

Research Article

An Energy Management Strategy for a Super-Mild Hybrid Electric Vehicle Based on a Known Model of Reinforcement Learning

Yanli Yin ¹, Yan Ran,¹ Liufeng Zhang,¹ Xiaoliang Pan,² and Yong Luo ³

¹School of Mechatronics & Automobile Engineering, Chongqing Jiao Tong University, Chongqing 400054, China

²Chongqing Changan Automobile Stock Co., Ltd., Chongqing 400054, China

³Key Laboratory of Advanced Manufacturing Technology for Automobile Parts, Ministry of Education, Chongqing University of Technology, Chongqing 400054, China

Correspondence should be addressed to Yanli Yin; cqu_ylyin@126.com

Received 9 December 2018; Revised 20 February 2019; Accepted 21 February 2019; Published 1 April 2019

Academic Editor: Yongji Wang

Copyright © 2019 Yanli Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For global optimal control strategy, it is not only necessary to know the driving cycle in advance but also difficult to implement online because of its large calculation volume. As an artificial intelligent-based control strategy, reinforcement learning (RL) is applied to an energy management strategy of a super-mild hybrid electric vehicle. According to time-speed datasets of sample driving cycles, a stochastic model of the driver's power demand is developed. Based on the Markov decision process theory, a mathematical model of an RL-based energy management strategy is established, which assumes the minimum cumulative return expectation as its optimization objective. A policy iteration algorithm is adopted to obtain the optimum control policy that takes the vehicle speed, driver's power demand, and state of charge (SOC) as the input and the engine power as the output. Using a MATLAB/Simulink platform, CYC_WVUCITY simulation model is established. The results show that, compared with dynamic programming, this method can not only adapt to random driving cycles and reduce fuel consumption of 2.4%, but also be implemented online because of its small calculation volume.

1. Introduction

With the increasing problems of global warming, air pollution, and energy shortage, hybrid electric vehicles (HEVs) have been extensively studied because of their potential to significantly improve fuel economy and reduce emissions.

Energy management strategies are critical for HEVs to achieve optimal performance and energy efficiency through power split control. Energy management strategies can be divided according to the way they are implemented: rule-based strategies and optimization-based strategies.

Optimization-based strategies form an important group of energy management strategies and can be divided into strategies based on instantaneous optimization, global optimization, model predictive control, and artificial intelligent.

Rule-based strategies are primarily based on human experience. The torque distribution of the engine and motor

is based on preset control rules, which are formulated by the steady-state MAP chart of the engine and motor. Reference [1] proposed an energy management strategy based on a logic threshold and a fuzzy algorithm for improving fuel economy. Wang et al. [2] developed algorithms for On/Off control, load tracking control, and bus voltage control and conducted a simulation. Rule-based control strategy is simple and can be easily implemented online; however, a static control strategy is not optimal in theory and it also does not consider the dynamic changes in working conditions.

An instantaneous optimal control strategy can ensure an optimum objective at every time step; however, it cannot guarantee an optimum objective design over the whole driving cycle. Compared with the logic threshold strategy, the calculation amount is large, but it can be implemented online. Reference [3] formulated the energy management strategy of an HEV based on driving style recognition. Reference

TABLE 1: Super-mild HEV parameters.

Component	Parameters	Value
Engine	Engine type	JL472Q1
	Rated power, kW	49±2.45
	Maximum torque, N•m	82±4.1
Motor	Rated power, kW	5
	Rated speed, r•min ⁻¹	1000
Battery	Rated capacity, A•h	6.5
	Rated voltage, V	80
Transmission	Minimum gear ratio	0.498
	Maximum gear ratio	4.04
	PGT parameter	3.875
	Fixed gear ratio	2.4704
Others	Final drive gear ratio	4.4059
	Motor reducing gear ratio	1.5

[4] combined K-means clustering algorithm with equivalent consumption minimization strategy to realize the energy management of the whole vehicle.

An MPC strategy can ensure an objective optimum design in the prediction domain and can be implemented online. Reference [5] proposed a stochastic model predictive control-based energy management strategy using the vehicle location, traveling direction, and terrain information of the area for HEVs operating in hilly regions with light traffic.

A global optimal control strategy can guarantee an objective optimum design over a given driving cycle by distributing the power of the engine and motor. However, it can only be implemented offline because of the large volume of calculations involved. Reference [6] proposed an improved dynamic programming (DP) control strategy for hybrid electric buses based on the state transition probability. DP was applied to an EVT-based HEV powertrain to realize its optimal control in [7]. A DP algorithm was also used for global optimization on the performance of a speed coupling ISG HEV in [8].

Intelligent energy management strategies include fuzzy logic, neural network, genetic algorithm, and machine learning-based control strategies. Energy management strategies based on machine learning include those based on supervised learning [9, 10], unsupervised learning [11, 12], and reinforcement learning.

RL is a data-driven approach that assumes the system as a black box, regardless of whether it is linear or nonlinear. As a type of self-adaptive optimal control method based on machine learning, RL has been widely applied to the learning control of several nonlinear systems.

Reference [13] proposed an energy management strategy for PHEB based on RL. Liu et al. [14] proposed an RL-enabled energy management strategy by using the speedy Q-learning algorithm, to accelerate the convergence rate in Markov Chain-based control policy computation. Reference [15] developed an online energy management controller for a plug-in HEV based on driving condition recognition and a

genetic algorithm. In [16], deep Q learning was adopted for energy management and the strategy was proposed and verified. RL was shown to derive model-free and adaptive control for energy management in [17]. Liu et al. [18] proposed a bi-level control framework that combined predictive learning with RL to formulate an energy management strategy.

RL can ensure global optimum over the driving cycle and does not require foreseeing the driving cycle. Compared with complex dynamic models, data-driven models can be implemented online because of the small calculation volume.

Aimed at super-mild HEV, [19] studied rule-based and instantaneous optimization methods to be applied to energy management strategies for super-mild HEVs. However, RL has not been reported to be applied to an energy management strategy of super-mild HEVs.

This study establishes an energy management strategy for super-mild HEVs based on a known model of RL. The optimal control results are obtained by a policy iteration algorithm. Based on the MATLAB/Simulink simulation platform, a simulation is conducted on the economic performance of the vehicle.

2. Super-Mild Hybrid Electric Vehicle Model

2.1. Structure and Main Parameters. A super-mild HEV is primarily composed of an engine, motor, and continuously variable transmission with reflux power. The continuously variable transmission with reflux power includes a metal belt continuously variable transmission, a fixed speed ratio gear transmission device, a planetary gear transmission device, wet clutches, a one-way clutch, and a brake. Its structure is shown in Figure 1. The main parameters are shown in Table 1.

