International Journal of Automotive Technology, Vol. 22, No. 5, pp. 1437–1452 (2021)
DOI 10.1007/s12239–021–0125–0

pISSN 1229–9138/ eISSN 1976–3832

A REVIEW OF OPTIMAL ENERGY MANAGEMENT STRATEGIES USING MACHINE LEARNING TECHNIQUES FOR HYBRID ELECTRIC VEHICLES

Changhee Song[1)], Kiyoung Kim[1)], Donghwan Sung[1)], Kyunghyun Kim[1)], Hyunjun Yang[1)], Heeyun Lee[1)], Gu Young Cho[2)] and Suk Won Cha[1,3)*]

[1)]Department of Mechanical Engineering, Seoul National University, Seoul 08826, Korea
[2)]Department of Mechanical Engineering, Dankook University, Gyeonggi 16890, Korea
[3)]Institute of Advanced Machines and Design, Seoul National University, Seoul 08826, Korea

(Received 15 February 2021; Revised 13 April 2021; Accepted 18 May 2021)

ABSTRACT–A hybrid electric vehicle (HEV) is defined as a vehicle that has two or more power sources, the hybrid electric vehicle is a representative eco-friendly vehicle because it can operate efficiently with each power source and requires only a small sized electric power source. However, it is not possible to develop high efficiency HEVs without an effective energy management system (EMS), a well-designed EMS is vital in HEVs because they need to manage two power sources. Motivated by this, there are continuing efforts being made to research and establish suitable energy management strategies in order to develop high efficiency HEVs. In the past, many energy management strategies for HEVs were developed based on optimal control theory. Recently, various kinds of machine learning technologies have been applied to HEV EMS development based on breakthroughs in the fields of machine learning and artificial intelligence (AI). Machine learning is a field of research that allows computers to perform arbitrary tasks guided by data rather than explicit programming. Machine learning can be classified into supervised learning, reinforcement learning (semi-supervised learning), and unsupervised learning depending on how the training data is structured. In this study, we look at cases and studies in which machine learning techniques from each category were used to develop HEV energy management strategies.

KEY WORDS : Hybrid electric vehicle, Energy management strategy, Machine learning, Artificial intelligence, Optimal control theory

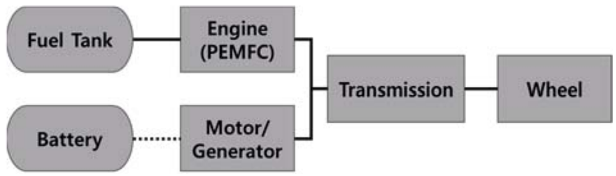
1. INTRODUCTION

Due to the dangers of global warming and climate change, emission regulations for vehicles are being stricter and stricter (Land Transport Guru, 2018). In accordance with these tightening regulations on vehicle emissions, there has been rapid progress made in vehicle electrification and hybridization (Wikipedia - Hybrid electric vehicle, 2020). Replacing internal combustion engine vehicles with HEVs is one of the most effective approaches to meeting the stringent emissions regulations (Plötz *et al.*, 2017).

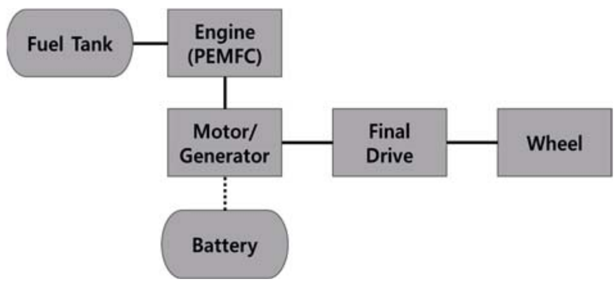
HEVs are defined as vehicles with more than one power source. In general, the power source of a hybrid vehicle consists of an internal combustion engine and a battery, HEVs can be further classified into parallel, series, and power-split architectures according to the powertrain structure, as shown in Figure 1 (Matsubara *et al.*, 2009; de Souza and Dedini, 2009; Ehsani *et al*., 2018). Figure 1 (a) shows a parallel HEV. Since parallel HEVs are designed to transmit the power of the electric motor and the internal combustion engine to the wheels, they have the advantage of being able to use smaller batteries (Montazeri-Gh *et al.*, 2006). Figure 1 (b) shows a series HEV. In series HEVs, the power required for propulsion of the vehicle is supplied only by the electric motor. Since series HEVs are powered only by the electric motor, they have the advantage of having mechanically simple systems (Dehghan *et al*., 2011). The structure of a power-split HEV is shown in Figure 1 (c). Power-split HEVs have the advantage that they can utilize both parallel and series configurations depending on the situation by using a planetary gear (PG) (Liu and Peng, 2008).

In recent years, along with the commercialization of fuel cell hybrid electric vehicles, studies on fuel cell hybrid electric vehicles are still being actively conducted (Xu *et al.*, 2020). The power sources of fuel cell hybrid electric vehicles (FCHEV) are generally composed of a fuel cell stack and battery. Like general HEVs, FCHEVs have Parallel, Series, and Power-Split configurations as shown in Figure 1. In FCEHV, PEMFC (Polymer Electrolyte Membrane Fuel Cell) plays the role of an internal combustion engine that is used as the main power source in general HEVs.

