Energy 312 (2024) 133384

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Deep reinforcement learning-based energy management system enhancement using digital twin for electric vehicles

Yiming Ye [a], Bin Xu [b], Hanchen Wang [b], Jiangfeng Zhang [a,*], Benjamin Lawler [a], Beshah Ayalew [a]

[a] Department of Automotive Engineering, Clemson University, Greenville, SC, 29607, USA
[b] School of Aerospace and Mechanical Engineering, The University of Oklahoma, Norman, OK, USA

ARTICLE INFO

Handling editor: X Ou

Index Terms:
Reinforcement learning
Digital twin
Energy management
Electric vehicle

ABSTRACT

Compared to conventional engine-based powertrains, electrified powertrain exhibit increased energy efficiency and reduced emissions, making electrification a key goal for the automotive industry. For a vehicle with a hybrid energy storage system, its performance and lifespan are substantially affected by the energy management system. Reinforcement learning-based methods are gaining popularity in vehicle energy management, but most of the literature in this area focuses on pure simulation, while hardware implementation is still limited. This paper introduces the digital twin methodology to enhance the reinforcement learning-based energy management system for battery and ultracapacitor electric vehicles. The digital twin model can exploit the bilateral interdependency between the virtual model and the actual system, which improves the control performance of the energy management system in real-time control. The physical model is established based on a hardware-in-the-loop simulation platform. In addition, battery degradation is also considered for prolonging the battery lifespan to reduce operating costs. The validation results of the trained reinforcement learning agent illustrate that the digital twin-enhanced Q-learning energy management system improves the energy efficiency by 7.08 % and reduces the battery degradation by 25.28 %.

1. Introduction

With the growing environmental and energy crises, stricter regulations on vehicle exhaust emissions are legislated, which will reduce vehicle emissions to around 100 g CO_2/km in 2025 [1]. Europe is the most aggressive region in setting the emissions target, with 95 g CO_2/km in future and further reductions of 15 % in 2025 and 37.5 % in 2030 [1]. To meet these emission reduction targets, vehicle electrification is one of the main technical routes of automotive manufacturers, where these electrified vehicles can be categorized into three groups: hybrid electric vehicles (HEV), plug-in hybrid electric vehicles (PHEV), and battery electric vehicles (BEV) [2–4].

One essential element of the electrified propulsion system in the above three groups of vehicles is the energy storage system (ESS). In the automotive industry, the most commonly used ESS is the lithium-ion battery (LIB) for its high energy density. However, the limited cycle life and power density weaken the performance of BEV [5]. On the contrary, the ultracapacitor (UC) has a significantly longer cycle life and a higher power density but a much lower energy density [6]. Another well-researched ESS is fuel cell. Although a fuel cell has a moderate power density and high energy density, it has the highest capital cost among the three ESSs described above [6]. Problems with hydrogen storage also restrict the application of fuel cells. Therefore, an ideal ESS that could satisfy all application needs does not exist so far. An increasingly common choice for balancing trade-offs between cycle life, cost, power density, and energy density is a hybrid energy storage system (HESS). The corresponding energy management strategy (EMS) is designed to take advantage of various ESSs while minimizing their drawbacks in order to achieve optimal energy management at the system level. The EMS coordinates the output power of different sources to satisfy the driver's demand under various driving situations. Ultimately, it will maximize energy efficiency and minimize battery degradation at the same time.

Most recently, reinforcement learning (RL) has attracted attention in the community of EMSs. The agent in the RL algorithm aims to find control variables that maximize cumulative reward. In the control problem of EMS, the environment consists of various factors that affect the vehicle's operating statuses, such as velocity, acceleration, and

* Corresponding author.
E-mail addresses: yimingy@clemson.edu (Y. Ye), binxu@ou.edu (B. Xu), haw110@ou.edu (H. Wang), jiangfz@clemson.edu (J. Zhang), bjlawle@clemson.edu (B. Lawler), beshah@clemson.edu (B. Ayalew).

https://doi.org/10.1016/j.energy.2024.133384
Received 1 May 2024; Received in revised form 1 August 2024; Accepted 5 October 2024
Available online 10 October 2024

powertrain dynamics. The RL agent is a power-split controller for the BEV equipped with HESS, which controls the LIB and UC's output power. The objective is to find a control sequence that minimizes battery degradation while maximizing energy efficiency. There exists some research on RL-based EMSs for electrified vehicles. Reference [7] adopts a temporal difference learning method to develop an optimal control for HEV, which has a relatively higher convergence rate and also achieves better performance under an environment without Markovian property. A data-driven RL EMS for the PHEV is proposed in Ref. [8], which allows charging within the trip to obtain a near-optimal result. Reference [9] designs an RL-based EMS combined with a geographic information system for PHEV, which aims to minimize the usage of batteries and the fuel consumption of an internal combustion engine. The trip distance is also taken as a state in the RL environment, and it shows the high correlation between the remaining driving distance and the total energy consumption [9]. To reduce energy usage and battery aging, reference [10] designs a method based on Q-learning. The proposed technique extends the driving range by 1.5–2% while delaying battery degradation by 13–20 % compared to a rule-based strategy. Model predictive control and RL are combined to establish EMS for an HEV, where the Q-learning algorithm is utilized to solve the energy management problem, and the speed predictor is created by fuzzy encoding and nearest neighbor prediction by historic driving information [11]. An RL-based EMS for HESS is developed with a transition probability matrix, which is evolved through the Kullback-Leibler divergence rate and forget factor [12]. A Q-learning-based EMS is established for HEV, which utilizes cloud computing to alleviate the computational burden of online training [13]. Reference [14] combines fuzzy logic control with the Q-learning algorithm to develop EMS for an HEV, which tunes the fuzzy logic parameters based on the Q value and adopts a neural network for the action-value function estimation. Additionally, a fuzzy logic controller and a RL algorithm are used to provide an online prediction EMS for a hybrid electric tracked vehicle [15], which minimizes the effects of prediction error. Reference [16] proposes an online RL-based EMS for HEV by Q-learning, which takes the weighted values of gas consumption and electric energy from the battery to form the reward function for the RL algorithm. Besides the table-based RL algorithms, network-based RL algorithms driven by deep neural networks are popular in EMS study. Reference [17] proposes a deep Q-network (DQN) based EMS for PHEV, indicating that dueling DQN converges faster. A hierarchical deep Q-learning algorithm (DQL-H) for HEVs is proposed that solves sparse reward problems and achieves optimal power distribution, resulting in better training efficiency and lower fuel consumption [18]. A deep Q-learning based EMS for hybrid battery systems in electric vehicles has been proposed, demonstrating improvements in energy loss reduction and safety enhancement compared to traditional methods [19]. A distributed deep reinforcement learning-based EMS is proposed by using DQN, asynchronous advantage actor-critic (A3C), and distributed proximal policy optimization, showing improved training efficiency and fuel economy [20]. A deep deterministic policy gradient (DDPG) approach has been presented that integrates expert knowledge to accelerate learning and improve fuel economy in power-split HEVs [21]. The DDPG has been introduced into the EMS to eliminate discretization errors [22]. The DDPG has been enrolled to establish the deep reinforcement learning (DRL) EMS, but the overestimation compensates for the performance. A twin delayed DDPG (TD3) based EMS has been proposed for electrified vehicle [23]. The results indicate that TD3 method has made better control performance than the ECMS-based EMS [23]. Reference [24] proposes the soft actor-critic DRL-based EMS for the hybrid electric bus to manage the energy split of different power sources, where the comparative study between the DQN-based and soft actor-critic-based EMS shows that the soft actor-critic EMS has a faster convergence speed and better optimization performance.

To improve the performance of electric vehicle (EV), the digital twin technology has been introduced into EMS study. Digital twin is a digital representation or model of a physical object, system, or process. It is designed to simulate the behavior and performance of the real-world counterpart, integrating real-time data from sensors and IoT devices. Digital twins enable simulation, analysis, and optimization of the physical systems they represent. The digital twin methodology gained recognition in a NASA program, where the digital twin served as a panoramic mapping of a physical body in the digital world [25]. Reference [26] proposed a performance digital twin simulation model of pure battery electric vehicles to predict and study the effects of different parameters on the performance characteristics of electric vehicles. Reference [27] introduced a digital twin model based on the conventional model in the dimension of temperature-energy consumption, which is used to predict the energy consumption of electric vehicles and verify the feasibility of the method. Reference [28] shows that digital twins can serve as effective soft-sensors for electric motors in the automotive industry, providing real-time monitoring and diagnosis while reducing the complexity and cost associated with traditional sensors. The study shows that digital twin technology, when integrated with a cloud-based battery management system, can provide accurate and reliable online state of charge and state of health estimation for battery systems; thus it has the potential to improve the efficiency, safety, and lifespan of batteries in various applications [29]. Reference [26] also shows that digital twin technology can be used to improve the estimation of health indicators for DC-DC converters. This approach has the potential to enhance the reliability and efficiency of power electronic systems, enable predictive maintenance, and reduce downtime [30]. Reference [31] shows that digital twin technology can effectively simulate and analyze the performance of hydrogen fuel cell hybrid electric vehicles, providing valuable insights into the impact of control strategies on energy efficiency and the optimal design of EVs. Reference [32] demonstrates that electric propulsion system simulation can serve as a solid foundation for developing a digital twin of an electric vehicle. This approach can be used for the design, testing, and optimization of electric vehicles, contributing to the development of more efficient and reliable transportation solutions. The above evidences show that the digital twin technology benefits the EV in many aspects.