2.2. Working Modes. There are four working modes of a super-mild HEV: only-motor mode, only-engine mode, engine-charging mode, and regenerative braking mode, as shown in Figure 2.

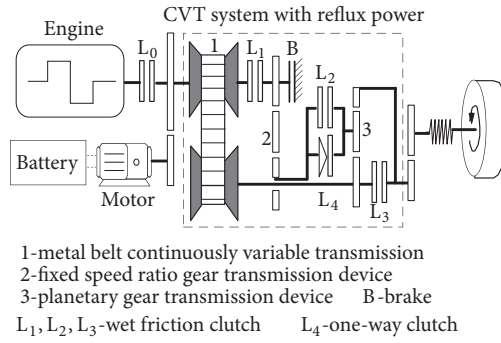


FIGURE 1: Super-mild HEV architecture.

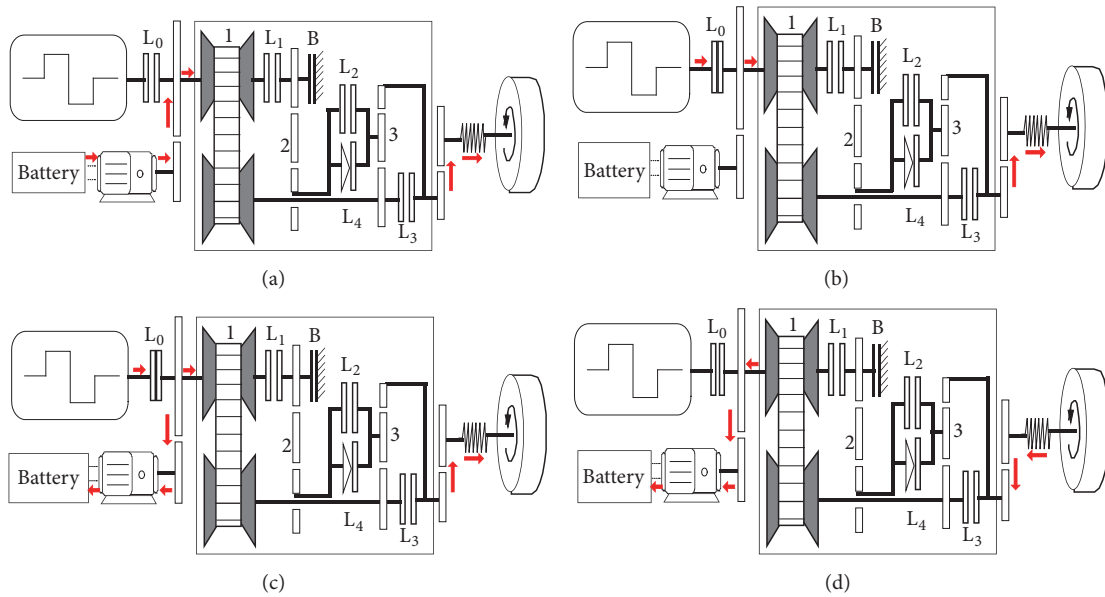


FIGURE 2: Working modes of super-mild HEV: (a) only-motor. (b) only-engine. (c) Engine-charging. (d) Regenerative braking.

2.3. Vehicle Dynamics Model

2.3.1. Power Demand Model. The power demand at the wheel is the power sum of rolling resistance, air resistance, and acceleration resistance:

$$\begin{aligned}
 P_{req} &= (F_f + F_w + F_j) v \\
 F_f &= fmg \\
 F_w &= \frac{C_D A v^2}{21.15} \\
 F_j &= \delta m \frac{dv}{dt}
 \end{aligned}
 \tag{1}$$

where F_f is the rolling resistance, F_w is the air resistance, F_j is the acceleration resistance, v is the vehicle speed, f is the rolling resistance coefficient, m is the vehicle mass, C_D is the air resistance coefficient, A is the frontal area, and δ is the conversion coefficient of rotation mass.

2.3.2. Engine Model. The engine is a highly nonlinear system and its working process is very complex. Therefore, engine data for a steady-state condition are obtained through experiment testing. Based on these data, an engine torque model and an effective fuel consumption model are established, as shown in Figures 3 and 4, respectively.

2.3.3. Battery Model. Based on a Ni-MH battery performance experiment, the electromotive force and internal resistance model are obtained as shown in Formulas (2) and (3):

$$E_{soc} = E_0 + \sum_1^5 E_i SOC^i
 \tag{2}$$

where E_0 is the electromotive constant of the battery, E_i is the fitting coefficients, SOC is the state of charge, and E_{soc} is the electromotive force under the current state:

$$R_{soc} = \delta_0 \left(R_0 + \sum_1^6 \lambda_i SOC^i \right)
 \tag{3}$$

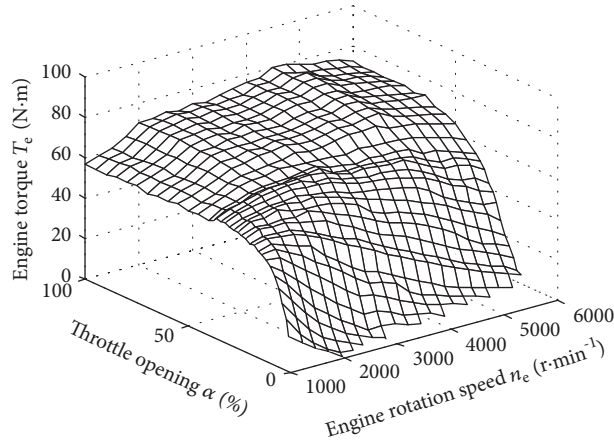


FIGURE 3: Engine torque model.

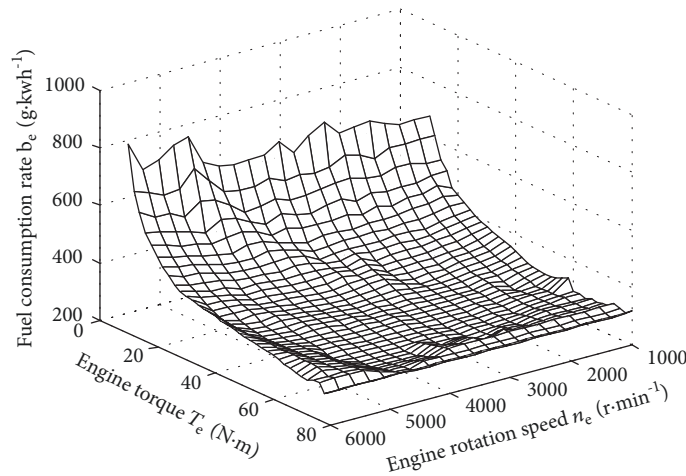


FIGURE 4: Engine fuel consumption rate model.

where R_0 is the internal resistance constant of battery, λ_i is the fitting coefficient, δ_0 is the compensation coefficient of internal resistance with the change in current, SOC is state of charge, and R_{soc} is internal resistance under the current state.

