Energy Proceedings
Vol 30, 2023
ISSN 2004-2965

Health-Aware Energy Management Strategy for Fuel Cell Hybrid Electric Vehicle Based on Soft Actor-Critic Algorithm

Weiqi Chen[1], Jiaxuan Zhou[1], Chunhai Wang[2], Xinfu Pan[3], Xinwei Fan[3], Jiankun Peng[1*]

1 School of Transportation, Southeast University, Nanjing 211102, China
2 Sky-well New Energy Automobile Group Co., Ltd, Nanjing 211102, China
3 CATARC Automotive Proving Ground Co., Ltd, Yancheng 224100, China

ABSTRACT

Energy management strategy plays an important role in improving fuel economy and prolonging life time for fuel cell hybrid electric vehicle. To keep charge margin and reduce overall driving cost which consists of fuel consumption and health degradation of power battery and fuel cell, this paper proposes a novel energy management strategy based on Soft Actor-Critic, a fully-continuous deep reinforcement learning algorithm. Numerous simulation experiments manifest that the proposed method can obtain excellent balance between charge-keeping and money-saving both in charge depleting and charge sustaining modes. Results suggest that running FCHEV in low charge for long time should be avoided.

Keywords: energy management, fuel cell, hybrid electric vehicle, state of health, soft actor-critic, multiple objective optimization

NONMENCLATURE

Abbreviations	
EMS	Energy Management Strategy
FCHEV	Fuel Cell Hybrid Electric Vehicle
DRL	Deep Reinforcement Learning
SAC	Soft Actor-Critic
SOC	State of Charge
SOH	State of Health
Symbols	
m	Total mass of vehicle
g	Gravity acceleration
a	Acceleration of vehicle
v	Velocity of vehicle
P_{req}	Requested power of motor
P_{bat}	Power of battery pack
P_{FCS}	Output power of fuel cell system
C_n	Nominal capacity of battery pack

1. INTRODUCTION

In context of global fossil energy crisis and demand for energy conservation and emission reduction, more and more automobile manufacturers turn their attention to hybrid electric vehicles, electric vehicles, and fuel cell hybrid electric vehicles [1]. FCHEV has advantages of no greenhouse gas emission, simple utilization, quiet operation and high efficiency [2], but suffers from low dynamic characteristics of output power [3]. Thus, a pack of lithium-ion power battery is equipped onboard as the auxiliary energy source to provide peak power and recover brake energy. However, the two energy sources are different in working characteristics, attenuation conditions etc [4]. Therefore, it is of great significance to develop an appropriate and effective energy management strategy for FCHEV to maximize their great economic potential.

As a key automotive technology, EMS plays an important role in the distribution of energy from fuel cell and battery pack, and thus leads to proper operation condition and reduction in running cost. The current EMS for FCHEV can classified as three types [5]: rule-based, optimization-based and learning-based.

The advantages of rule-based EMS are simplicity of design, ease of implementation, and low burden on computation. However, the design of rules relies heavily on engineering experience and comprehensive expertise, and cannot guarantee optimal performance [6]. Therefore, optimization-based EMS became one of the research focuses, and can be divided into global optimization and instantaneous optimization [7].

As a classical global optimization algorithm, dynamic programming (DP) is utilized to develop off-line EMS under the premise of knowing whole driving cycle in advance, but it is mainly used as benchmark due to heavy computation burden and global optimality [8]. With the study of on-line optimal control method, Pontryagin's minimum principle (PMP) and model predictive control (MPC) have been employed to develop instantaneous

* Corresponding author e-mail: jkpeng@seu.edu.cn

optimization-based EMS. Ouyang et al. [9] implemented a PMP-based EMS to reduce hydrogen consumption by 5.9% per 100 km. Xiaosong H et al. [10] proposed a MPC framework to minimize running cost of FCHEV. The EMS based on instantaneous optimization method have no heavy computational burden, and are able to be implemented in real-time application. However, instantaneous optimization is not equal to overall optimization, and the global optimal performance cannot be guaranteed [11].