Although DRL-based EMSs have made progress in the control of electrified vehicles, a significant barrier still stands in the way of real-world applicability. Due to the required large number of iterations in the learning process, the learning time is not applicable in real-time control, and it is not well addressed in existing studies. As mentioned before, the DRL agent aims to plan its behavior based on the interactions with a dynamic environment on a trial-and-error basis, which is a low efficient learning method and requires considerable experience collected in the interaction. From existing literature, the table-based RL EMSs have iterations varied from 2000 to 300000 [8,33,34], and the neural network-based RL EMSs have iterations vary from 15000 to 150000 [17,18]. A large number of iterations puts a heavy computing strain on the system and necessitates months' worth of dyno and road testing, which is expensive in terms of both time and money. This obstacle impedes the DRL-based EMS from being applied in adaptive and real-time optimal control for the electric vehicle. To address this issue, this paper adopts the digital-twin technology to liberate the optimal control potential of DRL-based EMS to achieve higher energy efficiency and prolong the battery life. To the best of authors' knowledge, this is the first time that DRL has been integrated with the digital twin technology to establish the EMS for the electric vehicle equipped with battery and UC. The main contributions of this work are summarized as follows.

1) It proposes a DRL-based EMS for a BEV HESS through digital twin technology to achieve adaptive real-time control.
2) Both energy efficiency and battery degradation of the BEV HESS are integrated into the reward of the DRL.
3) The efficacy of the proposed method is verified through testing with a hardware-in-the-loop (HIL) platform.

The remaining part of this paper is organized as follows. Section II presents the modeling of vehicle dynamics and propulsion systems. The detailed methodology is shown in Section III. Next, Section IV illustrates a case study for the proposed method and comparison with other popular EMSs, where the HIL tests are also presented. Finally, the conclusion and discussion of future work are given in Section V.

2. Modeling

2.1. Vehicle model

When the vehicle is in motion, the balance equation of the total force is shown as follows:

$$F_{trac} = F_{inertia} + F_{roll} + F_{aero} + F_{grade} \tag{1}$$

where F_{grade} is the gradient resistance, F_{aero} is the air resistance, F_{trac} is the traction force, F_{roll} is the rolling resistance, and $F_{inertia}$ is the inertia force.

The traction force is calculated as follows:

$$F_{trac} = F_{mg} - F_{brake} \tag{2}$$

where F_{mg} is the driving force of the electric motor, F_{brake} is the deceleration force, which consists of regeneration braking force and mechanical braking force.

The drag of rolling resistance can be calculated through the rolling resistance coefficient $f_{roll}(V_{veh}, P_{tire}, \ldots)$, mass of vehicle M_{veh}, and road grade δ.

$$F_{roll} = f_{roll}(V_{veh}, P_{tire}, \ldots) M_{veh} g \cos \delta \tag{3}$$

where vehicle speed V_{veh}, tire pressure P_{tire}, and temperature are a few of the variables that affect the rolling resistance coefficient. To simplify the problem, f_{roll} is set as constant in this paper.

The aerodynamic resistance can be calculated by a function of air drag coefficient C_d, vehicle front area A_f, air density ρ_a, and vehicle velocity V_{veh}, which is shown in (4)

$$F_{aero} = \frac{1}{2} \rho_a A_f C_d V_{veh}^2 \tag{4}$$

The gradient resistance can be calculated through the function of the road grade δ and vehicle mass, as shown in (5).

$$F_{grade} = M_{veh} g \sin \delta \tag{5}$$

The power of the electric motor is calculated from the vehicle dynamics and shown as follows:

$$T_{mg} = \frac{F_{trac} d_w}{2 i_0} \tag{6}$$

$$\omega_{mg} = \frac{2 V_{veh}}{d_w} i_0 \tag{7}$$

$$\eta_{mg} = f(T_{mg}, \omega_{mg}) \tag{8}$$

$$P_{mg} = T_{mg} \omega_{mg} \eta_{mg} \hat{s}ign(T_{mg} \omega_{mg}) \tag{9}$$

$$P_{mg} = P_{cap} + P_{bat} \tag{10}$$

where T_{mg} is the EM output torque, d_w is the wheel diameter, i_0 is the gear ratio of the final reducer, ω_{mg} is the EM rotating speed, η_{mg} is the EM efficiency that is obtained by an appropriate EM look-up table that depends on torque and rotating speed values, P_{cap} and P_{bat} are the power supplied to EM from UC and LIB, respectively.

Table 1 lists the vehicle parameters used in this paper.

2.2. Propulsion system model

Fig. 1 depicts the design of the hybrid propulsion system.

The power source of the electric drive system consists of LIB and UC, which are connected to the power bus in parallel. The rear axle is driven by the electric motor through a main reducing gear. The driver sends the power request to the optimal controller, which regulates the motor drive to control the electric motor output torque to meet the drivers' request. The optimal controller needs to determine the ratio of output power from the LIB and the UC. In the electric drive system, the UC is connected to the power bus through a DC/DC converter that eliminates the negative influence of voltage fluctuations associated with UC. In addition, the LIB is connected to the power bus by the DC/DC converter too, which allows for the matching of the current flow demanded from the battery pack. The electric motor is connected to the power bus with the motor controller. The efficiency of the motor drive is set as 92 % [35].

2.3. Battery model

The battery model used in this paper consists of the battery dynamic model and battery degradation model. The zero-order equivalent model is used for battery dynamic. As shown in Fig. 2, a series connection of a resistor and an ideal voltage source is used to establish the zero-order equivalent circuit model to represent the battery.

The voltage and current of the power load are calculated as follows:

$$U_L = U_{oc} - I_L R_{bat} \tag{11}$$

$$I_L = \frac{U_{oc} - \sqrt{U_{oc}^2 - 4 R_{bat} P_{bat}}}{2 R_{bat}} \tag{12}$$

where U_L is the load voltage, I_L is the current, U_{oc} is the battery open-circuit voltage, and P_{bat} is the battery output power.

The ampere-hour integral method features minimal application costs, excellent computation efficiency, and straightforward hardware implementation. As a result, the ampere-hour integration approach is used in this study for the purpose of calculating the battery SOC, as shown in (13):

$$SOC(t) = SOC_0 - \int_0^t I_L(t) dt \Big/ Q_{bat} \tag{13}$$

where SOC_0 is the initial value of SOC, and Q_{bat} means the rated capacity of the battery.

The zero-order model is preferred for its simplicity and computational efficiency, providing quick estimates with minimal resources. The accuracy of the zero-order model was evaluated using test data, with the SOC trajectories compared with the second-order model shown in Fig. 3. The comparison shows that the two SOC curves calculated from zero-order and second-order models coincide. Thus, this zero-order model is adopted for better computational efficiency.

During the LIB degradation process, the battery SOC, current, temperature, and ampere-hour throughput will influence the LIB aging. According to Ref. [30], the LIB degradation process is captured through

Table 1
Parameters of the vehicle model.

Parameters	Value
Curb weight (kg)	1778
Max weight (kg)	2180
Windward area (m^2)	2.34
Air drag coefficient	0.30
Wheelbase (mm)	2870
Wheel diameter (mm)	693.7
Top speed (km/h)	200
0–100 km/h time (s)	8
Grade ability (%)	30

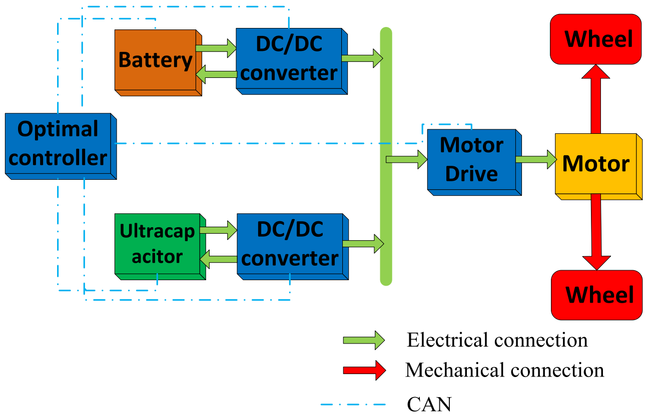

Fig. 1. The electric drive system diagram.

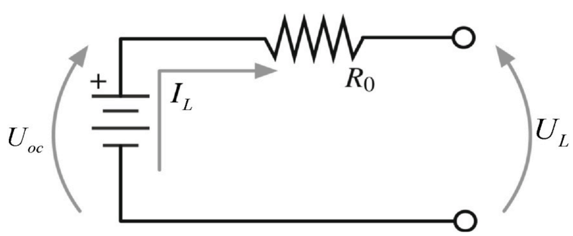

Fig. 2. The battery internal resistance model.