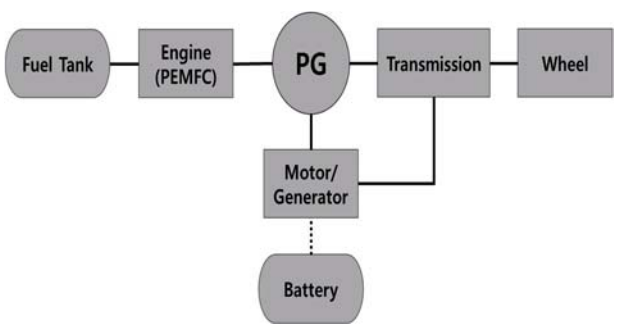
*****Corresponding author*. e-mail: swcha@snu.ac.kr



(a) Parallel configuration



(b) Series configuration



(c) Power-Split configuration

Figure 1. Hybrid electric vehicle configurations: (a) parallel configuration, (b) series configuration (c) power-split configuration.

Since the fuel cell stack of a FCHEV is vulnerable to stack degradation, it is necessary to consider deterioration of the fuel cell when developing the energy management strategy. Therefore, a multi-objective formulation of the problem that considers stack deterioration and fuel consumption of the fuel cell stack has been given in many studies on FCHEV energy management strategies (Mayur *et al.*, 2018; Li *et al.*, 2019a; Zheng *et al.*, 2014).

Since HEVs have more than a single power source, the efficiency of the overall system is greatly affected by the way the power required is distributed to each power source. For this reason, many studies have been conducted on energy management strategies for the HEVs.

A general energy management optimization problem for HEVs is defined in Equation (1). In Equation (1), $\dot{m}_{fc}$ is the instantaneous fuel consumption and L is a term that represents incurring economic costs such as the deterioration of parts. γ is the equivalent factor that equalizes L to the fuel consumption rate, and from the point of view of an optimization problem that considers only fuel consumption, γ becomes 0. Finally, x and u in Equation (1) represent the state and the control. The boundary conditions for Equation (1) are expressed in Equation (2), where $SOC(t_0)$ and $SOC(t_f)$ denote the initial state of charge (SOC) and final state of charge, respectively. A secondary condition of Equation (1) is expressed in Equation (3). In Equation (3), $P,\ \omega, T$ represent the power, rotational speed, and torque, respectively, while $(\cdot)_{eng}, (\cdot)_{bat}, (\cdot)_{mot}$ denote respective quantities relating to the engine, battery, and electric motor.

$$min \sum \dot{m}_{fc}\big(x(t), u(t)\big) + \gamma \cdot L\big(x(t), u(t)\big) \tag{1}$$

$$subject\ to$$

$$SOC\big(t_f\big) = SOC(t_0) \tag{2}$$

$$\begin{aligned} &P_{eng} \le P_{eng,max} \\ &P_{bat,min} \le P_{bat} \le P_{bat,max} \\ &\omega_{mot,min} \le \omega_{mot} \le \omega_{mot,max} \\ &T_{eng} \le T_{eng,max} \\ &\dots \end{aligned} \tag{3}$$

In other words, the goal of the HEV energy management optimization problem is to derive a strategy that minimizes the designed cost function while securing SOC-sustainability and satisfying all the physical constraints.

Many studies on energy management have produced strategies for HEVs that have been based on rule-based methods, optimal control theory, and reinforcement learning. In rule-based power distribution strategies, the driving modes are designed according to the driving conditions and are based on human experience or the results from optimal control theory (Son and Kim, 2016; Zhang and Tao, 2017; Geng *et al.*, 2019). Rule-based strategies have the advantage of being easily applicable to consumer vehicles because they require only a small amount of computation and utilize future information in a limited way when deriving the control values. However, the efficiency of the rule-base strategies is somewhat inferior to the results you might get through optimization. One of the most widely used theories in the study of HEV energy management is optimal control theory. The strategies based on optimal control theory can be divided into strategies based on dynamic programming that guarantee the global optimum solution and strategies based on real-time optimization theory. Strategies based on dynamic programming (DP) have the advantage that they can guarantee the global optimum solution, but have the disadvantage that they require heavy computations and rely on future driving information to derive the optimum control values (Pérez *et al*., 2006; Wang and Lukic, 2012). As such, DP-based strategies are mainly used to provide reference data for research or to derive the optimal size of the HEV elements as opposed to being directly applied in consumer vehicles (Vinot *et al.*, 2007; Pourabdollah *et al.*, 2014).

Among power distribution strategies that are based on real-time optimization theory, the most widely used theories are Pontryagin's minimum principle (PMP) and the equivalent consumption minimization strategy (ECMS) (Onori and Tribioli, 2015; Tang *et al*., 2015; Ou *et al*., 2018). The common goal of these two theories is to derive control values that minimize the cost function by considering the electrical energy and the fuel consumption at each time-step. Therefore, it is very important to design a co-state that equalizes electrical energy to fuel consumption properly for these two theories. Kim and Lee (2015) proved that when the co-state is designed properly, a PMP-based strategy can derive a result equivalent to that from a DP-based strategy that guarantees a global optimum solution. As such, a strategy based on real-time optimization theory requires only light computation and has the advantage that very high fuel efficiency can be expected if the co-state is designed properly. However, since the co-state has a large dependence on the driving cycle, it is very difficult to derive an appropriate co-state without future driving information.

