PLOS ONE

RESEARCH ARTICLE

Implementation of real-time energy management strategy based on reinforcement learning for hybrid electric vehicles and simulation validation

Zehui Kong, Yuan Zou*, Teng Liu

National Engineering Laboratory for Electric Vehicles, School of Mechanical Engineering, Institute of Technology, Beijing, China

* zouyuan@bit.edu.cn

OPEN ACCESS

Citation: Kong Z, Zou Y, Liu T (2017) Implementation of real-time energy management strategy based on reinforcement learning for hybrid electric vehicles and simulation validation. PLoS ONE 12(7): e0180491. https://doi.org/10.1371/journal.pone.0180491

Editor: Xiaosong Hu, Chongqing University, CHINA

Received: April 6, 2017

Accepted: June 15, 2017

Published: July 3, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the National Nature Science Foundation of China (Grant 51375044), University Science Introduction 111 Project (B12022), and Defense Basic Research Project (B20132010).

Competing interests: The authors have declared that no competing interests exist.

Abstract

To further improve the fuel economy of series hybrid electric tracked vehicles, a reinforcement learning (RL)-based real-time energy management strategy is developed in this paper. In order to utilize the statistical characteristics of online driving schedule effectively, a recursive algorithm for the transition probability matrix (TPM) of power-request is derived. The reinforcement learning (RL) is applied to calculate and update the control policy at regular time, adapting to the varying driving conditions. A facing-forward powertrain model is built in detail, including the engine-generator model, battery model and vehicle dynamical model. The robustness and adaptability of real-time energy management strategy are validated through the comparison with the stationary control strategy based on initial transition probability matrix (TPM) generated from a long naturalistic driving cycle in the simulation. Results indicate that proposed method has better fuel economy than stationary one and is more effective in real-time control.

1. Introduction

The hybrid electric vehicles (HEVs) are booming rapidly as a solution to the depletion of fossil fuel and severe pollution of air condition. Due to the cooperation of a battery pack and the internal combustion engine, the vehicle powertrain allows the engine to avoid operating at low load with poor efficiency, and the fuel economy and emission can be improved significantly. However, the flexibility of power split also makes energy management problem more challenging.

Energy management strategy (EMS) plays a crucial role in trade-off among performance, fuel economy and emission of HEVs. Numerous studies have been conducted in the energy management of HEVs [1]. Generally, energy management strategies of HEVs are classified into rule-based and optimization-based control strategy [2,3]. The rule-based strategy is widely used in practice due to the straightforward implementation and high computation efficiency. Jalil proposed a rule-based energy management strategy to determine the power split between the engine and battery by setting thresholds [4]. Trovão presented a new rule-based energy

management strategy integrating meta-heuristic optimization for a mutilevel EMS in a electric vehicle [5]. To further improve the performance of energy management system, an adaptive fussy logic controller was used to calibrate the operating points and key parameters to minize the fuel consumption according to the driving cycles [6]. However, the performance of any ruel-based strategy is highly dependent on the proper design of the control rules, which usually depends on the enginenering experience. Therefore, many researchers make more efforts to optimization-based energy strategy.

With a prior knowledge of the driving cycles, dynamic programming (DP) receives a optimal result and determines the best fuel economy. However, the real-time and robust performance of this strategy cannot be guaranteed [7]. Instead, DP is implemented offline and served as a benchmark to explore the potential of fuel economy [8]. To make on-line optimization possible, equivalent consumption minimization strategy (ECMS) and model predictive control (MPC) have been adopted to develop energy management [9,10]. ECMS is calculated based on the assumption that the variation of *SOC* (state of charge of the batttery) is neligible due to the slow dynamics compared to other dynamics in HEV [11]. The equivalence factor of ECMS has an important effect on the control performance. However, the optimal value should be determined offline according to a specific cycle [12]. MPC is a promising method for dynamic model due to the prediction ability in a finite future time-horizon. A MPC-based strategy is developed by predicting the road slope. The results show that the method not only maintains the battery *SOC* within its boundary, but also achieves better fuel economy [13]. A Pontryagain's minimum principle (PMP) is used to find the optimal energy management strategy through combining the power prediction based on the traffic information, such as the maximum acceleration, average velocity and maximum velocity [14]. However, the performance of MPC depends on the prediction accuracy heavily, and varying weather conditions and drving styles make it difficult to guarantee the accuracy.

The exsiting mehods mostly considered a single objective to optimize, such as fuel consuption, while disregarding many onther concerns. The convex multicriteria optimization approach was recently harnessed to optimize mutiple objectives of plug-in hybride electric vehicles, including the battery sizing, charging and on-road power management [15]. Ref. [16] studied the optimal tradeoff between the fuel-cell durability and hydrogen economy for a fuel-cell hybrid bus. Ref. [17] investgates the interactions among three control tasks, such as charging, on-road power management and battery degradation mitigation in samrt grid environment, aiming to minimize the daily operational expense of a PHEV. A high-efficiency convex programming framework is harnessed to minimize daily CO_2 emissions throgh integrating renewable energy and system-level hybrid powertrain optimization [18].

Several novel algorithms for energy management of HEVs have been developed to realize online optimization for multiple types of HEVs, such as the game theory [19], stochastic dynamic programming (SDP) [20], and reinforcement learning (RL) [21]. Wang optimized the power management problem for a hybrid electric vehicle based on SDP algorithm. However, heavy computation burden makes it difficult to implement online [22]. In numerous areas, RL is a heuristic learning method. Ref. [23] applied RL in the energy management strategy for an electric hybrid tracked vehicle, and compared the performance of RL and SDP. The results indicate that RL algorithm has a better performance and a shorter computation time. However, the RL doesn't update online and fails to maintain the optimal performance when the driving cycles vary.

This paper proposes a real-time energy management strategy for a small series hybrid electric tracked vehicle. A recursive algorithm is developed to compute the transition probability matrix (TPM) of power-request when the new statistical characteristics of online driving cycle are considerd. RL algorithm is applied to obtain the optimal control policy based on updating TPM of power-request at the regular intervals. A facing-forward simulation model is

estabalised to evaluate RL-based real-time energy management strategy through the comparison with RL-stationary control strategy. Simulation results show that the proposed method achieves better fuel economy than RL-stationary one and is feasible for real-time control.

2. Modelling of hybrid electric tracked vehicle

2.1 Vehicle configuration and parameters

The structure of the hybrid electric tracked vehicle is illustrated in Fig 1. An engine-generator-rectifier set (EGS) and a battery pack supply the electricity to dual motors, propelling the sprocket independently. The engine gives 50 kW maximum power and 93 Nm maximum torque within the speed range from 1200 r/min to 6200 r/min. The generator offers 107 Nm maximum torque within the speed range from 0 r/min to 6400 r/min and 40 kW maximum power. The 37.6 Ah lithium-ion phosphate battery pack gives 307 rated voltage. The essential parameters of the major sub-systems are listed in Table 1.

2.2 Vehicle and powertrain modelling

The vehicle is taken as a concentrated mass. The dynamical equation of the vehicle is expressed as

$$ma = F_{TR} - (mg\sin\theta + mgf\cos\theta + C_D A v_{ave}{}^2/21.15) \tag{1}$$

where F_{TR} means the tractive force; m is the curb weight, and a is the vehicle acceleration; g represents the gravity acceleration; θ means the road slope angle; f is the rolling resistance coefficient; C_D is the aerodynamic drag coefficient; A represents the front area of the vehicle; v_{ave} means the average speed of the two tracks and is determined by $v_{ave} = (v_1+v_2)/2$; v_1, v_2 are the speed of the two tracks, respectively.