In recent years, reinforcement learning-based EMS have been very popular for HEV with internal combustion engine, due to near-optimal performance and on-line application capacity [12]. But there are very few relevant reports about learning-based EMS for FCHEV. The first time that deep reinforcement learning is implemented in EMS of FCHEV is reported in [11]. They used deep Q-Network, and discretized power change of FCS into nine values as control actions, which brought two problems. The discretization process reduced control accuracy, and huge action space increased computation burden and may even cause dimensionality curse.

This paper is devoted to bridge the aforementioned research gaps and proposed a novel EMS for FCHEV based on Soft Actor-Critic, which is a well-proved fully continuous DRL method. By designing reward function, health performance of both power battery and fuel cell, fuel economy, and charge margin are taken into consideration to develop an optimal control strategy. The rest of this paper is organized as follows: section 2 describes powertrain configuration, fuel cell model, and power battery model; section 3 introduces SAC algorithm and design details of EMS; simulation results are analyzed in section 4; section 5 concludes this paper.

2. PAPER STRUCTURE

2.1 Model Description

2.1.1 Powertrain of FCHEV

The research object in this paper is a fuel cell hybrid electric bus, which is driven by an electric motor with peak power of 120 kW. As shown in *Fig. 1*, the power of motor comes from two parts: the fuel cell engine and the power battery pack. Main configuration of powertrain is listed in *Table 1*. The overall power demand P_{req} is:

$$\begin{cases} F = mgf\cos\vartheta + mg\sin\vartheta + \frac{AC_D v^2}{21.15} + ma \\ P_{req} = F \cdot v = T_{mot} \cdot W_{mot} \\ P_{req} = (P_{bat} + P_{FCS}) \cdot \eta_{inv} \end{cases} \quad (1)$$

where f is rolling resistance coefficient, ϑ denotes road slope, C_D is air resistance coefficient, and A is front window area. T_{mot} and W_{mot} denote torque and speed of traction motor respectively, and η_{inv} is efficiency of inverter. The traction motor is modeled by quasi-steady state method, and the efficiency map is illustrated in *Fig. 2*.

Table. 1 Main configuration of vehicle

Items	Parameters	Value
Vehicle	Curb weight	14500 kg
	Rolling resistance coefficient	0.0085
	Tire radius	0.466 m
	Air resistance coefficient	0.55
	Front windward area	8.16 m^2
	Velocity	[0, 69] km/h
	Acceleration	[-1.5, 0.7] m/s^2
Motor	Peak power	200 kW
	Efficiency	[0.85, 0.97]
FCS	Peak power	60 kW
DC-DC converter	Peak power	60 kW
	Efficiency	[0.90, 0.95]
Battery	Capacity	108.14 kWh

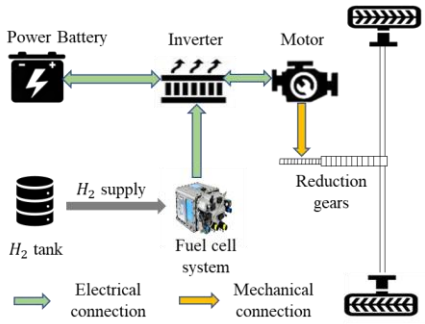

Fig. 1 Powertrain structure

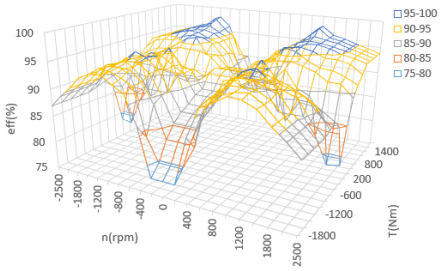

Fig. 2 Motor efficiency map

2.1.2 Fuel cell model

As the main power source of FCHEV, fuel cell system converts chemical energy of hydrogen and oxygen into electrical energy through electrochemical reaction. This paper uses physical and empirical model by considering physical laws and operating conditions. The hydrogen consumption rate of fuel cell stack can be calculated [13]:

$$\dot{m} = \frac{P_{FCS}}{\eta_{FCS} \cdot L_v} \quad (2)$$

where L_v is hydrogen lower heating value equaling to 120 kJ/g, and η_{FCS} is the efficiency of fuel cell stack.