The process to calculate the change in SOC is shown by Formulas (4) ~ (6):

$$\Delta SOC = -\frac{I_{bat}(t)}{Q_{bat}} \quad (4)$$

$$I_{bat}(t) = \frac{E_{soc} - \sqrt{E_{soc}^2 - 4P_{bat}R_{soc}}}{2R_{soc}} \quad (5)$$

Therefore,

$$\Delta SOC = -\frac{E_{soc} - \sqrt{E_{soc}^2 - 4P_{bat}R_{soc}}}{2R_{soc}Q_{bat}} \quad (6)$$

where P_{bat} is the battery power, I_{bat} is the battery current, and Q_{bat} is the battery capacity.

3. Stochastic Modeling of Driver's Power Demand

Traditionally, the driver's power demand is obtained according to a given driving cycle; however, in reality, the driving cycle is random. The discrete-time stochastic process is regarded as Markov decision process (MDP). In other words, the transition probability from the current state to the next state only depends on the current state and the selected action; this is independent of historical states. The power demand at next state only depends on the current power demand, which is independent of any previous state. To establish the transition probability of power demand, a large volume of data is required. In this study, time-speed datasets of the UDDS and ECE.EUDC driving cycles are adopted to calculate the power demand at each moment (Figure 5). The transition probability matrix of the driver's power demand is obtained using the maximum likelihood estimation method.

The steps to calculate the power demand, based on the driving cycle, are as follows:

$$P_{req} \in \{P_{req}^1, P_{req}^2, \dots, P_{req}^s\} \quad (7)$$

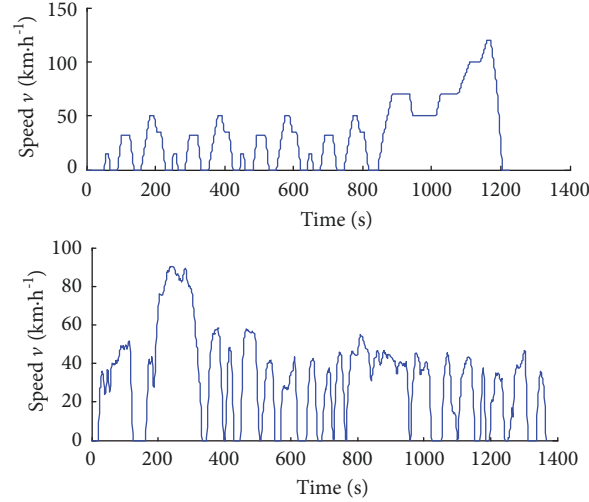


FIGURE 5: ECE_EUDC and UDSS driving cycles.

The transition probability P_{ij} is the probability from the current state P_{req}^i to the next state P_{req}^j :

$$P_{ij} = P \{ P_{req}(k+1) = P_{req}^j \mid P_{req}(k) = P_{req}^i \} \quad (8)$$

$$i, j = 1, 2, \dots, s, \text{ and } \sum_{j=1}^s P_{ij} = 1$$

According to the maximum likelihood estimation method, the power transition probability can be obtained by

$$P_{ij} = \frac{m_{ij}}{m_i} \quad i, j \in \{1, 2, \dots, s\} \quad (9)$$

where m_{ij} represents the number of times that the transition from P_{req}^i to P_{req}^j has occurred given the vehicle speed and m_i represents the total number of times that P_{req}^i has occurred at the vehicle speed; it is given by

$$m_i = \sum_{j=1}^s m_{ij} \quad (10)$$

Based on the data of ECE_EUDC and UDSS driving cycles, transition probability matrices for the power demand are obtained at given speeds of 10 km/h, 20 km/h, 30 km/h, 40 km/h, as shown in Figure 6.

4. Known-Model RL Energy Management Strategy

4.1. Addressing Energy Management by RL. From a macroscopic perspective, the energy management strategy of an HEV involves determining the driver's power demand according to the driver's operation (acceleration pedal or braking pedal) and distributing optimal power split between two power sources (engine and motor) on the premise of

guaranteeing dynamic performance. From a microscopic perspective, the energy management strategy can be abstracted as solving a sequential optimization decision problem. RL is a machine learning method based on Markov decision process, which can solve sequential optimization decision problems.

The process of solving sequential decision problem by RL is as follows. First, a Markov decision process is represented by the tuple (s, a, P, r, γ) , where s is the state set, a is the action set, P is the transition probability, r is the return function, and γ is the discount factor (Figure 7). Second, based on the Markov decision process, the vehicle controller of an HEV is regarded as an agent, the distribution of its engine power is regarded as action a , and the hybrid electric system except for the vehicle controller is regarded as the environment. In order to achieve minimum cumulative fuel consumption, the agent takes certain action to interact with the environment. After the environment accepts the action, the state begins to change and an immediate return is generated to feedback to the agent. The agent chooses the next action based on an immediate return and the current state of environment and, then, interacts with the environment again. The agent interacts with the environment continually, thus generating a considerable amount of data. RL utilizes the generated data to modify the action variable. Then, the agent interacts with the environment again and generates new data; this new data is utilized to further optimize the action variable. After several iterations, the agent will learn the optimal action that can complete the corresponding task; in other words, it will determine the decision sequence $\pi_0^* \rightarrow \pi_1^* \rightarrow \dots \rightarrow \pi_t^*$ (Figure 8), thereby solving the sequential decision problem.

4.2. Mathematical Model of an Energy Management Strategy Based on RL. Mathematical models of the action variable a , state variable s , and immediate return r are established according to RL theory. The SOC, vehicle speed v , and power demand P_{req} are taken as state variables; the engine power P_e is taken as the action variable, and the sum of equivalent fuel consumption of a one-step state transition and the SOC

