

REVIEW

Open Access



Reinforcement Learning-Based Energy Management for Hybrid Power Systems: State-of-the-Art Survey, Review, and Perspectives

Xiaolin Tang¹ , Jiaxin Chen¹, Yechen Qin², Teng Liu¹, Kai Yang¹, Amir Khajepour³ and Shen Li^{4*}

Abstract

The new energy vehicle plays a crucial role in green transportation, and the energy management strategy of hybrid power systems is essential for ensuring energy-efficient driving. This paper presents a state-of-the-art survey and review of reinforcement learning-based energy management strategies for hybrid power systems. Additionally, it envisions the outlook for autonomous intelligent hybrid electric vehicles, with reinforcement learning as the foundational technology. First of all, to provide a macro view of historical development, the brief history of deep learning, reinforcement learning, and deep reinforcement learning is presented in the form of a timeline. Then, the comprehensive survey and review are conducted by collecting papers from mainstream academic databases. Enumerating most of the contributions based on three main directions—algorithm innovation, powertrain innovation, and environment innovation—provides an objective review of the research status. Finally, to advance the application of reinforcement learning in autonomous intelligent hybrid electric vehicles, future research plans positioned as “Alpha HEV” are envisioned, integrating Autopilot and energy-saving control.

Keywords New energy vehicle, Hybrid power system, Reinforcement learning, Energy management strategy

1 Introduction

The future transportation system revolves around two major themes: Autopilot and energy-saving driving. New energy vehicles in China, including electric vehicles (EVs), plug-in hybrid electric vehicles (PHEVs), and fuel cell vehicles (FCVs), stand as the core carrier. Positioned at the forefront of automotive advancements, new energy

vehicles pave the way toward clean, green, and sustainable transportation [1].

EVs are powered by batteries as primary energy sources, with motors converting electrical energy into kinetic energy to propel the vehicle. As a result, the research focus lies on the advancement of motors, batteries, and electronic control systems [2]. While EVs hold immense potential, there are ongoing endeavors to address challenges such as enhancing driving range, developing fast charging solutions, ensuring safety measures, and establishing recycling and cascade utilization methods. Furthermore, the widespread promotion necessitates the development of infrastructure to support integration into daily life. FCVs utilize propulsion systems equipped with fuel cells and power batteries. Operating on hydrogen, fuel cell systems generate electricity through the electrochemical reaction between hydrogen

*Correspondence:

Shen Li

sli299@tsinghua.edu.cn

¹ College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400044, China

² School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China

³ Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

⁴ School of Civil Engineering, Tsinghua University, Beijing 100084, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

and oxygen, emitting water as a byproduct. This electro-chemical coupling positions FCVs as an environmentally friendly solution for the future [3]. Notably, FCVs present advantages for long-distance passenger or freight transportation, addressing the limitations of driving range in PEVs. Their quick refueling time, coupled with their eco-friendly nature and cleanliness, makes them an ideal choice for various applications, holding the promise of transforming transportation. However, FCVs are still facing challenges, including infrastructure, cost, and durability of fuel cells, hydrogen storage and distribution, hydrogen production, and recycling issues, resulting in the status not being ideal [4]. Although the above two types of cars have multiple sources of energy, only the electric motor serves as the power source, converting electrical energy into mechanical energy to propel the vehicle forward.

Hybrid electric vehicles (HEVs) represent an innovation integrating both gasoline and batteries as energy sources. By harnessing the power of internal combustion engines (ICEs) and motors, HEVs offer an efficient approach to propulsion. Moreover, the advent of rechargeable batteries has led to the development of PHEVs allowing for longer electric driving capabilities [5]. One of the mechanisms employed in HEVs to achieve energy savings lies in optimizing the operation of the engine within the high-efficiency range. Simultaneously, the motor mainly serves a key role in regenerative braking, converting braking energy into usable electricity. Nowadays, Hybrid powertrains can be mainly classified into three types: series, parallel, and hybrid, each offering unique advantages to suit diverse conditions [6]. Series HEVs can be likened to that of PEVs with a range extender. The ICE connects to the generator, converting mechanical energy into electric energy, and it allows the ICE to work within a high-efficiency range, and the motor serves as the sole power source for propulsion and regenerative braking. Parallel HEVs offer more complex and adaptable driving modes and can be subdivided into P0-P4 configurations based on the different positions of motors. Treating the P2 as an example, both the ICE and the motor can function independently to propel the HEV. When the demand is big, these two power sources can simultaneously deliver power through a mechanical structure. Hybrid HEVs stand as an excellent achievement of engineering with their sophisticated structure. Its primary essence lies in the power-split mechanism, engineered with planetary gears, with the Prius standing as a quintessential Hybrid HEV. One notable feature involves the integration of motors and generators, enabling simultaneous operations in driving and charging. Due to the intricate structure and technological challenges, only a handful of

manufacturers have successfully achieved the proficiency required to develop Hybrid HEVs [7].

The design of the HEV needs a multi-faceted endeavor, encompassing configuration screening, parameter matching, and energy management [8]. The configuration design shapes the dynamic interplay among each power and transmission component, considering factors like technical foundations, potential challenges, and user requirements. The parameter matching needs an extreme balance, as it not only influences the performance of the vehicle but also has implications for manufacturing costs [9]. Moreover, it is essential to ensure that the dynamic can satisfy minimum requirements while considering various scenes, including extreme environments. This attention to detail guarantees the capability of HEVs to perform exceptionally across various conditions. The energy management strategy (EMS) plays a core role in enhancing energy-saving performance. It efficiently distributes power flow while adhering to constraints, leading to optimizations in fuel economy, exhaust emissions, battery characteristics, and other objectives [10, 11]. Now, three types of EMS have been summarized and proposed: rule-based, optimization-based, and learning-based EMS [12]. Rule-based EMS relies on a series of experiences to determine power distribution among various power sources. It is computationally efficient and often implemented in real controllers. Rule-based EMS can be further categorized into deterministic rules and fuzzy rules. However, one main limitation is the requirement of extensive experimental data, as well as limited adaptability to random scenes. Optimization-based EMS transforms the EMS into an optimization problem. By defining an objective function and considering system constraints, these strategies determine the control sequence that corresponds to the target within the given environment, such as fuel consumption and lifespan. Optimization-based EMSs are divided into global and instantaneous optimization. Global optimization-based EMSs adopt solvers such as dynamic programming (DP) [13] and Pontryagin's minimum principle (PMP) [14], and the instantaneous optimization-based EMS use algorithms like equivalent consumption minimum strategy (ECMS) [15] and model predictive control (MPC) [16].

The birth of the learning-based EMS benefits from the development of artificial intelligence (AI), especially deep learning (DL) and reinforcement learning (RL). While some reviews have offered insights by categorizing algorithms and contributions belonging to RL-based EMSs [12, 17–27]. By contrast, especially for rule-based and optimization-based EMSs, many reviews have comprehensively revealed the novel research status. Because of the abundance of existing literature, this paper actively avoids the repetitive content of traditional EMSs, and the

latest survey and review focuses on the reinforcement learning (RL)-based EMS and aims to present a thorough and up-to-date review by enumerating all contributions and drawing from research experiences. Given the relatively short development time, the total number of literature within an acceptable range makes it feasible to list and summarize the achievements of all RL-based EMSs.

The main contributions and the remainder of the paper are organized as follows. For the macroscopic grasp of historical development, Section 2 summarizes a brief history, famous scholars, and important achievements of DL, RL, and deep reinforcement learning (DRL) in the form of a timeline, and this is also the first time that the development process is fully displayed in the form of figures. Section 3 summarizes all contributions of RL-based EMSs for hybrid power systems and provides a comprehensive review. It collects 266 papers from databases such as Web of Science, IEEE Xplore, and ScienceDirect, focusing on EV, energy management, and RL as keywords. The state-of-the-art status is analyzed based on innovations in algorithms, powertrains, and environments for further discussion. Section 4 envisions future

research aimed at developing an autonomous intelligent HEV, with "Alpha HEV" as the ultimate goal. Section 5 concludes with key opinions and insights.

2 Brief Development History of DL/RL/DRL

In this section, the timeline in Figure 1 presents a brief history of DL, RL, and DRL, including the significant achievements of notable scholars and forming a historical perspective that enhances comprehension of the evolution.

2.1 The Development History of DL

DL takes a leading position in the realm of machine learning (ML), representing a groundbreaking methodology aimed at uncovering intricate patterns and representations concealed within extensive datasets. Its objective is to replicate human-like analytical and learning capabilities, enabling machines to learn and grasp diverse forms of data, such as text, images, and sounds [29]. As shown in Figure 1, the roots of DL can be traced back to the 1940s when W.S. McCulloch and W. Pitts sought to simulate a neural reaction within the human brain when

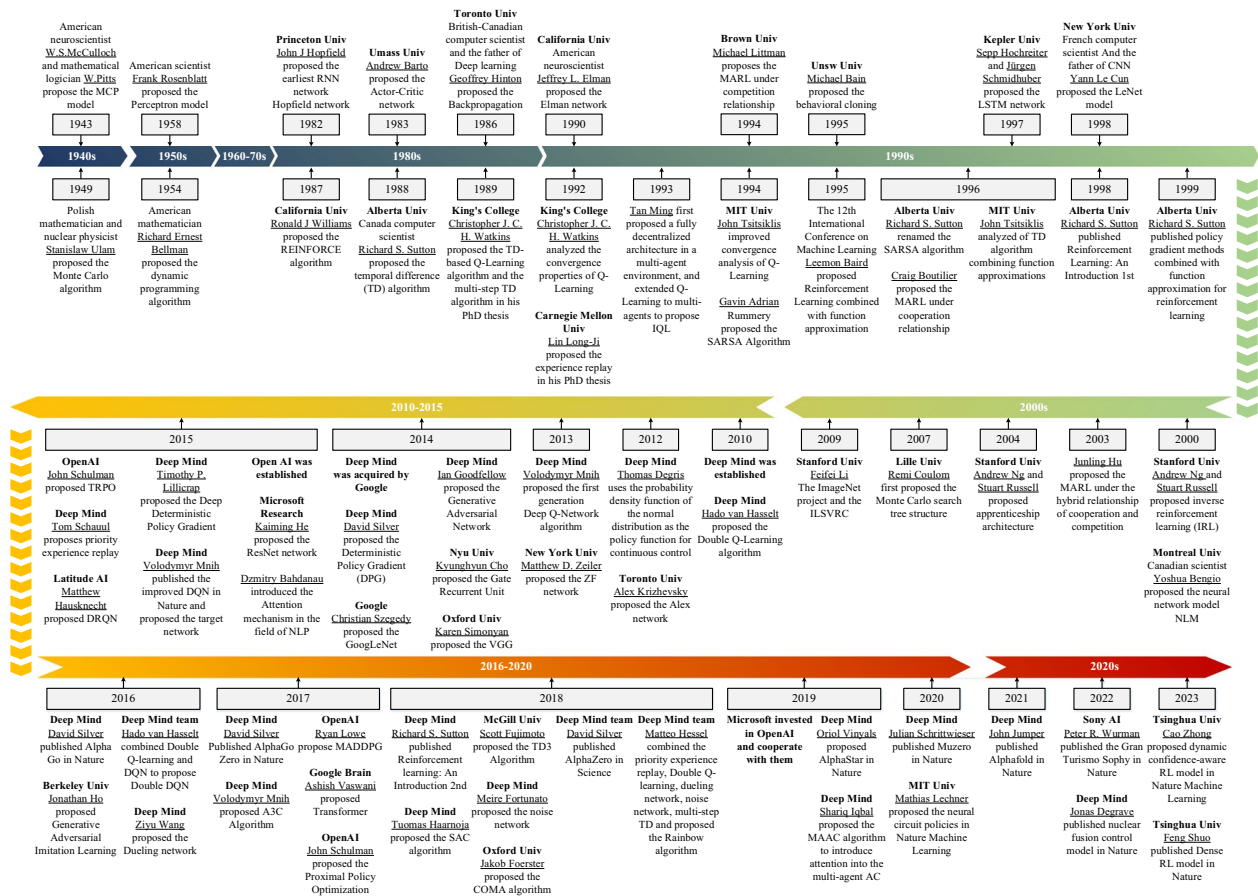


Figure 1 The timeline of the brief development history of DL/RL/DRL

processing information. They developed a simplified artificial neuron model known as MCP [30]. It encompassed the fundamental functions of basic neurons: linear weighting, summation, and non-linear activation. Expanding upon the foundational model, in 1958, Frank Rosenblatt proposed the Perceptron [31], a two-layer feedforward network based on the MCP. The Perceptron can be employed to classify binary linear problems by mapping input matrices to output values and making decisions based on thresholds and weights. By adopting the loss minimization and gradient descent, training could yield a linear plane for classification. However, it was proven that the Perceptron was limited to linear problems. In 1982, John J. Hopfield designed the Hopfield Network [32], considered the earliest recurrent neural network (RNN). It links the output of each neuron to the input of other neurons and forms an innovation that has been critical for the future. Another breakthrough came from backpropagation, which led to the development of a multilayer feedforward network known as Back Propagation (BP). Geoffrey Everest Hinton proposed the BP in 1986 [33]. It involved signal propagation, error backpropagation, and weight updates. The BP network addressed the limitations of Perceptron, enabling nonlinear classification and becoming the milestone in DL. Two other key contributions were attributed to the Elman network [34] and the LeNet network [35]. The Elman network proposed by Jeffrey Elman in 1990, functioned as a feedforward network with local memory units, local feedback connections, and a multilayer structure, and it aims for speech recognition. The LeNet network, proposed by Yann LeCun in 1998, was the first convolutional neural network (CNN). Although effects were limited by data and computing power, the LeNet can successfully recognize handwritten fonts. Long short-term memory (LSTM) [36], proposed by Sepp Hochreiter in 1997, solved the vanishing gradient and long-term dependencies and became a basic model for processing and forecasting events in time series data. The core elements of an LSTM cell consist of three gates: the input gate, the forget gate, and the output gate. These gates regulate the flow of information, allowing it to retain key information over long sequences. In 2000, a feedforward neural language model (NLM) [37] was proposed by Yoshua Bengio which employs neural networks to model the probability distribution of the natural text. It plays a key role in the field of natural language processing (NLP), particularly in tasks involving language generation and language understanding. Hence, Geoffrey Hinton, Yann Le Cun, and Yoshua Bengio are recognized as the "Big Three of Deep Learning."

In recent years, many amazing achievements have been witnessed, and the generative model ChatGPT,

developed by OpenAI, is regarded as the most famous product. Before that, the generative adversarial network (GAN) [38], proposed by Ian Goodfellow in 2014, utilizes the two-module framework consisting of one generator and one discriminator to achieve impressive outputs through mutual learning. The training of GAN becomes a confrontation process: the generator and the discriminator will engage in competition with each other to improve their capabilities. Simply speaking, the generator tries to generate more realistic data, while discriminators try to distinguish real data from generated data. Eventually, the performance of the generator steadily enhances, resulting in generated samples that resemble the distribution of real data. Therefore, the GAN performs well in many tasks, including image generation, image super-resolution, style transfer, etc. In the same year, Kyunghyun Cho proposed a gated recurrent unit (GRU) [39], a simplified LSTM with fewer parameters. It solves the challenges of long-term memory and gradients by employing gating units. The model incorporates reset and update gates, determining how data combines with previous memory and how much memory will be retained. Compared to the LSTM, the GRU is demonstrated that it is beneficial in improving training efficiency and is favored. Additionally, in 2017, Ashish Vaswani from Google Brain published and proposed the Transformer [40], which is a network structure with significant influence. The Transformer has profoundly influenced NLP by introducing the self-attention mechanism, enabling it to capture long-range dependencies effectively. Its performance across diverse NLP tasks, propagation of pre-training and fine-tuning methods, and expansion into domains beyond NLP highlight its wide-reaching influence on DL and its practical applications. Compared with RNNs and CNNs, the self-attention mechanism enables the model to process sequence data by simultaneously considering all positions in the input, while the ability of parallel computing makes it feasible to handle lengthy information.

