco-driving behavior—A systematic review of naturalistic driving studies[J]. *Accid Anal Prev* 2021;161: 106349.
- [34] Zhang H, Peng J, Dong H, et al. Hierarchical reinforcement learning based energy management strategy of plug-in hybrid electric vehicle for ecological car-following process[J]. *Appl Energy* 2023;333:120599.
- [35] Sohn C, Andert J, Jolovic D. An analysis of the tradeoff between fuel consumption and ride comfort for the pulse and glide driving strategy[J]. *IEEE Trans Veh Technol* 2020;69(7):7223–33.
- [36] Li G, Görges D. Ecological adaptive cruise control and energy management strategy for hybrid electric vehicles based on heuristic dynamic programming[J]. *IEEE Trans Intell Transp Syst* 2018;20(9):3526–35.
- [37] Wang Y, Wu Y, Tang Y, et al. Cooperative energy management and eco-driving of plug-in hybrid electric vehicle via multi-agent reinforcement learning[J]. *Appl Energy* 2023;332:120563.
- [38] Peng J, Fan Y, Yin G, et al. Collaborative optimization of energy management strategy and adaptive cruise control based on deep reinforcement learning[J]. *IEEE Trans Transp Electrif* 2022;9(1):34–44.
- [39] Chen W, Yin G, Fan Y, et al. Ecological Driving Strategy for Fuel Cell Hybrid Electric Vehicle Based on Continuous Deep Reinforcement Learning[C]//2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI). IEEE, 2022: 1–6.
- [40] Peng J, Chen W, Fan Y, et al. Ecological Driving Framework of Hybrid Electric Vehicle Based on Heterogeneous Multi Agent Deep Reinforcement Learning[J]. *IEEE Trans Transp Electrif* 2023.
- [41] Deng K, Fang T, Feng H, et al. Hierarchical eco-driving and energy management control for hydrogen powered hybrid trains[J]. *Energ Conver Manage* 2022;264: 115735.
- [42] Li J, Wu X, Xu M, et al. Deep reinforcement learning and reward shaping based eco-driving control for automated HEVs among signalized intersections[J]. *Energy* 2022;251:123924.
- [43] Jiang X, Zhang J, Li D. Eco-driving at signalized intersections: a parameterized reinforcement learning approach[J]. *Transportmetrica B: Transport Dynamics* 2023;11(1):1406–31.
- [44] Guo Q, Angah O, Liu Z, et al. Hybrid deep reinforcement learning based eco-driving for low-level connected and automated vehicles along signalized corridors[J]. *Transportation Research Part C: Emerging Technologies* 2021;124:102980.
- [45] Luo Q, Xun L, Cao Z, et al. Simulation analysis and study on car-following safety distance model based on braking process of leading vehicle[C]//2011 9th World Congress on Intelligent Control and Automation. IEEE 2011:740–3.
- [46] Yang D, Pu Y, Yang F, et al. Car-following model based on optimal distance and its characteristics analysis[J]. *J Northwest Transp Univ* 2012:47.
- [47] Deng K, Liu Y, Hai D, et al. Deep reinforcement learning based energy management strategy of fuel cell hybrid railway vehicles considering fuel cell aging[J]. *Energ Conver Manage* 2022;251:115030.
- [48] Raeesi M, Changizian S, Ahmadi P, et al. Performance analysis of a degraded PEM fuel cell stack for hydrogen passenger vehicles based on machine learning algorithms in real driving conditions[J]. *Energ Conver Manage* 2021;248:114793.
- [49] Huang Y, Hu H, Tan J, et al. Deep reinforcement learning based energy management strategy for range extend fuel cell hybrid electric vehicle[J]. *Energ Conver Manage* 2023;277:116678.
- [50] Lillicrap TP, Hunt JJ, Pritzel A, et al. Continuous control with deep reinforcement learning[J]. arXiv preprint arXiv:1509.02971; 2015.
- [51] Smith LN. Cyclical learning rates for training neural networks[C]//2017 IEEE winter conference on applications of computer vision (WACV). IEEE 2017:464–72.