The demand power to propel the vehicle, denoted by P_{dem}, is calculated by

$$\begin{cases} P_{dem} = F_{TR} \bullet v_{ave} + M \bullet \omega \\ M = \dfrac{1}{4} u_t mgL \\ u_t = u_{\max} \cdot (0.925 + 0.15R/B)^{-1} \\ R = \dfrac{v_{ave}}{|\omega|} \\ \omega = \dfrac{v_1 - v_2}{B} \end{cases} \tag{2}$$

where P_{dem} consists of two parts, straight power and steering power. M is the resisting yaw moment; ω represents the rotational speed; u_t is the lateral resistance coefficient; R is the turning radius; L is behalf of the contacting track width, and B means the tread of the vehicle; u_{max} is the maximum steering resistance coefficient with the radius of braking steering, $R = B/2$.

To analyze and evaluate the EGS's fuel economy, the equivalent electric circuit is established in Fig 2, which consists of the diesel engine, permanent magnet generator, and rectifier. The output voltage and electromagnetic torque, U_g and T_g of the generator are determined as [24]

$$\begin{cases} U_g = K_e\omega_g - K_x\omega_g I_g \\ T_g = K_e I_g - K_x I_g^2 \end{cases} \tag{3}$$

where ω_g is the angular velocity; $K_e\omega_g$ represents the electromotive force; K_x is the equivalent resistance coefficient and $K_x = 3PL^g/\pi$; L^g is the synchronous inductance of the armature, and P is the number of poles; I_g and U_g are the current and output voltage of the generator.

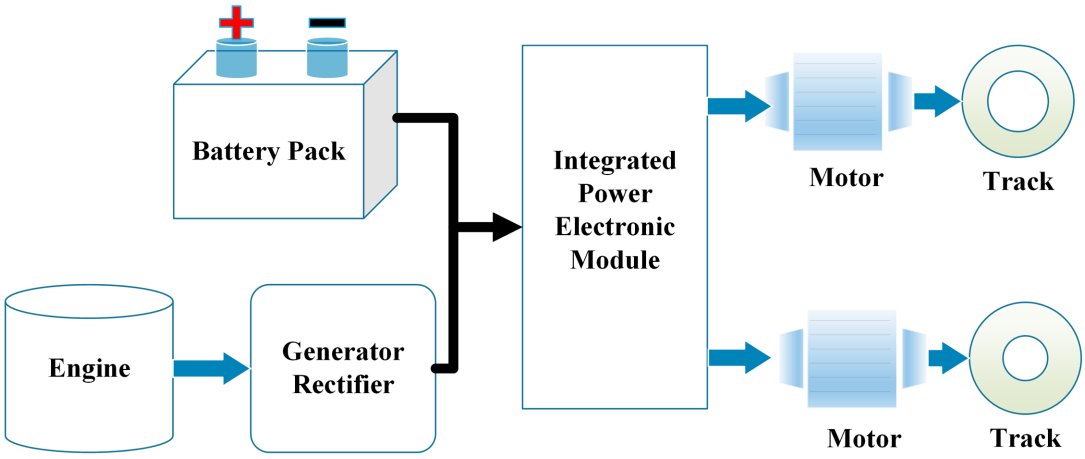

Fig 1. The powertrain structure.

https://doi.org/10.1371/journal.pone.0180491.g001

The fuel consumption of the vehicle is determined by the engine torque T_{eng} and speed n_{eng}. The speed of the engine is determined as follows.

$$\begin{cases} (J_{\text{eng}} + J_{\text{g}})\dfrac{\pi}{30}\dfrac{dn_{\text{g}}}{dt} = T_{\text{eng}} - T_{\text{g}} \\ n_{eng} = n_g \end{cases} \tag{4}$$

where n_g, n_{eng} denote the rotational speed of the generator and engine, respectively; J_{eng} and J_{g} are the moments of inertia for the engine and generator; T_{eng} is the torque of engine and T_g is the electromagnetic torque of the generator.

An internal-resistance model is used to reflect *SOC* dynamics as [25]

$$\begin{cases} \dfrac{d\text{SOC}}{dt} = -\dfrac{I_b}{C_b} \\ SOC_{\min} \le SOC(t) \le SOC_{\max} \end{cases} \tag{5}$$

where C_b is the battery capacity; I_b means the battery current which is calculated by

$$\begin{cases} I_b = \dfrac{(V_{OC} - \sqrt{{V_{OC}}^2 - 4R_{int}P_b})}{2R_{int}} \\ I_{b,\min} \le I_b(t) \le I_{b,\max} \end{cases} \tag{6}$$

Table 1. Parameters of the major sub-systems.

Symbol	Parameter	Value
m	Curb Weight	2500kg
r	Sprocket radius	0.15m
f	Rolling coefficient	0.05
A	Front area	$0.7m^2$
C_D	Air drag coefficient	0.6
i_0	Ratio from motors to sprockets	41/6
B	Tread of the vehicle	1.1m
K_e	Electromotive force parameter	$0.887V_{sard}^{-2}$
Kx	Equivalent resistance coefficient	$0.00105NmA^{-2}$

https://doi.org/10.1371/journal.pone.0180491.t001

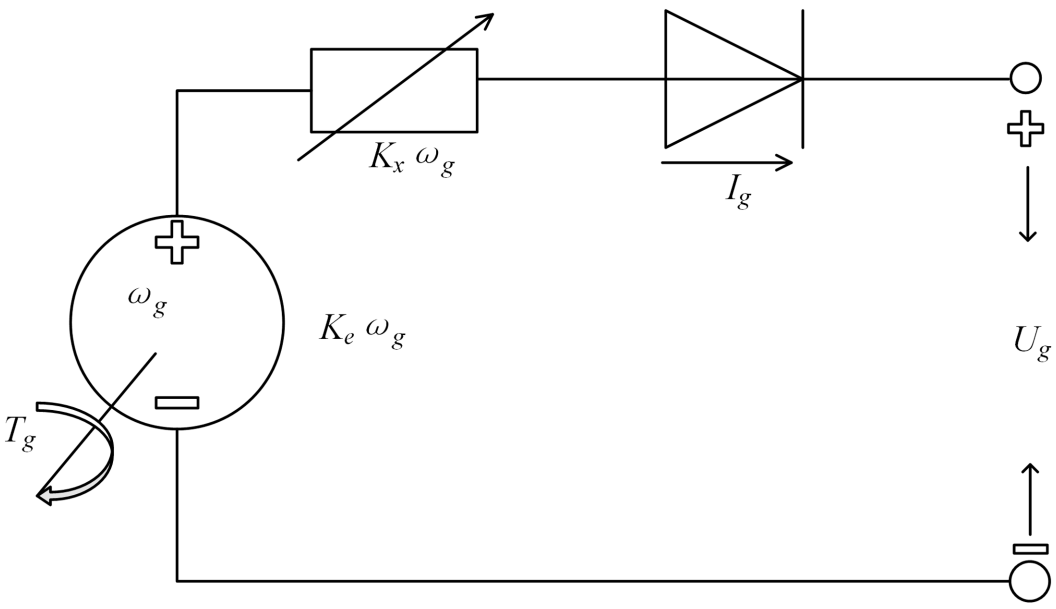

Fig 2. Equivalent electric circuit of the EGS.

https://doi.org/10.1371/journal.pone.0180491.g002

where V_{OC} is the open circuit voltage; R_{int} means the internal resistance of battery and P_b is the output power of the battery.