The relationships between power P_{FCS} and hydrogen consumption rate $\dot{m}$ and efficiency η_{FCS} are illustrated in *Fig. 3*.

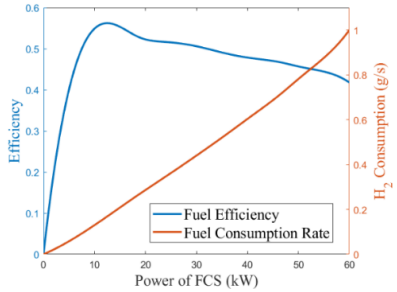

Fig. 3 H_2 consumption rate and efficiency map

Fuel cell degradation is mainly caused by four kinds of unfavorable driving conditions: load changing cycles, start-stop cycles, low-power load, and high-power load. Based on the contributions of Song et.al [14], discrete expressions for fuel cell degradation are as follows:

$$D_{FC} = \sum_{t=0}^{n} \left[d_{ss}(t) + d_{low}(t) + d_{high}(t) + d_{cha}(t) \right] \quad (3)$$

where D_{FC} (%) is the total performance degradation of fuel cell system, *n* is the number of time steps. d_{ss}, d_{low}, d_{high}, d_{cha} are the performance degradation caused by start-stop cycles, low-power load, high-power load, and load changing cycles at moment *t* respectively. Their accurate calculation method can be found in [15].

2.1.3 Power Battery Model

As the other energy source of FCHEV, the power battery pack is mainly utilized to provide peak power and store excess power. The equivalent circuit model is used to simulate battery pack [16]:

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

where V_{OC} is the open circuit voltage, I is load current, R is the internal resistance, SOC_0 is the initial value of SOC. *Fig. 4* describes the characteristics of power battery pack.

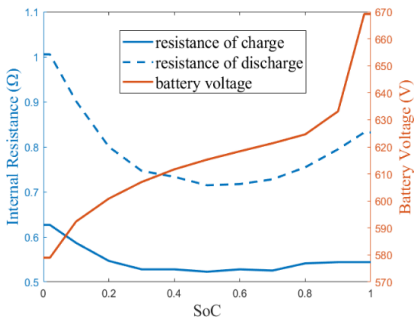

Fig. 4 Battery characteristic

The energy-throughput model [17] is adopted to evaluate performance degradation of power battery pack. The attenuation of SOH under multi-stresses is [18]:

$$\Delta SOH_t = -\frac{|I_t| \Delta t}{2N(c,T) C_n} \quad (5)$$

where N is the total number of cycles before the battery failure, and Δt is current duration. T is battery internal temperature which is assumed to be constant due to appropriate thermal management system. The C-rate (c) has significant impact on capacity loss, hence the Arrhenius equation is given as follows:

$$\Delta C_n = B(c) \cdot \exp\left(-\frac{E_a(c)}{RT} \right) \cdot Ah^z \quad (6)$$

where $\Delta C_n(\%)$ is loss of capacity, B denotes pre-exponential factor which is dependent on C-rate. Its value can be referred to *Table 2*. R is ideal gas constant, z is power-law factor, 0.55. Ah is the accumulated ampere-hour throughput, and $E_a(c)$ is the activation energy defined by:

$$E_a(c) = 31700 - 370.3 \cdot c \quad (7)$$

The life end of power battery is reached when its capacity drops by 20%, Ah and N can be derived as:

$$Ah(c) = \left[\frac{20}{B(c)} \cdot \exp\left(-\frac{E_a(c)}{RT} \right) \right]^{1/z} \quad (8)$$

$$N(c) = 3600 \cdot Ah(c,T) / C_n \quad (9)$$

Finally, the degradation can be calculated by Eq. (5).