DL comprises three elements in Figure 2: algorithms, data, and computing power. As the global academic community continues to propose advanced algorithms and networks, a notable contribution has been made by ImageNet, introduced by Prof. Feifei Li. The ImageNet serves as a visualization database for object recognition, encompassing over 20000 categories and more than 14 million annotated images [41]. ImageNet large-scale visual recognition challenge (ILSVRC) becomes one of the most esteemed competitions in computer vision (CV). Many outstanding networks, including AlexNet, ZFNet, VGG, GoogLeNet, and ResNet, have emerged, with 2010 acknowledged as the dawn of DL. The champion of ILSVRC in 2012 was AlexNet, a collaborative effort by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton [42]. The

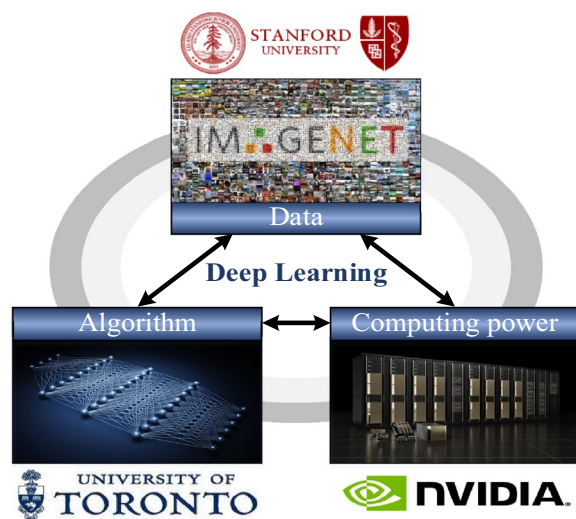


Figure 2 Three elements of deep learning [28]

core contributions lie in introducing the CNN composed of convolutional layers and fully connected layers, pioneering the adoption of ReLU as the activation function, devising the Dropout method to mitigate overfitting, adopting mini-batch gradient descent with momentum for convergence, utilizing data augmentation to combat overfitting, and employing the parallel computing of the NVIDIA graphics processing unit (GPU) to accelerate the training. In 2013, the champion ZFNet [43] mainly made modifications to the size, number, and convolutional strides of the kernel, and the next year witnessed an influence on the champion GoogLeNet [44] and the runner-up VGG [45]. GoogLeNet proposed the Inception structure, retaining more features within the input data. By eliminating the first two fully connected layers of AlexNet and employing average pooling, GoogLeNet reduced its parameter count to 5 million, a 12-fold reduction compared to AlexNet. Then, GoogLeNet designed auxiliary classifiers in intermediate layers to solve the vanishing gradient. VGG, proposed by a Visual Geometry Group at Oxford University, abandoned the large-scale kernels such as 11×11 and 5×5 , instead using some smaller 3×3 kernels to achieve a larger receptive field. Moreover, VGG eliminated a local response normalization (LRN) used by AlexNet. Generally, a deeper neural network enables the extraction of more sophisticated features. However, the increase in the number of layers leads to flaws such as a vast number of parameters and the risk of overfitting. In 2015, the ILSVRC championship ResNet [46], proposed by Kaiming He, addressed these challenges with a key contribution: the residual module. The core idea lies in introducing "skip connections," where the input is directly added to the output port, preserving

the original information and facilitating the gradient flow during backpropagation. Moreover, it introduced batch normalization to combat the vanishing gradient, mitigating reliance on initialization. At the same time, a unique initialization method was proposed specifically for the activation function ReLU.

For computing power, AI computing has followed several trends. Firstly, the widespread adoption of specialized hardware accelerators such as GPUs and Tensor Processing Units (TPUs) has significantly improved the computational efficiency of tasks. Secondly, the heterogeneous computing platforms integrating different processors have enabled more efficient computation. Additionally, major cloud computing providers offer specialized cloud services, such as Amazon and Google, which provide flexible computing resources and high-performance hardware infrastructure. Of course, GPU is currently the most commonly used and popular. The GPU has contributed to the rise in computing power. Under the background of developing DL, GPUs serve as indispensable tools akin to shovels in a gold rush. NVIDIA defined GPUs in 1999, and under the leader Jensen Huang, the specialized processors were defined for computationally intensive tasks. Compared to central processing units (CPUs), GPUs offer advantages in parallel computing and performance, and they have revolutionized the gaming market, redefined computer graphics, and transformed parallel computing. Consequently, GPUs are widely adopted, specifically in game engines and rendering, allowing for rapid calculations of elements such as geometry, light, and shadows, facilitating the creation of more realistic visual effects. In the realm of DL, there are a large number of matrix computations and tensor operations involved. The parallel computing of GPUs can significantly accelerate the training process and make it possible to deal with large-scale datasets and models. Nowadays, the latest NVIDIA DGX equipped GH200, A100 or H100 provides solutions for large-scale AI infrastructure, and the first DGX was donated to Open AI, the artificial intelligence team that developed ChatGPT. NVIDIA DGX SuperPOD has become a one-stop AI platform that can cope with challenging AI and high-performance workloads. In this era, computing power has become the engine to promote the development of AI.

2.2 The Development History of RL

The basic process of RL in Figure 3 contains two basic modules: the Agent and the Environment, along with three variables: state, action, and reward. A basic process of learning can be described as when the Agent, guided by the current strategy, outputs an action based on the state of the Environment, the Environment executes

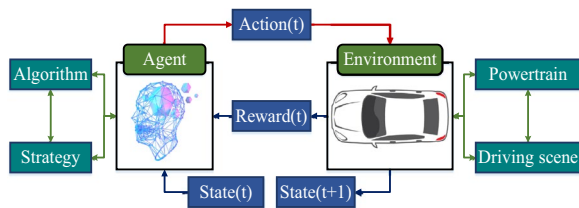


Figure 3 The basic process of reinforcement learning

the action, transitions to the next state, and generates a related reward. Relying on an instant reward, the Agent determines the loss and the gradient to update the current strategy. Through the iterative trial and error to carry out the above process, the Agent struggles for the optimal action corresponding to each state. When the convergence of the mean reward to its maximum, it means the acquisition of the optimal policy within the current environment [47]. After decades of development, the system of RL can be classified into three types based on the scope of application: DP, Monte Carlo (MC), and Temporal Difference (TD).

The DP-based RL belongs to the model-based and offline learning category, and it can in some cases be used to solve problems in discrete state and action spaces, where the agent tries to learn the policy to make optimal decisions in a given environment. The DP proposed by Richard Bellman [48] in 1954 aims to decompose complex large-scale problems into subproblems and combine subsolutions to construct the final optimal solution for the original problem. MC-based RL [49], falling under the model-free and offline learning category, was proposed by Stanislaw Ulam in 1949. The MC relies on data description, with a large of samples forming an accurate reflection. MC-based Reinforce [50], proposed by Ronald J. Williams in 1987, introduced the gradient descent to update policies. Faced with a large number of model-free tasks from the real world, it is difficult to solve them using DP, and the MC-based method heavily relying on sampling necessitates the completion of each episode before learning, making it challenging to satisfy the efficiency in applications. Richard Sutton proposed the TD algorithm in 1988 [51], a model-free and online learning-based category. Based on the indicator of on-policy and off-policy, TD-based RL contains SARSA (on-policy) [52] and Q-Learning (off-policy) [53], and a core difference lies in how to calculate the target prediction when updating the value function. SARSA was proposed by Gavin Adrian Rummery in 1994, and officially renamed by Sutton in 1996. Q-Learning originated from the work of Watkins, who proposed the TD-based Q-Learning algorithm and the multi-step TD in 1989 and analyzed the convergence in 1992 [54]. Q-learning has become

the core algorithm in RL and serves as a foundational achievement for the development of RL. Furthermore, the experience replay, employed in various algorithms, was proposed by Lin in 1992 [55]. Subsequently, research shifted towards function approximation. Leemon Baird [56] and John Tsitsiklis [57] delved into related studies in 1995 and 1996, respectively. Richard Sutton published the book "Reinforcement Learning: An Introduction" in 1998, which was regarded as the bible of RL, and then he analyzed the policy gradient integrated with function approximation in 1999 [58]. Additionally, Inverse RL (IRL) was introduced by Andrew Ng and Stuart Russell in 2000 [59] and is usually utilized to define the reward function, with an apprenticeship architecture published in 2004 [60]. In 2006, the Monte Carlo Search Tree (MCST) was proposed by Rémi Coulom [61], influencing the creation of AlphaGo.

The above are classic examples of traditional RL, laying a theoretical and algorithmic foundation for the development and application of DRL algorithms.

2.3 The Development History of DRL

For RL, early-stage limitations in data and computing power hindered the progress, and RL has to deal with stability and reliability problems. In the meantime, the trial-and-error of RL agents led to models getting stuck or failing to converge to optimal solutions, and the uncertainty and unreliability made RL challenging for applications. On the other hand, the table-based RL had severe limitations, such as the "Curse of Dimensionality" and "Discretization Error".

In recent years, there have been significant advancements, driven by the efforts of teams like DeepMind and OpenAI. DeepMind, founded by Demis Hassabis in 2010 and later acquired by Google in 2014, has played a crucial role in RL. It proposed various DRL algorithms suitable for different tasks. In 2013, Volodymyr Mnih from DeepMind proposed the first DRL algorithm called Deep Q-Network (DQN) [62]. The improved version with target networks was officially in 2015 [63], demonstrating superior control in Atari 2600. In the improved DQN, the main improvement is to first use the neural network to parametrically fit the original value table, and then suppress the instability in the training through the target network, and also use experience replay to effectively break the correlation problem between each training sample. Next, more algorithms were successively proposed, such as deep deterministic policy gradient (DDPG) by Timothy P. Lillicrap [64], prioritized experience replay (PRE) by Tom Schaul [65], trust region policy optimization (TRPO) by John Schulman [66], and deep recurrent q-network (DRQN) by Matthew Hausknecht [67]. In 2016, DeepMind made a breakthrough in the game

of Go named Alpha Go [68]. By integrating DL, RL, and MCST, Alpha Go defeated Fan Hui, Lee Sedol, and Ke Jie. More improved versions of AlphaGo Zero [69] and AlphaZero [70] obtained greater achievement in mastering board games. Other classic DRL algorithms or improvement measures have also been proposed, including Double DQN (DDQN) by Hasselt [71], dueling network by Wang [72], asynchronous advantage actor-critic (A3C) by Mnih [73], proximal policy optimization (PPO) by Schulman [74], soft actor-critic (SAC) by Haarnoja [75], twin delayed deep deterministic policy gradient (TD3) by Fujimoto [76], noise network by Fortunato [77], and rainbow by Hessel [78].

DRL has also been applied to other fields. In 2019, Oriol Vinyals proposed the AlphaStar achieved grandmaster-level performance in StarCraft II [79]. AlphaStar is a remarkable achievement that has captivated the world of Esports. This groundbreaking AI has mastered the control strategies of all races in the game. In 2020, MuZero [80] proposed by Julian Schrittwieser grasps considerable effects in Atari and board games such as Go, Chess, and Shogi. Recent advancements include AlphaFold proposed by John Jumper [81] in 2021, for predicting protein structures, as well as the application in the Gran Turismo (GT) Sport defeating top players on the PlayStation by Sony, named GT Sophy [82], and successfully controlling superheated plasma in nuclear fusion reactors under the collaboration of DeepMind and Swiss Federal Institute of Technology in Lausanne [83]. In 2023, scholars from Tsinghua University, Cao Zhong and Feng Shuo, made contributions to autopilot [84] and safety testing [85] relying on DRL algorithms.

3 The Survey and Review of RL-Based EMSs

3.1 The State-of-the-Art Survey of Research Status

As of July 21, 2023, a state-of-the-art survey was completed in major academic databases such as Web of Science, IEEE Xplore, and ScienceDirect, and a total of 266 papers have been searched about the keywords: electric vehicle, energy management, and RL. According to all current literature, the comprehensive survey about RL-based EMSs contains the universities and institutions, and contributions published in conference and journal papers. Through sorting and analysis, all contributions can be classified into algorithm innovation, powertrain innovation, and environmental innovation. Due to the length of the paper and the large amount of data, the most detailed content in the form of tables is uploaded at <https://github.com/KaysenC/Reinforcement-Learning-based-Energy-Management-for-Hybrid-Power-Systems>.

The detailed tables mainly include the following contents:

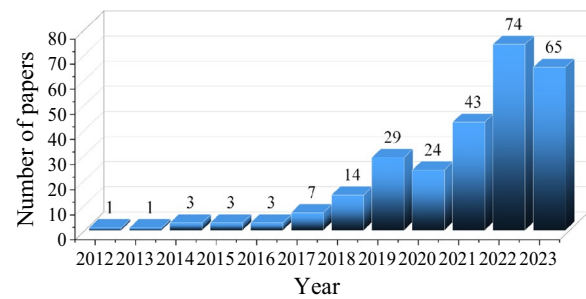


Figure 4 The total publication number over the years

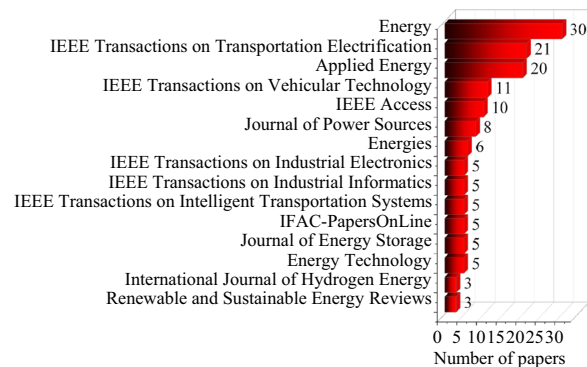


Figure 5 The top 15 journals along with the number of papers

- (1) Statistics on the earliest time and the number of results for RL-based EMSs for universities and institutions.
- (2) Statistics on authors, powertrains, algorithms, and contributions of conference papers.
- (3) Statistics on authors, powertrains, algorithms, and contributions of journal papers (algorithm innovation, powertrain innovation, and environmental innovation).

Moreover, the VISIO file containing the timeline depicted in Figure 1 has been uploaded, and we encourage scholars to contribute enhancements and rectifications.

Within the collection, there are 71 conference papers [19, 86–155] and 195 research papers [12, 17, 18, 20–27, 156–339]. Figure 4 reveals the representation of the publication number over the years, highlighting the development of RL-based EMSs for hybrid power systems. An obvious aspect is that the origin application of RL can be traced back to as early as 2012, and it commenced its significant development in 2018. Figure 5 reveals the top 15 journals with the number of papers, and the pioneering work of introducing RL into EMSs was completed by scholars from National Chiayi University, who focused on the hybrid electric bicycle [158]. Since then,

Table 1 Representative research achievements of algorithm innovation

Reference No.	Powertrain	Algorithm	Contribution
[174]	HETV	Fast Q learning	Hardware-in-loop; KL divergence to trigger the update; cloud computation;
[176]	Power split HE bus	DDPG	Integrating the terrain information; hybrid action space (engine and powertrain mode); dueling network; DP-based pre-training process;
[182]	PHE bus	Q learning	The optimal reference SOC trajectories as the expert experience; multiply driving cycle training method for the generalization performance; the RL-based agent is used to determine the co-state of the PMP-based EMS;
[187]	Power split HEV	DDPG	Transfer learning among four types of powertrains (Prius, power-split bus, series HEV, and series-parallel bus);
[194]	Multi-mode HEV	DQN	The co-state of the PMP-based EMS is determined by the DRL agent;
[198]	Parallel HEV	A3C DPPO	A3C-based and DPPO-based energy and emission management strategy; improving the learning efficiency in the multi-thread training process;
[204]	Parallel HEV	TD3	Heuristic rule-based local controller for eliminating irrational torque allocation; and environmental disturbances; hybrid experience replay consisting of offline computed optimal experience and online learned experience
[207]	Multi-mode HEV	A3C PPO	Online updating framework; generating probable cycles using historical data;
[208]	HETV	Double DQN	Modified prioritized experience replay; adaptive optimization AMSgrad;
[209]	Parallel HEV	DDPG	Deployment inefficiency, safety constraint, and simulation-to-real gap; hardware-in-the-loop; an offline cloud-based DRL framework;
[210]	Parallel HEV	DDPG	The equivalent factor of ECMS-based EMS is determined by the DRL; a safe exploration relying on modifying by including the heuristic domain knowledge within the ECMS-based for the DRL agent; hardware-in-the-loop;
[213]	PHEV	Q learning	KL divergence rate to update the TPM of the demand power;
[216]	Power split HEV	DQN	Generalization ability; coding and decoding multiple states; KL-divergence is used to guide the training; an auxiliary agent and a correlation agent;
[217]	Power split HEV	DQN	Hierarchical RL; transfer the optimal BSFC as the sub-goal, and avoid the 'blind' exploration by guiding to the direction of the sub-goal in the lower level;
[219]	HETV	SAC	The Munchausen SAC-based EMS to bootstrap and improve optimization; prioritized experience replay; DP-based early assisted training sample;
[220]	Parallel HEV	DDPG DQN	The multi-objective control aiming at the engine and gearbox; combined with the learning-based EMS and rule-based engine start-stop strategy; RL is not suitable for learning intermittent strategies;
[221]	Power split HE bus	DDPG	The softmax deep double deterministic policy gradients algorithm; an action masking technique for preventing invalid actions; transfer learning; double Q-learning network, Boltzmann softmax, and dual-actor;
[223]	Power split HEV	Dueling DDPG	Transfer learning; the adaptive parameter space noise to balance exploration and exploitation which is better than the action space noise; a novel real-time four-phase approach: modeling, pre-training, transferring, fine-tuning;
[229]	Serial HEV	SARSA	TPM of the power demand; the forgetting factor, KL divergence rate threshold value, and TPM updating interval to determine the update of the strategy;
[234]	Parallel HEV	TD3	DRL-based supplementary learning controller for a rule-based EMS; ape-X distributed architecture for improving the converging speed; the gap between simulation and real application; The actor in the vehicle and the learning in the cloud;
[246]	Serial HEV	Nash QL	Multi-agent reinforcement learning combines game theory and reinforcement learning; engine-generator set, battery, and supercapacitor; fuel economy, SOC and health of batteries and SOC of supercapacitor; the Nash equilibrium of multiple objectives;
[247]	Serial HEV	Dyna	Online-learning adaptive EMS; the real precious experience is used to train the policy and establish an interactive model; realizing the rapid and low-cost online learning.

this revolutionary field has expanded to other universities such as the University of Southern California, the University of Michigan, the University of California, the Beijing Institute of Technology, and Chongqing University, with their team developing a unique set of technical routes.