Reinforcement learning is formulated based on a Markov decision process (MDP), as shown in Figure 2 (Sutton and Barto, 1992). MDP consists of an environment and an agent. The agent receives a state from the environment as input and delivers an action to the environment as output. The environment receives the action from the agent and then delivers an updated state and a reward to the agent. That is, the agent receives the state and delivers an action based on this, which is a kind of control value, to the environment according to the defined policy. In reinforcement learning, the agent constructs a policy in a way that maximizes the expected value of the reward based on previous experiences. When applying reinforcement learning framework to the HEV energy management problem, the environment corresponds to the HEV and the agent corresponds to the strategy. It is very important to define the state, the action, and the reward in reinforcement learning. In general, the state includes features such as the demand power, the velocity, SOC deviation, and torque. The action adjusts the battery power or the engine torque. The reward is designed by equalizing fuel consumption and electrical energy (Lin *et al*., 2014; Liu *et al*., 2017; Xiong *et al*., 2018). Strategies based on reinforcement learning have the advantage that generalization performance is very good because they derive actions only from observable states without the need for future information (Xiong *et al*., 2018).

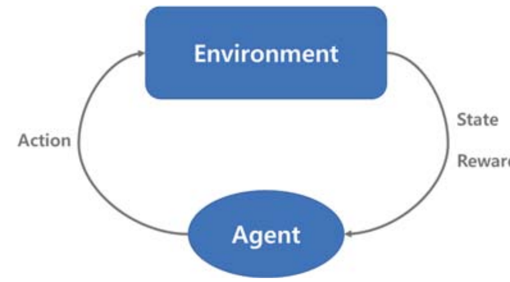


Figure 2. Illustration for reinforcement learning framework on the Markov decision process.

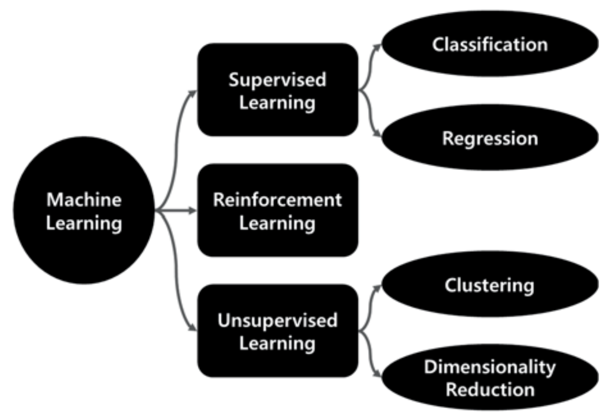


Figure 3. Classification of machine learning.

In recent years, due to rapid developments in the field of machine learning and AI, machine learning techniques are being actively utilized in the development of control strategies for HEVs. Machine learning is a research field that allows computers to learn to perform arbitrary tasks through the use of training data rather than explicit programming. Machine learning can be further classified into supervised learning, reinforcement learning (semi-supervised learning), and unsupervised learning according to the method used to organize the training data as shown in Figure 3.

In supervised learning, input data, expressed as a feature, and target data are paired to form the training data. A problem in which the dependent variable is expressed as categorical data is defined as a classification problem, and a problem in which the dependent variable is expressed as continuous data is defined as a regression problem. Supervised learning is the most widely used machine learning framework when developing power distribution strategies for HEVs. To carry out supervised learning, it is necessary to derive target data corresponding to the features. Optimal control theory is generally used in order to derive these labeled data for supervised learning-based HEV strategies. (Chia *et al.*, 2015; Lin *et al.*, 2015; Chen *et al*., 2018; Zheng *et al*., 2020).

Unsupervised learning consists of only input values without the target data. Representative algorithms for unsupervised learning include clustering and dimensionality reduction algorithms. Clustering refers to a machine learning technique that classifies data based on a similarity measure between features. Dimensional reduction algorithms are algorithms that reduce the number of feature dimensions and lower the co-linearity between features. Choi (2019) classified driving profiles using the k-means clustering algorithm and developed a power distribution strategy in which different co-state maps were applied according to how each driving profile was classified.

Reinforcement learning is also known as semi-supervised learning because the agent is trained to maximize the reward it receives through various experiences under conditions where the target data is not given explicitly. In recent years, the field of reinforcement learning has been rapidly developed by the fusion with deep learning, which is then called deep reinforcement learning (DRL). Taking advantage of the fact that deep learning can derive efficient feature representations for complex states or actions, DRL can effectively solve complex problems that cannot be solved using the existing reinforcement learning framework. (Mnih *et al*., 2015; Lillicrap *et al*., 2016). Since DRL derives power distribution strategies for HEVs using only observable states, many studies have used taken advantage of DRL to develop energy management strategies that are able to ensure generalization performance (Wang *et al*., 2019; Wu *et al*., 2018).

We summarized the general characteristics of the above three machine learning techniques, supervised learning, unsupervised learning, and reinforcement learning through Table 1. First, in terms of the necessity of pair-data composed of a pair of feature and target, supervised learning requires pair-data, and unsupervised learning only requires information on features. And since reinforcement learning proceeds based on state and reward derived from the environment, pair-data is not essential. Hyper-parameters are parameters that must be set by humans in a machine learning algorithm. In general, there are many hyperparameters in the order of reinforcement learning, supervised learning, and unsupervised learning. Reinforcement learning requires more hyper-parameters than the other two algorithms because reinforcement learning needs to define hyper-parameters such as exploit-exploration ratio, discount factor, and replay memory size as well as hyper-parameters for learning. Also, in reinforcement learning, because the agent is trained through various experiences and trial & error, reinforcement learning tends to require a large amount of computation compared to the other two algorithms. From a general point of view, reinforcement learning is more difficult to apply than supervised learning and unsupervised learning because both environment and agent must be configured, and appropriate action, state, and reward must be defined for effective learning. And, due to the cost and time consumed in configuring pair-data, supervised learning tends to be more difficult than unsupervised learning.