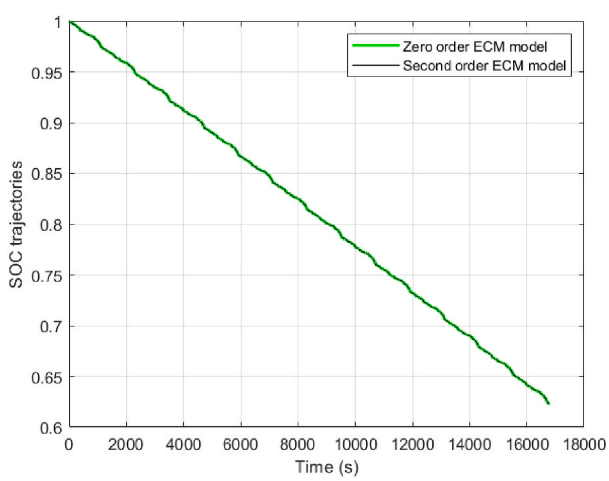

Fig. 3. Battery model comparison.

a semi-empirical model as follows:

$$Ah(n) = DOD(n)^{*}Ah \tag{14}$$

$$E_a(n) = 31500\text{-}152.5^{*}C\text{-}rate \tag{15}$$

$$B = \alpha^{*}SOC + \beta \tag{16}$$

$$Q_{C\text{-}rate}(n) = B(n)e^{\frac{-E_a(n)}{RT}}Ah^{z}(n) \tag{17}$$

where Ah is the ampere-hour throughput, n represents the n-th C-rate under consideration, $Ah(n)$ is the Ah throughput corresponding to the n-th C-rate, and $Q_{C\text{-rate}}(n)$ is the battery capacity loss caused by the n-th C-rate. α and β are defined for different SOC values as follows:

$$\begin{cases} \alpha = 2896.6, \beta = 7411.2, \textit{if} \quad SOC < 0.45 \\ \alpha = 2694.5, \beta = 6022.2, \textit{if} \quad SOC \geq 0.45 \end{cases} \tag{18}$$

The total LIB capacity degradation can be calculated as (19).

$$Q_{cycle} = \sum_{1}^{n} Q_{C\text{-rate}}(n) \tag{19}$$

In the long-run usage, the battery cycle is repeated on a daily basis. When the lithium-ion battery works during the i-th cycle, the Ah of each C-rate is accumulated from the 1st cycle to the i-th cycle, and the lithium-ion battery capacity degradation is calculated through the total Ah throughput of each C-rate. The final capacity at the end of each battery cycle is used as the capacity value at the starting point of the next battery cycle.

2.4. Ultracapacitor model

The output voltage and current of the UC are calculated through (20)–(21):

$$U_{UC} = U_{UC,OC} - I_{UC}R_{UC} \tag{20}$$

$$I_{UC} = \frac{P_{UC}}{U_{UC}} \tag{21}$$

where U_{UC} is the load voltage of UC, $U_{UC,OC}$ is the open-circuit voltage of UC, I_{UC} is the current of UC and R_{UC} is the resistance. The current can also be a function of SOC:

$$I_{UC} = C_{UC}U_{UC,max}SOC_{UC} \tag{22}$$

where C_{UC} is the capacity of UC, $U_{UC,max}$ is the maximum voltage of UC, and the SOC_{UC} is the SOC of UC. The SOC_{UC} can be expressed as follow:

$$SOC_{UC}(t) = SOC_{UC}(0) - \frac{\int_0^t I_{UC}(\tau)d\tau}{C_{UC}U_{UC,max}} \tag{23}$$

2.5. EM model

Fig. 4 displays the map of the EM efficiency. All EM-related losses are included in the efficiency map. The energy management system calculates the EM efficiency by coupling the EM torque demand with the rotational speed and sending the data into the efficiency map. The EM changes to a generator during the regenerative braking process to replenish the LIB and UC, which can increase energy efficiency.

2.6. DC/DC converter model

In this paper, the UC and LIB are integrated into the power system by DC/DC converter, respectively. The output power and current, which are depicted in Fig. 5, can be considered as a function of the efficiency of the DC/DC converter.

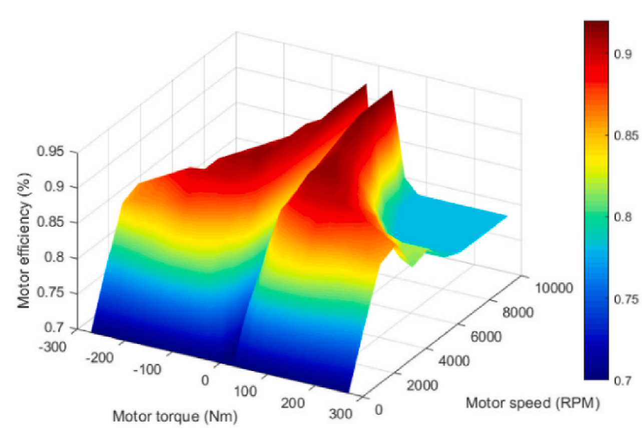

Fig. 4. Efficiency map of EM.

3. Methodology

3.1. Q-learning-based EMS

Q-leaning is a kind of off-policy temporal difference learning algorithm from reinforcement learning [36]. The temporal difference learning strategy integrates the features of Dynamic Programming and the Monte Carlo method. In the Q-learning EMS, the environment contains the electric vehicle plant model and any factors engaging in the driving process. The Q-learning agent will learn and update through interactions with the environment. The value function estimation can be updated in the training process by a temporal difference strategy. Therefore, the agent does not have to renew the optimal policy. According to the aforementioned features, the optimal selection of control strategy can be extracted from the value function of each action when the system is at a specific state. The Q value update function is the essential content of the Q-learning algorithm, which is shown in (24),

$$Q_{now}(S_t, A_t) = Q_{past}(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \, max_a Q_{past}(S_{t+1}, a) - Q_{past}(S_t, A_t) \right] \tag{24}$$

where Q_{now} is the Q value of the current state S_t and action A_t that need to be updated in the step t. Q_{past} is the Q value of the current state S_t and action A_t at the current step t. $max_a Q_{past}(S_{t+1}, a)$ is the maximum Q value of the next state S_{t+1} among all possible states. R_{t+1} is the instantaneous reward from the environment. α is the learning rate. Learning rate defines the weight of an old Q value that accounts for the new Q value. γ is the discount factor which indicates the importance of future rewards.

The agent in the RL algorithm aims to find actions that maximize cumulative reward. Typically, a trial-and-error search is adopted to establish the mapping relationship between the observations with the optimal control actions. The state is detected from the environment as observations, and then specific actions are taken to control the model in the trial-and-error search. Finally, the rewards are updated according to the control effect. To ensure the training process achieve the best result, the coordination of exploration and exploitation is essential. The exploration enables the agent to absorb knowledge in the environment, and the exploitation allows the agent to utilize existing knowledge to select the most valuable action. The exploration and exploitation are balanced by the ε-greedy action search algorithm in most cases. Because the RL agent's environment is considered a Markov Decision Process with no aftereffects, the probability distribution of the future state depends on the present state rather than the order of the events. The agent gets updated by interacting with the environment. Fig. 6 depicts the process of interaction that occurs during the training phase. During the interaction, the agent can select appropriate actions at different states to affect the reward for updating the agent. In the control problem of EMS, the environment consists of various factors that affect the vehicle operating statuses, such as velocity, acceleration, and powertrain dynamics. For the BEV equipped with HESS, the RL agent is a power-split controller, which controls the LIB and UC's output power. The objective is to find a control sequence that minimizes battery degradation while maximizing energy efficiency. The vehicle speed and torque demand are selected as states in the training endeavor. The action is UC output power. The reward is a function of energy consumption and battery degradation, which is shown in (25),

$$R = -\omega(E_{bat} + E_{UC}) - (1-\omega)Q_l + \beta \tag{25}$$

where E_{bat} is the energy consumption of the battery, E_{UC} is the energy consumption of the UC, Q_l is the battery degradation of the LIB, ω is the weight factor to control the trade-off between energy saving and battery aging, β is a bias constant. The aim of the EMS is to find a solution to a minimization problem, but the Q-learning algorithm is presented in a way for solving a maximization problem. Therefore, the energy consumption and battery degradation are set to be negative. However, if the reward is negative, the selected action will vanish because the ε-greedy selection policy will choose actions with the greatest reward value. Thus, the offset value β needs to be added to make the reward a positive function.