It is key to note that the achievements of RL in other fields like Egames and Autopilot, have yielded numerous notable results published in prestigious journals such as

Nature and Science. Therefore, there is potential for RL-based EMSs of hybrid power systems, with contributions extending beyond optimization, adaptability, and generalization. More scholars are investigating the RL-based EMS for FCVs, HEVs, EVs equipped with supercapacitors or hybrid battery systems, (HEBs), hybrid electric tracked vehicles (HETVs), etc., and Q-learning is considered the most popular. Subsequently, Qi et al. [169]

introduced the use of DRL and defined the third category of learning-based EMSs. Liu [164–167, 177] made numerous contributions and proposed RL-based EMSs that minimize fuel consumption across various conditions with the help of mathematical theories like transition probability matrix (TPM) and Kullback-Leibler (KL) divergence. He et al. [165] also proposed a predictive EMS, combining speed prediction and RL, and the proposed strategy was validated by hardware-in-the-loop (HIL). Qi et al. [95] employed the DQN to learn EMS based on historical mileage information, while Li et al. [162] used an actor-critic (AC) architecture for continuous state and action spaces. In the next few years, classic algorithms and improved approaches like Q-Learning and experience replay have usually been adopted, as have popular algorithms like DQN, Double DQN, Dueling DQN, DDPG, and priority experience replay [161]. More recently, some improved algorithms such as SAC, PPO, TD3, A3C, and transfer learning (TL) have been tentatively applied. Additionally, AMSGrad, Fast Q-Learning, NAG-Adam, and Munchausen SAC have been tried and utilized to improve efficiency. Meanwhile, multi-agent reinforcement learning (MARL) gained more attention. For some typical scenarios like car following and traffic flow, MARL, like multi-agent deep deterministic policy gradient (MADDPG), facilitated cruising driving and energy management. Scholars also tried to combine RL agents with rule-based or optimization-based EMSs. Relying on the stability of PMP/ECMS, researchers begin to employ RL to adjust adaptive parameters such as the co-state or equivalent factor (EF), as well as LSTM and learning vector quantization (LVQ) networks are utilized to improve accuracy and efficiency within the MPC algorithm.

According to statistics on the current status, the year 2018 marked the end of the initial stage and the beginning of the development stage of learning-based EMSs. The following content summarizes and reviews all of the contributions after 2018, focusing on journal papers.

3.2 The Comprehensive Review of Research Status

Between 2019 and July 2023, a total of 166 journal papers were published, focusing on contributions categorized into algorithm, powertrain, and environment, and Tables 1, 2, and 3 present representative papers for each of these categories.

3.2.1 Algorithm Innovation

Algorithm innovation often plays a pivotal role throughout, indicating that when improving efficiency and addressing inherent flaws.

First of all, emerging researchers made contributions by employing various RL algorithms. Subsequently,

individuals delved beyond Q-Learning, exploring alternative algorithms like SARSA, Dyna-H, and DDPG. Furthermore, challenges posed by the "curse of dimensionality" and "discretization error" prompted scholars to pivot towards DRL algorithms. The quest for algorithmic innovation represents a significant advancement in the field, fostering a dynamic and vibrant research environment. For instance, Fast Q-Learning in Ref. [174], DDPG in Ref. [176], Dyna-H in Ref. [177], Dueling structure in Ref. [178], distributed DRL with A3C and DPPO in Ref. [198], TD3 in Ref. [196], SAC in Ref. [211], and Nash Q-Learning of MARL in Ref. [246]. Techniques like PER have gained more attention, and the adoption of TL has commenced. These contributions aim to enhance the training efficiency, solve flaws like overestimation, and achieve more efficient nonlinear fitting of value functions. For the current research, Guo et al. [183], Lee et al. [185], Lian et al. [187], Wang et al. [221], and Xu et al. [223] designed TL-based EMSs, and Lian et al. [187] analyzed the transfer process in detail for four hybrid power systems. Scholars have used the latest algorithms more frequently, meaning the advantages such as TD3 and MARL are grasped. Furthermore, when the reward function contains multiple items, Lv et al. [214] used IRL to determine the suitable weight of each item.

Moreover, one branch is dedicated to enhancing efficiency through self-designed methods. For example, Li et al. [176] improved the exploration by storing optimal results based on DP-based EMSs in the experience pool. Many researchers have utilized heuristic experience to guide the RL agent in the action space, like the brake-specific fuel consumption (BSFC) curve [186] and battery characteristics, or focused on updating the TPM by discriminative mechanisms, like KL divergence [174, 213, 216, 229], and induced matrix norm (IMN) [201] for modeling the environment and triggering the update. Some results have been achieved by combining rule-based and learning-based policies, capitalizing on strengths, and compensating for limitations. Tang et al. [220] merged learning-based EMSs with the rule-based engine start-stop, controlling the working period of the engine and enabling it to work efficiently when required. Xu et al. [199] adopted the ECMS-based EMS and heuristic control to pre-initialize the Q-table as the warm start, and Wu et al. [244] utilized a rule-based mode control to eliminate unreasonable exploration. The above are all auxiliary improvements to the RL agent in the control process.

Then, RL is regarded as a controller for key parameters in traditional EMSs. It involves selecting the co-state for PMP-based EMSs and the EF for ECMS-based EMSs to promote adaptability in stochastic environments. Guo et al. [182], Lee et al. [193], and Hu et al. [209, 210] made

Table 2 Representative research achievements of powertrain innovation

Reference No.	Powertrain	Algorithm	Contribution
[250]	FCV	Q learning	Fuel cell, battery, and supercapacitor; shrink state-action space based on an adaptive fuzzy filter; speedy Q-Learning; ECMS-based multi-objective optimization considering lifespan and fuel efficiency of fuel cell; RL-based splitting strategy for battery and fuel cell;
[252]	EV	DQN	Hybrid battery systems (high-energy and a high-power battery pack); energy loss minimization and electrical and thermal safety enhancement;
[253]	HE bus	SAC	Thermal safety and degradation of onboard lithium-ion battery; the over-temperature penalty and multi-stress-driven degradation cost of battery are introduced;
[256]	PHEV	DQN	The dual-mode engine with homogeneous charge compression ignition (HCCI) and spark ignition (SI); DP obtains optimal combustion mode and SOC reference trajectories at the cloud. DRL-based EMS and combustion mode with prioritized experience replay at the powertrain; makes full use of HCCI combustion mode and avoids frequent switching of combustion modes;
[270]	EV	SAC	Battery and supercapacitor; slower convergence rate, brittle training stability, and dissatisfactory optimization; parallel computing; DP-based expert knowledge is embedded;
[280]	HETV	DDPG	The lateral dynamics; and steering resistance on the energy distribution are considered; the multidimensional matrix framework; the pyramid-like network; hardware-in-the-loop;
[281]	Electric hydraulic HEV	TD3	A self-adaptive electric-hydraulic ratio under different driving cycles; a DRL-based EMS with a rule-based mode switching strategy;

Table 3 Representative research achievements of environmental innovation

Reference No.	Powertrain	Algorithm	Contribution
[296]	Serial HEV	DDPG	History of cumulative trip information to obtain the space-domain-indexed SOC trajectory; offline training and online application;
[302]	Power split PHEV	Actor-Critic	Route planning and power management; the minimum energy consumption route;
[304]	HE bus Power split HEV	DDPG	SUMO-based Multi-states of Traffic Information (the state of the surrounding vehicles and signal lights); transfer learning; the prior knowledge trained by HEB is transferred to Prius;
[307]	Power split HEV	DDPG	Connected traffic environment (distance headway, fuel consumption, and terrain); DDPG-based reference speed planning in the car following scenarios; A-ECMS-based EMS;
[311]	Parallel HEV	DQN	Five roads are collected (dry asphalt, wet asphalt, snow, dry cobblestone, and wet cobblestone); the time-varying driving environment (driving images, slope, speed, and the number of passengers.); the stereoscopic control network for the high-dimensional task (engine, transmission, motor);
[317]	FCV	SAC	Fuel cell, battery, and supercapacitor; the attention-based LSTM-based velocity and load power predictor by standard and real driving cycles;
[320]	Power split HEV	DDPG	The K-means and principal component analysis-based specific driving cycle is constructed by naturalistic data; expert-assisted EMS with brake-specific fuel consumption curve; battery aging model;
[323]	Parallel HEV	DDPG DQN	Processor-in-the-loop; lane-level high-definition map-driven integrated control; Map modeling by Google Earth and Google map; the velocity and steering of the vehicle layer and EMS of the powertrain;
[326]	Power split HEV	DDPG	Connected HEVs; MPC-based speed control for maintaining a safe distance and ensuring riding comfort; prioritized experience replay; the expert knowledge-based EMS;
[331]	Serial HEV	MADDPG	Ecological driving in the car following scenario in SUMO; Heterogeneous Multi-Agent DRL; MADDPG-based adaptive cruise control and EMS; prioritized experience replay;
[332]	Power split HEV	DDPG	Collaborative optimization of energy management strategy and adaptive cruise control; the noise on the action space (Ornstein-Uhlenbeck action noise and soft-max action noise).

main contributions similar to the above idea. Various problems in the simulation have been pointed out. Hu et al. [209] identified several main challenges like deployment inefficiency, safety constraints, and the gap between virtual simulation and the real world, and they are

incorporating data from both real and simulated environments to guide RL agents.

In addition, the specific parameters and settings that affect the training process are analyzed. Xu et al. [223] discuss the impact of introducing noise in action and parameter spaces. Other scholars explore novel

optimizers like AMSgrad [181] or delve into hyperparameters [190] such as discretization, and experience pool. Wang et al. [243] provide a comparison of 13 RL-based EMSs, analyzing various aspects like reward functions, computational efficiency, and convergence.

Finally, the safety of RL has received significant attention. During the training, RL agents learn the optimal strategy by trial and error while balancing the exploration-exploitation. Empirical evidence suggests that RL overlooks the dynamic of powertrains when generating actions, resulting in abrupt changes in actions. To address this problem, measures such as the penalty term for unreasonable actions [196], designing the coach mechanism to ensure training safety [202], and the rule-based controller to eliminate unreasonable distribution have been employed to constrain control actions [244]. Zhou et al. [204] and Hu et al. [210, 231] are taking a heuristic rule to eliminate irrational allocation and ensure safe exploration. Wang et al. [221] adopt the action masking technology to prevent unreasonable actions. Therefore, it is crucial to give enough attention to the safety of RL to ensure their practical deployment, as achievements realized at the simulation level may not effectively translate into real-world environments.

3.2.2 Powertrain Innovation

Powertrain innovation means promoting the diversification of powertrains and the realization of modeling schemes. This goes beyond the traditional focus on fuel economy and SOC, and the target aims at achieving multi-objective optimization, such as efficiency, temperature, and lifespan.

Firstly, as current research progresses, literature reflects significant diversity in terms of hybrid power systems, such as EVs with hybrid battery systems (high power battery and high energy battery) or supercapacitors, FCVs with fuel cells and batteries, and the three-energy system with fuel cells, batteries, and supercapacitors. Therefore, power distribution can also appear in EVs and FCVs, which means that EMSs are not only suitable for gasoline-electric hybrid systems. As to some special-purpose vehicles, some also regard them as targets, such as the rail transit adopted by Yang et al. [255], hybrid construction vehicles (HCVs) targeted by Zhang et al. [272], and electric-hydraulic HEVs targeted by Zhang et al. [293, 294]. Deng et al. [258] proposed the RL-based EMS that minimizes hydrogen consumption and fuel cell aging costs for the unique fuel cell railway vehicle. In addition, for many researchers from the Beijing Institute of Technology, HETVs from special vehicles and series/parallel HEVs from public transportation are studied, and the difficulty of control is also significantly increased [280].

Another item lies in focusing on temperature and lifespan. Many efforts have been dedicated to alleviating degradation and extending life through the utilization of various models. When the capacity decays to 80% of the initial capacity, the battery is treated as scrapped. Li et al. [252] built equivalent circuit models, electro-thermal models, and aging models for hybrid battery systems equipped with high-energy and high-power batteries. Zhang et al. [292] focused on the lithium-plating suppressed effect and designed the hybrid particle swarm optimization to complete the parameter identification. Haskara et al. [262] and Deng et al. [279] took temperature as the main goal and realized the temperature management of the cabin by adding heating ventilation air conditioning. Wu et al. [253] introduced the over-temperature and multi-stress-driven degradation costs.

Next, some scholars merged their expertise from unique domains and introduced specialized models, which made the modeling method closer to actual components and reflected many effects. Zhang et al. [256, 271] performed research on the dedicated dual-mode combustion engine with the spark ignition (SI) and homogeneous charge compression ignition (HCCI) modes. Wang et al. [267] took a waste heat recovery system based on the organic Rankine cycle. Hong et al. [281] and Zhang et al. [294] completed a training process of power distribution and mode switching strategy by using the RL for electro-hydraulic hybrid power systems.

Finally, some scholars make partial contributions based on improving the dynamic model, which is also the essential direction that is currently lacking in the EMS. Han et al. [280] added the lateral dynamics of the vehicle and introduced the steering resistance into the design of the EMS, resulting in the vehicle model getting rid of the status of just focusing on the longitudinal dynamics in the past. In this regard, there is still a lot of work to be done. For a real car, an ideal strategy, the high-fidelity dynamics model, an experienced driver, and a smooth driving environment are all factors to focus on.

3.2.3 Environmental Innovation

Environmental innovation represents the advancement of EMSs, involving the integration of technologies from more fields to enhance eco-driving. The development of autopilot and communication paves the way for energy-saving control for intelligent connected HEVs (ICHEVs) [340].

First, SOC planning and velocity prediction are primarily revolved. By using historical data or so-called connected information such as vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I), to gain insights into the environment, the future short-term SOC and velocity trajectories could be planned. For SOC planning,

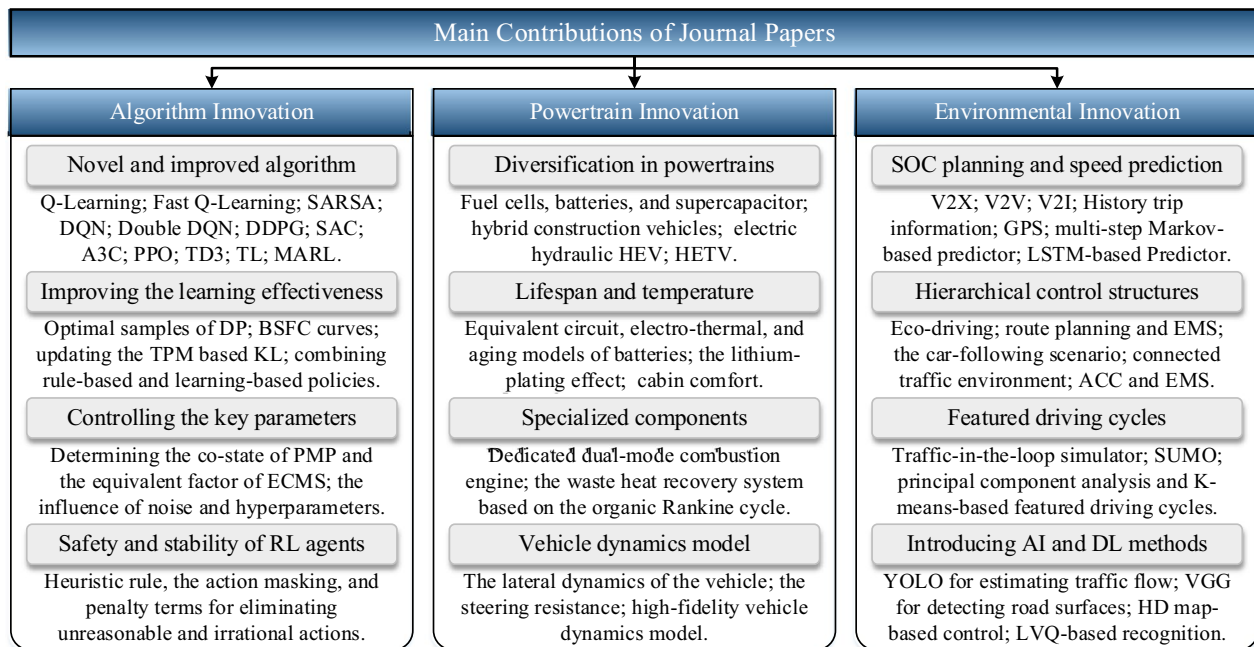


Figure 6 The main direction of the contributions of journal papers

there are local SOC trajectories designed by Guo et al. [295] and the space-domain-indexed SOC trajectory obtained by Li et al. [296] through a history of cumulative trip information. Similarly, Zhang et al. [301] are trying to employ GPS to complete the global planning of the SOC. Another item is the short-term prediction of speed, which not only allows RL to grasp future information but also researchers to integrate RL with the MPC framework. For speed prediction, Chen et al. [299], Yang et al. [322], and Wang et al. [334] adopted the multi-step Markov chain as predictors, and Liu et al. [329] and Kim et al. [317] utilized LSTM as the predictive tool. Studies have also demonstrated the advantages of the LSTM in velocity prediction, ensuring both accuracy and efficiency.

Additionally, the hierarchical structure under connected environments starts to be focused on, and more factors like driving conditions and driving styles reflecting randomness and personalized influences, are usually taken into account. Zhang et al. [302] researched eco-driving with route planning in the environment and energy management in the system. Li et al. [307] assumed that DDPG is used in the connected traffic environment to realize the reference speed planning in the car-following scene, and the A-ECMS is utilized for energy management. Peng et al. [331, 332] and Zhang et al. [338] both aimed at eco-driving and combined adaptive cruise control (ACC) with EMSs based on RL to achieve co-optimization in terms of velocity and power distribution.

Moreover, in the car-following scenario, maintaining a safe following distance and maintaining driving comfort has also become one of the main goals.