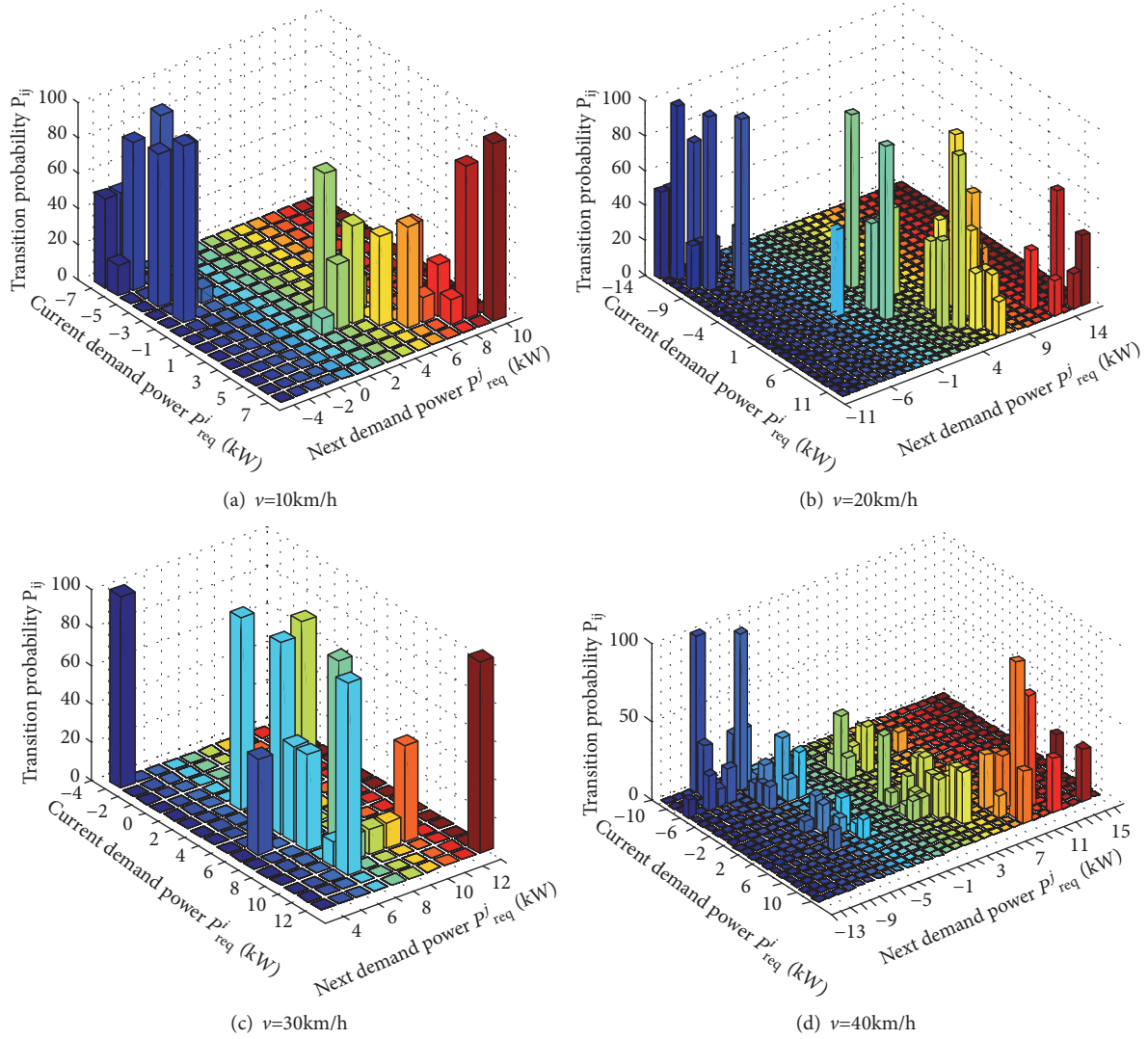


FIGURE 6: Power demand transition probability matrix.

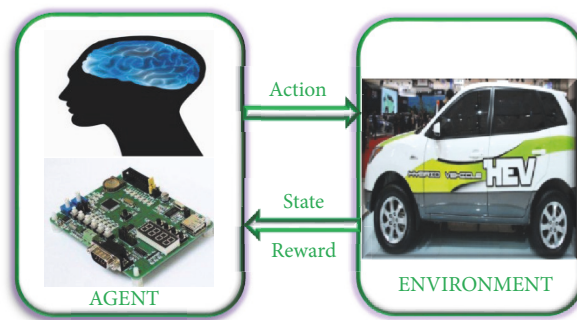


FIGURE 7: Schematic of reinforcement learning.

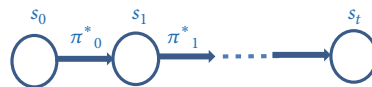


FIGURE 8: Policy decision sequence.

PI Algorithm

1. Input: state transition probability P_{ij} , return function r , discount factor γ .
Initialization value function: $V(s)=0$
Initialization policy π_0
2. Repeat $k=0,1,\dots$
3. find V^{π_k} policy evaluation
4. $\pi_{k+1}(s) = \arg \min_a q^{\pi_k}(s, a)$ policy improvement
5. Until $\pi_{k+1} = \pi_k$
6. Output: $\pi^* = \pi_k$

ALGORITHM 1: Flowchart of the PI algorithm.

penalty is taken as the immediate return r . λ' is a factor of the equivalent fuel consumption, m_{fuel} is the fuel consumption, m_{ele} is the electricity consumption, α is the penalty factor of the battery, and SOC_{ref} is reference value of the SOC:

$$\begin{aligned} s &= [SOC, v, P_{req}] \\ a &= P_e \\ r &= m_{fuel} + \lambda' m_{ele} + \alpha (SOC - SOC_{ref})^2 \end{aligned} \quad (11)$$

In an infinite time domain, the Markov decision process will solve the problem of determining a sequence of decision policies that can predict the minimum cumulative return expectation of a random process called the optimal state value function $V^*(s)$:

$$V^*(s) = \min_{\pi} E \left[\sum_{t=0}^{\infty} \gamma^t r_{t+1} \right] \quad (12)$$

where $\gamma \in [0, 1]$ is the discount factor, r_{t+1} represents the return value at $t+1$ time, and π represents the control policy.

Meanwhile, the control and state variables must also meet the following constraints:

$$\begin{aligned} SOC_{min} &\leq SOC \leq SOC_{max} \\ v_{min} &\leq v \leq v_{max} \\ P_{emin} &\leq P_e \leq P_{emax} \\ P_{mmin} &\leq P_m \leq P_{mmax} \end{aligned} \quad (13)$$

where the subscript min and max represent the maximum and minimum threshold values for the state of charge, speed, and power, respectively.

4.3. Policy Iterative Algorithm. For Formula (12), according to the Berman equation

$$\begin{aligned} V^*(s) &= \min_a \sum_{s'} \pi(s, a) P(s, a, s') \\ &\cdot (r(s, a, s') + \gamma V^*(s')) \end{aligned} \quad (14)$$

where s' represents the state at the next time step.

TABLE 2: Simulation results based on ECE_EUDC.

Control Policy	Fuel Consumption (L/100 km)
Rule-based	5.368
Instantaneous optimization	5.256
DP	4.952
RL	5.069

The purpose of the solution is to determine the optimal policy π^* :

$$\begin{aligned} \pi^*(s) &= \arg \min_a \sum_{s'} \pi(s, a) P(s, a, s') \\ &\cdot (r(s, a, s') + \gamma V^*(s')) \end{aligned} \quad (15)$$

A policy iterative (PI) algorithm is used to solve the problem of the random process. It involves a policy estimation step and a policy improvement-step.