This rest of this paper is organized as follows. Section 2 introduces representative HEV structures and describes the characteristics of each HEV structure as well as related studies. In Section 3, we examine rule-based strategies and optimal control theory-based research in detail. Many machine learning-focused papers have used optimal control theory or rule-based strategies to generate training data or perform comparative studies, so looking at these studies is important to understand machine learning-based research. In Section 4, we look at various HEV strategies based on machine learning. Section 4 is divided into three parts according to the type of machine learning used. In the first part, we describe HEV power distribution strategies that use supervised learning. In the second and third parts, we discuss energy management strategies based on deep reinforcement learning and on unsupervised learning. Finally, we will summarize and conclude this study in section 5.

2. CONFIGURATION OF HYBRID ELECTRIC VEHICLES

There are several kinds of HEVs. There are series HEVs, parallel HEVs and power-split HEVs. Each structure has its own characteristics, advantages and disadvantages.

The serial type HEV structure was created to improve the range of EVs due to the low energy density of batteries.

HEVs with a parallel structure can be driven using both an engine and an electric motor, here the electric motor can compensate for disadvantages of the traditional internal combustion engine.

The power split HEV is a structure that adds one more electric motor and drives the vehicle using one engine and two motors. By properly controlling the planetary gear, often by using both a mechanical and an electrical couple, a series-type structure and a parallel-type structure can be

Table 1. Characteristics of three machine learning algorithms.

	Supervised learning	Unsupervised learning	Reinforcement learning
The necessity for pair-data	Need	Not needed	Not needed
Number of hyper-parameters	Medium	Small	Large
Computation time	Medium	Medium	Large
Difficulty level	Medium	Easy	Difficulty

implemented (Miller, 2006). Each structure has advantages and disadvantages, and their usages differ, but in overall point of view, it is a structure that supplies steady power to the internal combustion engine and supplements dynamic power through an electric motor (Ehsani *et al*., 2007).

2.1. Series Hybrid Electric Vehicle

In series HEVs, two power sources drive one electric motor. The first power source is an internal combustion engine. A generator is connected to an electrical coupler to supply energy in one direction to charge a battery. The other power source connected to the electrical coupler is the battery pack, energy is allowed to travel in both directions between the electric motor and battery. The electric motor is also connected to an electrical coupler, the motor not only receives energy to drive the vehicle, but it can also generate energy through regenerative braking. The biggest advantage of a series HEV is that there is no mechanical connection between the engine and the wheels, so the engine can operate independently of any power required to drive the wheels. Since the operating conditions that are efficient for an electric motor are not as sensitive as those for a combustion engine, there is no great need for a transmission.

However, series type HEVs have the disadvantage of suffering high losses due to frequent energy conversion. There is another disadvantage in that the weight and cost of the vehicle are increased by the generator and a high output electric motor is required.

Since serial HEVs require high electric motor output, they are not widely used as general passenger vehicles but are mainly used as commercial vehicles (Chambon *et al.*, 2017).

2.2. Parallel Hybrid Electric Vehicle

Similar to traditional internal combustion engine vehicles, in parallel HEVs, the engine drives the wheels through a mechanical coupler. The driving forces from the two power sources are combined by also having an electric motor mechanically coupled to the drive shaft to assist with extra driving force. Parallel type HEVs have the advantage not needing to convert power from the engine to electricity, in contrast to serial type HEVs, as such there are fewer losses in parallel HEVs. In addition, parallel HEVs do not require a generator and can use a smaller size electric motor than series HEVs because in series HEVs, the electric motor alone handles the driving power. However, since parallel type HEVs have complex mechanical structures and their engines are mechanically coupled to the drive shaft, they cannot choose to operate the engine only under the highest efficiency conditions, unlike in serial type HEVs (Gao *et al*., 1997).

There are two types of mechanical coupling: torque couplings and speed couplings. In torque couplings, the mechanical coupling combines the torque of the engine and electric motor and transmits it to the wheels. Therefore, the torque of the engine and electric motor can be applied independently. However, since the engine and the electric motor are coupled in a fixed ratio to the vehicle, the speed the wheels turn at is constrained. In the case of speed couplings, similarly, the engine and electric motor can rotate at different speeds but the speed coupler allows the two speeds to be combined and transmitted to the wheel, however, the torque supplied to the wheels is then constrained.

2.3. Power-Split Hybrid Electric Vehicle

It is also possible to have both a torque coupler and speed coupler in a parallel HEV and use them according to the situation. This gives more freedom in choosing how the engine and motor operate, which can lead to higher performance and efficiency. For example, while in the low speed range, the torque coupler can be used to achieve high acceleration (or climbing ability). Conversely, while in the high-speed range, the speed coupler can be used so the engine can drive in its optimum conditions. However, this method has the limitation that only one of the two methods can be selected. One way that Toyota has proposed to overcome this problem is to add another drive source. In their method one engine and two electric motors are connected and combined by a planetary gear and an inverter, this structure is called a power-split HEV or called a series/parallel type (Liu and Peng, 2008, Pei *et al.*, 2018). If the planetary gear is properly arranged, it can operate like a parallel HEV by mechanically connecting the engine and electric motor with the drive shaft, or it can be operated like a series HEV that generates power for the batteries using the engine and motor. The driving power is transmitted by the former through a mechanical path and by the latter through an electrical path. The transmission used in this type is called an e-CVT because it decouples the speed of the wheels and engine through electronic control and enables continuous transmission even when using a planetary gear (Hoeijimakes *et al.*, 2006).