Another area is the construction of featured driving cycles. While most EMSs use standard driving cycles, the features in real-world scenarios, influenced by many factors such as traffic signals, traffic flow, pavement properties, and weather, are overlooked. Therefore, research has begun constructing featured driving cycles based on real data, using techniques such as principal component analysis (PCA) and K-means, which provide a realistic and intuitive reflection of velocity. He et al. [304] constructed the traffic environment containing information on surrounding vehicles and signal lights in the SUMO. Chang et al. [310], Tang et al. [320], and Huang et al. [327] adopted PCA and clustering algorithms to form the featured driving cycle. For the ramp scenario, Lin et al. [318] proposed a DDPG-based merging controller. Yan et al. [321] proposed DRL-based launch control to select the appropriate start time for reducing frequent starting and stopping through the traffic intersection. Moreover, Chen et al. [324] employed a traffic-in-the-loop simulator under various urban scenarios.

Next, more AI and ML technologies have been integrated. For instance, object detection algorithms, you only look once (YOLO), are utilized to identify traffic signals and estimate traffic flow according to the number of surrounding cars by Wang et al. [309], and Tang et al. [319] employ the YOLO to detect the leading car

and measure the following distance in the car-following scene. In addition, considering the impact of different road surfaces on driving safety, Chen et al. [311] trained the VGG16 neural network to identify and estimate the optimal slip rate for safe braking. Chen et al. [323] also built a lane-level map through the route and geographic data from Google Maps and Google Earth and adopted multiple DRL agents to achieve integrated control of the ICHEV. Moreover, driving condition recognition is often mentioned, and learning vector quantization (LVQ)-based recognition is employed by Chang et al. [310], Fang et al. [312], and Liu et al. [328], while Yang et al. [322] employees probabilistic neural networks for pattern recognition.

Finally, MARL algorithms and cloud and edge computing platforms have emerged as the burgeoning direction. Wang et al. [335] used independent SAC belonging to the MARL to research eco-driving and EMSs. Peng et al. [331] proposed a similar idea and used MARL to achieve the ACC and EMS tasks for eco-driving. Furthermore, Both Hu et al. [306] and Li et al. [308] proposed a training concept of cloud platforms and edge computing, which will be an inevitable method in the future. For a generalized strategy, large-scale computing devices on cloud platforms can satisfy the requirements for speed and computing power, and for personalized strategies, edge devices assigned to individuals, such as NVIDIA Jetson, are ideal training machines.

3.3 Discussion on RL-Based EMSs

Figure 6 summarizes the majority of the research contributions, forming the most intuitive expression, which is beneficial for researchers to quickly grasp the mainstream.

However, the enhancement of research popularity and the increase in the number of literature could only represent the positive aspect of the development of RL-based EMSs, but there are also some difficulties worth discussing. Next, the discussion on RL-based EMSs for hybrid power systems will also be carried out from three aspects: algorithm, data, and computing power.

- (1) Algorithm: In terms of algorithm improvements, TD3, SAC, and MARL are currently the most popular modules. They are committed to refining the training process of RL by enhancing various improvements to networks, amplifying both exploration and exploitation, and expanding the scale of training modes. Meanwhile, targeting online scenarios, the on-policy-based PPO that does not rely on experience pools has also achieved significant accomplishments. In fact, if we liken RL agents to students, researchers act as teachers responsible

for educating neural networks. Currently, most literature focuses on gradually improving training schemes in offline simulation and overcoming inherent flaws. This is like saying that teachers should designate different plans for different teaching contents in daily classes, and they should match the abilities and personalities of different students to achieve better guidance. Therefore, for the development of RL-based EMSs, several challenges should be encountered:

- a) Selecting the appropriate algorithm for different tasks.
- b) Adjusting neural network structures, state spaces, action spaces, reward functions, and hyperparameters.
- c) Verifying the generalization, safety, and robustness of trained agents in offline scenarios.
- d) Bridging the gap between simulation environments and the real world during offline training.
- e) Addressing the limitation of depending solely on offline simulation for modeling and training, the research aims to accomplish the ongoing updating of RL-based EMSs in the complex real-world while ensuring control safety.

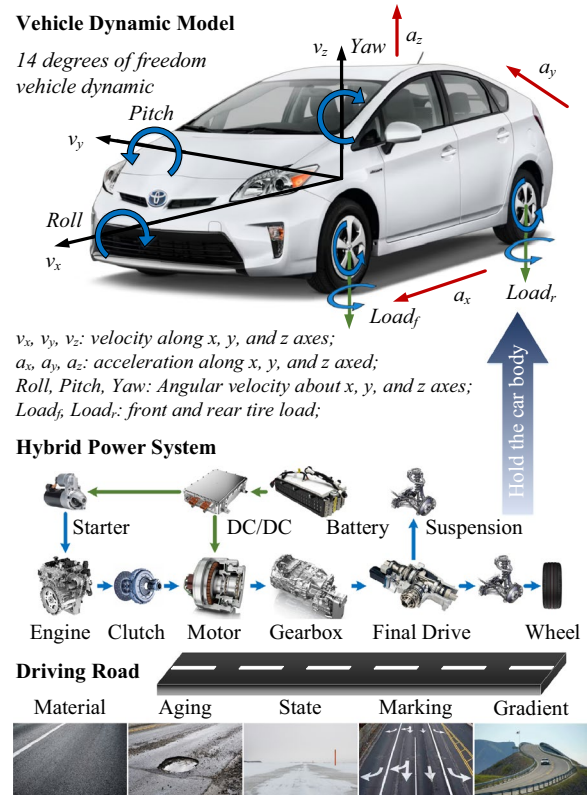


Figure 7 The modeling approach for ICHEVs

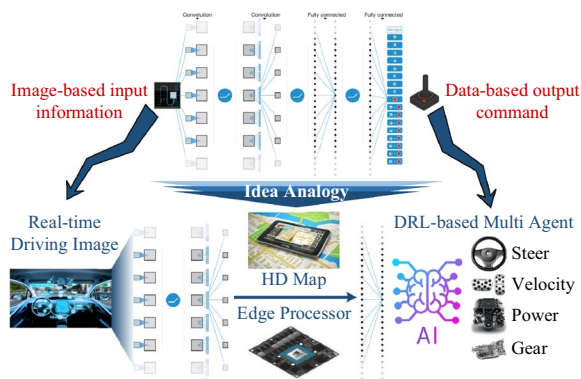


Figure 8 The integrated control framework for ICEVs

- f) The most debated topic in the AI community revolves around the necessity for trained models to comply with human morals and law. Otherwise, it may result in the emergence of "Terminators" rooted in silicon-based life forms as agents evolve.

Nowadays, progress is focused on addressing stages such as generalization and safety, online learning, and handling gaps. If RL is going to be deployed on real HEVs, there is still much work to be finished.

- (2) Data: As mentioned earlier, RL is one of the subsets of ML, distinct from DL in that it does not heavily manipulate labeling. It relies solely on a defined reward to assess the quality of outputs relative to inputs. Therefore, both DL and RL as sample-driven forms strive to fit complex and abstract relationships between all inputs and outputs. Regarding RL-based EMSs, the most controversial aspects lie in the degree of modeling of HEVs in offline training environments and the effectiveness of simulating driving conditions. Similar to autonomous driving, the environment in which hybrid power systems operate is complex, dynamic, and subject to various influencing factors such as temperature, aging, wear and tear, and potential accidents. Thus, if RL agents are to be trained for HEVs or autonomous driving, the dynamic nature of the environment and the "long tail" scenarios that cannot be exhaustively traversed pose significant challenges in training data collection. Currently, researchers from the University of Zurich have for the first time applied DRL to real unmanned aerial vehicles [341] and achieved championship-level effect in drone

races against humans. They utilized residual models of dynamics to compensate for the samples in the simulation environment. However, it is noted that trained DRL agents fail when faced with different lighting or collisions leading to drone crashes, indicating a lack of robustness comparable to human drivers.

- (3) Computing power: Although MathWorks has released a toolkit for RL, Python-based frameworks like TensorFlow and PyTorch remain the primary modeling environments for DRL agents. By operating in the form of tensor on GPUs, training efficiency can be significantly enhanced. In practice, for RL-based EMSs, the demand for GPU computing power is not particularly evident, as the main architecture typically consists of fully connected networks, and the input state is represented as tensor-form data. Instead, the reliance is more on Simulink-based powertrains and rendered 3D scenes of driving environments like CARLA and NVIDIA Driven sim. However, achieving end-to-end energy-saving autonomous driving with vehicle visualization will be a challenge for RL. Drawing from developments in robotics, training multiple agents in a 3D training environment will require substantial computing power. The NVIDIA Omniverse platform and the NVIDIA Isaac Sim provide tools for robot simulation and data generation, offering realistic and physically accurate virtual environments for developing, testing, and managing robots. Additionally, creating physically accurate large-scale simulations will enable the development of real-world digital twins, necessitating extensive support, such as NVIDIA OVX to accelerate AI-enabled workloads.

Finally, as stated in the tenets of the two AI teams, Deep Mind, and OpenAI: "Solve intelligence. Use it to make the world a better place" and DRL is the essential key that opens the door to the future era of AI.

4 The Future of Autonomous Intelligent HEVs

The advancement of autonomous driving and RL algorithms presents an enticing opportunity for developing autonomous intelligent HEVs, which may be named the "Alpha HEV", and we have forecasted the dependable technical strategy to attain complete control through RL.

Firstly, it completely gets rid of the backward simulation, the crude method of calculating the demand power and the dynamic, and easy models of engines and batteries. Not only to strengthen the modeling of each component, such as the engine, motor, battery, gearbox, clutch, shaft, brake, etc. but also to improve the vehicle model as

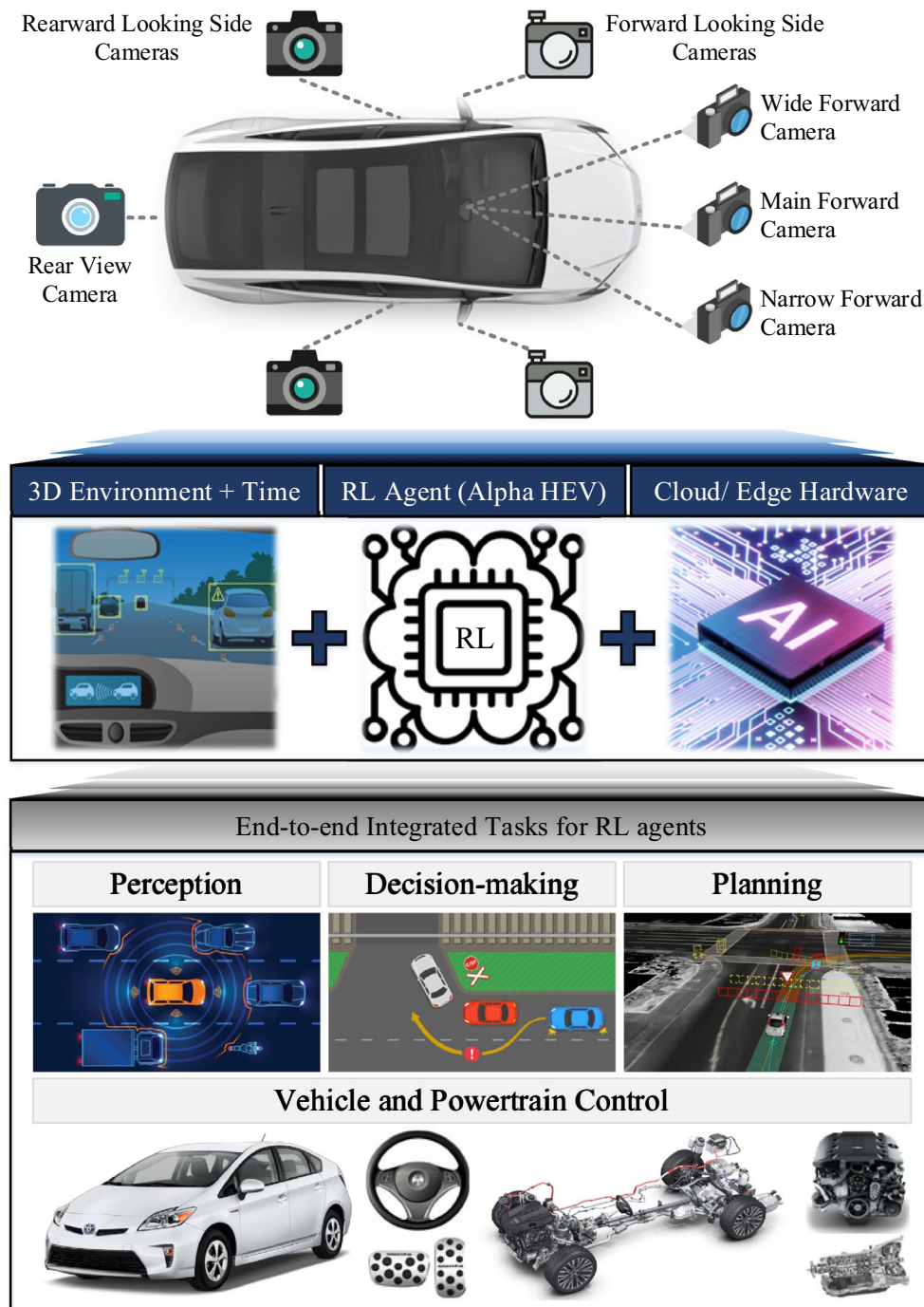


Figure 9 The RL-based autonomous intelligent HEV

a whole. Incorporating a forward simulation that includes a driver, the vehicle body should also be equipped with more degrees of freedom, and essential parts such as suspensions, tires, and body must be included. These enhancements enable the vehicle to respond realistically to external environments and internal systems, resulting

in accurate and reliable simulations. Furthermore, the modeling method of road surfaces becomes necessary. Beginning with basic features like slope, curvature, and road signs, more advanced factors should be involved like road materials, road aging, and dynamic variations influenced by weather. These elements not only impact

energy-efficient driving but also play a key role in safety, comfort, and other aspects, as depicted in Figure 7.

Then, achieving intelligent control through RL requires a fusion of multi-modal information. Commands in Figure 8 for the car encompass a wide range of functionalities, including acceleration, braking, steering, and managing the engine and gearbox within a powertrain. The intricate control facilitates efficient driving and the integration of various components for better performance. Drawing inspiration from an analogy of DQN in the Atari 2600, HEVs can also benefit from visual perception and high-definition maps from autopilot systems. This involves adding vehicle vision to RL agents, allowing them to navigate the path forward. Relying on vehicle vision and HD maps is reliable for perceiving surrounding vehicles and obtaining path information, and the ultimate goal is to achieve integrated control by using multi-agents, effectively allowing RL to fully take over all the controlled components.

Finally, there are more strict requirements for perception, decision-making, planning [342], software calibration, and hardware computing power. Taking inspiration from the Full Self-Driving (FSD) of Tesla, a three-dimensional perception space is constructed from real-time images captured by eight cameras, allowing for a comprehensive understanding of the surroundings in Figure 9, and this perception space is extended into the temporal dimension to ensure accurate perception of temporarily obscured objects. Then, a path is determined and planned, and the vehicle is controlled to track the route while the powertrain is managed at the same time. As an even more ambitious idea, a large DRL agent that acts as both a driver and an engineer is trained to handle all tasks solely from real-time images, and the perception and decision-making about autonomous driving will become cognition about the driving environment and passenger needs. While this notion presents significant challenges, it holds the potential to revolutionize autonomous driving. In conclusion, the deployment of RL agents in the "Alpha HEV" hinges on end-to-end autopilot functionality, resilient perception systems, and sophisticated decision-making algorithms, requiring collaboration among industry leaders to achieve safe and energy-saving autopilot.

5 Conclusions

This paper provides a state-of-the-art survey and review of the status in the field of RL-based EMSs for hybrid power systems. Firstly, it begins by tracing the development history of DL, RL, and DRL and highlighting many milestones and famous scholars. As the painter, we very much welcome and thank all subsequent scholars for

their corrections and more comprehensive additions to this timeline in Figure 1. Then, the focus shifts to shifts to RL-based EMSs, where a total of 266 papers have been collected as of July 21, 2023. The detailed tables summarizing all of the content have been uploaded at <https://github.com/KaysenC/Reinforcement-Learning-based-Energy-Management-for-Hybrid-Power-Systems> based on scholars, years, target powertrains, algorithms, and contributions. Moreover, statistical information is presented to illustrate the annual growth of research papers, providing valuable insights into the evolving interest in the field. Then, a comprehensive review of the current status is completed, with a novel emphasis on algorithm innovation, powertrain innovation, and environmental innovation. At the same time, difficulties that may arise in the subsequent development stages of RL-based EMSs are discussed from the aspects of algorithms, data, and computing power.

The ultimate goal is named "Alpha HEV," an autonomous intelligent HEV, and three main directions are highlighted: enhancing modeling, full takeover by DRL, and cognitive-oriented energy-saving autonomous driving.

In all, this paper summarizes the latest research status and presents the promising outlook for DRL-based HEVs. The pursuit of autonomous intelligent HEVs holds great potential to revolutionize the automotive industry, leading to efficient and environmentally friendly vehicles.