The calculation process is shown in Algorithm 1.

In the policy evaluation step, for a control policy $\pi_k(s)$ (the subscript k represents the number of iterations), the corresponding state value function $V^{\pi_k}(s)$ is calculated, as shown in

$$\begin{aligned} V^{\pi_k}(s) &= \sum_a \pi_k(s, a) \sum_{s'} P(s, a, s') (r(s, a, s') + \gamma V^{\pi_k}(s')) \end{aligned} \quad (16)$$

In the policy improvement step, the improved policy $\pi_{k+1}(s)$ is determined through a greedy strategy, as shown in

$$\begin{aligned} \pi_{k+1}(s) &= \arg \min_a q^{\pi_k}(s, a) \\ q^{\pi_k}(s, a) &= r(s, a) + \gamma \sum_{s'} P(s, a, s') V^{\pi_k}(s') \end{aligned} \quad (17)$$

During policy iteration, policy evaluation and policy improvement are performed alternately until the state value function and the policy converge.

The policy iteration algorithm is adopted, obtaining the optimum control policy that takes the vehicle speed, driver's power demand, and SOC as the input and the engine power as the output. Figure 9 shows the optimized engine powers at vehicle speeds of 10 km/h, 20 km/h, 30 km/h, and 40 km/h.

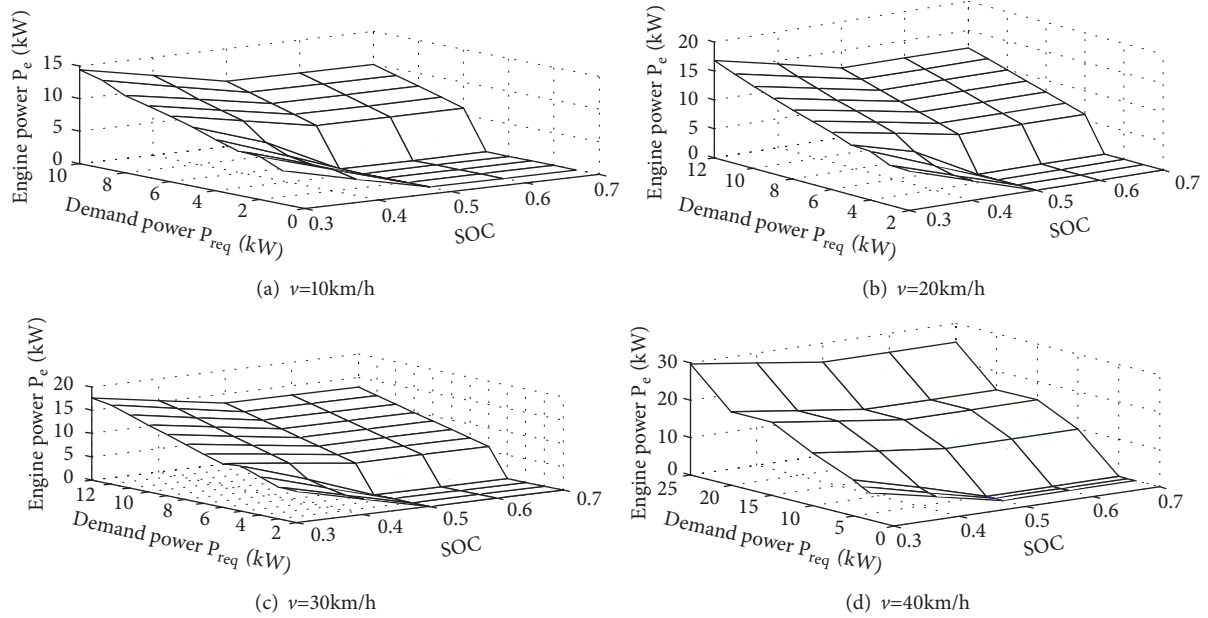


FIGURE 9: Optimization charts of engine power.

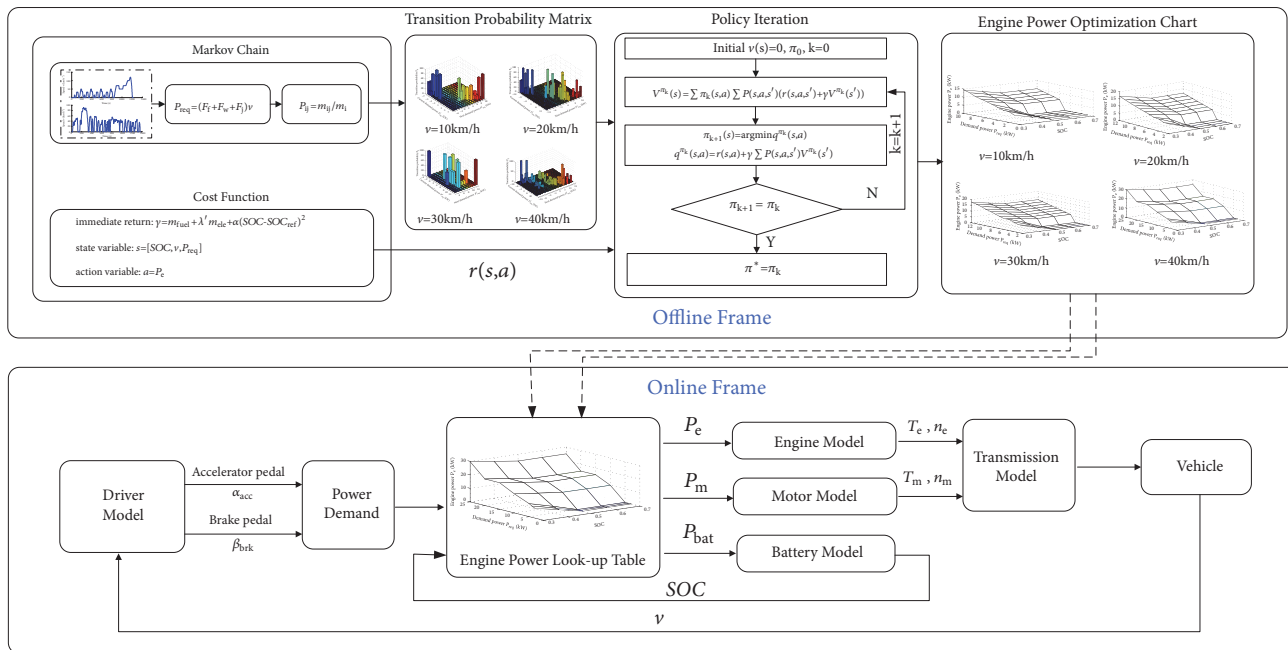


FIGURE 10: Implementation frame of the energy management strategy.