Table. 2 Reference value of B(c)

c	0.5	2	6	10
B(c)	31630	21681	12934	15512

2.2 Health-Aware Ems Based on SAC

2.2.1 Soft actor-critic algorithm

The SAC is one of the most popular off-policy DRL methods with soft policy iteration. It is based on actor-critic framework, in which the actor network output a stochastic policy to enhance exploration. Unlike original actor-critic architecture, the SAC agent maximizes the information entropy of actions apart from conventional cumulative rewards. The state-action value function is given by the soft Bellman equation [19]:

$$Q(s_t, a_t) = r_t + \gamma E_{s_{t+1}, a_{t+1}} \begin{bmatrix} Q(s_{t+1}, a_{t+1}) \\ -\alpha \log(\pi(a_t | s_t)) \end{bmatrix} \quad (10)$$

where s_t, a_t, r_t are the state, action, reward with respect to step t respectively, and s_{t+1}, a_{t+1} are the state and action after state transition. γ is discount factor, and E denotes mathematical expectation. α is the temperature factor to adjust the relative importance of the entropy term versus the reward, and it is tuned automatically through neural network. $\pi(a_t | s_t)$ is the policy to be learned, and the optimal policy is defined as:

$$\pi^* = \arg\max_{\pi} \sum_t E_{(s_t,a_t)\sim\rho_\pi}[r_t - \alpha \log(\pi(a_t|s_t))] \quad (11)$$

Neural networks are employed to approximate the Q-value function, and the policy can be modeled as a Gaussian distribution with mean and covariance given by neural networks. Thus, the actor and critic can be optimized by stochastic gradient descent during back propagation. The Q-value function parameters θ can be trained to minimize the soft Bellman residual:

$$J_Q(\theta) = E_{(s_t,a_t,r_t,s_{t+1})\sim M}$$

$$\frac{1}{2}\left[Q(s_t,a_t) - \left[r_t + \gamma \begin{pmatrix} Q'^{(s_{t+1},\pi(s_{t+1}))} \\ -\alpha \log(\pi(a_{t+1}|s_{t+1})) \end{pmatrix}\right]\right]^2 \quad (12)$$

where M is experience replay pool, and (s_t, a_t, r_t, s_{t+1}) are minibatches sampled from it randomly. The target critic network Q' is utilized to accelerate and stabilize training process, and its parameter θ' is updated softly:

$$\theta' \leftarrow (1-\tau)\theta' + \tau\theta \quad (13)$$

where τ is step factor to control update amplitude.

The policy network parameters ϕ is updated by minimizing:

$$J_\pi(\phi) = E_{s_t\sim M}\left[E_{a_t\sim\pi_\phi}\begin{bmatrix}\alpha \log(\pi(a_t|s_t)) \\ -Q(s_t,a_t)\end{bmatrix}\right] \quad (14)$$

Temperature factor α is regulated automatically, its gradients is computed with the following objective:

$$J(\alpha) = E_{E_{a_t\sim\pi_t}}[-\alpha log\pi_t(a_t|s_t) - \alpha\bar{H}] \quad (15)$$

where target entropy $\bar{H}$ is the opposite number of action dimension, i.e., -1 in this paper.

2.2.2 States and actions

The states vector contains important information to decision-making, and is inputted into Q-value function and policy network. It is defined as:

$$s = [SOC, SOH, a, v, P_{FCS}] \quad (16)$$

While the continuous control action is output power of fuel cell engine: $P_{FCS} \in [0,60]kW$.

2.2.3 Reward function

At each moment t, the SAC agent observes current state s_t, then executes the action a_t from the policy network, and obtains a numeric reward r_t from the environment. Afterwards, the interactive scene steps into next state. A fine-designed reward function is of great significance to guide agent to learn optimal policy.