Abbreviations

A3C	Asynchronous advantage actor-critic
AC	Actor-critic
ACC	Adaptive cruise control
ADVISOR	Advanced vehicle simulation
AI	Artificial intelligence
BP	Back propagation
BSFC	Brake-specific fuel consumption curve
CNN	Convolutional neural network
CPU	Central processing unit
CV	Computer vision
DDPG	Deep deterministic policy gradient
DDQN	Double DQN
DL	Deep learning
DP	Dynamic programming
DQL	Deep Q-learning
DQN	Deep Q-network
DRL	Deep reinforcement learning
DRQN	Deep recurrent Q-network
ECMS	Equivalent consumption minimum strategy
EF	Equivalent factor
EMS	Energy management strategy
ERDEV	Extended-range delivery electric vehicles
FCV	Fuel cell vehicle
FSD	Full self-driving
GAN	Generative adversarial network
GPS	Global positioning system
GPU	Graphics processing unit
GRU	Gated recurrent unit
GT	Gran turismo
HCCI	Homogeneous charge compression ignition
HEB	Hybrid electric bus

HETV	Hybrid electric tracked vehicle
HEV	Hybrid electric vehicle
HIL	Hardware-in-the-loop
ICE	Internal combustion engine
ICHEV	Intelligent connected HEV
ILSVRC	ImageNet large-scale visual recognition challenge
IMN	Induced matrix norm
IRL	Inverse RL
KL	Kullback-Leibler
LSTM	Long short-term memory
LVQ	Learning vector quantization
MADDPG	Multi-agent deep deterministic policy gradient
MARL	Multi-agent reinforcement learning
MC	Monte Carlo
MCST	Monte Carlo Search Tree
ML	Machine learning
MPC	Model predictive control
NLP	Natural language processing
NML	Neural language model
PCA	Principal component analysis
PEV	Pure electric vehicle
PHEV	Plug-in hybrid electric vehicle
PMP	Pontryagin's minimum principle
PPO	Proximal policy optimization
PRE	Prioritized experience replay
RL	Reinforcement learning
RNN	Recurrent neural network
SAC	Soft actor-critic
SI	Spark ignition
TD	Temporal difference
TD3	Twin delayed deep deterministic policy gradient
TL	Transfer learning
TPM	Transition probability matrix
TRPO	Trust region policy optimization
V2I	Vehicle-to-infrastructure
V2V	Vehicle-to-vehicle
YOLO	You only look once

Acknowledgements

Not applicable.

Authors' Contributions

JC reviewed existing papers and drafted the manuscript. KY conducted a literature review. XT proposed the outlines. AK and SL assisted in the manuscript review. TL and YQ aided in manuscript revisions. All authors read and approved the final manuscript.

Funding

Supported by National Natural Science Foundation of China (Grant Nos. 52222215, 52072051), Fundamental Research Funds for the Central Universities in China (Grant No. 2023CDJXY-025), and Chongqing Municipal Natural Science Foundation of China (Grant No. CSTB2023NSCQ-JQX0003).

Data availability

The data used in this study are described in detail within this manuscript and are available upon request in electronic format. The datasets are referenced within the text and corresponding citations are provided for reader access. Readers are encouraged to utilize these data for further research and analysis within reasonable bounds, while maintaining appropriate citation and data usage standards. (<https://github.com/KaysenC/Reinforcement-Learningbased-Energy-Management-for-Hybrid-Power-Systems>) For further inquiries regarding data availability or access to datasets, please contact the authors.

Declarations

Competing Interests

The authors declare no competing financial interests.

Received: 19 December 2023 Revised: 25 March 2024 Accepted: 2 April 2024

Published online: 17 May 2024

References

- [1] Z Liu, H Hao, X Cheng, et al. Critical issues of energy efficient and new energy vehicles development in China. *Energy Policy*, 2018, 115: 92-97.
- [2] H He, F Sun, Z Wang, et al. China's battery electric vehicles lead the world: Achievements in technology system architecture and technological breakthroughs. *Green Energy and Intelligent Transportation*, 2022: 100020.
- [3] X Zhao, L Wang, Y Zhou, et al. Energy management strategies for fuel cell hybrid electric vehicles: Classification, comparison, and outlook. *Energy Conversion and Management*, 2022, 270: 116179.
- [4] Z Li, A Khajepour, J Song. A comprehensive review of the key technologies for pure electric vehicles. *Energy*, 2019, 182: 824-839.
- [5] M F M Sabri, K A Danapalasingam, M F Rahmat. A review on hybrid electric vehicles architecture and energy management strategies. *Renewable and Sustainable Energy Reviews*, 2016, 53: 1433-1442.
- [6] D D Tran, M Vafaiepour, Baghdadi M El, et al. Thorough state-of-the-art analysis of electric and hybrid vehicle powertrains: Topologies and integrated energy management strategies. *Renewable and Sustainable Energy Reviews*, 2020, 119: 109596.
- [7] Y Cao, M Yao, X Sun. An overview of modelling and energy management strategies for hybrid electric vehicles. *Applied Sciences*, 2023, 13(10): 5947.
- [8] H Pei, X Hu, Y Yang, et al. Designing multi-mode power split hybrid electric vehicles using the hierarchical topological graph theory. *IEEE Transactions on Vehicular Technology*, 2020, 69(7): 7159-7171.
- [9] X Hu, J Han, X Tang, et al. Powertrain design and control in electrified vehicles: A critical review. *IEEE Transactions on Transportation Electrification*, 2021, 7(3): 1990-2009.
- [10] B HomChaudhuri, R Lin, P Pisu. Hierarchical control strategies for energy management of connected hybrid electric vehicles in urban roads. *Transportation Research Part C: Emerging Technologies*, 2016, 62: 70-86.
- [11] M A Hannan, F A Azidin, A Mohamed. Hybrid electric vehicles and their challenges: A review. *Renewable and Sustainable Energy Reviews*, 2014, 29: 135-150.
- [12] A S Mohammed, S M Atnaw, A O Salau, et al. Review of optimal sizing and power management strategies for fuel cell/battery/super capacitor hybrid electric vehicles. *Energy Reports*, 2023, 9: 2213-2228.
- [13] J Peng, H He, R Xiong. Rule based energy management strategy for a series-parallel plug-in hybrid electric bus optimized by dynamic programming. *Applied Energy*, 2017, 185: 1633-1643.
- [14] S Zhang, X Hu, S Xie, et al. Adaptively coordinated optimization of battery aging and energy management in plug-in hybrid electric buses. *Applied Energy*, 2019, 256: 113891.
- [15] H Li, A Ravey, A N'Diaye, et al. Online adaptive equivalent consumption minimization strategy for fuel cell hybrid electric vehicle considering power sources degradation. *Energy Conversion and Management*, 2019, 192: 133-149.
- [16] F Zhang, X Hu, R Langari, et al. Energy management strategies of connected HEVs and PHEVs: Recent progress and outlook. *Progress in Energy and Combustion Science*, 2019, 73: 235-256.
- [17] X Hu, T Liu, X Qi, et al. Reinforcement learning for hybrid and plug-in hybrid electric vehicle energy management: Recent advances and prospects. *IEEE Industrial Electronics Magazine*, 2019, 13(3): 16-25.
- [18] R Ostadian, J Ramoul, A Biswas, et al. Intelligent energy management systems for electrified vehicles: Current status, challenges, and emerging trends. *IEEE Open Journal of Vehicular Technology*, 2020, 1: 279-295.
- [19] Q Feiyan, L Weimin. A review of machine learning on energy management strategy for hybrid electric vehicles. *2021 6th Asia Conference on Power and Electrical Engineering (ACPEE)*, 2021: 315-319.
- [20] T Liu, W Tan, X Tang, et al. Driving conditions-driven energy management strategies for hybrid electric vehicles: A review. *Renewable and Sustainable Energy Reviews*, 2021, 151: 111521.

- [21] C Song, K Kim, D Sung, et al. A review of optimal energy management strategies using machine learning techniques for hybrid electric vehicles. *International Journal of Automotive Technology*, 2021, 22: 1437-1452.
- [22] A H Ganesh, B Xu. A review of reinforcement learning based energy management systems for electrified powertrains: Progress, challenge, and potential solution. *Renewable and Sustainable Energy Reviews*, 2022, 154: 111833.
- [23] R Venkatasathish, C Dhanamjayulu. Reinforcement learning based energy management systems and hydrogen refuelling stations for fuel cell electric vehicles: An overview. *International Journal of Hydrogen Energy*, 2022, 47(64): 27646-27670.
- [24] M Al-Saadi, M Al-Greer, M Short. Reinforcement learning-based intelligent control strategies for optimal power management in advanced power distribution systems: A survey. *Energies*, 2023, 16(4): 1608.
- [25] J Gan, S Li, C Wei, et al. Intelligent learning algorithm and intelligent transportation-based energy management strategies for hybrid electric vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [26] D Qiu, Y Wang, W Hua, et al. Reinforcement learning for electric vehicle applications in power systems: A critical review. *Renewable and Sustainable Energy Reviews*, 2023, 173: 113052.
- [27] D Xu, C Zheng, Y Cui, et al. Recent progress in learning algorithms applied in energy management of hybrid vehicles: A comprehensive review. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2023, 10(1): 245-267.
- [28] NVIDIA. DGX Platform. Available: <https://www.nvidia.com/en-us/data-center/dgx-platform/>.
- [29] Y LeCun, Y Bengio, G Hinton. Deep learning. *Nature*, 2015, 521(7553): 436-444.
- [30] W S McCulloch, W Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 1943, 5: 115-133.
- [31] F Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 1958, 65(6): 386.
- [32] J J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 1982, 79(8): 2554-2558.
- [33] D E Rumelhart, G E Hinton, R J Williams. Learning representations by back-propagating errors. *Nature*, 1986, 323(6088): 533-536.
- [34] J L Elman. Finding structure in time. *Cognitive Science*, 1990, 14(2): 179-211.
- [35] Y LeCun, L Bottou, Y Bengio, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [36] S Hochreiter, J Schmidhuber. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735-1780.
- [37] Y Bengio, R Ducharme, P Vincent, et al. A Neural Probabilistic Language Model. *Journal of Machine Learning Research*, 2003, 3: 1137-1155.
- [38] J Goodfellow, J Pouget-Abadie, M Mirza, et al. Generative adversarial networks. *Communications of the ACM*, 2020, 63(11): 139-144.
- [39] K Cho, Merriënboer B Van, D Bahdanau, et al. On the properties of neural machine translation: Encoder-decoder approaches. 2014. arXiv preprint <https://arxiv.org/abs/1409.1259>
- [40] A Vaswani, N Shazeer, N Parmar, et al. Attention is All You Need. *Advances in Neural Information Processing*, 2017, 30: 5998-6008.
- [41] J Deng, W Dong, R Socher, et al. Imagenet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 248-255.
- [42] A Krizhevsky, I Sutskever, G E Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84-90.
- [43] M D Zeiler, R Fergus. Visualizing and understanding convolutional networks. *Computer Vision—ECCV 2014: 13th European Conference*, 2014: 818-833.
- [44] C Szegedy, W Liu, Y Jia, et al. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1-9. <https://doi.org/10.48550/arXiv.1409.4842>.
- [45] K Simonyan, A Zisserman. Very deep convolutional networks for large-scale image recognition. 2014. arXiv preprint <https://arxiv.org/abs/1409.1556>
- [46] K He, X Zhang, S Ren, et al. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.
- [47] K Arulkumaran, M P Deisenroth, M Brundage, et al. Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, 2017, 34(6): 26-38.
- [48] R Bellman. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 1954, 60(6): 503-515.
- [49] N Metropolis, S Ulam. The monte carlo method. *Journal of the American Statistical Association*, 1949, 44(247): 335-341.
- [50] R J Williams. *Reinforcement-learning connectionist systems (Technical Report NU-CCS-87-3)*. Boston, MA: Northeastern University, College of Computer Science, 1987.
- [51] R S Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 1988, 3: 9-44.
- [52] G Arummery, M Niranjan. *On-line Q-learning using connectionist systems*. Cambridge, 1994.
- [53] C J C H Watkins. *Learning from delayed rewards*. Dissertation, Cambridge University, 1989.
- [54] C J C H Watkins, P Dayan. Q-learning. *Machine Learning*, 1992, 8: 279-292.
- [55] L J Lin. *Reinforcement learning for robots using neural networks*. Carnegie Mellon University, 1992.
- [56] L Baird. Residual algorithms: Reinforcement learning with function approximation. *Machine Learning Proceedings 1995*, 1995: 30-37.
- [57] J Tsitsiklis, Roy B Van. Analysis of temporal-difference learning with function approximation. *Advances in Neural Information Processing Systems*, 1996, 9.
- [58] R S Sutton, D McAllester, S Singh, et al. Policy gradient methods for reinforcement learning with function approximation. *Advances in Neural Information Processing Systems*, 1999, 12.
- [59] A Y Ng, S Russell. Algorithms for inverse reinforcement learning. *ICML*, 2000, 1: 2.
- [60] P Abbeel, A Y Ng. Apprenticeship learning via inverse reinforcement learning. *Proceedings of the Twenty-First International Conference on Machine Learning*, 2004:1. <https://doi.org/10.1145/1015330.1015430>.
- [61] R Coulom. Efficient selectivity and backup operators in Monte-Carlo tree search. *International Conference on Computers and Games*, 2006: 72-83.
- [62] V Mnih, K Kavukcuoglu, D Silver, et al. Playing atari with deep reinforcement learning. 2013. <https://arxiv.org/abs/1312.5602>
- [63] V Mnih, K Kavukcuoglu, D Silver, et al. Human-level control through deep reinforcement learning. *Nature*, 2015, 518(7540): 529-533.
- [64] T P Lillicrap, J J Hunt, A Pritzel, et al. Continuous control with deep reinforcement learning. 2015. arXiv preprint <https://arxiv.org/abs/1509.02971>
- [65] T Schaul, J Quan, I Antonoglou, et al. Prioritized experience replay. 2015. arXiv preprint <https://arxiv.org/abs/1511.05952>
- [66] J Schulman, S Levine, P Abbeel, et al. Trust region policy optimization. *International Conference on Machine Learning*, 2015: 1889-1897.
- [67] M Hausknecht, P Stone. Deep recurrent q-learning for partially observable mdps. *2015 AAAI Fall Symposium Series*, 2015.
- [68] D Silver, A Huang, C J Maddison, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, 2016, 529(7587): 484-489.
- [69] D Silver, J Schrittwieser, K Simonyan, et al. Mastering the game of go without human knowledge. *Nature*, 2017, 550(7676): 354-359.
- [70] D Silver, T Hubert, J Schrittwieser, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 2018, 362(6419): 1140-1144.
- [71] Hasselt H Van, A Guez, D Silver. Deep reinforcement learning with double q-learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2016, 30(1). <https://doi.org/10.1609/aaai.v30i1.10295>.
- [72] Z Wang, T Schaul, M Hessel, et al. Dueling network architectures for deep reinforcement learning. *International Conference on Machine Learning*, 2016: 1995-2003.
- [73] V Mnih, A P Badia, M Mirza, et al. Asynchronous methods for deep reinforcement learning. *International Conference on Machine Learning*, 2016: 1928-1937.
- [74] J Schulman, F Wolski, P Dhariwal, et al. Proximal policy optimization algorithms. 2017. arXiv preprint <https://arxiv.org/abs/1707.06347>