Through use of a continuous transmission, higher efficiency and performance than that seen in normal series or parallel HEVs can be achieved, and the advantages of both systems can be obtained. The representative vehicle to which this system was applied is the Toyota Prius, it was later also applied to the Camry and Lexus (Hashimoto, 2009).

3. CONVENTIONAL ENERGY MANAGEMENT STRATEGY

Hybrid electric vehicles have two or more power sources and their efficiency varies depending on the operating conditions of each unit, as such, it is very important how much power is distributed between the power sources under the same driving conditions and vehicle power demands. Due to their importance, there is a great variety and depth of research into power distribution strategies for HEVs.

This chapter will cover the general concepts and the give details of research conducted into rule-based strategies and optimal control theory. In Section 3.1, rule-based strategies are introduced, in Section 3.2, optimal control theory-based strategies are introduced.

3.1. Rule-Based Energy Management Strategies

The rule-based energy management strategies determine control based on experience (Hofman *et al*., 2007). This does not produce optimal energy strategies because this approach can only reflect human intuition, accumulated data. Various automobile manufacturers have used rule-based energy management because it requires only light computation and is easy to apply in algorithms on consumer vehicles (Zhang *et al*., 2015; Wang *et al*., 2020a). Rule based control strategies can be largely divided into Deterministic Rule-based strategies and Fuzzy Rule-based strategies. Deterministic methods include Thermostatic control and the Power follower method, while Fuzzy methods can be subdivided into Conventional, Adaptive, and Predictive methods.

3.1.1. Thermostat strategy

The thermostat strategy, mainly used in serial hybrid vehicles, keeps the SOC within a specified range by turning the engine on or off in a predefined way (Shabbir and Evangelou, 2019). Since the engine is only operated when the operating conditions give the highest efficiency, the efficiency of the engine and generator is high, but we cannot expect high overall system efficiency due to energy conversion losses.

3.1.2. Power follower

The power follower strategy maintains the SOC and applies several control methods according to the state of the vehicle so that the engine can operate optimally (Shabbir and Evangelou, 2019). The vehicle operation mode is determined according to factors such as the SOC of the battery and the power required by the engine and clutch, the engine is controlled to operate in optimum conditions. This method is mainly used in parallel HEVs and power-split HEVs (Yang and Zhu, 2016).

3.1.3. Conventional fuzzy rule-based strategy

These methods control HEV by applying basic fuzzy logic. The input is fuzzified through a membership function and the output values are inferred based on rules. Control is performed by defuzzing the inferred value. In the inference process, human expertise and intuition are used but the optimization results cannot be guaranteed to give optimal efficiency. However, many studies have been conducted using these methods because they have the advantage that HEV energy management is possible with only light computation (Denis *et al*., 2015). In research related to the development of HEV power distribution strategies, fuzzy rule-based strategies have been designed in consideration of the vehicle's required power, vehicle speed, and SOC (Abdelsalam and Cui, 2012).

3.1.4. Adaptive fuzzy rule-based strategy

In this methodology, parameters applied in the existing fuzzy method are adaptively changed according to the situation. The optimization of parameters is performed using various methodologies, in some studies, an adaptive fuzzy logic controller has been derived through training a Neural Network and using PMP results to provide optimal control (Tian *et al*., 2017). In addition, some studies have improved the efficiency of HEVs by considering the driver's driving style, current driving mode, and external driving factors as input values for the fuzzy logic (Langari and Won, 2005).

3.2. Optimal Control Theory based Energy Management Strategies

3.2.1. Dynamic programming (DP)

DP is one of the methods that can find the global optimal value in a discrete problem by considering the constraints based on the Bellman equation. The optimum of the objective function in each step is calculated using the optimum value from the last step to the next step. Redundant calculations can be avoided because each step refers to previously calculated optimal values. The formula is expressed as shown in Equation (4), where J, x, u represent the designed cost function, state variable, and control variable, respectively.

$$C^*_{t,N}(x_t,u_t) = min\ [J_{t,t+1}(x_t,u_t) + J^*_{t+1,N}(x_{t+1})\ ]$$
$$J^*_{t,N} = \min_{u_t}[C^*_{t,N}(x_t,u_t)] \tag{4}$$

DP cannot be applied to real-time control of vehicles due to a large amount of computation, but it has been used in many studies developing the HEV's power distribution strategies because it can guarantee global optimization. The results of DP are generally used as a reference for newly developed energy management strategies (Lin *et al*., 2003). It is also used to analyze the DP results to create a rule that approximates the optimal value (Lee *et al*., 2017; Wang *et al*., 2018).

In order to supplement the low applicability of DP, a power distribution strategy for HEVs using stochastic dynamic programming (SDP) was developed. In SDP, a transition matrix that maps the probability from the current state to the next state is constructed through driving data of the actual vehicles. And a policy that can maximize the expected value of the objective function through the configured transition matrix is derived based on the Bellman equation (Leroy *et al*., 2014; Vagg *et al*., 2015; Lee *et al*., 2016). Unlike DP, SDP has the advantage that its applicability to real vehicles is very high because it derives control values from the current state.