2.3 Optimal energy management problem formulation

Cost function is the trade-off of the fuel consumption and *SOC* variation, ensuring that the final *SOC* of battery stays at the same level of the initial value, expressed as follows

$$J = \int_{t_0}^{t_f} (\dot{fuel}(t) + \beta(SOC(t) - SOC_{ref})^2)dt \tag{7}$$

where β represents the penalty factor, which is a positive weighting factor; *fuel* is the fuel consumption at time *t*. The rate of flowing fuel mass $\dot{fuel}(t)$ is determined by engine torque T_{eng} and rotational speed n_{eng} based on BSFC (braked specific fuel consumption) map, typically obtained through a bench test. The engine's torque is used to regulate the power split between the EGS and battery to minimize the total fuel consumption.

3. Real-time energy management strategy

3.1 Online updating transition probability matrix (TPM)

The power-request of the vehicle is calculated according to the Eq (2). Using the maximum likelihood estimation and nearest method, the transition probability matrix of the power-request is described as

$$p_{ij} = \frac{N_{ij}}{N_{oi}} = \frac{N_{ij}(k)/k}{N_{oi}(k)/k} = \frac{F_{ij}(k)}{F_{oi}(k)} \tag{8}$$

where N_{ij} denotes the transition numbers from state x_i to state x_j, and N_{oi} is the total transition numbers initiated from state x_i; k means the number of transition; F_{ij} is the total frequency rate of transition event f_{ij} from state x_i to state x_j and F_{oi} is the frequency rate of transition event f_i

initiated from state x_i. $f_{ij}(t) = 1$ if a transition is occurred from x_i to x_j at time instant t; $f_i(t) = 1$ if there occurs a transition initiated from x_i at time instant t, otherwise, these take zero values. Then $\sum_{t=1}^{k} f_{ij}(t) = N_{ij}(k)$ and $\sum_{t=1}^{k} f_i(t) = N_i(k)$. The frequency rate F_{ij} and F_i are calculated as follows:

$$\begin{cases} F_{ij}(k) = \dfrac{N_{ij}(k)}{k} = \dfrac{1}{k}\sum_{t=1}^{k} f_{ij}(t) \\ F_{oi}(k) = \dfrac{N_{oi}(k)}{k} = \dfrac{1}{k}\sum_{t=1}^{k} f_i(t) \end{cases} \tag{9}$$

and the recursive expressions are deduced [26].

$$\begin{cases} F_{ij}(k) = \dfrac{1}{k}\sum_{t=1}^{k} f_{ij}(t) = \dfrac{1}{k}[(k-1)F_{ij}(k-1) + f_{ij}(k)] \\ \qquad = F_{ij}(k-1) + \dfrac{1}{k}[f_{ij}(k) - F_{ij}(k-1)] \\ F_{oi}(k) = \dfrac{1}{k}\sum_{t=1}^{k} f_i(t) = \dfrac{1}{k}[(k-1)F_{oi}(k-1) + f_i(k)] \\ \qquad = F_{oi}(k-1) + \dfrac{1}{k}[f_i(k) - F_{oi}(k-1)] \end{cases} \tag{10}$$

where $1/k$ is replaced by a constant ψ ranging from 0 to 1, namely forgetting factor to determine the effective memory depth of historic driving cycle, and the forgetting factor ψ is set to be 0.01 in this paper.

By substituting the expression (10) into (8), the recursive algorithm of TPM is derived for online learning as follows:

$$p_{ij} = \frac{F_{ij}(k)}{F_{oi}(k)} = \frac{F_{ij}(k-1) + \psi[f_{ij}(k) - F_{ij}(k-1)]}{F_{oi}(k-1) + \psi[f_i(k) - F_{oi}(k-1)]} \tag{11}$$

3.2 RL-based real-time energy management strategy

The driving schedule is considered as a finite Markov decision process, which comprises a set of state variables $s_t \in S = \{SOC(t), n_g(t) | 0.6 \le SOC(t) \le 0.9, 0 \le n_g(t) \le 6400\}$, a set of actions $a_t \in A = \{T_{eng}(t) | 0 \le T_{eng}(t) \le 93\}$, a reward function $r_t \in Reward = \{fuel(s_t, a_t)\}$.

Corresponding to the state s and action a, the optimal value of the state s_t is defined as the expected finite discounted sum of the rewards, as follows [27]:

$$\begin{aligned} V^*(s_t) &= \min_{\pi} E\Big(\sum_{t=t_0}^{t=t_f} \gamma^t r_t\Big) \\ &= \min_{a_t}\Big(r_t(s_t, a_t) + \gamma \sum_{s_t' \in S} p_{s_t a_t s_t'} V^*(s_t')\Big) \quad \forall s_t \in S \end{aligned} \tag{12}$$

where π is a control policy, $\gamma \in [0, 1]$ is a discount factor; p means the probability of the occurrence of a transition from state s_t to s_t' under action a_t. And the optimal control policy can be

decided by the function

$$\begin{aligned}\pi^*(s_t) &= \arg\min_{a_t} E\left(\sum_{t=t_0}^{t=t_f}\gamma^t r_t\right)\\ &= \arg\min_{a_t}\left(r_t(s_t,a_t)+\gamma\sum_{s_t'\in S}p_{s_t a_t,s_t'}V^*(s_t')\right)\end{aligned} \tag{13}$$

In addition, the Q value $Q(s_t, a_t)$ and optimal value Q^* corresponding to state s_t and action a_t are defined as follows

$$\begin{cases} Q(s_t,a_t) = r_t(s_t,a_t)+\gamma\sum_{s_t'\in S}p_{s_t a_t,s_t'}Q(s_t',a_t')(s',a_t') \\ Q^*(s_t,a_t) = r_t(s_t,a_t)+\gamma\sum_{s_t'\in S}p_{s_t a_t,s_t'}\min_{a'}Q^*(s_t',a_t') \end{cases} \tag{14}$$

Finally, the updating rule of Q-learning and optimal control strategy are defined as [28, 29]:

$$\begin{cases} Q(s_t,a_t) := Q(s_t,a_t)+\alpha(r_t+\gamma\min_{a'}Q(s_t'+a_t')-Q(s_t,a_t)) \\ \pi^*(s_t) = \arg\min_{a_t} Q^*(s_t,a_t) \end{cases} \tag{15}$$

where $\gamma\in[0, 1]$, $\alpha\in[0, 1]$ are respectively discount factor and decayed factor in Q-learning. The computational flowchart of RL is listed in the pseud code format as shown in Fig 3.