- [75] T Haarnoja, A Zhou, P Abbeel, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *International Conference on Machine Learning*, 2018: 1861-1870.
- [76] S Fujimoto, H Hoof, D Meger. Addressing function approximation error in actor-critic methods. *International Conference on Machine Learning*, 2018: 1587-1596.
- [77] M Fortunato, M G Azar, B Piot, et al. Noisy networks for exploration. 2017. arXiv preprint <https://arxiv.org/abs/1706.10295>
- [78] M Hessel, J Modayil, Hasselt H Van, et al. Rainbow: Combining improvements in deep reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, 32(1), <https://doi.org/10.1609/aaai.v32i1.11796>.
- [79] O Vinyals, I Babuschkin, W M Czarnecki, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 2019, 575(7782): 350-354.
- [80] J Schrittwieser, I Antonoglou, T Hubert, et al. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 2020, 588(7839): 604-609.
- [81] J Jumper, R Evans, A Pritzel, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 2021, 596(7873): 583-589.
- [82] P R Wurman, S Barrett, K Kawamoto, et al. Outracing champion Gran Turismo drivers with deep reinforcement learning. *Nature*, 2022, 602(7896): 223-228.
- [83] J Degraeve, F Felici, J Buchli, et al. Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature*, 2022, 602(7897): 414-419.
- [84] Z Cao, K Jiang, W Zhou, et al. Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning. *Nature Machine Intelligence*, 2023, 5(2): 145-158.
- [85] S Feng, H Sun, X Yan, et al. Dense reinforcement learning for safety validation of autonomous vehicles. *Nature*, 2023, 615(7953): 620-627.
- [86] R Abdelhedi, A Lahyani, A C Ammari, et al. Reinforcement learning-based power sharing between batteries and supercapacitors in electric vehicles. *2018 IEEE International Conference on Industrial Technology (ICIT)*, 2018: 2072-2077.
- [87] H Chaoui, H Gualous, L Boulon, et al. Deep reinforcement learning energy management system for multiple battery based electric vehicles. *2018 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2018: 1-6.
- [88] Y Fang, C Song, B Xia, et al. An energy management strategy for hybrid electric bus based on reinforcement learning. *The 27th Chinese Control and Decision Conference (2015 CCDC)*, 2015: 4973-4977.
- [89] R C Hsu, S M Chen, W Y Chen, et al. A reinforcement learning based dynamic power management for fuel cell hybrid electric vehicle. *2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS)*, 2016: 460-464.
- [90] S A Kouche-Biyouki, S M A Naseri-Javarehshk, A Noori, et al. Power management strategy of hybrid vehicles using sarsa method. *Electrical Engineering (ICEE)*, 2018: 946-950.
- [91] X Lin, Y Wang, P Bogdan, et al. Reinforcement learning based power management for hybrid electric vehicles. *2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2014: 33-38.
- [92] C Liu, Y L Murphey. Power management for plug-in hybrid electric vehicles using reinforcement learning with trip information. *2014 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2014: 1-6.
- [93] C Liu, Y L Murphey. Analytical greedy control and Q-learning for optimal power management of plug-in hybrid electric vehicles. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017: 1-8.
- [94] T Liu, C Yang, C Hu, et al. Reinforcement learning-based predictive control for autonomous electrified vehicles. *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018: 185-190.
- [95] X Qi, Y Luo, G Wu, et al. Deep reinforcement learning-based vehicle energy efficiency autonomous learning system. *2017 IEEE Intelligent Vehicles Symposium (IV)*, 2017: 1228-1233.
- [96] S Yue, Y Wang, Q Xie, et al. Model-free learning-based online management of hybrid electrical energy storage systems in electric vehicles. *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*, 2014: 3142-3148.
- [97] P Zhao, Y Wang, N Chang, et al. A deep reinforcement learning framework for optimizing fuel economy of hybrid electric vehicles. *2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2018: 196-202.
- [98] C H Zheng, C M Lee, Y C Huang, et al. Adaptive optimal control algorithm for maturing energy management strategy in fuel-cell/Li-ion-capacitor hybrid electric vehicles. *2013 9th Asian Control Conference (ASCC)*, 2013: 1-7.
- [99] A Biswas, P G Anselma, A Emadi. Real-time optimal energy management of electrified powertrains with reinforcement learning. *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019: 1-6.
- [100] T Gole, A Hange, R Dhar, et al. Reinforcement learning based energy management in hybrid electric vehicle. *2019 International Conference on Power Electronics*, 2019: 1-5.
- [101] D He, Y Zou, J Wu, et al. Deep Q-learning based energy management strategy for a series hybrid electric tracked vehicle and its adaptability validation. *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019: 1-6.
- [102] A Heimrath, J Froeschl, R Rezaei, et al. Reflex-augmented reinforcement learning for operating strategies in automotive electrical energy management. *2019 International Conference on Computing*, 2019: 62-67.
- [103] J Hofstetter, H Bauer, W Li, et al. Energy and emission management of hybrid electric vehicles using reinforcement learning. *IFAC-PapersOn-Line*, 2019, 52(29): 19-24.
- [104] S Inuzuka, F Xu, B Zhang, et al. Reinforcement learning based on energy management strategy for HEVs. *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2019: 1-6.
- [105] Keyser A De, G Crevecoeur. Integrated offline reinforcement learning for optimal power flow management in an electric dual-drive vehicle. *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2019: 1305-1310.
- [106] Y Li, J Tao, K Han. Rule and Q-learning based hybrid energy management for electric vehicle. *2019 Chinese Automation Congress (CAC)*, 2019: 51-56.
- [107] R Liessner, A M Dietermann, B Bäker. Safe deep reinforcement learning hybrid electric vehicle energy management. *Agents and Artificial Intelligence: 10th International Conference*, 2019: 161-181.
- [108] R Liessner, A Lorenz, J Schmitt, et al. Simultaneous electric powertrain hardware and energy management optimization of a hybrid electric vehicle using deep reinforcement learning and Bayesian optimization. *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2019: 1-6.
- [109] R Liessner, J Schmitt, A Dietermann, et al. Hyperparameter optimization for deep reinforcement learning in vehicle energy management. *ICAART*, 2019: 134-144.
- [110] N P Reddy, D Padeloup, M K Zadeh, et al. An intelligent power and energy management system for fuel cell/battery hybrid electric vehicle using reinforcement learning. *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019: 1-6.
- [111] I Sanusi, A Mills, G Konstantopoulos, et al. Power management optimisation for hybrid electric systems using reinforcement learning and adaptive dynamic programming. *2019 American Control Conference (ACC)*, 2019: 2608-2613.
- [112] P Wang, Y Li, S Shekhar, et al. A deep reinforcement learning framework for energy management of extended range electric delivery vehicles. *2019 IEEE Intelligent Vehicles Symposium (IV)*, 2019: 1837-1842.
- [113] P Wang, Y Li, S Shekhar, et al. Actor-critic based deep reinforcement learning framework for energy management of extended range electric delivery vehicles. *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2019: 1379-1384.
- [114] P Wang, Y Li, S Shekhar, et al. Uncertainty estimation with distributional reinforcement learning for applications in intelligent transportation systems: A case study. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019: 3822-3827.
- [115] A Biswas, P G. Anselma, A Rathore, et al. Comparison of three real-time implementable energy management strategies for multi-mode electrified powertrain. *2020 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2020: 514-519.
- [116] T Liu, X Tang, J Chen, et al. Transferred energy management strategies for hybrid electric vehicles based on driving conditions recognition. *2020 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2020: 1-6.
- [117] P Wang, Y Li, S Shekhar, et al. Risk-aware energy management of extended range electric delivery vehicles with implicit quantile

- network. *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, 2020: 772-778.
- [118] P Wang, Y Li, S Shekhar, et al. Uncertainty-aware energy management of extended range electric delivery vehicles with Bayesian ensemble. *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020: 1556-1562.
- [119] Y Wu, Y Liu, Z Chen, et al. Reinforcement energy management strategy for a plug-in hybrid electric vehicle considering state-of-charge constraint. *2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2020: 282-287.
- [120] Z Zhu, Y Liu, M Canova. Energy management of hybrid electric vehicles via deep Q-networks. *2020 American Control Conference (ACC)*, 2020: 3077-3082.
- [121] S Y Chen, H Y Lo, T Y Tsao, et al. Energy management system for a hybrid electric vehicle using reinforcement learning. *2021 IEEE International Symposium on Product Compliance Engineering-Asia (ISPCE-ASIA)*, 2021: 1-2.
- [122] K Deng, D Hai, H Peng, et al. Deep reinforcement learning based energy management strategy for fuel cell and battery powered rail vehicles. *2021 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2021: 1-6.
- [123] L Guo, Z Li, R Outbib. Reinforcement learning based energy management for fuel cell hybrid electric vehicles. *IECON 2021-47th Annual Conference of the IEEE Industrial Electronics Society*, 2021: 1-6.
- [124] S Hu, X Wu, J Li. Adaptive energy management strategy based on deep reinforcement learning for extended-range electric vehicles. *2021 5th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2021: 1-6.
- [125] R Huang, H He, X Meng, et al. A novel hierarchical predictive energy management strategy for plug-in hybrid electric bus combined with deep reinforcement learning. *2021 International Conference on Electrical*, 2021: 1-5.
- [126] R Huang, H He, X Meng, et al. Energy management strategy for plug-in hybrid electric bus based on improved deep deterministic policy gradient algorithm with prioritized replay. *2021 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2021: 1-6.
- [127] X Meng, Q Li, G Zhang, et al. Double Q-learning-based energy management strategy for overall energy consumption optimization of fuel cell/battery vehicle. *2021 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2021: 1-6.
- [128] S Nethagani, A S Yadav, S Kanagala, et al. Machine learning based Energy management system for better optimisation of power in electric vehicles. *2021 5th International Conference on Electronics*, 2021: 335-339.
- [129] J Tao, G Chen, R Gao. Neural network and reinforcement learning based energy management strategy for battery/supercapacitor HEV. *2021 China Automation Congress (CAC)*, 2021: 5623-5628.
- [130] Z Wei, H Ruan, H He. Battery thermal-conscious energy management for hybrid electric bus based on fully-continuous control with deep reinforcement learning. *2021 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2021: 1-5.
- [131] Y Ye, J Zhang, B Xu. A fast Q-learning energy management strategy for battery/supercapacitor electric vehicles considering energy saving and battery aging. *2021 International Conference on Electrical*, 2021: 1-6.
- [132] C Zheng, W Li, Y Xiao, et al. A deep deterministic policy gradient-based energy management strategy for fuel cell hybrid vehicles. *2021 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2021: 1-6.
- [133] A Biswas, Y Wang, A Emadi. Effect of immediate reward function on the performance of reinforcement learning-based energy management system. *2022 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2022: 1021-1026.
- [134] F Chen, P Mei, H Xie, et al. Reinforcement learning-based energy management control strategy of hybrid electric vehicles. *2022 8th International Conference on Control*, 2022: 248-252.
- [135] W Chen, G Yin, Y Fan, et al. Ecological driving strategy for fuel cell hybrid electric vehicle based on continuous deep reinforcement learning. *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2022: 1-6.
- [136] R Ghaderi, M Kandidayeni, L Boulon, et al. Power allocation of an electrified vehicle based on blended reinforcement learning with fuzzy logic. *2022 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2022: 1-5.
- [137] L Guo, Z Li, R Outbib. Fuzzy rule value reinforcement learning based energy management strategy for fuel cell hybrid electric vehicles. *IECON 2022-48th Annual Conference of the IEEE Industrial Electronics Society*, 2022: 1-7.
- [138] L Guo, Z Li, R Outbib. A lifetime extended energy management strategy for fuel cell hybrid electric vehicles via self-learning fuzzy reinforcement learning. *2022 10th International Conference on Systems and Control (ICSC)*, 2022: 161-167.
- [139] L Han, K Yang, X Zhang, et al. Energy management strategy for hybrid electric vehicles based on double Q-learning. *International Conference on Mechanical Design and Simulation (MDS 2022)*, 2022, 12261: 639-648.
- [140] S Hou, X Liu, H Yin, et al. Reinforcement learning-based energy optimization for a fuel cell electric vehicle. *2022 4th International Conference on Smart Power & Internet Energy Systems (SPIES)*, 2022: 1928-1933.
- [141] Y Lin, L Chu, J Hu, et al. DRL-ECMS: An adaptive hierarchical equivalent consumption minimization strategy based on deep reinforcement learning. *2022 IEEE Intelligent Vehicles Symposium (IV)*, 2022: 235-240.
- [142] Y Lin, L Chu, J Hu, et al. An intelligent energy management strategy for plug-in hybrid electric vehicle inspired from monte carlo tree search. *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022: 811-816.
- [143] Y Shen, F Yi, Y Fan, et al. Fuel cell bus energy management based on deep reinforcement learning in NGSIM high-speed traffic scenario. *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2022: 1-6.
- [144] Z Wang, J Xie, M Kang, et al. Energy management for a series-parallel plug-in hybrid electric truck based on reinforcement learning. *2022 13th Asian Control Conference (ASCC)*, 2022: 590-596.
- [145] Y Wu, R Lian, Y Wang, et al. Benchmarking deep reinforcement learning based energy management systems for hybrid electric vehicles. *CAA International Conference on Artificial Intelligence*, 2022: 613-625.
- [146] J Xu, Z Li, L Gao, et al. A comparative study of deep reinforcement learning-based transferable energy management strategies for hybrid electric vehicles. *2022 IEEE Intelligent Vehicles Symposium (IV)*, 2022: 470-477.
- [147] Y Ye, B Xu, J Zhang, et al. Reinforcement learning-based energy management system enhancement using digital twin for electric vehicles. *2022 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2022: 1-6.
- [148] Z Niu, H He, Y Wang, et al. Energy management optimization for connected hybrid electric vehicle with offline reinforcement learning. *2022 IEEE 12th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2022: 103-106.
- [149] C Zhang, W Cui, N Cui. Deep reinforcement learning based multi-objective energy management strategy for a plug-in hybrid electric bus considering driving style recognition. *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2022: 1-6.
- [150] C Zhang, W Cui, Y Du, et al. Energy management of hybrid electric vehicles based on model predictive control and deep reinforcement learning. *2022 41st Chinese Control Conference (CCC)*, 2022: 5441-5446.
- [151] K Zhang, J Ruan, Z Ye, et al. Energy management strategy based on constructing a fitting driving cycle for pure electric vehicles. *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2022: 1-6.
- [152] X Li, Y Zhang, Y Peng, et al. Reinforcement learning-based energy management for plug-in hybrid electric vehicles. *2023 9th International Conference on Electrical Engineering*, 2023: 1-6.
- [153] Y Xu, G Han, D Zhang, et al. Joint energy management and eco-routing for electric vehicles with hybrid energy storage systems. *2023 4th Information Communication Technologies Conference (ICTC)*, 2023: 374-378.
- [154] P Yadav, V K Saini, A S Al-Sumaiti, et al. Intelligent energy management strategies for hybrid electric transportation. *2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT)*, 2023: 1-7.
- [155] O Yazar, S Coskun, L Li, et al. Actor-critic TD3-based deep reinforcement learning for energy management strategy of HEV. *2023 5th International Congress on Human-Computer Interaction*, 2023: 1-6, <https://doi.org/10.1109/HORA58378.2023.10156727>.
- [156] J Cao, R Xiong. Reinforcement learning-based real-time energy management for plug-in hybrid electric vehicle with hybrid energy storage system. *Energy Procedia*, 2017, 142: 1896-1901.

- [157] Z Chen, H Hu, Y Wu, et al. Energy management for a power-split plug-in hybrid electric vehicle based on reinforcement learning. *Applied Sciences*, 2018, 8(12): 2494.
- [158] R C Hsu, C T Liu, D Y Chan. A reinforcement-learning-based assisted power management with QoR provisioning for human–electric hybrid bicycle. *IEEE Transactions on Industrial Electronics*, 2011, 59(8): 3350–3359.
- [159] Y Hu, W Li, K Xu, et al. Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning. *Applied Sciences*, 2018, 8(2): 187.
- [160] Z Kong, Y Zou, T Liu. Implementation of real-time energy management strategy based on reinforcement learning for hybrid electric vehicles and simulation validation. *PloS one*, 2017, 12(7): e0180491.
- [161] Y Li, H He, J Peng, et al. Energy management strategy for a series hybrid electric vehicle using improved deep Q-network learning algorithm with prioritized replay. *DEStech Transactions on Environment*, 2018, 978(1): 1–6.
- [162] Y Li, H He, J Peng, et al. Power management for a plug-in hybrid electric vehicle based on reinforcement learning with continuous state and action spaces. *Energy Procedia*, 2017, 142: 2270–2275.
- [163] T Liu, G Du, Y Zou, et al. Fast learning-based control for energy management of hybrid electric vehicles. *IFAC-PapersOnLine*, 2018, 51(31): 595–600.
- [164] T Liu, X Hu. A bi-level control for energy efficiency improvement of a hybrid tracked vehicle. *IEEE Transactions on Industrial Informatics*, 2018, 14(4): 1616–1625.
- [165] T Liu, X Hu, S E Li, et al. Reinforcement learning optimized look-ahead energy management of a parallel hybrid electric vehicle. *IEEE/ASME Transactions on Mechatronics*, 2017, 22(4): 1497–1507.
- [166] T Liu, B Wang, C Yang. Online Markov chain-based energy management for a hybrid tracked vehicle with speedy Q-learning. *Energy*, 2018, 160: 544–555.
- [167] T Liu, Y Zou, D Liu, et al. Reinforcement learning of adaptive energy management with transition probability for a hybrid electric tracked vehicle. *IEEE Transactions on Industrial Electronics*, 2015, 62(12): 7837–7846.
- [168] T Liu, Y Zou, D Liu, et al. Reinforcement learning-based energy management strategy for a hybrid electric tracked vehicle. *Energies*, 2015, 8(7): 7243–7260.
- [169] X Qi, G Wu, K Boriboonsomsin, et al. Data-driven reinforcement learning-based real-time energy management system for plug-in hybrid electric vehicles. *Transportation Research Record*, 2016, 2572(1): 1–8.
- [170] J Wu, H He, J Peng, et al. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Applied Energy*, 2018, 222: 799–811.
- [171] R Xiong, J Cao, Q Yu. Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. *Applied Energy*, 2018, 211: 538–548.
- [172] J Yuan, L Yang, Q Chen. Intelligent energy management strategy based on hierarchical approximate global optimization for plug-in fuel cell hybrid electric vehicles. *International Journal of Hydrogen Energy*, 2018, 43(16): 8063–8078.
- [173] Y Zou, T Liu, D Liu, et al. Reinforcement learning-based real-time energy management for a hybrid tracked vehicle. *Applied Energy*, 2016, 171: 372–382.
- [174] G Du, Y Zou, X Zhang, et al. Intelligent energy management for hybrid electric tracked vehicles using online reinforcement learning. *Applied Energy*, 2019, 251: 113388.
- [175] X Han, H He, J Wu, et al. Energy management based on reinforcement learning with double deep Q-learning for a hybrid electric tracked vehicle. *Applied Energy*, 2019, 254: 113708.
- [176] Y Li, H He, A Khajepour, et al. Energy management for a power-split hybrid electric bus via deep reinforcement learning with terrain information. *Applied Energy*, 2019, 255: 113762.
- [177] T Liu, X Hu, W Hu, et al. A heuristic planning reinforcement learning-based energy management for power-split plug-in hybrid electric vehicles. *IEEE Transactions on Industrial Informatics*, 2019, 15(12): 6436–6445.
- [178] X Qi, Y Luo, G Wu, et al. Deep reinforcement learning enabled self-learning control for energy efficient driving. *Transportation Research Part C: Emerging Technologies*, 2019, 99: 67–81.
- [179] H Tan, H Zhang, J Peng, et al. Energy management of hybrid electric bus based on deep reinforcement learning in continuous state and action space. *Energy Conversion and Management*, 2019, 195: 548–560.
- [180] Y Yin, Y Ran, L Zhang, et al. An energy management strategy for a super-mild hybrid electric vehicle based on a known model of reinforcement learning. *Journal of Control Science and Engineering*, 2019, <https://doi.org/10.1155/2019/9259712>.
- [181] G Du, Y Zou, X Zhang, et al. Deep reinforcement learning based energy management for a hybrid electric vehicle. *Energy*, 2020, 201: 117591.
- [182] H Guo, F Zhao, H Guo, et al. Self-learning energy management for plug-in hybrid electric bus considering expert experience and generalization performance. *International Journal of Energy Research*, 2020, 44(7): 5659–5674.
- [183] X Guo, T Liu, B Tang, et al. Transfer deep reinforcement learning-enabled energy management strategy for hybrid tracked vehicle. *IEEE Access*, 2020, 8: 165837–165848.
- [184] H Lee, C Kang, Y I Park, et al. Online data-driven energy management of a hybrid electric vehicle using model-based Q-learning. *IEEE Access*, 2020, 8: 84444–84454.
- [185] H Lee, C Song, N Kim, et al. Comparative analysis of energy management strategies for HEV: Dynamic programming and reinforcement learning. *IEEE Access*, 2020, 8: 67112–67123.
- [186] R Lian, J Peng, Y Wu, et al. Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle. *Energy*, 2020, 197: 117297.
- [187] R Lian, H Tan, J Peng, et al. Cross-type transfer for deep reinforcement learning based hybrid electric vehicle energy management. *IEEE Transactions on Vehicular Technology*, 2020, 69(8): 8367–8380.
- [188] C Liu, Y L Murphey. Optimal power management based on Q-learning and neuro-dynamic programming for plug-in hybrid electric vehicles. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(6): 1942–1954.
- [189] B Xu, X Hu, X Tang, et al. Ensemble reinforcement learning-based supervisory control of hybrid electric vehicle for fuel economy improvement. *IEEE Transactions on Transportation Electrification*, 2020, 6(2): 717–727.
- [190] B Xu, D Rathod, D Zhang, et al. Parametric study on reinforcement learning optimized energy management strategy for a hybrid electric vehicle. *Applied Energy*, 2020, 259: 114200.
- [191] G Du, Y Zou, X Zhang, et al. Heuristic energy management strategy of hybrid electric vehicle based on deep reinforcement learning with accelerated gradient optimization. *IEEE Transactions on Transportation Electrification*, 2021, 7(4): 2194–2208.
- [192] H Lee, S W Cha. Energy management strategy of fuel cell electric vehicles using model-based reinforcement learning with data-driven model update. *IEEE Access*, 2021, 9: 59244–59254.
- [193] H Lee, S W Cha. Reinforcement learning based on equivalent consumption minimization strategy for optimal control of hybrid electric vehicles. *IEEE Access*, 2020, 9: 860–871.
- [194] W Lee, H Jeoung, D Park, et al. A real-time intelligent energy management strategy for hybrid electric vehicles using reinforcement learning. *IEEE Access*, 2021, 9: 72759–72768.
- [195] X Lin, B Zhou, Y Xia. Online recursive power management strategy based on the reinforcement learning algorithm with cosine similarity and a forgetting factor. *IEEE Transactions on Industrial Electronics*, 2020, 68(6): 5013–5023.
- [196] Z E Liu, Q. Zhou, Y. Li, et al. An intelligent energy management strategy for hybrid vehicle with irrational actions using twin delayed deep deterministic policy gradient. *IFAC-PapersOnLine*, 2021, 54(10): 546–551.
- [197] C Qi, Y Zhu, C Song, et al. Self-supervised reinforcement learning-based energy management for a hybrid electric vehicle. *Journal of Power Sources*, 2021, 514: 230584.
- [198] X Tang, J Chen, T Liu, et al. Distributed deep reinforcement learning-based energy and emission management strategy for hybrid electric vehicles. *IEEE Transactions on Vehicular Technology*, 2021, 70(10): 9922–9934.
- [199] B Xu, J Hou, J Shi, et al. Learning time reduction using warm-start methods for a reinforcement learning-based supervisory control in hybrid electric vehicle applications. *IEEE Transactions on Transportation Electrification*, 2020, 7(2): 626–635.
- [200] N Yang, L Han, C Xiang, et al. Energy management for a hybrid electric vehicle based on blended reinforcement learning with backward