The integration of both energy efficiency and battery degradation into the reward function is a vital contribution of this work. This dual-objective reward function is designed to balance the trade-offs between energy savings and battery health, which is critical for EVs' long-term sustainability and performance. By including energy consumption (E_{bat} and E_{UC}) in the reward function, the Q-learning agent is incentivized to find control strategies that minimize the overall energy usage. This leads to enhanced energy efficiency, reducing operational costs and extending the driving range of the EV. The inclusion of battery degradation in the reward function addresses the longevity of the LIB. Battery degradation is influenced by various factors such as charge/discharge cycles and depth of discharge. By penalizing actions that lead to higher degradation, the Q-learning agent is encouraged to adopt strategies that mitigate battery wear and tear, thereby prolonging battery life and reducing replacement costs. The weight factor in the reward function allows for tuning the balance between energy efficiency and battery health. A higher value emphasizes energy savings, while a lower value focuses more on reducing battery degradation. This flexibility is essential for adapting the EMS to different operational priorities and scenarios. By integrating these aspects into the reward function, the proposed RL-based EMS achieves a comprehensive optimization that not only enhances the immediate energy performance but also safeguards the long-term health of the battery. This holistic approach addresses the critical challenges of EV energy management and sets a baseline example for further improvement of reinforcement learning applications in EVs.

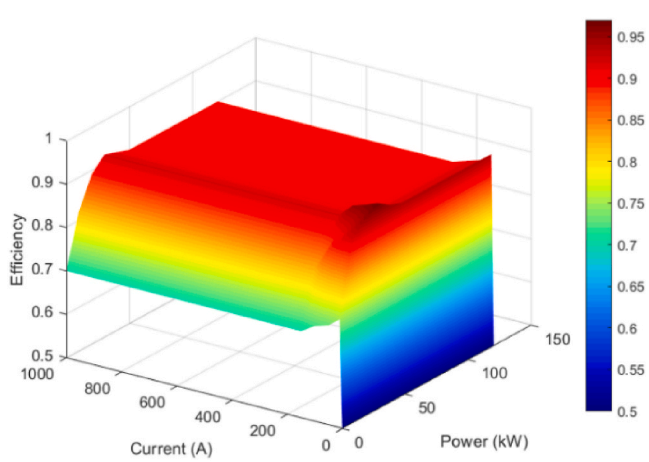

Fig. 5. Efficiency map of the DC/DC converter.

3.2. Digital twin-enhanced Q-learning EMS

A digital twin is a virtual representation of a physical system, process, or product. It serves as a bridge between the physical and digital worlds by using real-time data, simulation, machine learning, and other advanced technologies to gain insights, optimize performance, and make informed decisions. In the digital twin system, the data of the physical model is collected, processed, and integrated to provide an accurate and up-to-date representation of the physical system in the virtual space. By simulating different scenarios and control strategies, digital twins can be used to optimize performance, energy efficiency, and resource utilization. The digital twin diagram is shown in Fig. 7. The physical model is the basis of the virtual model development. The virtual model is used to simulate the physical system. It also provides control to the physical model.

In most cases, Q-learning-based EMSs are pre-trained through

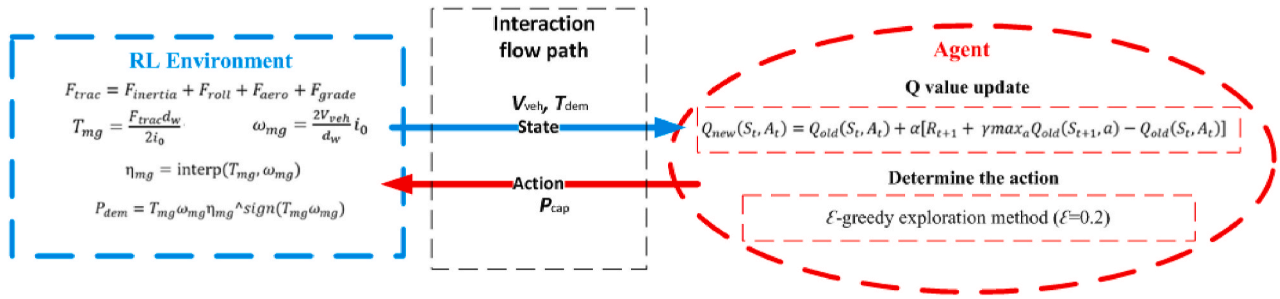

Fig. 6. Q-leaning training flow.

standard driving cycles in simulation [37]. Although the driving cycle is used to model real traffic conditions, it can only partially represent actual traffic characteristics [38]. Thus, the pre-trained Q-learning agent can provide optimal control strategies in certain situations but cannot handle the uncertainty and abruptness in real traffic. This paper adopts the digital twin methodology to enhance the conventional reinforcement learning-based EMS. When the electric vehicle equipped with pre-trained Q-learning EMS is in driving mode, the digital twin model runs simultaneously in the virtual space. The collected data are uploaded to the data transmitter and downloaded by the digital twin model. The virtual model will run with the physical model's data and use the historical data to update the EMS. The updated Q-learning agent will upload to the data transmitter and download by the physical model to replace the old EMS and improve the control performance.

The digital twin system integrates real-time environmental data, such as temperature and humidity, into its simulations. Sensors on the vehicle monitor these parameters, transmitting the data to the digital twin to adjust the virtual model. The system includes temperature-dependent models of battery performance, accounting for efficiency, capacity, and degradation variations at different temperatures. For example, the model simulates increased battery degradation at high temperatures and reduced capacity at low temperatures. By leveraging historical data from past extreme weather conditions, the digital twin continuously improves its accuracy in predicting and managing battery performance. Real-time monitoring and feedback allow for immediate adjustments, ensuring the digital twin remains accurate as environmental conditions change.

The digital twin system employs high-precision sensors and data acquisition systems to collect real-time data from the physical model. These sensors measure critical parameters such as battery temperature, SOC, power output, and environmental conditions, ensuring accuracy through calibration. Advanced data synchronization protocols, including time-stamping, align data streams from various sensors for coherent processing by the digital twin. Redundant communication channels and error correction algorithms address data transmission delays and packet loss, ensuring critical data is reliably transmitted. These measures maintain the digital twin system's accuracy and

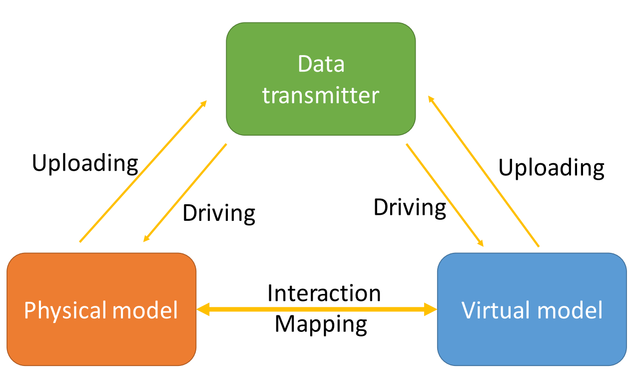

Fig. 7. Digital twin interaction diagram.

responsiveness. Extensive HIL tests validate the system's performance under varying network conditions, confirming its high accuracy and reliability.

3.3. Digital twin-enhanced DRL-based EMS

Although the RL-based EMS has achieved progress in the optimal control of EVs' energy management, it still has insolvable gaps in this field. The table-based RL algorithms only be able to handle discrete action and state spaces in the EMS. In the previous section, the state and action variables of the Q-learning-based EMS are discrete, which means the set of states and actions are limited. Its performance is dependent on the discretization of both the environmental states and the action space. The training is complicated drastically with the dimension buildup of state and action space, i.e., the strategy suffers from the so-called "curse of dimensionality'' [2]. However, in many applications, such as robotics and energy management, discretization is not desirable, as they have a negative impact on the quality of the solution and, at the same time, require large amounts of memory and computing power in the case of a fine discretization. Compared with Q-learning, the DRL method uses multi-layer neural networks to approximate the Q-matrix, enabling a noticeable improvement toward the continuous state space. Therefore, this paper adopts the DDPG to establish the DRL-based EMS for the EV equipped with HESS. DDPG is a model-free, policy-based reinforcement learning algorithm used to solve continuous control tasks. DDPG is an off-policy algorithm that combines concepts from DQN and policy gradient methods. It utilizes deep neural networks to approximate the policy (actor) and the action-value function (critic). Since the DDPG uses an actor-critic architecture, it has better convergence performance than Q-learning. The critic in DDPG learns an action-value function, which helps to reduce the variance in the policy gradient estimation. This can bring in more stable learning and improved convergence compared to Q-learning, particularly in environments with high variance. The TD3 is also be used to establish the EMS of EVs. However, for the tasks in this paper, DDPG is preferred over TD3 for model adaptation due to its simpler architecture, making it easier to fine-tune and customize for new tasks. It has less computational effort and better generalization ability across different environments owing to its straightforward design. Additionally, the reduced complexity of DDPG simplifies the process of adaptation, enhancing the efficiency of the proposed method.

When combined with the digital twin technology, the DDPG is more scalable than Q-learning for problems with large or continuous action spaces. The digital twin model has higher fidelity, and the virtual space contains more information than the conventional RL environment, such as real-time traffic data. That means the action spaces grows rapidly, which dramatically increases Q-learning's computational complexity. But DDPG avoids this issue and can handle more complex problems with larger action spaces by directly learning a policy. Besides, the DDPG allows for more efficient exploration in continuous action spaces compared to Q-learning. In Q-learning, exploration is typically achieved through ε-greedy exploration, which can be inefficient in continuous action spaces. DDPG uses noise added to the output of the actor-network for exploration, such as Ornstein-Uhlenbeck noise in this paper, which can lead to more effective exploration strategies in continuous action spaces.