Through adopting Q-learning algorithm above, the stationary control policy is derived based on initial TPM generated from a long naturalistic driving cycle as shown in Fig 4. This paper proposes a real-time energy management strategy, aiming at improving the adaptability to change of power-request characteristics. Firstly, the updating TPM of power-request is acquired according to Eq (11) in real-time. Then at set intervals, the control policy is

Initialize $Q(s, a)$ arbitrarily

Repeat (for each episode):

  Initialize s

  Repeat (for each step of episode)

    Choose a from s using policy derived from Q (ε-greedy)

    Take action a, observe r, s'

    $Q(s,a)=Q(s,a)+\alpha\left(r+\gamma\min_{a'}Q(s'+a')-Q(s,a)\right)$

    $s=s'$

  until s is terminal

Fig 3. The pseud code of *Q*-learning.

https://doi.org/10.1371/journal.pone.0180491.g003

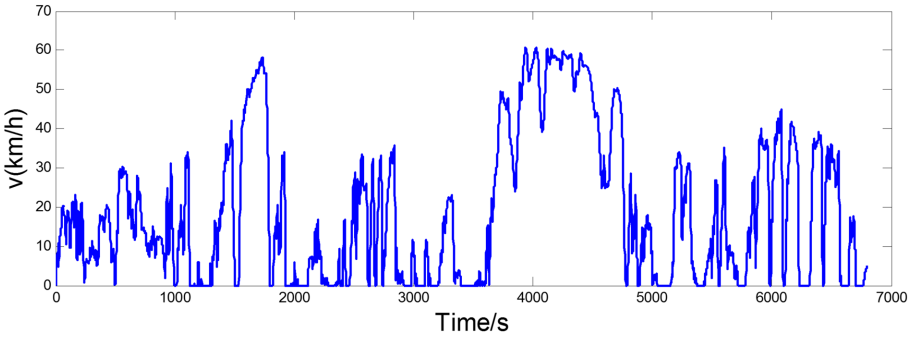

Fig 4. A long naturalistic driving cycle.

https://doi.org/10.1371/journal.pone.0180491.g004

recalculated by using *Q*-learning algorithm, adapting the controller to the varying driving schedules. Table 2 lists the fuel consumption and calculation time in cases of choosing different time interval. The results show that the fuel consumptions have no big difference, with 50 s, 100 s, 200 s and 300 s selected as the time interval. When fuel consumptions are similar, the case of longer time interval can reduce the updating number of control policy and calculation time. Therefore, the time interval of updating control policy is set to be 300 s in this paper. And the design process of the real-time energy management strategy is depicted in Fig 5.

4. Simulation and validation

The simplified electric coupling model of powertrain for hybrid electric tracked vehicle is given in Fig 6, where the power provided from EGS and battery should keep balance with the power-request of two driving motors. Based on the electric coupling model and powertrain model in Section 2, a facing-forward detailed model for the hybrid electric tracked vehicle is established in the Simulink environment, consisting of the engine-generator model, power battery pack, two motors, vehicle dynamic model and controller as shown in Fig 7. The proportional-integral driver model is adopted to adjust the toques of both motors to follow the target driving cycles. The controller determines the throttle according to the control map obtained through RL algorithm when the states of vehicle feedbacks to the controller. The RL-based real-time energy management strategy and stationary energy management strategy are applied into the facing-forward simulation model for the same driving cycle, respectively. The control map remains unchanged in the case of stationary strategy while the control map is updated every 300s in the case of real-time strategy. The initial values of the state variables n_g and *SOC* are taken as 1200 rpm and 0.75, respectively.

Fig 8 shows the experimental driving schedule used in the simulation to validate the proposed method. The *SOC* trajectories and power split between EGS and battery for the experimental driving schedule are illustrated in Fig 9 and Fig 10. Final *SOC* of two methods were

Table 2. Fuel consumption.

Time interval (s)	Fuel consumption(g)	Relative increase	Calculation time (h)	Updating number
50	430.2	—	9.5	19
100	430.8	0.1%	4.5	9
200	432.0	0.4%	2	4
300	432.5	0.5%	1.5	3
400	445.8	3.6%	1	2

https://doi.org/10.1371/journal.pone.0180491.t002

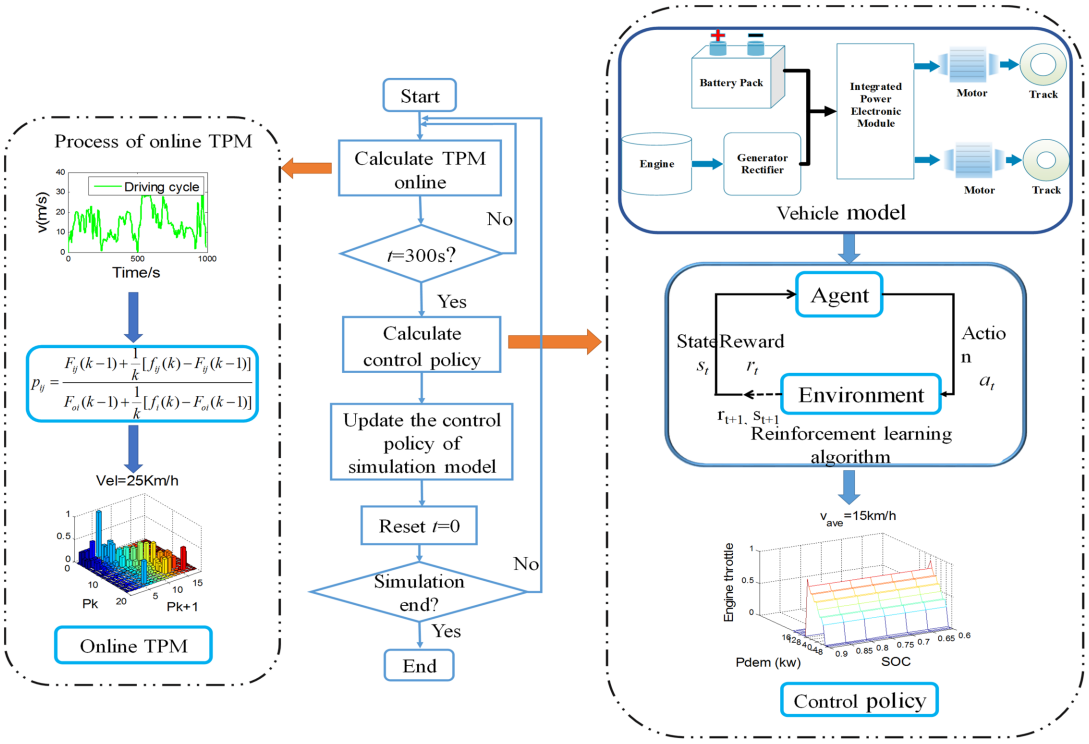

Fig 5. Process of real-time energy management strategy design.

https://doi.org/10.1371/journal.pone.0180491.g005

close to initial *SOC* value due to the final constraints for *SOC* value. During the first 300 seconds, two control strategies are based on the same TPM, so *SOC* trajectories change almost in the same way. However, when the driving condition changes, the control policy is recalculated at 300 s for real-time control strategy. And the same process is triggered at 600 s and 900 s. Similarly, the power split changes correspondingly as shown in Fig 10. For the sake of eliminating the influence of the deviation between the final *SOC* on fuel consumption, an *SOC*-correction method [30] is utilized to compensate for the fuel consumption. Table 3 shows the fuel consumption of two methods after *SOC* correction. The fuel consumption of RL-based real-time strategy is 6% lower than that of stationary strategy. Fig 11 depicts the working points of