- focusing and prioritized sweeping. *IEEE Transactions on Vehicular Technology*, 2021, 70(4): 3136-3148.
- [201] N Yang, L Han, C Xiang, et al. An indirect reinforcement learning based real-time energy management strategy via high-order Markov chain model for a hybrid electric vehicle. *Energy*, 2021, 236: 121337.
- [202] H Zhang, J Peng, H Tan, et al. A deep reinforcement learning-based energy management framework with lagrangian relaxation for plug-in hybrid electric vehicle. *IEEE Transactions on Transportation Electrification*, 2020, 7(3): 1146-1160.
- [203] J Zhang, X Jiao, C Yang. A double-deep Q-network-based energy management strategy for hybrid electric vehicles under variable driving cycles. *Energy Technology*, 2021, 9(2): 2000770.
- [204] J Zhou, S Xue, Y Xue, et al. A novel energy management strategy of hybrid electric vehicle via an improved TD3 deep reinforcement learning. *Energy*, 2021, 224: 120118.
- [205] Q Zhou, D Zhao, B Shuai, et al. Knowledge implementation and transfer with an adaptive learning network for real-time power management of the plug-in hybrid vehicle. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(12): 5298-5308.
- [206] R Zou, L Fan, Y Dong, et al. DQL energy management: An online-updated algorithm and its application in fix-line hybrid electric vehicle. *Energy*, 2021, 225: 120174.
- [207] A Biswas, P G Anselma, A Emadi. Real-time optimal energy management of multimode hybrid electric powertrain with online trainable asynchronous advantage actor-critic algorithm. *IEEE Transactions on Transportation Electrification*, 2021, 8(2): 2676-2694.
- [208] G Du, Y Zou, X Zhang, et al. Energy management for a hybrid electric vehicle based on prioritized deep reinforcement learning framework. *Energy*, 2022, 241: 122523.
- [209] B Hu, J Li. A deployment-efficient energy management strategy for connected hybrid electric vehicle based on offline reinforcement learning. *IEEE Transactions on Industrial Electronics*, 2021, 69(9): 9644-9654.
- [210] B Hu, J Li. An adaptive hierarchical energy management strategy for hybrid electric vehicles combining heuristic domain knowledge and data-driven deep reinforcement learning. *IEEE Transactions on Transportation Electrification*, 2021, 8(3): 3275-3288.
- [211] T Li, W Cui, N Cui. Soft actor-critic algorithm-based energy management strategy for plug-in hybrid electric vehicle. *World Electric Vehicle Journal*, 2022, 13(10): 193.
- [212] W Li, J Ye, Y Cui, et al. A speedy reinforcement learning-based energy management strategy for fuel cell hybrid vehicles considering fuel cell system lifetime. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2021: 1-14.
- [213] X Lin, K Zhou, L Mo, et al. Intelligent energy management strategy based on an improved reinforcement learning algorithm with exploration factor for a plug-in PHEV. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(7): 8725-8735.
- [214] H Lv, C Qi, C Song, et al. Energy management of hybrid electric vehicles based on inverse reinforcement learning. *Energy Reports*, 2022, 8: 5215-5224.
- [215] C Maino, A Mastropietro, L Sorrentino, et al. Project and development of a reinforcement learning based control algorithm for hybrid electric vehicles. *Applied Sciences*, 2022, 12(2): 812.
- [216] C Qi, C Song, F Xiao, et al. Generalization ability of hybrid electric vehicle energy management strategy based on reinforcement learning method. *Energy*, 2022, 250: 123826.
- [217] C Qi, Y Zhu, C Song, et al. Hierarchical reinforcement learning based energy management strategy for hybrid electric vehicle. *Energy*, 2022, 238: 121703.
- [218] M Sun, P Zhao, X Lin. Power management in hybrid electric vehicles using deep recurrent reinforcement learning. *Electrical Engineering*, 2022, 104(3): 1459-1471.
- [219] W Sun, Y Zou, X Zhang, et al. High robustness energy management strategy of hybrid electric vehicle based on improved soft actor-critic deep reinforcement learning. *Energy*, 2022, 258: 124806.
- [220] X Tang, J Chen, H Pu, et al. Double deep reinforcement learning-based energy management for a parallel hybrid electric vehicle with engine start-stop strategy. *IEEE Transactions on Transportation Electrification*, 2021, 8(1): 1376-1388.
- [221] K Wang, R Yang, Y Zhou, et al. Design and improvement of SD3-based energy management strategy for a hybrid electric urban bus. *Energies*, 2022, 15(16): 5878.
- [222] B Xu, X Tang, X Hu, et al. Q-learning-based supervisory control adaptability investigation for hybrid electric vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(7): 6797-6806.
- [223] J Xu, Z Li, G Du, et al. A transferable energy management strategy for hybrid electric vehicles via dueling deep deterministic policy gradient. *Green Energy and Intelligent Transportation*, 2022, 1(2): 100018.
- [224] N Yang, L Han, C Xiang, et al. Real-time energy management for a hybrid electric vehicle based on heuristic search. *IEEE Transactions on Vehicular Technology*, 2022, 71(12): 12635-12647.
- [225] B Zhang, Y Zou, X Zhang, et al. Online updating energy management strategy based on deep reinforcement learning with accelerated training for hybrid electric tracked vehicles. *IEEE Transactions on Transportation Electrification*, 2022, 8(3): 3289-3306.
- [226] J Zhou, Y Xue, D Xu, et al. Self-learning energy management strategy for hybrid electric vehicle via curiosity-inspired asynchronous deep reinforcement learning. *Energy*, 2022, 242: 122548.
- [227] J Zhou, J Zhao, L Wang. An energy management strategy of power-split hybrid electric vehicles using reinforcement learning. *Mobile Information Systems*, 2022, <https://doi.org/10.1155/2022/9731828>.
- [228] M Acquarone, C Maino, D Misul, et al. Influence of the reward function on the selection of reinforcement learning agents for hybrid electric vehicles real-time control. *Energies*, 2023, 16(6): 2749.
- [229] L Bo, L Han, C Xiang, et al. A real-time energy management strategy for off-road hybrid electric vehicles based on the expected SARSA. *Proceedings of the Institution of Mechanical Engineers*, 2023, 237(2-3): 362-380.
- [230] L Guo, Z Li, R Outbib, et al. Function approximation reinforcement learning of energy management with the fuzzy REINFORCE for fuel cell hybrid electric vehicles. *Energy and AI*, 2023, 13: 100246.
- [231] B Hu, Y Xiao, S Zhang, et al. A data-driven solution for energy management strategy of hybrid electric vehicles based on uncertainty-aware model-based offline reinforcement learning. *IEEE Transactions on Industrial Informatics*, 2022.
- [232] B Hu, S Zhang, B Liu. A hybrid algorithm combining data-driven and simulation-based reinforcement learning approaches to energy management of hybrid electric vehicles. *IEEE Transactions on Transportation Electrification*, 2023, <https://doi.org/10.1109/TTE.2023.3266734>.
- [233] D Hu, H Xie, K Song, et al. An apprenticeship-reinforcement learning scheme based on expert demonstrations for energy management strategy of hybrid electric vehicles. *Applied Energy*, 2023, 342: 121227.
- [234] Y Hu, H Xu, Z Jiang, et al. Supplementary learning control for energy management strategy of hybrid electric vehicles at scale. *IEEE Transactions on Vehicular Technology*, 2023.
- [235] M Hua, C Zhang, F Zhang, et al. Energy management of multi-mode plug-in hybrid electric vehicle using multi-agent deep reinforcement learning. 2023. arXiv preprint <https://arxiv.org/abs/2303.09658>
- [236] R Huang, H He. A novel data-driven energy management strategy for fuel cell hybrid electric bus based on improved twin delayed deep deterministic policy gradient algorithm. *International Journal of Hydrogen Energy*, 2023.
- [237] Y Huang, H Hu, J Tan, et al. Deep reinforcement learning based energy management strategy for range extend fuel cell hybrid electric vehicle. *Energy Conversion and Management*, 2023, 277: 116678.
- [238] K Li, C Jia, X Han, et al. A novel minimal-cost power allocation strategy for fuel cell hybrid buses based on deep reinforcement learning algorithms. *Sustainability*, 2023, 15(10): 7967.
- [239] Y Liu, Y Wu, X Wang, et al. Energy management for hybrid electric vehicles based on imitation reinforcement learning. *Energy*, 2023, 263: 125890.
- [240] Z Liu, Q Zhou, Y Li, et al. Safe deep reinforcement learning-based constrained optimal control scheme for HEV energy management. *IEEE Transactions on Transportation Electrification*, 2023.
- [241] A Mousa. Extended-deep Q-network: A functional reinforcement learning-based energy management strategy for plug-in hybrid electric vehicles. *Engineering Science and Technology*, 2023, 43: 101434.
- [242] J Ruan, C Wu, Z Liang, et al. The application of machine learning-based energy management strategy in a multi-mode plug-in hybrid electric vehicle, part II: Deep deterministic policy gradient algorithm design for electric mode. *Energy*, 2023, 269: 126792.

- [243] H Wang, Y Ye, J Zhang, et al. A comparative study of 13 deep reinforcement learning based energy management methods for a hybrid electric vehicle. *Energy*, 2023, 266: 126497.
- [244] C Wu, J Ruan, H Cui, et al. The application of machine learning based energy management strategy in multi-mode plug-in hybrid electric vehicle, part I: Twin delayed deep deterministic policy gradient algorithm design for hybrid mode. *Energy*, 2023, 262: 125084.
- [245] F Yan, J Wang, C Du, et al. Multi-objective energy management strategy for hybrid electric vehicles based on TD3 with non-parametric reward function. *Energies*, 2023, 16(1): 74.
- [246] N Yang, L Han, R Liu, et al. Multi-objective intelligent energy management for hybrid electric vehicles based on multi-agent reinforcement learning. *IEEE Transactions on Transportation Electrification*, 2023.
- [247] N Yang, L Han, X Zhou, et al. Online-learning adaptive energy management for hybrid electric vehicles in various driving scenarios based on Dyna framework. *IEEE Transactions on Transportation Electrification*, 2023.
- [248] Q Zhou, J Li, B Shuai, et al. Multi-step reinforcement learning for model-free predictive energy management of an electrified off-highway vehicle. *Applied Energy*, 2019, 255: 113755.
- [249] A Lahyani, R Abdelhedi, A C Ammari, et al. Reinforcement learning based adaptive power sharing of battery/supercapacitors hybrid storage in electric vehicles. *Energy Sources*, 2020: 1-22.
- [250] H Sun, Z Fu, F Tao, et al. Data-driven reinforcement-learning-based hierarchical energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles. *Journal of Power Sources*, 2020, 455: 227964.
- [251] Y Zhou, L Huang, X Sun, et al. A long-term energy management strategy for fuel cell electric vehicles using reinforcement learning. *Fuel Cells*, 2020, 20(6): 753-761.
- [252] W Li, H Cui, T Nemeth, et al. Deep reinforcement learning-based energy management of hybrid battery systems in electric vehicles. *Journal of Energy Storage*, 2021, 36: 102355.
- [253] J Wu, Z Wei, W Li, et al. Battery thermal-and health-constrained energy management for hybrid electric bus based on soft actor-critic DRL algorithm. *IEEE Transactions on Industrial Informatics*, 2020, 17(6): 3751-3761.
- [254] B Xu, J Shi, S Li, et al. Energy consumption and battery aging minimization using a Q-learning strategy for a battery/ultracapacitor electric vehicle. *Energy*, 2021, 229: 120705.
- [255] Z Yang, F Zhu, F Lin. Deep-reinforcement-learning-based energy management strategy for supercapacitor energy storage systems in urban rail transit. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 22(2): 1150-1160.
- [256] H Zhang, Q Fan, S Liu, et al. Hierarchical energy management strategy for plug-in hybrid electric powertrain integrated with dual-mode combustion engine. *Applied Energy*, 2021, 304: 117869.
- [257] Y Cheng, G Xu, Q Chen. Research on energy management strategy of electric vehicle hybrid system based on reinforcement learning. *Electronics*, 2022, 11(13): 1933.
- [258] K Deng, Y Liu, D Hai, et al. Deep reinforcement learning based energy management strategy of fuel cell hybrid railway vehicles considering fuel cell aging. *Energy Conversion and Management*, 2022, 251: 115030.
- [259] Z Fu, H Wang, F Tao, et al. Energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles using deep reinforcement learning with action trimming. *IEEE Transactions on Vehicular Technology*, 2022, 71(7): 7171-7185.
- [260] X Guo, X Yan, Z Chen, et al. Research on energy management strategy of heavy-duty fuel cell hybrid vehicles based on dueling-double-deep Q-network. *Energy*, 2022, 260: 125095.
- [261] L Han, K Yang, T Ma, et al. Battery life constrained real-time energy management strategy for hybrid electric vehicles based on reinforcement learning. *Energy*, 2022, 259: 124986.
- [262] I Haskara, B Hegde, C Chang. Reinforcement learning based EV energy management for integrated traction and cabin thermal management considering battery aging. *IFAC-PapersOnLine*, 2022, 55(24): 348-353.
- [263] H Hu, C Lu, J Tan, et al. Effective energy management strategy based on deep reinforcement learning for fuel cell hybrid vehicle considering multiple performance of integrated energy system. *International Journal of Energy Research*, 2022, 46(15): 24254-24272.
- [264] J Li, H Wang, H He, et al. Battery optimal sizing under a synergistic framework with DQN-based power managements for the fuel cell hybrid powertrain. *IEEE Transactions on Transportation Electrification*, 2021, 8(1): 36-47.
- [265] W Shi, Y Huangfu, L Xu, et al. Online energy management strategy considering fuel cell fault for multi-stack fuel cell hybrid vehicle based on multi-agent reinforcement learning. *Applied Energy*, 2022, 328: 120234.
- [266] X Tang, H Zhou, F Wang, et al. Longevity-conscious energy management strategy of fuel cell hybrid electric vehicle based on deep reinforcement learning. *Energy*, 2022, 238: 121593.
- [267] X Wang, R Wang, G Shu, et al. Energy management strategy for hybrid electric vehicle integrated with waste heat recovery system based on deep reinforcement learning. *Science China Technological Sciences*, 2022, 65(3): 713-725.
- [268] B Xiao, W Yang, J Wu, et al. Energy management strategy via maximum entropy reinforcement learning for an extended range logistics vehicle. *Energy*, 2022, 253: 124105.
- [269] B Xu, Q Zhou, J Shi, et al. Hierarchical Q-learning network for online simultaneous optimization of energy efficiency and battery life of the battery/ultracapacitor electric vehicle. *Journal of Energy Storage*, 2022, 46: 103925.
- [270] D Xu, Y Cui, J Ye, et al. A soft actor-critic-based energy management strategy for electric vehicles with hybrid energy storage systems. *Journal of Power Sources*, 2022, 524: 231099.
- [271] H Zhang, S Liu, N Lei, et al. Learning-based supervisory control of dual mode engine-based hybrid electric vehicle with reliance on multivariate trip information. *Energy Conversion and Management*, 2022, 257: 115450.
- [272] W Zhang, J Wang, Z Xu, et al. A generalized energy management framework for hybrid construction vehicles via model-based reinforcement learning. *Energy*, 2022, 260: 124849.
- [273] Y Zhang, C Zhang, R Fan, et al. Twin delayed deep deterministic policy gradient-based deep reinforcement learning for energy management of fuel cell vehicle integrating durability information of powertrain. *Energy Conversion and Management*, 2022, 274: 116454.
- [274] C Zheng, W Li, W Li, et al. A deep reinforcement learning-based energy management strategy for fuel cell hybrid buses. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2022, 9(3): 885-897.
- [275] C Zheng, D Zhang, Y Xiao, et al. Reinforcement learning-based energy management strategies of fuel cell hybrid vehicles with multi-objective control. *Journal of Power Sources*, 2022, 543: 231841.
- [276] S Ahmadian, M Tahmasbi, R Abedi. Q-learning based control for energy management of series-parallel hybrid vehicles with balanced fuel consumption and battery life. *Energy and AI*, 2023, 11: 100217.
- [277] W Chen, J Peng, J Chen, et al. Health-considered energy management strategy for fuel cell hybrid electric vehicle based on improved soft actor critic algorithm adopted with Beta policy. *Energy Conversion and Management*, 2023, 292: 117362.
- [278] H Cui, J Ruan, C Wu, et al. Advanced deep deterministic policy gradient based energy management strategy design for dual-motor four-wheel-drive electric vehicle. *Mechanism and Machine Theory*, 2023, 179: 105119.
- [279] L Deng, S Li, X Tang, et al. Battery thermal-and cabin comfort-aware collaborative energy management for plug-in fuel cell electric vehicles based on the soft actor-critic algorithm. *Energy Conversion and Management*, 2023, 283: 116889.
- [280] R Han, R Lian, H He, et al. Continuous reinforcement learning-based energy management strategy for hybrid electric-tracked vehicles. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 2021, 11(1): 19-31.
- [281] J Hong, T Zhang, Z Zhang, et al. Investigation of energy management strategy for a novel electric-hydraulic hybrid vehicle: Self-adaptive electric-hydraulic ratio. *Energy*, 2023, 278: 127582.
- [282] H Hu, W Yuan, M Su, et al. Optimizing fuel economy and durability of hybrid fuel cell electric vehicles using deep reinforcement learning-based energy management systems. *Energy Conversion and Management*, 2023, 291: 117288.
- [283] R Huang, H He, M Gao. Training-efficient and cost-optimal energy management for fuel cell hybrid electric bus based on a novel distributed deep reinforcement learning framework. *Applied Energy*, 2023, 346: 121358.