It can be seen from Figure 9 that only the motor works when the SOC is high and the power demand is low; when both the SOC and the power demand are high, only the engine works, and, when the SOC is low, the engine is in the charging mode.

Figure 10 illustrates the offline and online implementation frames of the energy management strategy. In the offline part, the power demand transition probability matrix is obtained by the Markov Chain. Mathematical models

of the state variable s , action variable a , and immediate return r are derived according to RL theory. The policy iteration algorithm is employed to determine engine power optimization tables. In the online part, the driver's power demand is obtained by the opening of driver's accelerator pedal and brake pedal. Then, the power of the engine and motor is distributed by looking up the offline engine power optimization tables. Finally, the power is transferred to the wheels by the transmission system.

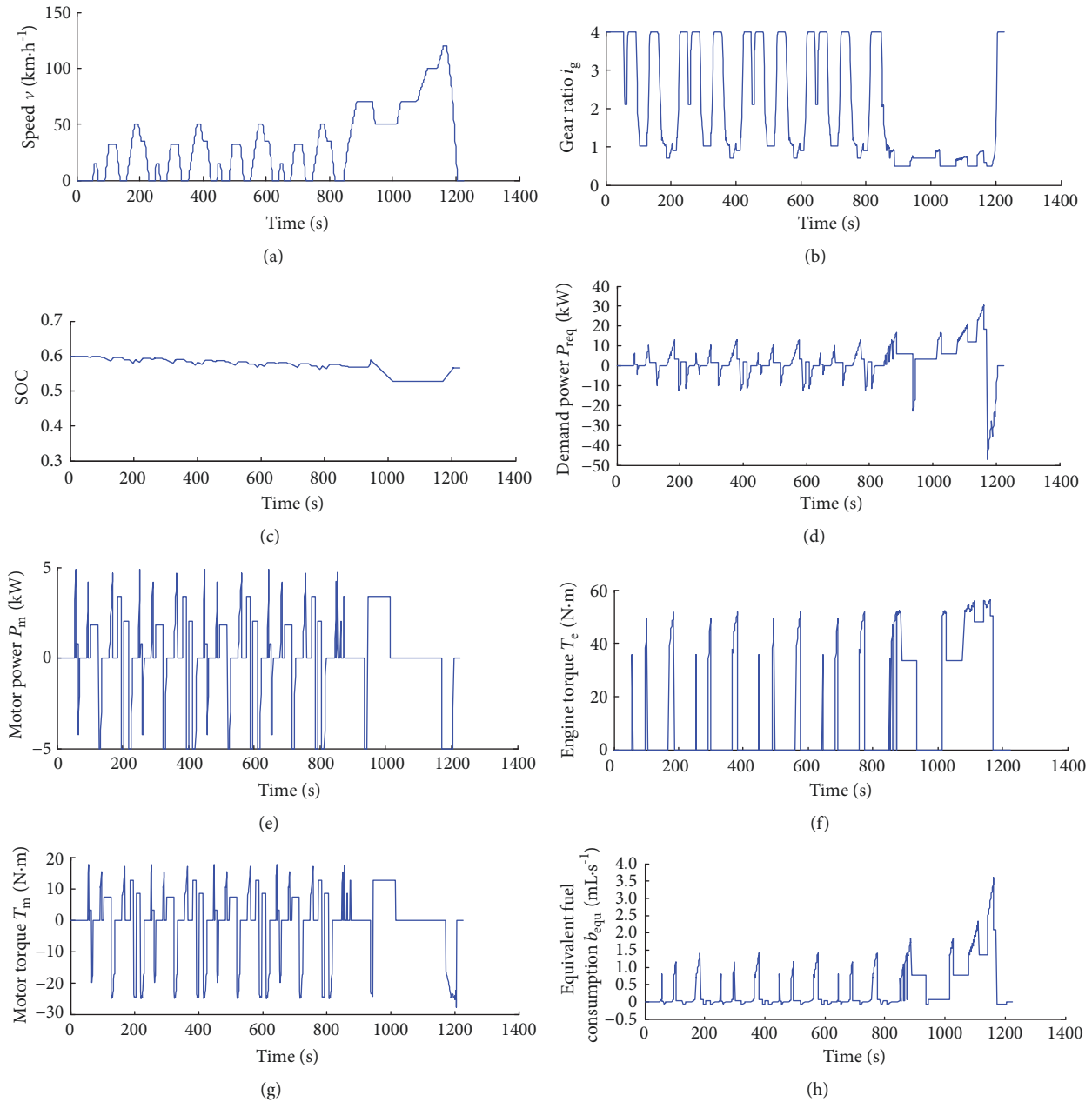


FIGURE 11: ECE_EUDC simulation results based on RL.

5. Simulation Experiment and Results

A simulation is conducted on a MATLAB/ Simulink platform, taking ECE_EUDC driving cycle as the simulation driving cycle and setting the initial SOC as 0.6. The energy management strategy based on a known model of RL is adopted, simulating the vehicle operation status online. Results are shown in Figure 11.

From Figure 11, we can see that the following parameters change with time: gear ratio of transmission i_g , SOC of the battery, power demand P_{req} , motor power P_m , engine torque T_e , motor torque T_m , and instantaneous equivalent fuel consumption b_{equ} . From Figure 11(b), it can be seen that

the gear ratio varies continuously from 0.498 to 4.04; this is primarily because of the use of a CVT with reflux power, which, unlike AMT, can realize power without interruption. In Figure 11(c), SOC of the battery is seen to change from the initial value of 0.6 to the final value of 0.5845; $\Delta SOC = 0.0155$; this can meet the HEV SOC balance requirement before and after the cycle ($-0.02 \leq \Delta SOC \leq 0.02$). The power demand change curve of the vehicle is obtained based on ECE_EUDC, as shown in Figure 11(d). The power distribution curves of the motor and the engine can be obtained according to the engine power optimization control tables, as shown in Figures 11(e) and 11(f). The motor power is negative, which indicates that the motor is in the generating state.