In this paper, the DDPG is combined with the digital twin. The diagram of digital twin-enhanced DDPG-EMS is shown in Fig. 8. The actor-network in the DDPG interacts with the EV digital twin model in the virtual space and stores the transactions s_t, a_t, r_t, s_{t+1} in experience replay buffer. The experience buffer randomly samples mini-batch s_i, a_i, r_i, s_{i+1} and feed into actor and critic network. The Critic target network calculates the expected target return y_i by using action $\mu'(s_{i+1})$ given by the Actor target network.

$$y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1})) \tag{26}$$

where γ is the discount factor, Q' is Critic target network and μ' Actor target network. With the target Q value, the Critic loss can be expressed by,

$$L = \frac{1}{N}\sum\nolimits_i (y_i - Q(s_i, a_i))^2 \tag{27}$$

where N is the number of the mini batches. With the help of Critic network, the Actor policy is updated by the sampled policy gradient, given by

$$\nabla_{\theta^\mu} J = \frac{1}{N}\sum\nolimits_i [\nabla_a Q(s_i, \mu(s_i)) \nabla_{\theta^\mu} \mu(s_i|\theta^\mu)] \tag{28}$$

where θ^μ is the parameters of online actor network. In order to improve the learning stability, the target network updates softly and slowly with a small hyperparameter ε:

$$\theta' \leftarrow \varepsilon\theta + (1-\varepsilon)\theta' \tag{29}$$

$$\theta^{\mu'} \leftarrow \varepsilon\theta^\mu + (1-\varepsilon)\theta^{\mu'} \tag{30}$$

where θ' is parameters of the Critic target network, θ is parameters of the Critic online network, and $\theta^{\mu'}$ is parameters of Actor target network.

As shown in Algorithm 1, a high-fidelity digital twin of the physical model is mapped in the virtual space, accurately representing the powertrain components, battery system, driving conditions, and environmental factors that affect energy usage and driving range. The action of the DDPG agent is the ratio between the output power of LIB and UC. The velocity and torque demand still be chosen as the states. But it is different from the conventional DDPG-based EMS, whose speed and torque demand are designed on the fixed driving cycle. The states in the proposed method are collected from the physical model to improve the adaptability of the EMS in real-world traffic situations.

Besides, to fully exploit the scalability of DDPG, the battery SOC is introduced into the state space. This helps the EMS absorb more information from the environment to improve control accuracy. The reward function for the DDPG is the same as the Q-learning, which considers energy consumption and battery degradation. Once the DDPG agent is trained, evaluate its performance by simulating the control policy in the digital twin environment. The obtained results are also compared to existing control strategies or benchmarks. If the performance is unsatisfactory, refine the DDPG algorithm or reward function and retrain the agent. After achieving satisfactory performance in the digital twin environment, the learned control policy is deployed in the physical model.

4. Results and analysis

4.1. Results of digital twin enhanced DRL-based EMS

The physical model is constructed based on the HIL platform shown in Fig. 9.

The trained DRL agent is deployed to dSPACE SCALEXIO. The information on the driving cycle is held by the host PC and transmitted to the SCALEXIO, and then the vehicle model calculates the current torque demand. The trained RL agent receives the current state and decides the action. The motor-generator set controller receives the action command to control the motor and generator. The computer for the virtual space has Intel(R) Core (TM) i7-9750H CPU @ 2.60 GHz, NVIDIA GeForce RTX 2060 GPU, and 16.0 GB RAM.

The power output of the electric drive system is shown in Fig. 9. In the vehicle acceleration stage, the UC provides the highest power to fulfill the large power request, but the UC cannot store enough energy for a long time operation. Thus, the Q-learning EMS takes advantage of the battery, using the battery as the continuous power source. When the vehicle is decelerating, the UC absorbs considerable power from the

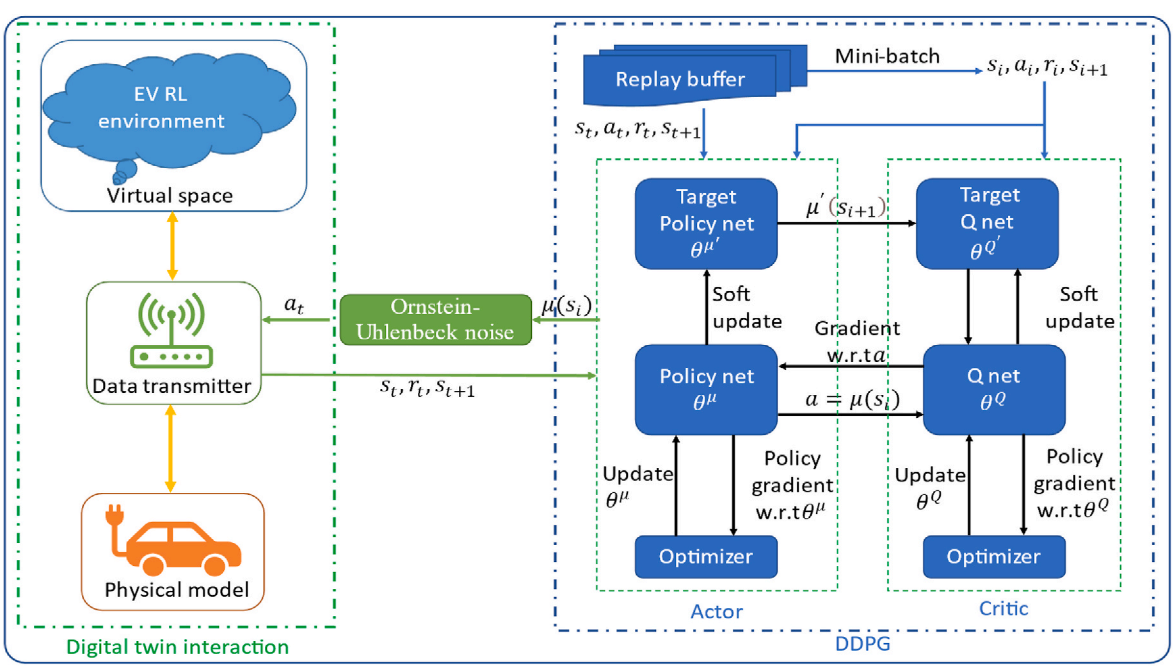

Fig. 8. Digital twin-enhanced DDPG-EMS.

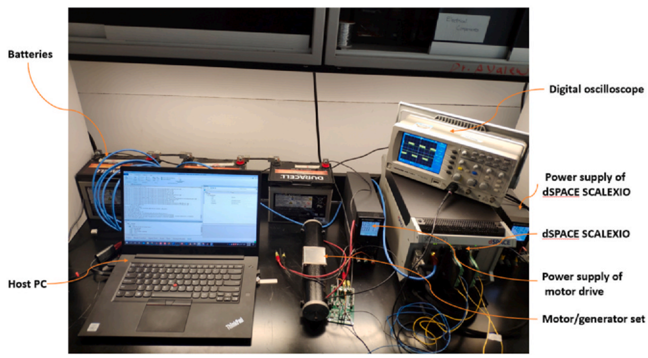

Fig. 9. HIL platform.

regenerative braking. If the UC is fully charged, the remaining part of the regenerated energy will charge the battery. The process indicates that the UC can significantly shave the peaks of battery charging and discharging power. Thus, battery degradation will be reduced.

In the pre-train phase, the EPA Urban Dynamometer Driving Schedule (UDDS) is adopted to initialize the DRL agent at the starting point. The results of the pre-trained DDPG-based EMS are show in Fig. 10.

From Fig. 10, it can be seen that the DDPG-based EMS can fully exploit the UC to alleviate the stress of LIB. In the vehicle acceleration stage, the UC provides the highest power to fulfill the large power request, but the UC cannot store enough energy for a long time operation. Thus, the DDPG-based EMS takes advantage of the battery, using the battery as the continuous power source. When the vehicle is decelerating, the UC absorbs large negative power from the regenerative braking. If the UC is fully charged, the remaining part of the regenerated energy will charge the battery. The process indicates that the UC can

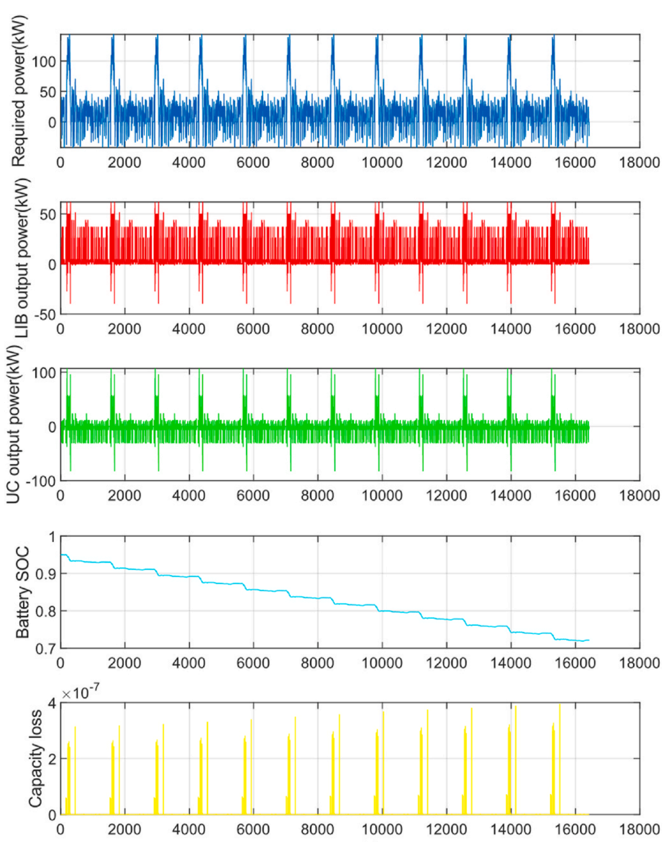

Fig. 10. Results of conventional DDPG-based EMS.

significantly shave the peaks of battery charging and discharging power. Thus, battery degradation will be reduced.