There are three primary optimization objectives of energy management: 1) save fuel; 2) reduce degradation of fuel cell and power battery; 3) keep SOC margin. Therefore, the reward function is designed as follows:

$$r_t = -\begin{bmatrix}\beta_1\dot{m}(t) + \beta_2 D_{FC}(t) + \beta_3\Delta SOH(t) \\ +\omega[SOC(t) - SOC_{tar}]\end{bmatrix} \quad (17)$$

where $\beta_1, \beta_3, \beta_3$ are H_2 price, replacement cost of fuel cell system, and replacement cost of power battery pack respectively. And $\dot{m}(t)$, $D_{FC}(t)$, $\Delta SOH(t)$ denote H_2 consumption, fuel cell degradation, power battery degradation at time step t respectively. The weight coefficient ω determines the relative importance of the money cost versus SOC value. SOC_{tar} is target value of SOC, and is dependent on charge mode as follows:

$$SOC_{tar} = \begin{cases} SOC_0 - 0.2, charge\ depleting\ mode \\ SOC_0, charge\ sustaining\ mode \end{cases} \quad (18)$$

2.3 Simulation Results

Since there is an appropriate value of ω waiting for exploration, numerous simulation experiments were executed in this chapter firstly to obtain balance of money cost versus SOC value in charge sustaining (CS) mode. Then, the previously determined ω is tested in charge depleting (CD) mode. Note that the experiments were implemented under the China typical urban driving cycle (CTUDC) as shown in *Fig. 5*.

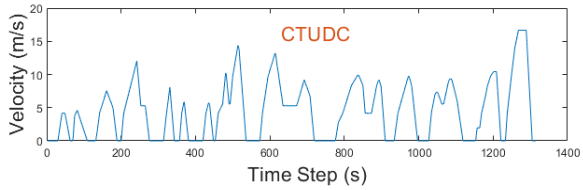

Fig. 5 China Typical Urban Driving Cycle

2.3.1 Exploration of weight coefficient

Given the CTUDC and vehicle configurations, the requested power curve of vehicle can be calculated at each moment as shown in *Fig. 6*. And it is the energy to be managed by the proposed method.

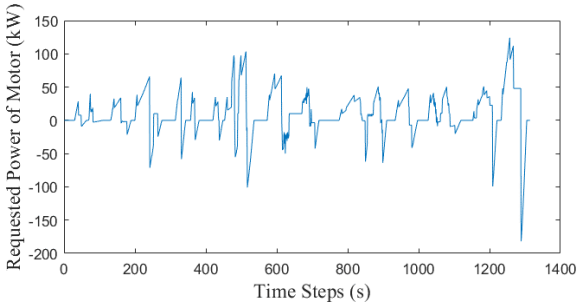

Fig. 6 Requested power curve of the bus

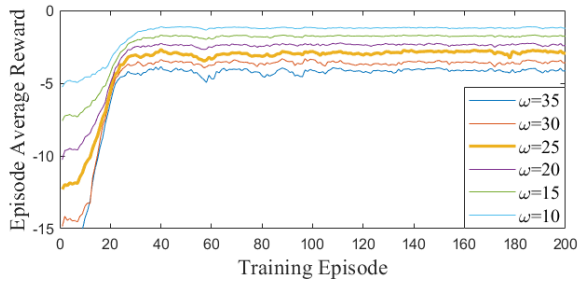

Fig. 7 Reward curves of different ω

The core idea of deep reinforcement learning is to guide the agent to learn a policy with the maximum

expectation of discounting reward. Thus, the reward curve can indicate the training performance. As shown in *Fig. 7*, regardless of the ω value, the proposed strategy can converge rapidly and stably.

The goal of energy management is to reduce the overall driving cost as much as possible. Unlike other types of HEV with internal combustion engine, the performance of fuel cell system onboard vehicle is easier to degrade and more expensive, which should be considered into EMS. Thus, the overall driving cost consists of three parts: hydrogen consumption, power battery degradation, and fuel cell degradation. The price of hydrogen is 55 RMS per kilograms, the replacement cost of power battery pack and fuel cell stack are 20000 RMB and 300000 RMB respectively.