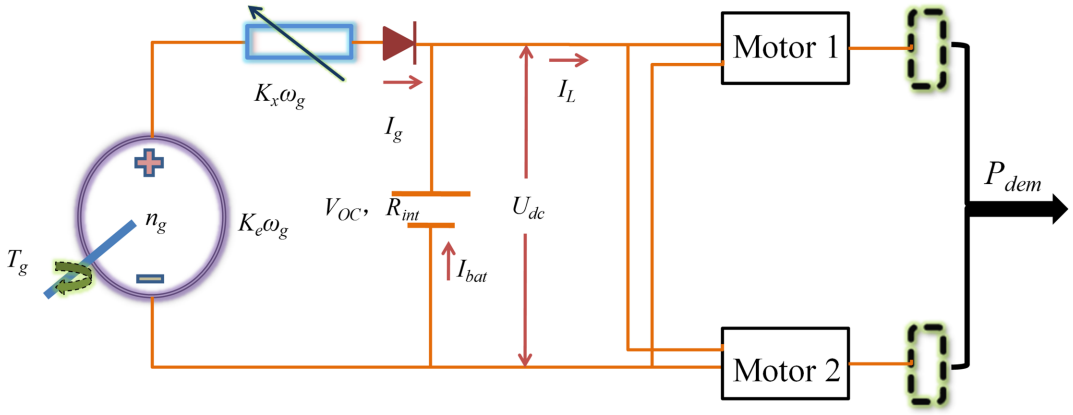

Fig 6. The simplified electric coupling model.

https://doi.org/10.1371/journal.pone.0180491.g006

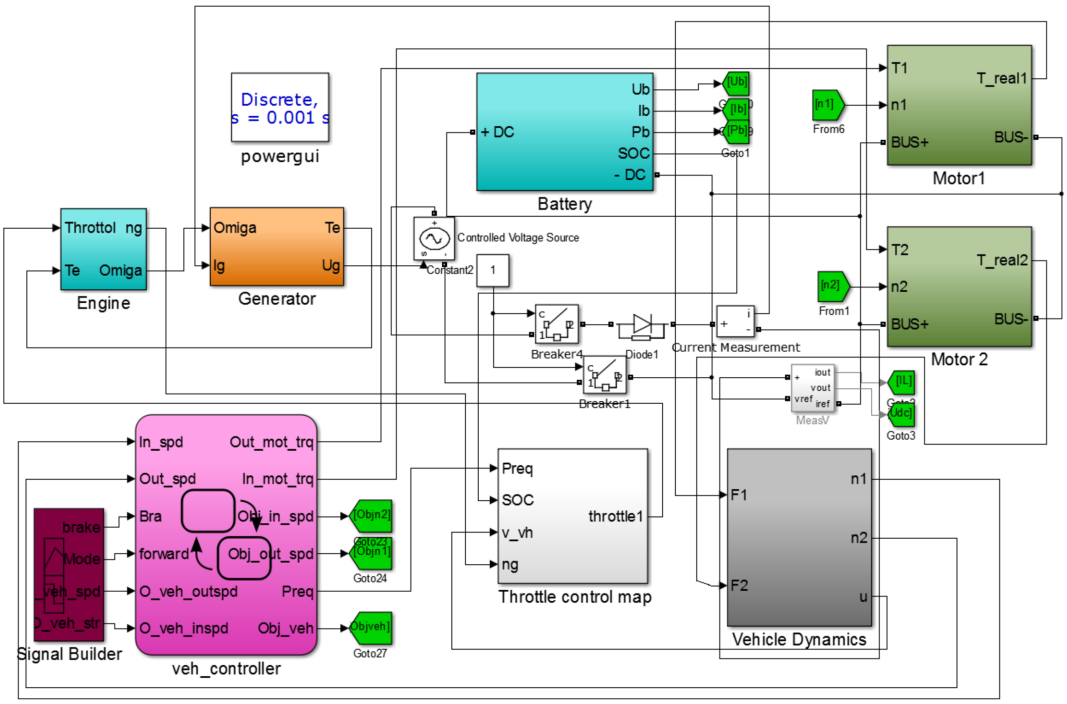

Fig 7. The facing-forward simulation model.

https://doi.org/10.1371/journal.pone.0180491.g007

the engine for the experimental driving schedule. There are more points in the optimal fuel consumption area in the real-time strategy than that of stationary one.

To further explore the adaptability and robustness of RL-based real-time strategy, the same procedure is performed for the validation driving schedule in the field test (shown in Fig 12). The *SOC* trajectories and working points of engine are shown in Fig 13 and Fig 14. Because of the constraint for *SOC* value, the final *SOC* values are close to the initial *SOC* value. Table 4 lists the fuel consumption after *SOC* correction. Due to utilizing the newest characteristics of driving schedule, the RL-based real-time control strategy can reduce about 8% fuel consumption than stationary one. The reason why the fuel improvement is bigger in the later simulation

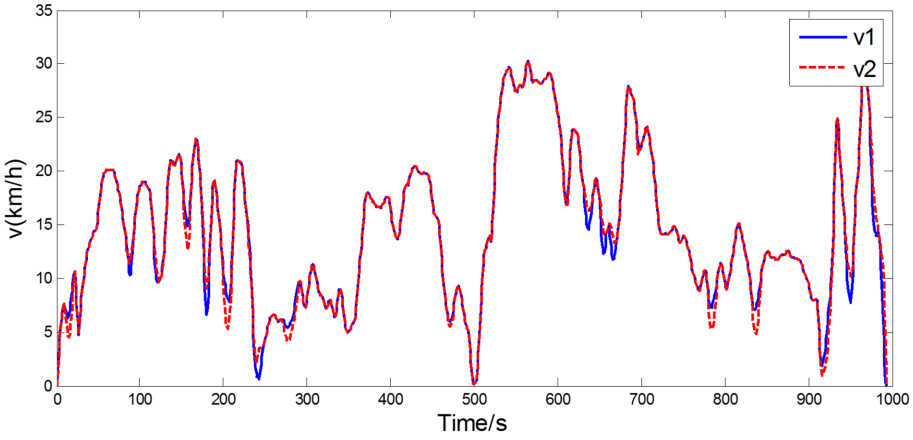

Fig 8. The experimental driving schedule.

https://doi.org/10.1371/journal.pone.0180491.g008

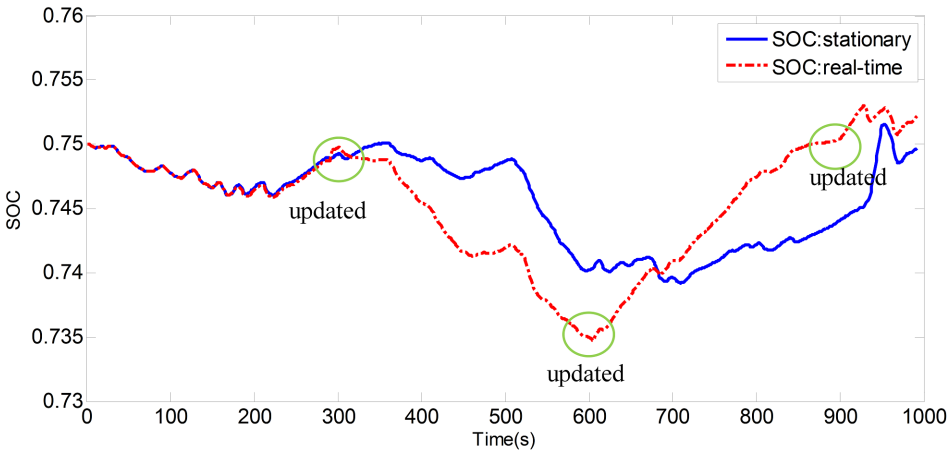

Fig 9. *SOC* trajectory in different strategies.

https://doi.org/10.1371/journal.pone.0180491.g009

than the former is that the driving condition during the whole validation driving schedule is similar, as presented in Fig 12, and the control policy based on updating TPM at 600 s has already included the statistical information of whole driving schedule. The results suggest that the real-time control strategy has a good performance of robustness and enable to adapt to different driving cycles better.