- [284] R Huang, H He, X Zhao, et al. Longevity-aware energy management for fuel cell hybrid electric bus based on a novel proximal policy optimization deep reinforcement learning framework. *Journal of Power Sources*, 2023, 561: 232717.
- [285] C Jia, K Li, H He, et al. Health-aware energy management strategy for fuel cell hybrid bus considering air-conditioning control based on TD3 algorithm. *Energy*, 2023, 283: 128462.
- [286] H Lu, F Tao, Z Fu, et al. Battery-degradation-involved energy management strategy based on deep reinforcement learning for fuel cell/battery/ultracapacitor hybrid electric vehicle. *Electric Power Systems Research*, 2023, 220: 109235.
- [287] F Tao, H Gong, Z Fu, et al. Terrain information-involved power allocation optimization for fuel cell/battery/ultracapacitor hybrid electric vehicles via an improved deep reinforcement learning. *Engineering Applications of Artificial Intelligence*, 2023, 125: 106685.
- [288] C Wang, R Liu, A Tang, et al. A reinforcement learning-based energy management strategy for a battery-ultracapacitor electric vehicle considering temperature effects. *International Journal of Circuit Theory and Applications*, 2023.
- [289] Z Wei, Y Ma, N Yang, et al. Reinforcement learning based power management integrating economic rotational speed of turboshaft engine and safety constraints of battery for hybrid electric power system. *Energy*, 2023, 263: 125752.
- [290] Y Ye, J Zhang, S Pilla, et al. Application of a new type of lithium-sulfur battery and reinforcement learning in plug-in hybrid electric vehicle energy management. *Journal of Energy Storage*, 2023, 59: 106546.
- [291] C Zhang, T Li, W Cui, et al. Proximal policy optimization based intelligent energy management for plug-in hybrid electric bus considering battery thermal characteristic. *World Electric Vehicle Journal*, 2023, 14(2): 47.
- [292] D Zhang, S Li, Z Deng, et al. Lithium-plating suppressed and deep deterministic policy gradient based energy management strategy. *IEEE Transactions on Transportation Electrification*, 2023.
- [293] Z Zhang, T Zhang, J Hong, et al. Energy management strategy of a novel parallel electric-hydraulic hybrid electric vehicle based on deep reinforcement learning and entropy evaluation. *Journal of Cleaner Production*, 2023, 403: 136800.
- [294] Z Zhang, T Zhang, J Hong, et al. Double deep Q-network guided energy management strategy of a novel electric-hydraulic hybrid electric vehicle. *Energy*, 2023, 269: 126858.
- [295] H Guo, G Wei, F Wang, et al. Self-learning enhanced energy management for plug-in hybrid electric bus with a target preview based SOC plan method. *IEEE Access*, 2019, 7: 103153-103166.
- [296] Y Li, H He, J Peng, et al. Deep reinforcement learning-based energy management for a series hybrid electric vehicle enabled by history cumulative trip information. *IEEE Transactions on Vehicular Technology*, 2019, 68(8): 7416-7430.
- [297] J Wu, Y Zou, X Zhang, et al. An online correction predictive EMS for a hybrid electric tracked vehicle based on dynamic programming and reinforcement learning. *IEEE Access*, 2019, 7: 98252-98266.
- [298] Y Wu, H Tan, J Peng, et al. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. *Applied Energy*, 2019, 247: 454-466.
- [299] Z Chen, C Mi, B Xia, et al. Energy management of power-split plug-in hybrid electric vehicles based on simulated annealing and Pontryagin's minimum principle. *Journal of Power Sources*, 2014, 272: 160-168.
- [300] T Liu, B Tian, Y Ai, et al. Parallel reinforcement learning-based energy efficiency improvement for a cyber-physical system. *IEEE/CAA Journal of Automatica Sinica*, 2020, 7(2): 617-626.
- [301] H Zhang, J Peng, H Tan, et al. Tackling SOC long-term dynamic for energy management of hybrid electric buses via adaptive policy optimization. *Applied Energy*, 2020, 269: 115031.
- [302] Q Zhang, K Wu, Y Shi. Route planning and power management for PHEVs with reinforcement learning. *IEEE Transactions on Vehicular Technology*, 2020, 69(5): 4751-4762.
- [303] J Cao, H He, D Wei. Intelligent SOC-consumption allocation of commercial plug-in hybrid electric vehicles in variable scenario. *Applied Energy*, 2021, 281: 115942.
- [304] H He, Y Wang, J Li, et al. An improved energy management strategy for hybrid electric vehicles integrating multistates of vehicle-traffic information. *IEEE Transactions on Transportation Electrification*, 2021, 7(3): 1161-1172.
- [305] W He, Y Huang. Real-time energy optimization of hybrid electric vehicle in connected environment based on deep reinforcement learning. *IFAC-PapersOnLine*, 2021, 54(10): 176-181.
- [306] B Hu, J Li. An edge computing framework for powertrain control system optimization of intelligent and connected vehicles based on curiosity-driven deep reinforcement learning. *IEEE Transactions on Industrial Electronics*, 2020, 68(8): 7652-7661.
- [307] J Li, X Wu, S Hu, et al. A deep reinforcement learning based energy management strategy for hybrid electric vehicles in connected traffic environment. *IFAC-PapersOnLine*, 2021, 54(10): 150-156.
- [308] W Li, H Cui, T Nemeth, et al. Cloud-based health-conscious energy management of hybrid battery systems in electric vehicles with deep reinforcement learning. *Applied Energy*, 2021, 293: 116977.
- [309] Y Wang, H Tan, Y Wu, et al. Hybrid electric vehicle energy management with computer vision and deep reinforcement learning. *IEEE Transactions on Industrial Informatics*, 2020, 17(6): 3857-3868.
- [310] C Chang, W Zhao, C Wang, et al. A novel energy management strategy integrating deep reinforcement learning and rule based on condition identification. *IEEE Transactions on Vehicular Technology*, 2022, 72(2): 1674-1688.
- [311] J Chen, H Shu, X Tang, et al. Deep reinforcement learning-based multi-objective control of hybrid power system combined with road recognition under time-varying environment. *Energy*, 2022, 239: 122123.
- [312] Z Fang, Z Chen, Q Yu, et al. Online power management strategy for plug-in hybrid electric vehicles based on deep reinforcement learning and driving cycle reconstruction. *Green Energy and Intelligent Transportation*, 2022, 1(2): 100016.
- [313] W Han, X Chu, S Shi, et al. Practical application-oriented energy management for a plug-in hybrid electric bus using a dynamic SOC design zone plan method. *Processes*, 2022, 10(6): 1080.
- [314] H He, R Huang, X Meng, et al. A novel hierarchical predictive energy management strategy for plug-in hybrid electric bus combined with deep deterministic policy gradient. *Journal of Energy Storage*, 2022, 52: 104787.
- [315] D Hu, Y Zhang. Deep reinforcement learning based on driver experience embedding for energy management strategies in hybrid electric vehicles. *Energy Technology*, 2022, 10(6): 2200123.
- [316] R Huang, H He, X Zhao, et al. Battery health-aware and naturalistic data-driven energy management for hybrid electric bus based on TD3 deep reinforcement learning algorithm. *Applied Energy*, 2022, 321: 119353.
- [317] D Kim, S Hong, S Cui, et al. Deep reinforcement learning-based real-time joint optimal power split for battery-ultracapacitor-fuel cell hybrid electric vehicles. *Electronics*, 2022, 11(12): 1850.
- [318] Y Lin, J McPhee, N L Azad. Co-optimization of on-ramp merging and plug-in hybrid electric vehicle power split using deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 2022, 71(7): 6958-6968.
- [319] X Tang, J Chen, K Yang, et al. Visual detection and deep reinforcement learning-based car following and energy management for hybrid electric vehicles. *IEEE Transactions on Transportation Electrification*, 2022, 8(2): 2501-2515.
- [320] X Tang, J Zhang, D Pi, et al. Battery health-aware and deep reinforcement learning-based energy management for naturalistic data-driven driving scenarios. *IEEE Transactions on Transportation Electrification*, 2021, 8(1): 948-964.
- [321] M Yan, G Li, M Li, et al. Hierarchical predictive energy management of fuel cell buses with launch control integrating traffic information. *Energy Conversion and Management*, 2022, 256: 115397.
- [322] D Yang, L Wang, K Yu, et al. A reinforcement learning-based energy management strategy for fuel cell hybrid vehicle considering real-time velocity prediction. *Energy Conversion and Management*, 2022, 274: 116453.
- [323] J Chen, S Li, K Yang, et al. Deep reinforcement learning-based integrated control of hybrid electric vehicles driven by lane-level high definition map. *IEEE Transactions on Transportation Electrification*, 2023.

- [324] Z Chen, S Wu, S Shen, et al. Co-optimization of velocity planning and energy management for autonomous plug-in hybrid electric vehicles in urban driving scenarios. *Energy*, 2023, 263: 126060.
- [325] N Cui, W Cui, Y Shi. Deep reinforcement learning based PHEV energy management with co-recognition for traffic condition and driving style. *IEEE Transactions on Intelligent Vehicles*, 2023, <https://doi.org/10.1109/TIV.2023.3235110>.
- [326] J Guo, J Wang, Q Xu, et al. Deep reinforcement learning-based hierarchical energy control strategy of a platoon of connected hybrid electric vehicles through cloud platform. *IEEE Transactions on Transportation Electrification*, 2023.
- [327] R Huang, H He. Naturalistic data-driven and emission reduction-conscious energy management for hybrid electric vehicle based on improved soft actor-critic algorithm. *Journal of Power Sources*, 2023, 559: 232648.
- [328] R Liu, C Wang, A Tang, et al. A twin delayed deep deterministic policy gradient-based energy management strategy for a battery-ultracapacitor electric vehicle considering driving condition recognition with learning vector quantization neural network. *Journal of Energy Storage*, 2023, 71: 108147.
- [329] X Liu, Y Wang, K Zhang, et al. Energy management strategy based on deep reinforcement learning and speed prediction for power-split hybrid electric vehicle with multidimensional continuous control. *Energy Technology*, 2023, 11(8): 2300231.
- [330] P Mei, H Karimi, H Xie, et al. A deep reinforcement learning approach to energy management control with connected information for hybrid electric vehicles. *Engineering Applications of Artificial Intelligence*, 2023, 123: 106239.
- [331] J Peng, W Chen, Y Fan, et al. Ecological driving framework of hybrid electric vehicle based on heterogeneous multi agent deep reinforcement learning. *IEEE Transactions on Transportation Electrification*, 2023, <https://doi.org/10.1109/TTE.2023.3278350>.
- [332] J Peng, Y Fan, G Yin, et al. Collaborative optimization of energy management strategy and adaptive cruise control based on deep reinforcement learning. *IEEE Transactions on Transportation Electrification*, 2022, 9(1): 34-44.
- [333] H Sun, F Tao, Z Fu, et al. Driving-behavior-aware optimal energy management strategy for multi-source fuel cell hybrid electric vehicles based on adaptive soft deep-reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(4): 4127-4146.
- [334] Y Wang, W Li, Z Liu, et al. An energy management strategy for hybrid energy storage system based on reinforcement learning. *World Electric Vehicle Journal*, 2023, 14(3): 57.
- [335] Y Wang, Y Wu, Y Tang, et al. Cooperative energy management and eco-driving of plug-in hybrid electric vehicle via multi-agent reinforcement learning. *Applied Energy*, 2023, 332: 120563.
- [336] X Wu, J Li, C Su, et al. A deep reinforcement learning based hierarchical eco-driving strategy for connected and automated HEVs. *IEEE Transactions on Vehicular Technology*, 2023.
- [337] N Yang, S Ruan, L Han, et al. Reinforcement learning-based real-time intelligent energy management for hybrid electric vehicles in a model predictive control framework. *Energy*, 2023, 270: 126971.
- [338] H Zhang, J Peng, H Dong, et al. Hierarchical reinforcement learning based energy management strategy of plug-in hybrid electric vehicle for ecological car-following process. *Applied Energy*, 2023, 333: 120599.
- [339] K Zhang, J Ruan, T Li, et al. The effects investigation of data-driven fitting cycle and deep deterministic policy gradient algorithm on energy management strategy of dual-motor electric bus. *Energy*, 2023, 269: 126760.
- [340] K Yang, B Li, W Shao, et al. Prediction failure risk-aware decision-making for autonomous vehicles on signalized intersections. *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [341] E Kaufmann, L Bauersfeld, A Loquercio, et al. Champion-level drone racing using deep reinforcement learning. *Nature*, 2023, 620(7976): 982-987.
- [342] K Yang, X Tang, J Li, et al. Uncertainties in onboard algorithms for autonomous vehicles: Challenges, mitigation, and perspectives. *IEEE Transactions on Intelligent Transportation Systems*, 2023.

Xiaolin Tang is currently a professor at *College of Mechanical and Vehicle Engineering, Chongqing University, China*. He received the Ph.D. degree in Mechanical Engineering from *Shanghai Jiao Tong University, China*, in 2015.

Jiaxin Chen is pursuing the Ph.D. degree at *College of Mechanical and Vehicle Engineering, Chongqing University, China*.

Ye Chen Qin is currently an associate professor at *School of Mechanical Engineering, Beijing Institute of Technology, China*. He received the Ph.D. degree from *Beijing Institute of Technology, China*, in 2016.

Teng Liu received the PhD degree from *Beijing Institute of Technology, China*, in 2017.

Kai Yang is pursuing the Ph.D. degree at *College of Mechanical and Vehicle Engineering, Chongqing University, China*.

Amir Khajepour is currently a Professor of mechanical and mechatronics engineering at the *University of Waterloo, Canada*.

Shen Li is currently a research associate at *Tsinghua University, China*. He received the Ph.D. degree from *University of Wisconsin-Madison, USA*, in 2018.