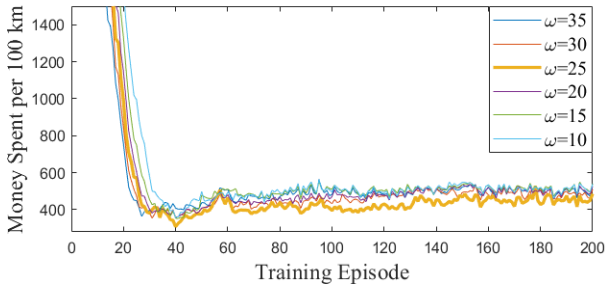

Fig. 8 Money spent per 100km of different ω

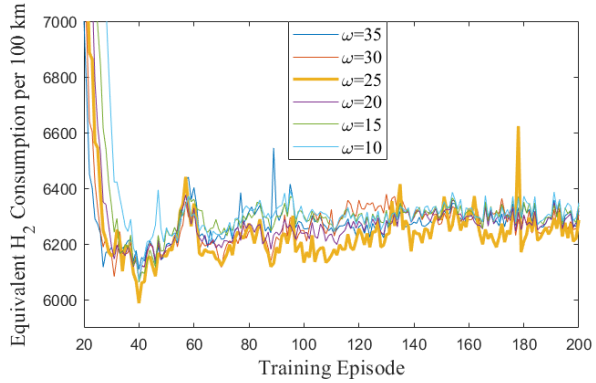

Fig. 9 Equivalent H_2 cost per 100km of different ω

Table 3 Comparison of different ω (SOC_0=0.5)

ω	Final SOC	H_2 Cost (g)	Battery Pack SOH	Fuel Cell System SOH	Money Cost (￥/100km)
10	0.2773	6071.1	0.999865	0.9999496	356.52
15	0.2747	6063.4	0.999866	0.9999498	353.44
20	0.2753	6065.0	0.999867	0.9999490	357.77
25	0.2606	5988.8	0.999868	0.9999570	308.22
30	0.2775	6076.3	0.999865	0.9999488	360.15
35	0.2856	6109.9	0.999821	0.9999435	408.17

Fig 8 illustrates that the most economic policy can be learned when $\omega = 25$, and Fig. 9 supports this point with the minimal equivalent hydrogen consumption per 100 km curve. As we can see in *Table 3*, with the initial SOC equaling to 0.5, when $\omega = 25$, the equivalent hydrogen consumption is 5988.8g, which saves 2% compared to that of $\omega = 35$; the overall money spent is 308.22 RMB, which is 75.5% of that when $\omega = 35$. Both the least equivalent hydrogen and overall money spent manifest that a near-optimal policy is realized under current circumstance, and 25 is determined as the preferred value of ω in later experiments.

2.3.2 Charge depleting mode

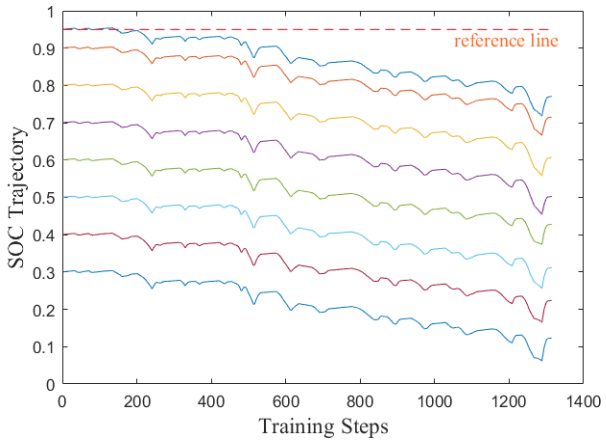

Fig. 10 SOC trajectory of different initial SOC, CD mode

Table 4 Comparison of different SOC_0 (ω=25, CD mode)

SOC_0	SOC consumption	Money Cost (￥/100km)
0.95	0.1793	323.86
0.9	0.1849	314.99
0.8	0.1932	331.58
0.7	0.1978	330.88
0.6	0.1721	355.42
0.5	0.1881	308.22
0.4	0.1766	359.62
0.3	0.1771	484.65
Mean	0.1836	351.15

Continue Table 4

SOC_0	H_2 Cost (g)	Battery Pack SOH	Fuel Cell System SOH
0.95	6043.6	0.999904	0.9999526
0.9	6014.7	0.999898	0.9999542
0.8	6048.4	0.999892	0.9999519
0.7	6032.3	0.999886	0.9999522
0.6	6083.6	0.999880	0.9999498
0.5	5988.8	0.999868	0.9999570
0.4	6051.5	0.999838	0.9999503
0.3	6164.0	0.999673	0.9999396
Mean	6053.4	0.999855	0.9999508