3.2.2. Equivalent cost minimization strategy (ECMS)

The equivalent cost minimization strategy (ECMS) is a methodology to derive a control value that minimizes the designed cost function through equalization of the fossil fuel energy consumed by the engine with the electric energy consumed by batteries (Musardo *et al*., 2005). The cost function from a typical ECMS can be expressed as shown in Equation (5), where $\dot{m}_{fc}$ is the fuel consumption rate, $s\dot{o}c$ is the rate of change of the battery SOC, and λ is the equivalent factor that equalizes the battery SOC to the fuel consumption rate.

$$J = \dot{m}_{fc} + \lambda \cdot s\dot{o}c \tag{5}$$

The key part of ECMS is to estimate the optimal co-state that will minimize fuel consumption while maintaining SOC. In the early days of ECMS research, many studies simply estimated the equivalent factor of a constant value (Paganelli *et al*., 2001, 2002). However, it is difficult to accurately predict such an equivalent factor given the various driving environment experienced by the vehicle. Therefore, there have been research efforts to update the equivalent factor according to the current state variables of the vehicle (A-ECMS)(Razaei *et al*., 2017), or to predict the future driving situation and calculate the equivalent factor accordingly (Musardo *et al*., 2005; Zeng *et al*., 2018). In recent years, studies have looked to derive the equivalent factor for the ECMS using reinforcement learning that takes into account the various vehicle driving environments (Lee *et al*., 2020).

3.2.3. Pontryagin's minimum principle (PMP)

Pontryagin's Minimum Principle (PMP) is an optimization theory based on the Euler-Lagrange equation and aims to derive a control value that minimizes the system's Hamiltonian, which can be seen as a kind of cost function. For HEV energy management, the general Hamiltonian is constructed as shown in Equation (6) (Kim *et al*., 2010; Hou *et al*., 2014).

$$\mathrm{H} = \dot{m}_{fc}(P_{bat}(t), t) + p(t)S\dot{O}C(SOC(t), P_{bat}(t)) \tag{6}$$

PMP has the disadvantage that the control values it derives cannot guarantee being the global optimal solution because the instantaneous optimal control value is derived at each step. Recently, research has been conducted into increasing computation efficiency by applying an MPC framework based on PMP (Xie *et al*., 2019). In order to develop a real-time energy management system (EMS), a closed-form solution was derived using only some of the essential conditions from the PMP, this real-time control technique does not require a co-state variable (Nguyen *et al*., 2018). In addition, a technique for optimizing the fuel consumption and regenerative braking using PMP in the power system of the fuel cell hybrid system has been introduced (Li *et al*., 2019b).

3.2.4. Particle swarm optimization (PSO)

Particle swarm optimization (PSO) is a relatively recently developed optimization technique developed by Kennedy and Eberhart (1995). This is an algorithm developed under the inspiration of the social system. Multiple particles in the solution space are optimized through simulation results for themselves and peers. This PSO algorithm can be said to be similar to a genetic algorithm that optimizes a fitness function through crossover and mutation of chromosomes in population. PSO has the advantage that the algorithm is simple and there are few parameters to be considered in the development of the algorithm. Therefore, PSO has the advantage of being easy to apply to real-time HEV energy management (Wu *et al*., 2008a; Wu *et al*., 2008b; Hwang and Chen, 2020).

4. ENERGY MANAGEMENT STRATEGY USING MACHINE LEARNING FRAMEWORK

Due to advances in the field of machine learning, machine learning is being actively used to develop energy power distribution strategies for HEVs. Machine learning approaches can be divided into supervised learning, reinforcement learning, and unsupervised learning. In this study, HEV power distribution strategies based on machine learning are divided up according to these classifications. Section 4.1 will look at the research on HEV energy management strategies that use supervised learning, Section 4.2 will explain HEV energy management strategies that use reinforcement learning. In particular, Section 4.2 will focus on research into deep reinforcement learning that applies deep neural networks to the reinforcement learning framework. Finally, in Section 4.3, we will look at research into unsupervised learning that is used to develop HEV power distribution strategies.

4.1. Energy Management Strategies Based on the Supervised Learning Framework

Supervised learning is the machine learning technique that is most actively used to develop HEV energy management strategies. A typical framework for developing HEV energy management strategies through supervised learning is shown in Figure 4. HEV power distribution strategies are developed by generating target data corresponding to a feature to create training data for the supervised learning process, models such as neural networks, support vector machines, and random forest mode systems are then trained with this data.