Although the conventional DDPG-based EMS has achieved nearly optimal results for a specific driving cycle used in pre-training, the control performance is not optimal under real driving conditions. Thus, the digital twin is integrated into the proposed EMS to utilize the up-to-date traffic condition to improve the adaptivity of the DDPG-based EMS. In this paper, the Worldwide Harmonized Light Vehicles Test Procedure (WLTP) is adopted to mimic real-world driving conditions. The comparison of conventional DDPG-based EMS and the proposed method is shown as follows:

Fig. 11 illustrates the SOC of conventional DDPG-based EMS and the digital twin-enhanced DDPG-based EMS. The comparison shows that the proposed method outperforms the conventional DDPG-based EMS. According to the final SOC value, the energy efficiency of the digital twin-enhanced DDPG-based EMS is 17.08 % higher than that of the single DDPG-based EMS. As for the battery degradation, the proposed method also has better control performance to alleviate the battery capacity loss. Fig. 12 indicates the battery aging during driving in the WLTP. The total capacity losses in the simulation of DDPG-based EMS are 3.41e-07 % and 1.32e-07 % without and with the digital twin, respectively. The results show that the digital twin-enhanced DDPG-based EMS has better adaptability and control performance. The Agents of both methods are pre-trained through UDDS. When deployed in the physical model under the WLTP, the proposed method achieves higher energy efficiency and lower battery degradation.

4.2. *Comparative study with other EMSs*

The EV HIL simulations equipped with HESS under different EMSs are conducted in this paper. In order to analyze the effect of the digital twin-enhanced DDPG-based EMS, the digital twin-enhanced Q-learning-based EMS, Q-learning-based EMS, and rule-based EMS are simulated in the same environment to study the control performance and computational efficiency. The main control logic and rules of the chosen rule-based EMS are as follows: 1)Power Demand Split: During vehicle acceleration, the UC is prioritized to provide the peak power demand due to its high power density and rapid response capability. This helps in reducing the stress on the battery. During steady-state driving, the battery is used as the primary power source to supply the continuous power demand. 2) During regenerative braking, the UC is prioritized to absorb the regenerative energy first. If the UC is fully charged, the excess energy is directed to the battery. 3)The SOC of both the battery and the UC is continuously monitored. If the UC SOC falls below a predefined

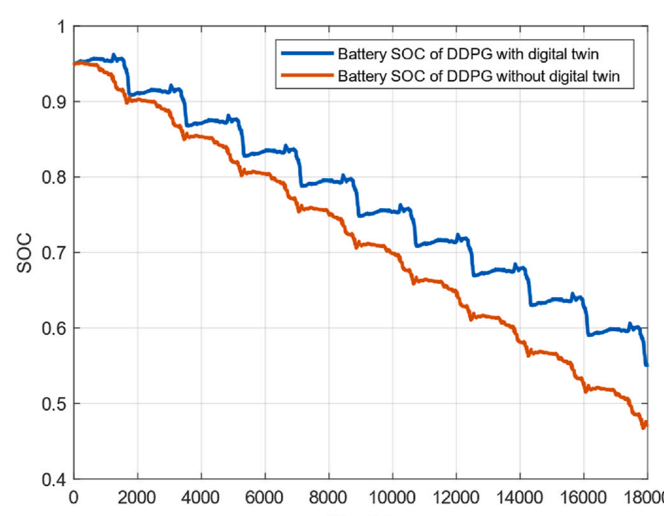

Fig. 11. SOC trajectories.

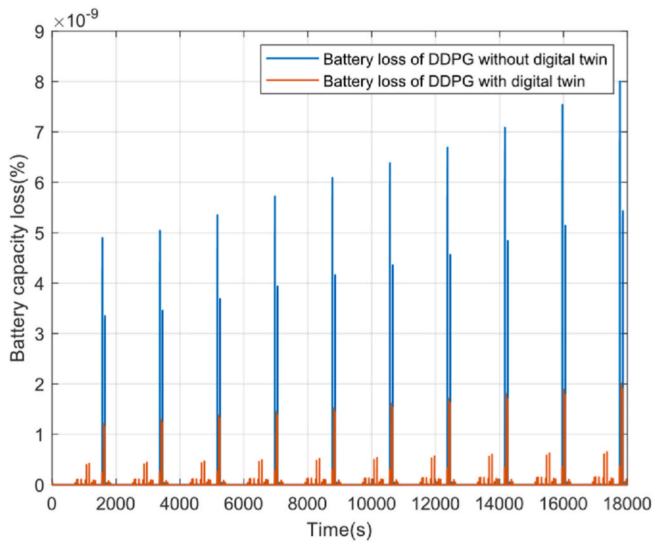

Fig. 12. Battery capacity loss trajectories.

threshold, the battery is used to charge the UC. If the battery SOC falls below a critical level, the vehicle operates in a mode that minimizes power consumption to preserve battery life. 4) The UC is used to level the load on the battery, absorbing power surges and providing power during high demand periods to prevent excessive battery cycling and degradation.

The above rule-based EMS is chosen as a baseline to compare with the proposed digital-twin enhanced EMS. The chosen rule-based EMS represents a realistic and practical benchmark commonly used in industry. Although designing a more effective rule-based EMS is possible, its design and implementation would require extensive simulations and tuning, potentially reducing generalizability. Our proposed method aims to provide a robust and adaptable solution that performs optimally across diverse conditions without extensive rule-tuning. The comparison of the proposed DDPG with the rule-based EMS below shows the advantages of the learning-based approach, including its adaptability and superior performance. The output power of LIB and UC of the HESS under different EMSs is illustrated in Fig. 13.

The output power comparison presents the UC output power and LIB output in the simulation driving cycle under four EMSs. From the results, the digital twin-enhanced DDPG-based EMS achieves the best control performance since the proposed method takes advantage of the digital twin to infuse real-time traffic information. In Fig. 13(a), the UC provides the most peak power for charging and discharging, and the LIB delivers continuous power to maintain the vehicle driving and avoid excessive charging and discharging so that the battery degradation is reduced. The rule-based EMS is designed through expert experiences and knowledge, which is suitable for limited situations only. If the heuristic rules cannot cover the application scenarios, the control performance cannot be guaranteed. As shown in Fig. 13(d), the UC is engaged in the charging process only, and the peak charging power is always supplied by the LIB. Furthermore, the LIB also provides continuous power for driving; thus, this rule-based strategy is inefficient in reducing battery degradation. The digital twin-enhanced Q-learning and conventional Q-learning both achieve a near-optimal control performance. Both RL EMSs can exploit the advantages of HESS that use the SC to handle most peak power to alleviate the degradation of LIB. However, there is a slight difference between the digital twin-enhanced Q-learning EMS and conventional Q-learning, which is presented in Fig. 13(b) and (c). The digital twin-enhanced Q-learning EMS enables UC to cover all peak power, but under the conventional Q-learning EMS, the LIB absorbs the peak charging energy. To further evaluate the control performance of the proposed method, the LIB degradation trajectories are shown in

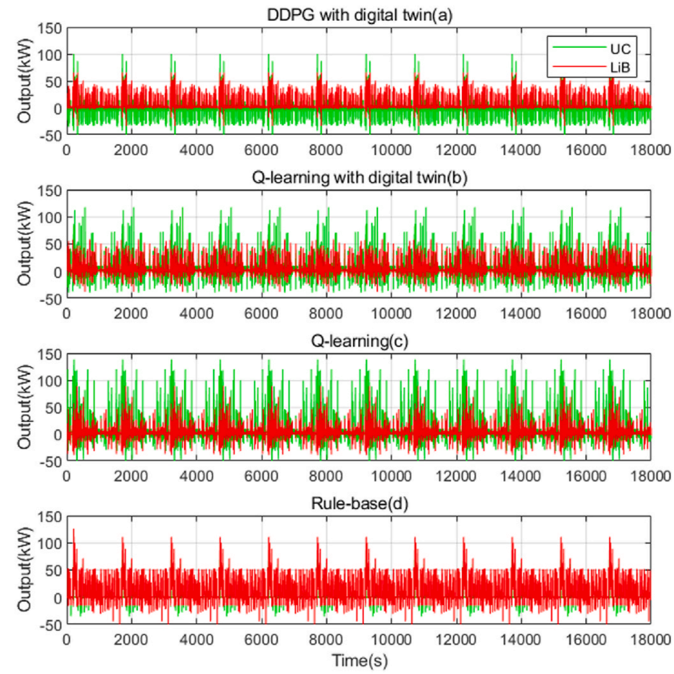

Fig. 13. EMS comparison.