Charge depleting mode is a very common operation mode during driving when SOC is sufficient. Since the weight coefficient is determined, set the initial value of SOC as [0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3] respectively, and other configurations are the same. *Fig. 10* shows SOC trajectories of different initial SOC values in CD mode, from where we see steady and uniform decrease of

charge. *Table 4* shows SOC cost and money cost in different initial SOC. The average money cost per 100 km is 351.15 RMB, and average equivalent hydrogen consumption is 6053.4 g, and average SOC cost is 0.1836.

As we can see in *Table 4*, the most economical driving strategy is when SOC_0 is 0.95 and 0.9, and the money spent increase gradually with the reduce of initial SOC. This is mainly because that the EMS prefers to use fuel cell system to drive the bus when SOC is low, and the replacement cost of fuel cell stack is much more expensive than that of power battery pack. This phenomenon suggests that the power battery system should be preferred to drive the vehicle when the charge is sufficient.

2.3.3 Charge sustaining mode

The hybrid electric bus should maintain certain charge margin during driving to cope with emergencies such as running out of fuel, especially when SOC is low. Thus, EMS performance in CS mode should also be taken into consideration. Keep the configuration as previous experiments, and modify SOC reward function in Equation (18). Figures in *Fig. 11* shows that the proposed strategy can maintain the SOC value around initial value whatever SOC_0 is. The average consumption of SOC in CS mode is -0.0125, which also means that the SOC is kept close to its initial value.

However, the average money cost per 100 km is 858.39 RMB, which is 2.44 times as much as CD mode, and equivalent hydrogen consumption is 7373.2g, which is 1.22 times as much as CD mode does, as the data listed in *Table 5*. This is because that fuel cell system is used intensively in order to maintain charge margin while satisfying the power demand of driving. And this leads to more hydrogen consumption and more serious degradation of fuel cell system which is much more expensive. And the SOH of power battery and fuel cell both performs worser than that of CD mode.

These results suggest that running the FCHEV in low charge mode for long time should be avoided, as this can cause much more degradation of both power battery and fuel cell, and thus increase the cost of operation and maintenance.

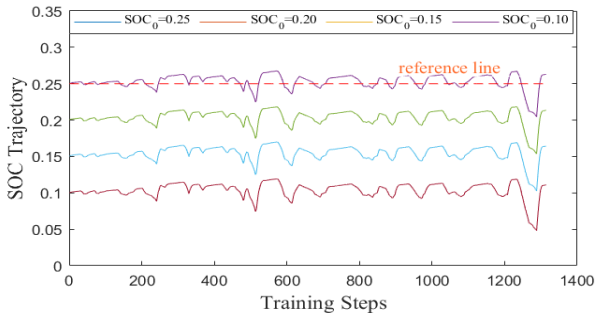

Fig. 11 SOC trajectory of different initial SOC, CS mode

Table 5 Comparison of different SOC_0 (ω=25, CS mode)

SOC_0	SOC consumption	Money Cost (￥/100km)
0.25	-0.0125	815.52
0.2	-0.0129	837.36
0.15	-0.0137	877.82
0.1	-0.0107	902.89
Mean	-0.0125	858.39

Continue Table 5

SOC_0	H_2 Cost (g)	Battery Pack SOH	Fuel Cell System SOH
0.25	7383.6	0.999631	0.9999090
0.2	7408.2	0.999568	0.9999089
0.15	7357.2	0.999484	0.9999069
0.1	7343.9	0.999445	0.9999045
Mean	7373.2	0.999532	0.9999073

2.4 Conclusions

A novel health-aware energy management strategy for FCHEV base on SAC is proposed for the first time in this paper. Keeping charge margin and reducing overall driving cost are the two goals in the multi-objective optimization problem, where overall driving cost consists of hydrogen consumption and health degradation of both lithium-ion power battery and fuel cell system. After numerous explorations for weight coefficient, the trained strategy performs well both in charge sustaining and charge depleting modes. The main conclusions are as follows:

(1) The health state of power battery pack and fuel cell system should be taken into consideration, due to their disadvantages of easy degradation and expensive cost.