Therefore, one of the most important elements of supervised learning is defining the target values that correspond to a feature. In many studies, the target has been defined as future driving information and energy management strategies have then been developed based on predicted driving information (Sun *et al*., 2014a; Xie *et al*., 2017; He *et al*., 2017; Sun *et*

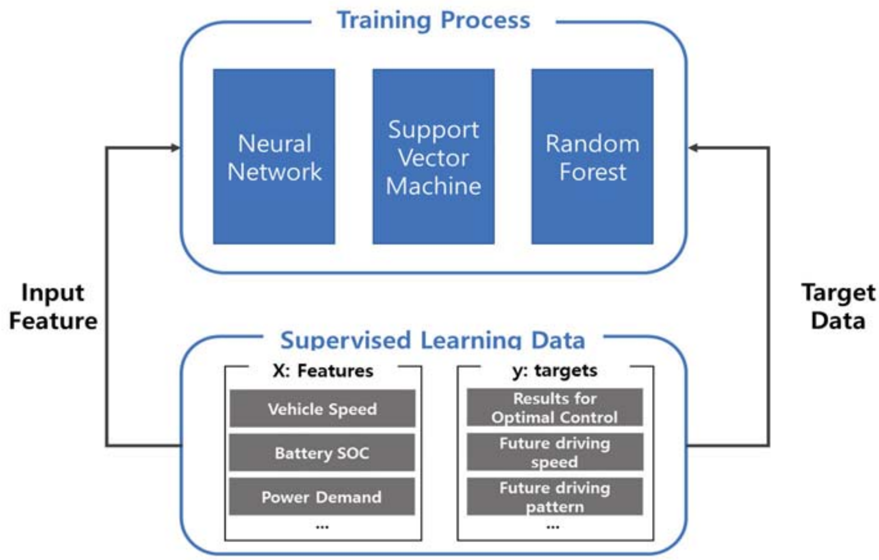


Figure 4. Framework for supervised learning algorithm in the energy management for HEV.

al., 2014b; Asadi and Vahidi 2010; Song *et al*., 2014; Lin *et al*., 2004; Langari and Won, 2005; Zhang and Deng, 2016; Park *et al*., 2009; Niu, 2015; Liu *et al*., 2019; Gaikwad *et al*., 2019).

This predictive model mainly carries out optimal control based on prediction results derived through machine learning algorithms. Xie *et al*. (2017) derived a probability distribution and a transfer matrix through a Markov decision process to predict the vehicle's velocity. In that study, the power distribution strategy was developed by formulating a model predictive control (MPC) algorithm that derives control values from the predicted speed through the transfer matrix and DP results for the prediction horizon. Sun *et al*. (2014a) constructed various model-based velocity predictors and compared the performance of each velocity predictor through the results of the MPC algorithm. In that study, an exponentially varying velocity predictor, a Markov- chain-based velocity predictor, and a neural network-based velocity predictor were developed. The neural network-based velocity predictor outperformed the other two velocity predictors. In particular, the radial basis function neural network (RBF-NN) proved to be superior to the recurrent neural network (RNN) that is efficient for time series problems and backpropagation neural network (BP-NN).

Another typical way to develop a supervised learning-based power distribution strategy is to construct a supervised learning data set through simulation results based on optimal control theory before using this data set train a model. Machine learning models that are trained on data generated based on optimal control theory have a great advantage in that they can compensate for the low generalization performance of the model through optimal control theory (Harold *et al*., 2020; Murphey *et al*., 2012; Venditti, 2016; Xie *et al*., 2018; Feldkamp *et al*., 2009; Han *et al*., 2020; Son *et al*., 2018; Zhuang *et al*., 2017; Chen *et al*., 2018).

Murphey *et al*. (2012) constructed an energy management system that determines battery power and engine rotational speed through a hierarchical neural network structure. In that paper, the output values of the neural network, which predict the traffic congestion level and driving trends, were used as the input features of the neural networks that determine the battery power and rotational speed of the engine. The training data for the neural networks that determine the battery power and engine speed was created from the simulation results of dynamic programming. Xie *et al*. (2018) used a neural network system to predict the equivalent factor that equalizes the fuel consumption and battery SOC change in the ECMS algorithm. It is very important to derive the optimal equivalent factor from the ECMS algorithm because the ECMS algorithm can produce high fuel efficiency that is at the same level as DP where the optimal equivalent factor is guaranteed. However, the optimal equivalent factor can only be derived when all driving information including future states are available. As the equivalent factor is sensitive to the driving cycle, it is very difficult to predict the optimal equivalent factor. In Xie's study, the training data was constructed so that the optimal equivalent factor, which is the target value, corresponds to the input features including the demanding power, the battery SOC, and the ratio of the distance traveled to the total distance. The neural network model trained using this training data predicts the optimal equivalent factor using the current power demands, the battery SOC, and the distance ratio. Zhuang *et al*. (2017) developed an SVM model that can predict the optimal operating mode by deriving the operating mode corresponding to vehicle speed and torque demands using DP simulation results.

Therefore, HEV energy management strategies based on

Table 2. Summary of EMS based on supervised learning algorithm.

Refs	Algorithm	Features	Targets	Problem Categories
Murphey *et al*. (2012)	Hierarchical NN	Battery SOC, vehicle speed, demanding power, output value of NN	Batter power Engine speed	Regression
Sun *et al*. (2014a)	RBF-NN	Recent vehicle speed sequence	Future vehicle Velocity sequence	Regression
Zhang *et al*. (2016)	NN	Driving characteristics	Driving pattern	Classification
Xie *et al*. (2017)	MDP	Neighboring state	Future vehicle velocity sequence	Regression
Zhuang *et al*. (2017)	SVM	Vehicle speed, driver's torque demand	Operating mode	Classification
Xie *et al*. (2018)	NN	Battery SOC, demanding power, distance ratio	Optimal co-state	Regression

supervised learning can be classified by their features and targets, as well as by the algorithms that map targets from features. Table 2 shows information on major papers organized by feature, target, and algorithm. Supervised learning is divided into regression problems and classification problems, EMSs for HEVs utilize both classification and regression approaches. In general, a regression algorithm is used for the prediction of speed, engine torque, and battery power, while the classification model is used for the prediction of driving patterns, operating mode, and gear ratio.