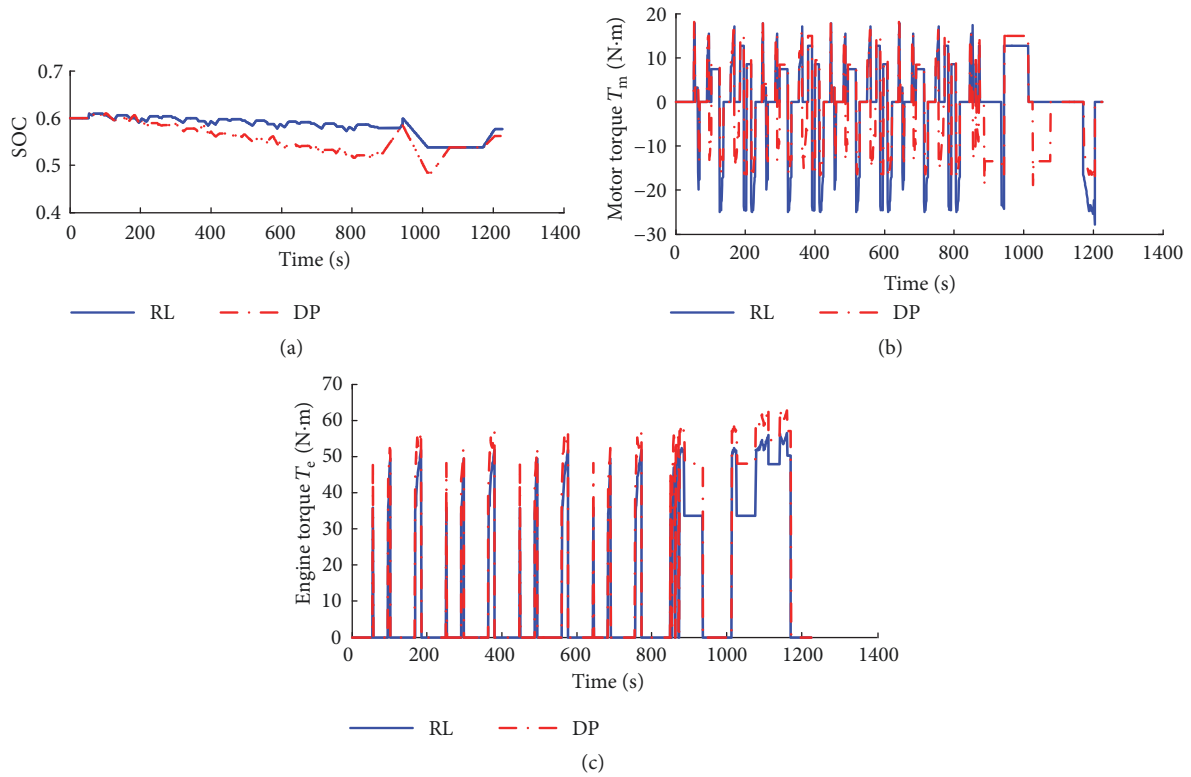


FIGURE 12: Comparison between DP and RL optimization results based on ECE_EUDC.

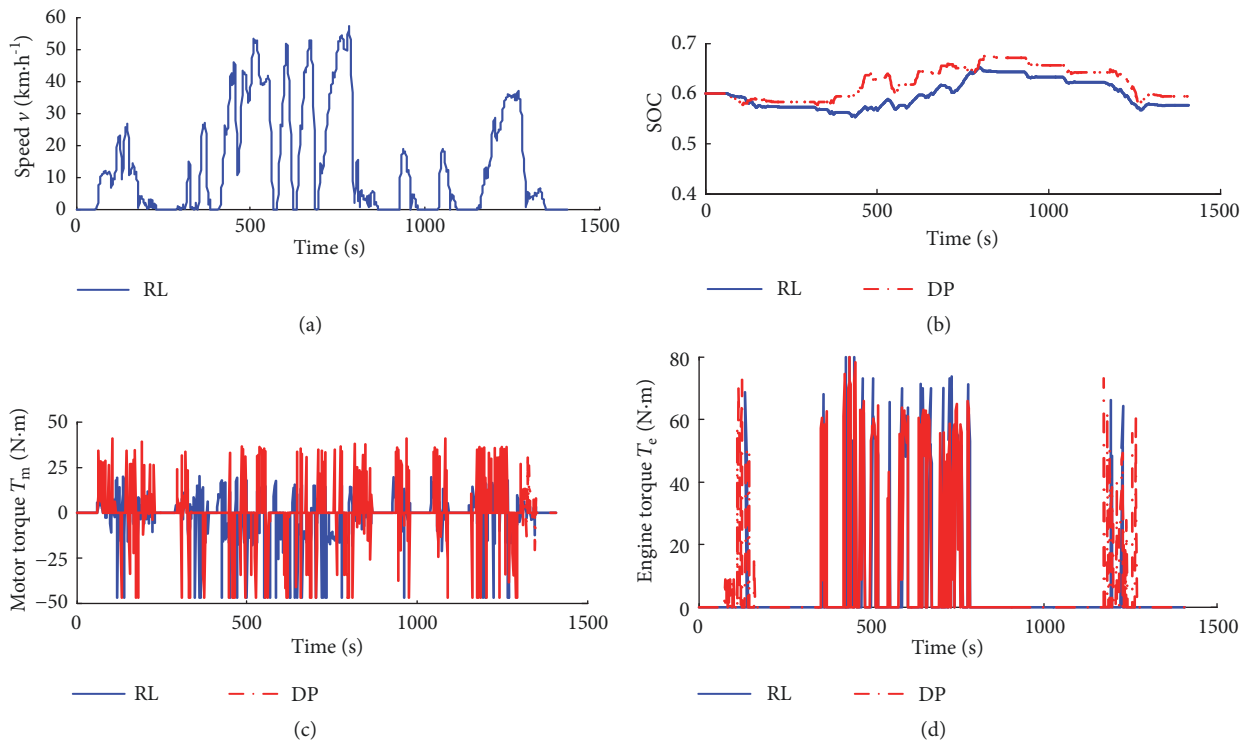


FIGURE 13: Comparison between DP and RL optimization results based on CYC_WVUCITY.

TABLE 3: Simulation results based on CYC_WVUCITY.

Control Policy	Fuel Consumption (L/100 km)	Computation time(s) ^a	
		offline	online
DP	4.983	5280	—
RL	4.864	4300	35

^aA 2.5GHz microprocessor with 8GB RAM was used.

In order to validate the optimal performance of the RL control strategy, vehicle simulation tests based on DP and RL were carried out. Figure 12 shows a comparison of the optimization results by the two control methods based on the ECE_EUDC driving cycle. The solid line indicates the optimization results of RL and the dotted line indicates the results of DP. Figure 12(a) shows the SOC optimization trajectories of the DP- and RL-based strategies. The trend of the two curves is essentially the same. The difference between the final SOC and the initial SOC is within 0.02, and the SOC is relatively stable. Compared with RL, the SOC curve of DP-based strategy fluctuates greatly; this is primarily related to the distribution of the power source torque. Figures 12(b) and 12(c) indicate engine and motor torque distribution curves based on DP and RL strategies. The engine torque curves essentially coincide; however, the motor torque curves are somewhat different. This is primarily reflected in the torque distribution when the motor is in the generation state.