(2) Under simulation condition, the average money costs per 100 km are 351.15 RMB in CD mode and 858.39 RMB in CS mode respectively; and the average hydrogen consumption are 6053.4 g in CD mode and 7373.2 g in CS mode respectively.

(3) Simulation results suggest that the power battery system should be preferred to drive the FCHEV when the charge is sufficient, and avoid running the vehicle in low charge for long time.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China (Grant No.52072074), the Fundamental Research Funds for the Central Universities (Grant No.2242021R40007), Jiangsu Province Technology Project (Grant No.BE2021067) and cooperative project of CATARC Automotive Proving Ground Co., Ltd.

REFERENCE

[1] Hu X, Zhang X, Tang X, Lin X. Model predictive control of hybrid electric vehicles for fuel economy, emission reductions, and inter-vehicle safety in car-following scenarios. Energy 2020; 196:117101.

[2] He H, Jia C, Li J. A new cost-minimizing power-allocating strategy for the hybrid electric bus with fuel cell/battery health-aware control[J]. International Journal of Hydrogen Energy, 2022.

[3] Quan S, Wang Y X, Xiao X, et al. Real-time energy management for fuel cell electric vehicle using speed prediction-based model predictive control considering performance degradation[J]. Applied Energy, 2021, 304: 117845.

[4] Song K, Chen H, Wen P, Zhang T, Zhang B, Zhang TJEA. A comprehensive evaluation framework to evaluate energy management strategies of fuel cell electric vehicles. 2018. p. 960e73. 292.

[5] Kandidayeni M, Trovão J P, Soleymani M, et al. Towards health-aware energy management strategies in fuel cell hybrid electric vehicles: A review[J]. International Journal of Hydrogen Energy, 2022.

[6] Min D, Song Z, Chen H, et al. Genetic algorithm optimized neural network based fuel cell hybrid electric vehicle energy management strategy under start-stop condition[J]. Applied Energy, 2022, 306: 118036.

[7] Hu X, Han J, Tang X, et al. Powertrain design and control in electrified vehicles: A critical review[J]. IEEE Transactions on Transportation Electrification, 2021, 7(3): 1990-2009.

[8] 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]. Applied Energy, 2017, 185: 1633-1643.

[9] Xu L, Ouyang M, Li J, et al. Application of Pontryagin's Minimal Principle to the energy management strategy of plugin fuel cell electric vehicles[J]. International Journal of Hydrogen Energy, 2013, 38(24): 10104-10115.

[10] Xu 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 transactions on power electronics, 2019, 35(1): 382-392.

[11] 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.

[12] Wu J, He H, Peng J, et al. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus[J]. Applied energy, 2018, 222: 799-811.

[13] Lin W S, Zheng C H. Energy management of a fuel cell/ultracapacitor hybrid power system using an adaptive optimal-control method[J]. Journal of Power Sources, 2011, 196(6): 3280-3289.

[14] 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.

[15] Wang Y, Advani S G, Prasad A K. A comparison of rule-based and model predictive controller-based power management strategies for fuel cell/battery hybrid vehicles considering degradation[J]. International Journal of Hydrogen Energy, 2020, 45(58): 33948-33956.

[16] 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 Transactions on Transportation Electrification, 2022.

[17] Ebbesen S, Elbert P, Guzzella L. Battery state-of-health perceptive energy management for hybrid electric vehicles[J]. IEEE Transactions on Vehicular technology, 2012, 61(7): 2893-2900.

[18] Wu J, Wei Z, Liu K, et al. Battery-involved energy management for hybrid electric bus based on expert-assistance deep deterministic policy gradient algorithm[J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 12786-12796.

[19] Haarnoja T, Zhou A, Hartikainen K, et al. Soft actor-critic algorithms and applications[J]. arXiv preprint arXiv:1812.05905, 2018.