Fig. 14.

The total battery capacity loss in one round of simulation under rule-based, conventional Q-learning, digital twin-enhanced Q-learning, and digital twin-enhanced DDPG-based EMS is 1.57e-06 %, 6.48e-07 %, 5.81e-07 %, and 1.32e-07 %, respectively. The digital twin-enhanced DDPG-based EMS achieves the best results in reducing LIB

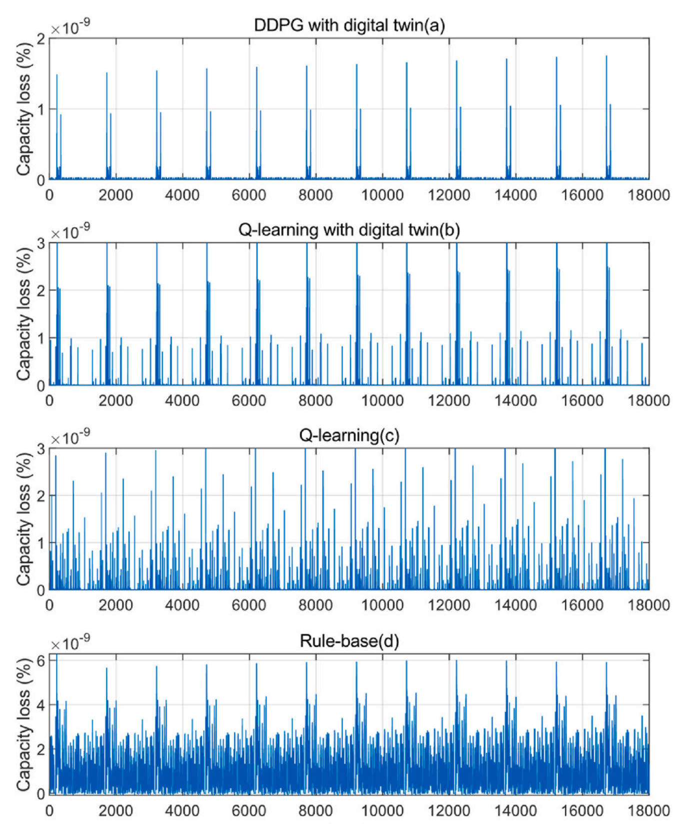

Fig. 14. Battery capacity loss comparison.

degradation, which is 67.99 % less than that of the digital twin-enhanced Q-learning-based EMS. Comparing Fig. 14(a) and (b), the LIB degradation pattern of the digital twin-enhanced Q-learning-based EMS and digital twin-enhanced DDPG-based EMS are similar, but the value of the digital twin-enhanced Q-learning-based EMS is much higher. This is also reflected in Fig. 14 (a) and (d). The LIB is mostly working in discharging mode to drive the vehicle and is rarely held for charging under the control of digital twin-enhanced DDPG-based EMS. However, the digital twin-enhanced Q-learning-based EMS makes the LIB participate in both the charging and discharging processes, so the cumulative battery aging is larger than that of digital twin-enhanced DDPG-based EMS. The LIB degradation of conventional Q-learning is 11.53 % higher than that of the digital twin-enhanced Q-learning-based EMS. The rule-based EMS cannot fully exploit the advantages of HESS, and the LIB has the most severe degradation among all EMSs. From the results of battery degradation, The digital twin technology enables the RL-based EMS to exploit the potential for the optimal solution fully. Both of the DDPG and the Q-learning methods outperform the conventional counterpart. The digital twin-enhanced DDPG-based EMS also achieves a better result for alleviating battery degradation.

The energy stored by the UC occupies only a tiny portion of the electric drive system, so the SOC of the battery is taken as an indicator of the stored energy to evaluate both the battery and UC. The SOC trajectories comparison is shown in Fig. 15. According to the LIB SOC trajectories, the energy efficiency of all RL-based EMS is higher than that of rule-based EMS. The SOC starting point of all EMS is 0.95, and the final SOC of the rule-base EMS is 0.3627, which is the lowest in all tested EMS. The digital twin-enhanced DDPG-based EMS has the highest energy efficiency, whose final SOC value is 0.5851. The digital twin-enhanced Q-learning-based EMS also achieves good control performance, and its final SOC value is 0.5509, which is lower than the proposed method by 6.21 %. The final SOC of conventional Q-learning-based EMS is 0.4708, which is much lower than the Q-learning with the digital twin. This phenomenon is the same as we discussed in the last subsection, and the comparison between the digital twin-enhanced DDPG and conventional DDPG shows that the digital twin can enhance the performance of RL-based EMS.

The proposed method leverages digital twin technology to offload RL training tasks to the virtual space, reducing the onboard system's computational load. This allows the on-board system to apply pre-trained models, significantly lowering computational requirements during operation. The RL training and optimization are performed in the cloud, while traditional RL based EMS relies on onboard resources, which makes it challenging to meet the relevant computational requirements. The computational times for different methods are shown in Table 2, from which it is obvious that our proposed digital twin-enhanced DDPG EMS has much less computing time than conventional Q-learning.

5. Conclusion

This paper proposes an enhanced DDPG-based EMS using a digital twin model for an electric vehicle equipped with battery and ultracapacitor. The propulsion system model is presented, and the configuration of the digital twin model is explained. By leveraging deep reinforcement learning and digital twin technologies, advanced EMSs for electric vehicles can be developed and deployed. This will lead to significant improvements in energy efficiency, driving range, and reducing battery degradation, ultimately benefiting both consumers and the environment. The results show that the digital twin model can update conventional DDPG-based EMS to improve the total energy efficiency by 7.08 % and reduce the battery degradation by 25.28 %. These benefits will reduce the overall operating cost of the electric vehicle.

However, the enhanced Q-learning EMS does not consider the change in the number of passengers in the vehicle. The variation in the number of passengers can cause a significant change in the total weight of the vehicle, which may consume more energy under the same traffic conditions. Also, the HIL platform is not based on a full-scale vehicle, and the data have been downscaled to fit the parameters of the HIL components. A more completed HIL platform will be established and applied in future work.

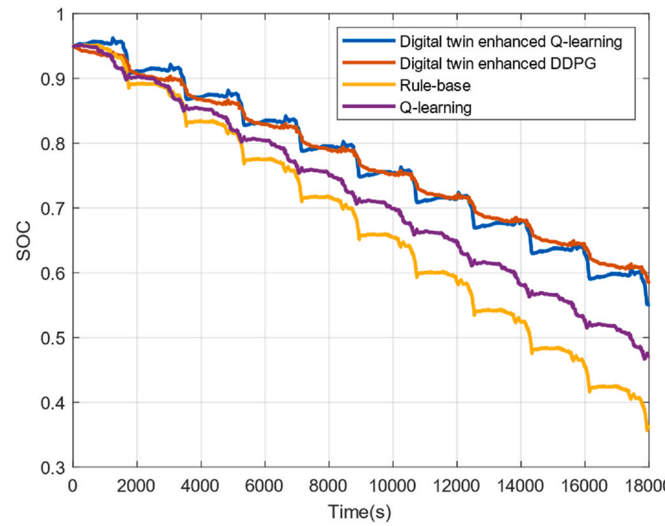

Fig. 15. SOC trajectories comparison.

Table 2
Computation time.

Method	Value (s)
Rule-base	5.26
Conventional Q-learning	96.38
Digital twin-enhanced DDPG	33.26

CRediT authorship contribution statement

Yiming Ye: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Bin Xu:** Writing – review & editing, Investigation, Data curation. **Hanchen Wang:** Writing – review & editing, Validation. **Jiangfeng Zhang:** Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Benjamin Lawler:** Writing – review & editing. **Beshah Ayalew:** Writing – review & editing.

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.

References

[1] Continental Automotive. Worldwide emission standards and related regulations: Passenger cars, light & medium duty vehicles. Hanover (Germany): Continental AG. 2019.

[2] Xu B, et al. Parametric study on reinforcement learning optimized energy management strategy for a hybrid electric vehicle. Appl Energy Feb. 2020;259:114200. https://doi.org/10.1016/j.apenergy.2019.114200.