Table 2 shows the equivalent fuel consumption obtained through DP and RL optimization. Compared with that obtained by DP (4.952L), the value of fuel consumption obtained by RL is 2.3% higher. The reason for this is that DP only ensures a global optimum at a given driving cycle (ECE_EUDC), whereas RL optimizes the result for a series of driving cycles in an average sense, thereby realizing a cumulative expected value optimum. Compared with rule-based (5.368L) and instantaneous optimization (5.256L) proposed in the literature [19], RL decreased the fuel consumption by 5.6% and 3.6%, respectively.

In order to validate the adaptability of RL to random driving cycle, CYC_WVUCITY is also selected as a simulation cycle. The optimization results are shown in Figure 13. Figure 13(b) shows the variation curve of the SOC of the battery based on DP and RL control strategies. The solid line indicates the optimization results of RL and the dotted line indicates the results of DP. The final SOC value for the DP and RL strategies is 0.59 and 0.58, respectively, with the same initial value of 0.6. The SOC remained essentially stable. Figures 13(c) and 13(d) show the optimal torque curves of the motor and the engine based on DP and RL strategies. It can be seen from the figure that the changing trend of the torque curves based on the two strategies is largely the same. In comparison, the change in the torque obtained by DP fluctuates greatly. Table 3 demonstrates the equivalent fuel consumption of the two control strategies. Compared with DP, RL saves fuel by 2.4%; it can adapt with random cycle perfectly. Meanwhile, the computation time based on DP and RL is recorded, which include offline and online computation time. The offline computation time of DP is 5280s and that of RL is 4300s. Due to large calculation volume and the

driving cycle unknown in advance, DP cannot be realized online, while RL is not limited to a certain cycle, which can be realized online by embedding the engine power optimization tables into the vehicle controller. The online operation time of RL is 35s based on CYC_WVUCITY simulation cycle.

6. Conclusion

(1) We established a stochastic Markov model of the driver's power demand based on datasets of the UDDS and ECE_EUDC driving cycles.

(2) An RL energy management strategy was proposed which takes SOC, vehicle speed, and power demand as state variables, engine power as power as the action variable, and minimum cumulative return expectation as the optimization objective.

(3) Using the MATLAB/Simulink platform, a CYC_WVUCITY simulation model was established. The results show that, compared with DP, RL could reduce fuel consumption by 2.4% and be implemented online.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research is supported by Scientific and Technology Research Program of Chongqing Municipal Education Commission (Grant no. KJQN201800718) and Chongqing Research Program of Basic Research and Frontier Technology (Grant no. cstc2017jcyjAX0053).

References

- [1] G. Zhao and Z. Chen, "Energy management strategy for series hybrid electric vehicle," *Journal of Northeastern University*, vol. 34, no. 4, pp. 583–587, 2013.
- [2] W. Wang and X. Cheng, "Forward simulation on the energy management system for series hybrid electric bus," *Automotive Engineering*, vol. 35, no. 2, pp. 121–126, 2013.
- [3] D. Qin and S. Zhan, "Management strategy of hybrid electrical vehicle based on driving style recognition," *Journal of Mechanical Engineering*, vol. 52, no. 8, pp. 162–169, 2016.

- [4] S. Zhan and D. Qin, "Energy management strategy of HEV based on driving cycle recognition using genetic optimized K-means clustering algorithm," *Chinese Journal of Highway and Transport*, vol. 29, no. 4, pp. 130–137, 2016.
- [5] X. Zeng and J. Wang, "A Parallel Hybrid Electric Vehicle Energy Management Strategy Using Stochastic Model Predictive Control With Road Grade Preview," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 6, pp. 2416–2423, 2015.
- [6] X. Tang and B. Wang, "Control strategy for hybrid electric bus based on state transition probability," *Automotive Engineering Magazine*, vol. 38, no. 3, pp. 263–268, 2016.
- [7] Z. Ma and S. Cui, "Applying dynamic programming to HEV powertrain based on EVT," in *Proceedings of the 2015 International Conference on Control, Automation and Information Sciences*, vol. 22, pp. 37–41, 2015.
- [8] Y. Yang and P. Ye, "A Research on the Fuzzy Control Strategy for a Speed-coupling ISG HEV Based on Dynamic Programming optimization," *Automotive Engineering*, vol. 38, no. 6, pp. 674–679, 2016.
- [9] Z. Chen, C. C. Mi, J. Xu, X. Gong, and C. You, "Energy management for a power-split plug-in hybrid electric vehicle based on dynamic programming and neural networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 4, pp. 1567–1580, 2014.
- [10] C. Sun, S. J. Moura, X. Hu, J. K. Hedrick, and F. Sun, "Dynamic traffic feedback data enabled energy management in plug-in hybrid electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 3, pp. 1075–1086, 2015.
- [11] B. Bader, "An energy management strategy for plug-in hybrid electric vehicles," *Materia*, vol. 29, no. 11, 2013.
- [12] Z. Chen, L. Li, B. Yan, C. Yang, C. Marina Martinez, and D. Cao, "Multimode Energy Management for Plug-In Hybrid Electric Buses Based on Driving Cycles Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2811–2821, 2016.
- [13] Z. Chen and Y. Liu, "Energy Management Strategy for Plug-in Hybrid Electric Bus with Evolutionary Reinforcement Learning method," *Journal of Mechanical Engineering*, vol. 53, no. 16, pp. 86–93, 2017.
- [14] T. Liu, B. Wang, and C. Yang, "Online Markov Chain-based energy management for a hybrid tracked vehicle with speedy Q-learning," *Energy*, vol. 160, pp. 544–555, 2018.
- [15] L. Teng, "Online Energy Management for Multimode Plug-in Hybrid Electric Vehicles," *IEEE Transactions on Industrial Informatics*, vol. 2018, no. 10, pp. 1–9, 2018.
- [16] J. Wu, H. He, J. Peng, Y. Li, and Z. Li, "Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus," *Applied Energy*, vol. 222, pp. 799–811, 2018.
- [17] T. Liu, X. Hu, S. E. Li, and D. Cao, "Reinforcement Learning Optimized Look-Ahead Energy Management of a Parallel Hybrid Electric Vehicle," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 4, pp. 1497–1507, 2017.
- [18] T. Liu and X. Hu, "A Bi-Level Control for Energy Efficiency Improvement of a Hybrid Tracked Vehicle," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1616–1625, 2018.
- [19] Y. Yin, *Study on Optimization Matching and Simulation for Super-Mild HEV*, Chongqing university, Chongqing, China, 2011.



Hindawi

Submit your manuscripts at
www.hindawi.com