The three different driving schedules were adopted to validate the robustness of the RL-based real-time energy management strategy again, as shown in Fig 15. Table 5 lists the fuel consumption after *SOC*-correction. The real-time method, which utilizes the newest driving characteristics, gives superior fuel economy to the stationary one.

5. Conclusion

This paper proposes a real-time energy management strategy based on reinforcement learning for a hybrid electric tracked vehicle. A recursive algorithm for the transition probability matrix

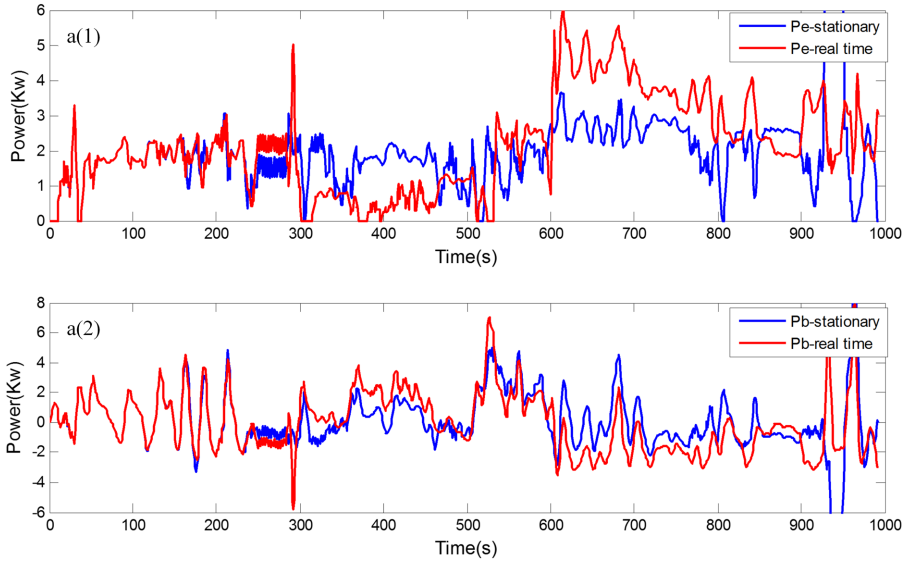

Fig 10. Power split. a1 is power of the engine, and a2 is the power of the battery.

https://doi.org/10.1371/journal.pone.0180491.g010

Table 3. Fuel consumption for the experimental driving schedule.

Algorithms	Fuel consumption(g)	Relative increase	Final SOC
RL online	430.856	—	0.7497
Stationary	456.707	6%	0.7522

https://doi.org/10.1371/journal.pone.0180491.t003

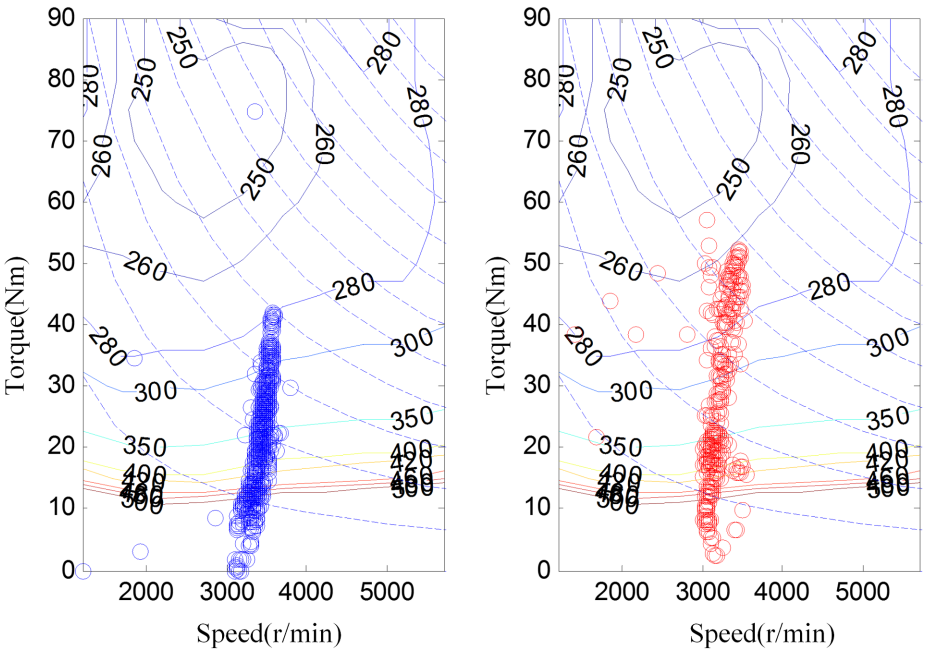

Fig 11. Working points of the engine.

https://doi.org/10.1371/journal.pone.0180491.g011

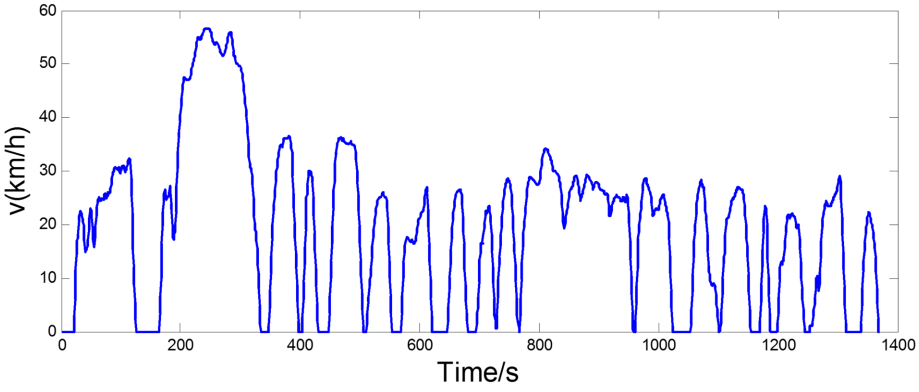

Fig 12. The validation driving schedule.

https://doi.org/10.1371/journal.pone.0180491.g012

(TPM) is developed to make use of new statistical characteristics of online driving schedule. Based on the updating transition probability matrix (TPM), the control policy is calculated and updated at regular intervals to adapt to different driving conditions. A detailed facing-forward simulation model including the engine-generator model, battery model and vehicle