[3] Ye Y, Zhang J, Pilla S, Rao AM, Xu B. Application of a new type of lithium-sulfur battery and reinforcement learning in plug-in hybrid electric vehicle energy

management. J Energy Storage Mar. 2023;59:106546. https://doi.org/10.1016/j.est.2022.106546.
[4] Geng L, Lu Z, Guo X, Zhang J, Li X, He L. Coordinated operation of coupled transportation and power distribution systems considering stochastic routing behaviour of electric vehicles and prediction error of travel demand. IET Gener Transm Distrib 2021;15(14):2112–26. https://doi.org/10.1049/gtd2.12161.
[5] Manthiram A. A reflection on lithium-ion battery cathode chemistry. Nat Commun Mar. 2020;11(1). https://doi.org/10.1038/s41467-020-15355-0. Art. no. 1.
[6] Rahman AU, Ahmad I, Malik AS. Variable structure-based control of fuel cell-supercapacitor-battery based hybrid electric vehicle. J Energy Storage Jun. 2020; 29:101365. https://doi.org/10.1016/j.est.2020.101365.
[7] Lin X, Wang Y, Bogdan P, Chang N, Pedram M. Reinforcement learning based power management for hybrid electric vehicles. In: 2014 IEEE/ACM international conference on computer-aided design (ICCAD); Nov. 2014. p. 33–8. https://doi.org/10.1109/ICCAD.2014.7001326.
[8] Qi X, Wu G, Boriboonsomsin K, Barth MJ, Gonder J. Data-driven reinforcement learning–based real-time energy management system for plug-in hybrid electric vehicles. Transp. Res. Rec. Jan. 2016;2572(1):1–8. https://doi.org/10.3141/2572-01.
[9] Liu C, Murphey YL. Power management for plug-in hybrid electric vehicles using reinforcement learning with trip information. In: 2014 IEEE transportation electrification conference and expo (ITEC); Jun. 2014. p. 1–6. https://doi.org/10.1109/ITEC.2014.6861862.
[10] Xu B, Shi J, Li S, Li H, Wang Z. Energy consumption and battery aging minimization using a Q-learning strategy for a battery/ultracapacitor electric vehicle. Energy Aug. 2021;229:120705. https://doi.org/10.1016/j.energy.2021.120705.
[11] Liu T, Hu X, Li SE, Cao D. Reinforcement learning optimized look-ahead energy management of a parallel hybrid electric vehicle. IEEE ASME Trans Mechatron Aug. 2017;22(4):1497–507. https://doi.org/10.1109/TMECH.2017.2707338.
[12] Zou Y, Liu T, Liu D, Sun F. Reinforcement learning-based real-time energy management for a hybrid tracked vehicle. Appl Energy Jun. 2016;171:372–82. https://doi.org/10.1016/j.apenergy.2016.03.082.
[13] Ye Y, Zhang J, Xu B. A fast Q-learning energy management strategy for battery/supercapacitor electric vehicles considering energy saving and battery aging. In: 2021 international conference on electrical, computer and energy technologies (ICECET). Cape Town, South Africa: IEEE; Dec. 2021. p. 1–6. https://doi.org/10.1109/ICECET52533.2021.9698682.
[14] Hu Y, Li W, Xu H, Xu G. An online learning control strategy for hybrid electric vehicle based on fuzzy Q-learning. Energies Oct. 2015;8(10). https://doi.org/10.3390/en81011167. Art. no. 10.
[15] Wu J, Zou Y, Zhang X, Liu T, Kong Z, He D. An online correction predictive EMS for a hybrid electric tracked vehicle based on dynamic programming and reinforcement learning. IEEE Access 2019;7:98252–66. https://doi.org/10.1109/ACCESS.2019.2926203.
[16] Xu B, Malmir F, Filipi Z. Real-time reinforcement learning optimized energy management for a 48V mild hybrid electric vehicle. SAE Tech. Pap. Apr. 2019. https://doi.org/10.4271/2019-01-1208.
[17] Qi X, Luo Y, Wu G, Boriboonsomsin K, Barth M. Deep reinforcement learning enabled self-learning control for energy efficient driving. Transp. Res. Part C Emerg. Technol. Feb. 2019;99:67–81. https://doi.org/10.1016/j.trc.2018.12.018.
[18] Qi C, et al. Hierarchical reinforcement learning based energy management strategy for hybrid electric vehicle. Energy Jan. 2022;238:121703. https://doi.org/10.1016/j.energy.2021.121703.
[19] Li W, et al. Deep reinforcement learning-based energy management of hybrid battery systems in electric vehicles. J Energy Storage Apr. 2021;36:102355. https://doi.org/10.1016/j.est.2021.102355.
[20] Tang X, Chen J, Liu T, Qin Y, Cao D. Distributed deep reinforcement learning-based energy and emission management strategy for hybrid electric vehicles. IEEE Trans Veh Technol Oct. 2021;70(10):9922–34. https://doi.org/10.1109/TVT.2021.3107734.
[21] Lian R, Peng J, Wu Y, Tan H, Zhang H. Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle. Energy Apr. 2020;197:117297. https://doi.org/10.1016/j.energy.2020.117297.
[22] Wu J, He H, Peng J, Li Y, Li Z. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. Appl Energy Jul. 2018;222:799–811. https://doi.org/10.1016/j.apenergy.2018.03.104.
[23] Zhang Y, et al. Pre-optimization-assisted deep reinforcement learning-based energy management strategy for a series–parallel hybrid electric truck. Energy Sep. 2024; 302:131628. https://doi.org/10.1016/j.energy.2024.131628.
[24] Wu J, Wei Z, Li W, Wang Y, Li Y, Sauer DU. Battery thermal- and health-constrained energy management for hybrid electric bus based on soft actor-critic DRL algorithm. IEEE Trans. Ind. Inform. Jun. 2021;17(6):3751–61. https://doi.org/10.1109/TII.2020.3014599.
[25] The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles," Structures, Structural Dynamics, and Materials and Co-located Conferences. Accessed: April. 13, 2023. [Online]. Available: https://arc.aiaa.org/doi/10.2514/6.2012-1818.
[26] Xie Y, et al. Microsimulation of electric vehicle energy consumption and driving range. Appl Energy Jun. 2020;267:115081. https://doi.org/10.1016/j.apenergy.2020.115081.
[27] Zhang Z, Zou Y, Zhou T, Zhang X, Xu Z. Energy consumption prediction of electric vehicles based on digital twin technology. World Electr. Veh. J. Sep. 2021;12(4): 160. https://doi.org/10.3390/wevj12040160.
[28] Toso F, Torchio R, Favato A, Carlet PG, Bolognani S, Alotto P. Digital twins as electric motor soft-sensors in the automotive industry. In: 2021 IEEE international workshop on metrology for automotive (MetroAutomotive); Jul. 2021. p. 13–8. https://doi.org/10.1109/MetroAutomotive50197.2021.9502885.
[29] Li W, Rentemeister M, Badeda J, Jöst D, Schulte D, Sauer DU. Digital twin for battery systems: cloud battery management system with online state-of-charge and state-of-health estimation. J Energy Storage Aug. 2020;30:101557. https://doi.org/10.1016/j.est.2020.101557.
[30] Peng Y, Zhao S, Wang H. A digital twin based estimation method for health indicators of DC–DC converters. IEEE Trans Power Electron Feb. 2021;36(2): 2105–18. https://doi.org/10.1109/TPEL.2020.3009600.
[31] Bartolucci L, Cennamo E, Cordiner S, Mulone V, Pasqualini F, Boot MA. Digital twin of a hydrogen fuel cell hybrid electric vehicle: effect of the control strategy on energy efficiency. Int. J. Hydrog. Energy Dec. 2022. https://doi.org/10.1016/j.ijhydene.2022.11.283.
[32] Sidorov KM, Grishchenko AG, Sidorov BN, Stroganov VI. Electric propulsion system simulation as basis for the electric vehicle digital twin development. In: 2021 systems of signals generating and processing in the field of on board communications; Mar. 2021. p. 1–6. https://doi.org/10.1109/IEEECONF51389.2021.9416111.
[33] Xiong R, Cao J, Yu Q. Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. Appl Energy Feb. 2018;211:538–48. https://doi.org/10.1016/j.apenergy.2017.11.072.
[34] Cao J, Xiong R. Reinforcement learning-based real-time energy management for plug-in hybrid electric vehicle with hybrid energy storage system. Energy Proc Dec. 2017;142:1896–901. https://doi.org/10.1016/j.egypro.2017.12.386.
[35] Ye Y, Wang H, Xu B, Zhang J. An imitation learning-based energy management strategy for electric vehicles considering battery aging. Energy Nov. 2023;283: 128537. https://doi.org/10.1016/j.energy.2023.128537.
[36] Bertsekas D. Reinforcement learning and optimal control. Athena Scientific; 2019.
[37] Ye Y, Zhao X, Zhang J. Driving cycle electrification and comparison. Transp. Res. Part Transp Environ Oct. 2023;123:103900. https://doi.org/10.1016/j.trd.2023.103900.
[38] Zhao X, Ye Y, Ma J, Shi P, Chen H. Construction of electric vehicle driving cycle for studying electric vehicle energy consumption and equivalent emissions. Environ Sci Pollut Res Oct. 2020;27(30):37395–409. https://doi.org/10.1007/s11356-020-09094-4.