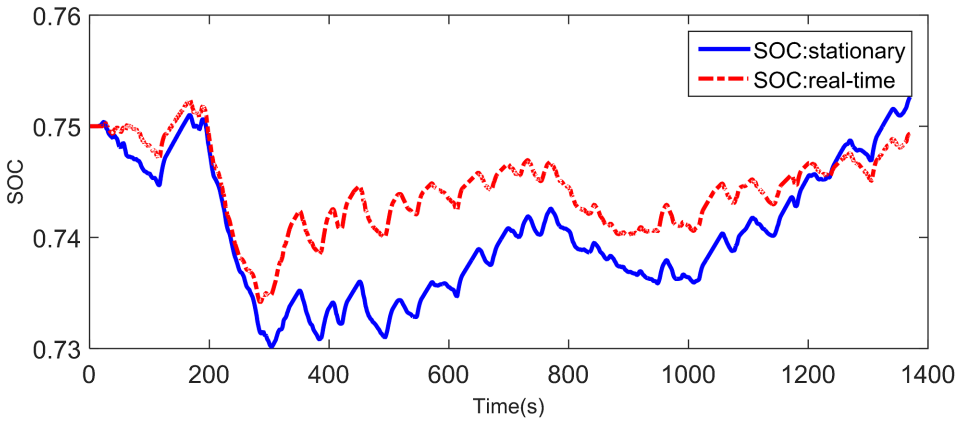

Fig 13. SOC trajectories for the validation driving schedule.

https://doi.org/10.1371/journal.pone.0180491.g013

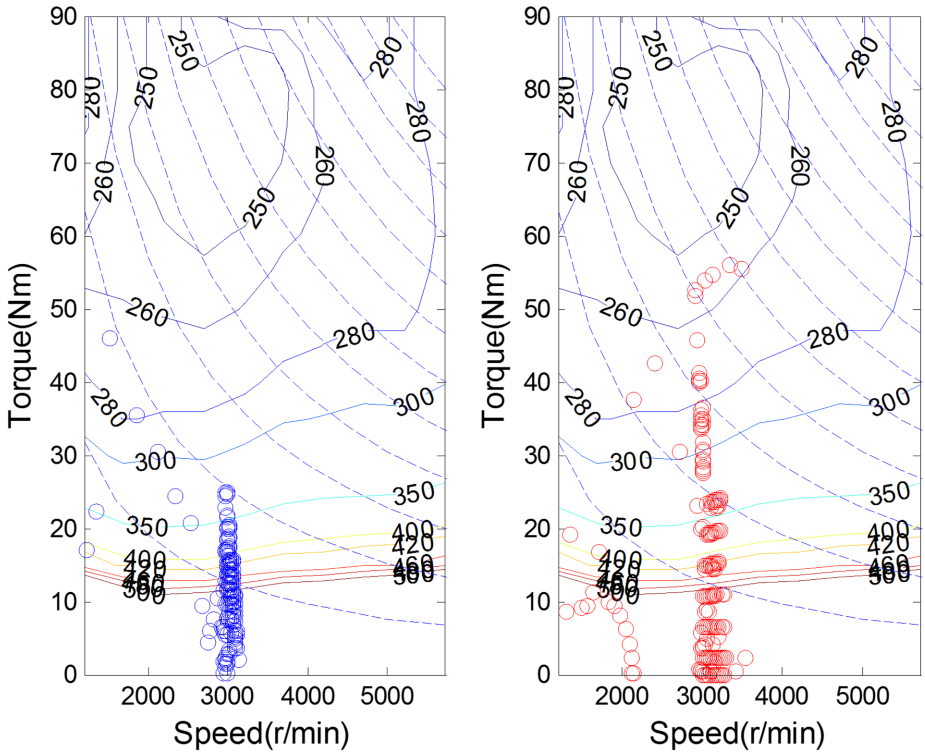

Fig 14. Working points of the engine.

https://doi.org/10.1371/journal.pone.0180491.g014

Table 4. Fuel consumption for the validation driving schedule.

Algorithms	Fuel consumption(g)	Relative increase	Final SOC
RL online	584.275	—	0.7526
Stationary	635.865	8%	0.7493

https://doi.org/10.1371/journal.pone.0180491.t004

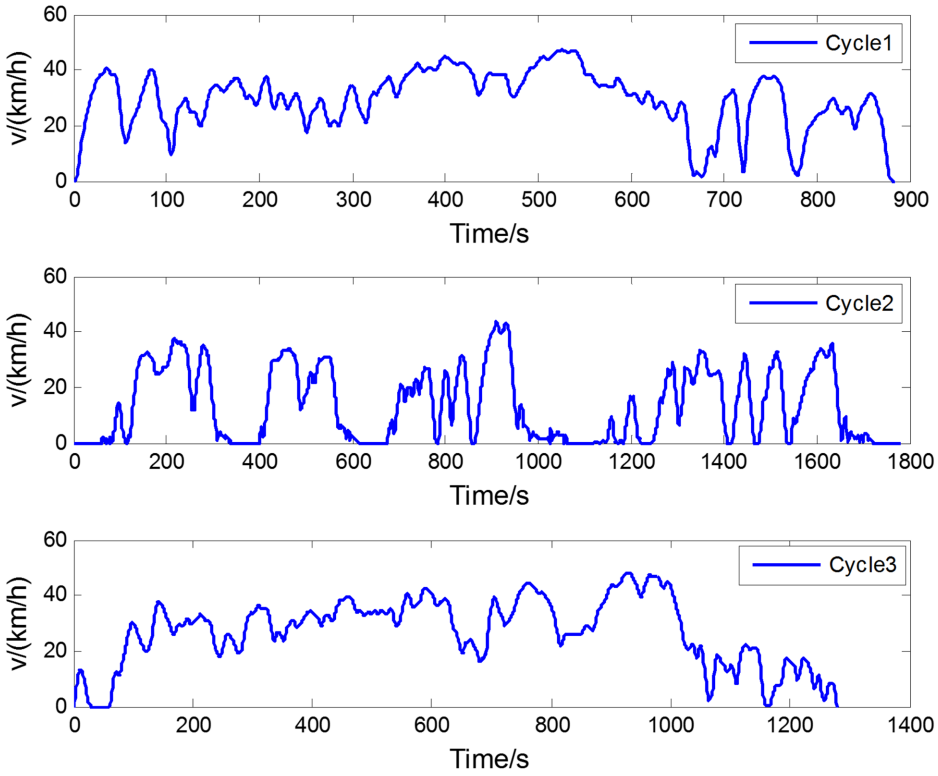

Fig 15. The three different driving schedules.

https://doi.org/10.1371/journal.pone.0180491.g015

Table 5. Fuel consumption for three different driving schedules.

Cycle	Algorithms	Fuel consumption(g)	Relative increase	Final SOC
Cycle1	RL online	903.6	—	0.6958
	Stationary	960.4	6.3%	0.7012
Cycle2	RL online	1097.3	—	0.6923
	Stationary	1176.1	7.2%	0.6915
Cycle	RL online	1245.5	—	0.7021
	Stationary	1320.8	6.0%	0.6902

https://doi.org/10.1371/journal.pone.0180491.t005

dynamical model is built. In order to validate the effectiveness and adaptability of real-time control policy, the simulation for two driving schedules are operated. The results indicate that real-time energy management strategy has a superior fuel economy than stationary one, and is more effective in real-time control requirement.

Supporting information

S1 Dataset. The data for long naturalistic driving schedule.
(MAT)

S2 Dataset. The data for experimental driving schedule.
(MAT)

S3 Dataset. The data for validation driving schedule.
(MAT)

S4 Dataset. The data for another three driving schedules.
(RAR)

Acknowledgments

This work was supported by the National Nature Science Foundation of China (Grant 51375044), University Talent Introduction 111 Project (B12022), and Defense Basic Research Project (B20132010).

Author Contributions

Conceptualization: ZHK YZ TL.

Funding acquisition: YZ.

Methodology: ZHK YZ TL.

Project administration: ZHK YZ.

Software: ZHK YZ TL.

Supervision: ZHK YZ.

Validation: ZHK YZ TL.

Visualization: ZHK YZ.

Writing – original draft: ZHK YZ.

Writing – review & editing: ZHK YZ.

References

1. Castaings A, Lhomme W, Trigui R, Bouscayrol A. Comparison of energy management strategies of a battery/supercapacitors system for electric vehicle under real-time constraints. Appl Energy. 2016; 163:190–200.
2. Panday A, Bansal HO. A Review of Optimal Energy Management Strategies for Hybrid Electric Vehicle. Int J Vehicular Technol. 2014; 2014:1–19.
3. Hou C, Ouyang M, Xu L, Wang H. Approximate Pontryagin's minimum principle applied to the energy management of plug-in hybrid electric vehicles. Appl Energy. 2014; 115:174–189
4. Jalil N, Kheir NA, Salman M. A rule-based energy management strategy for a series hybrid vehicle. In: Proceedings of the 1997 American Control Conference (Cat. No.97CH36041); 1997; 1:689–693.
5. Trovão JP, Pereirinha PG, Jorge HM, Antunes CH. A multi-level energy management system for multi-source electric vehicles–An integrated rule-based meta-heuristic approach. Appl Energy. 2013; 105:304–318.
6. John R. I., Hissel D. A survey-based type-2 fuzzy logic system for energy management in hybrid electrical vehicles. Inf Sci. 2012; 190(3):192–207.
7. Lin CC, Peng H, Grizzle JW, Kang JM. Power management strategy for a parallel hybrid electric truck. IEEE Trans Control Syst Technol. 2003; 11(6):839–849.
8. Delprat S, Lauber J, Guerra TM, Rimaux J. Control of a parallel hybrid powertrain: Optimal control. IEEE Trans Vehicular Technol. 2004; 53(3):872–881.
9. Bassam A., Phillips A., Turnock S., Wilson P. An improved energy management strategy for a hybrid fuel cell/battery passenger vessel. Int J Hydrogen Energy, 2016 Aug. https://doi.org/10.1016/j.ijhydene.2016.08.049
10. Xiang C. L., Feng D., Wang W. D., He W., Qi Y. L. MPC-based energy management with adaptive Markov-chain prediction for a dual-mode hybrid electric vehicle. Sci China Technol Sci, 2017:1–12.

11. Hemi H, Ghouili J, Cheriti A. A real time energy management for electrical vehicle using combination of rule-based and ECMS. IEEE Electrical Power Energy Conf (EPEC) 2014; p. 1–6;
12. Zhang P, Yan F, Du C. A comprehensive analysis of energy management strategies for hybrid electric vehicles based on bibliometrics. Renewable Sustainable Energy Rev. 2015; 48:88–104.
13. Zeng X, Wang J. A Parallel Hybrid Electric Vehicle Energy Management Strategy Using Stochastic Model Predictive Control With Road Grade Preview. IEEE Trans Control Syst Technol. 2015; 23 (6):2416–2423.
14. Feng T, Yang L, Gu Q, Hu Y, Yan T, Yan B. A Supervisory Control Strategy for Plug-In Hybrid Electric Vehicles Based on Energy Demand Prediction and Route Preview. IEEE Trans Vehicular Technol.2015; 64(5):1691–1700.
15. Hu X., Moura S. J., Murgovski N., Bo E., Cao D. Integrated Optimization of Battery Sizing, Charging, and Power Management in Plug-In Hybrid Electric Vehicles. IEEE Trans Control Syst Technol. 2016; 24(3):1036–1043.
16. Hu X., Jiang J., Bo E., Cao D. Advanced Power-Source Integration in Hybrid Electric Vehicles: Multicriteria Optimization Approach. IEEE Trans Ind Electron. 2015; 62(12):7847–7858.
17. Hu X, Martinez C M, Yang Y. Charging, power management, and battery degradation mitigation in plug-in hybrid electric vehicles: A unified cost-optimal approach. Mech Syst Signal Process. 2017; 87:4–16.
18. Hu X, Zou Y, Yang Y. Greener plug-in hybrid electric vehicles incorporating renewable energy and rapid system optimization. Energy. 2016; 111:971–980.
19. Yin H, Zhao C, Li M, Ma C, Chow M. A Game Theory Approach to Energy Management of An Engine-Generator/Battery/Ultracapacitor Hybrid Energy System. IEEE Trans Industr Inform. 2016; 63(7):4266–4277.
20. Li L, Yan B, Yang C, Zhang Y, Chen Z, Jiang G. Application-Oriented Stochastic Energy Management for Plug-in Hybrid Electric Bus With AMT. IEEE Trans Vehicular Technol.2016; 65(6):4459–4470.
21. Lin X, Wang Y, Bogdan P, Chang N, Pedram M. Reinforcement Learning Based Power Management for Hybrid Electric Vehicles. I CCAD IEEE ACM Int Conf Comput Aided Des. 2014; p. 32–38.
22. Wang Y, Sun Z. SDP-based Extremum Seeking Energy Management Strategy for a Power-Split Hybrid Electric Vehicle. In: Proc Am Control Conf. 2012; p. 553–558.
23. Liu T, Zou Y, Liu D, Sun F. Reinforcement Learning of Adaptive Energy Management With Transition Probability for a Hybrid Electric Tracked Vehicle. IEEE Trans Ind Electron. 2015; 62(12):7837–7846.
24. Ryan MJ, Lorenz RD. A "power-mapping" variable-speed control technique for a constant-frequency conversion system powered by a IC engine and PM generator. IEEE Ind Appl Soc. 2002; 4: 2376–2382.
25. Plett GL. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs —Part 2. Modeling and identification. J Power Sources. 2004; 134(2):262–276.
26. Filev DP, Kolmanovsky I. Generalized Markov Models for Real-Time Modeling of Continuous Systems. IEEE Trans Fuzzy Syst. 2014; 22(4):983–998
27. Liu C, Murphey Y L. Power management for Plug-in Hybrid Electric Vehicles using Reinforcement Learning with trip information. IEEE Trans Electrification Conf Expo. 2014:1–6.
28. Fang, Y., Song, C., Xia, B., Song, Q. An energy management strategy for hybrid electric bus based on reinforcement learning. IEEE Conf Decis Contro 2015:4973–4977.
29. Qi X, Wu G, Boriboonsomsin K, Barth MJ. A Novel Blended Real-time Energy Management Strategy for Plugin Hybrid Electric Vehicle Commute Trips. IEEE Int Conf Intell Trans Syst; 2015. p. 1002–1007.
30. Barillas J. K., Li J., Günther C., Danzer M. A. A comparative study and validation of state estimation algorithms for Li-ion batteries in battery management systems. Appl Energy 2015; 155:455–462.