Supervised learning has been the most active research field in machine learning, this field is developing rapidly around tackling image recognition problems and natural language processing problems. Due to the rapid development of supervised learning, the group of HEV energy management strategies applying supervised learning algorithms is expected to continue to expand.

4.2. Energy Management Strategies using Reinforcement Learning

Reinforcement learning provides a framework to map actions that maximize the expected value of a reward from a given state without using explicit target data. Recently, deep neural networks have been introduced in the field of reinforcement learning, leading to rapid progress being made in this field (Mnih *et al*., 2015). Mnih *et al*. (2015) investigated using Q-learning, a model-free algorithm for reinforcement learning, for the EMS of HEVs. They proved that you can effectively solve problems that cannot be solved with existing Q-learning, due to the complexity of the states involved, by applying deep neural networks to Q-learning. The Q-values in Q-learning represent the quality of an action given the current state, the Q-value is expressed as shown in Equation (7). In Equation (7), π, s, a, and R represent the policy, state, action, and reward, respectively, while ρ represents a discount factor with a value between 0 and 1. Q-values can be optimized by a Bellman optimality equation such as that shown in Equation (8), the optimal policy can be expressed as shown in Equation (9).

$$Q_\pi = E_\pi[R_{t+1} + \rho R_{t+2} + \cdots \mid S_t = s,\ A_t = a] \tag{7}$$

$$Q \leftarrow Q + \alpha \cdot \left(r_t + \rho \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right) \tag{8}$$

$$\pi^*(s) = \operatorname*{argmax}_a Q^*(s, a) \tag{9}$$

In DRL, the Q-value is estimated using the deep neural network, this type of the DRL algorithm is known as a deep Q-network (DQN). In a DQN, the weights of the deep neural networks are trained to minimize the loss according to Equation (10) (Mnih *et al.*, 2015). The DQN algorithm consists of two networks, the main network and the target network, for learning stability. In Equation (10), θ^Q represents the weights of the main network, $\theta^{Q'}$ is the weights of the target network.

$$Loss(\theta^Q) = [y_t - Q(s, a|\theta^Q)]$$
$$\because y_t = r_{t+1} + \rho \max_{a_{t+1}} Q(S_{t+1}, a_{t+1}|\theta^{Q'}) \tag{10}$$

The biggest drawback of the DQN is that since the actions are derived through the value-evaluation algorithm, only discretized action can be derived. To overcome these shortcomings, Lillicrap *et al*. (2016) developed a deep deterministic policy gradient (DDPG) model that can derive continuous value-type actions using a policy gradient algorithm. The DDPG consists of neural networks with an actor-critic structure, where the actor network derives an action expressed by a continuous value, and the critic network evaluates the validity of the action derived by the actor network. As shown in Equation (11), the gradient for the loss of the actor network is derived by the policy gradient method, where θ^A and θ^c represent the weights of the actor network and the weights of the critic network, respectively.

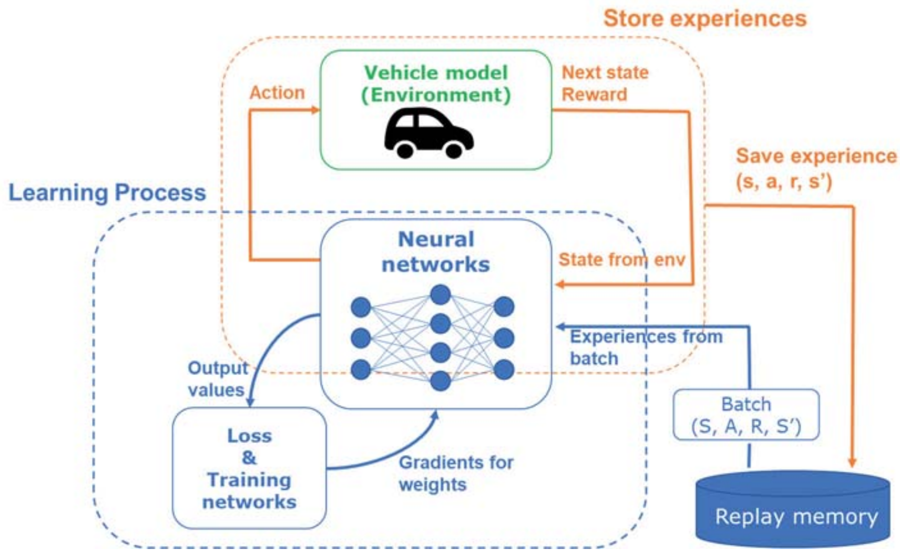


Figure 5. Schematic diagram of DRL algorithm in HEV power distribution strategy development.

$$\nabla_{\theta^A} J = \frac{1}{N}\sum \nabla_{\theta^\mu} Q(s, \mu(s|\theta^A)|\theta^C)$$
$$= \frac{1}{N}\sum \nabla_a Q(s, a|\theta^C)\nabla_{\theta^A}\mu(s|\theta^A) \quad (11)$$

The loss for the critical network is expressed in Equation (12), where $\theta^{A'}$ and $\theta^{C'}$ represent the weights of the target actor network and the weights of the target critic network, respectively.

$$Loss(\theta^C) = [y_t - Q(s, a|\theta^C)]$$
$$y_t = r_{t+1} + \rho Q(s_{t+1}, \mu(s_{t+1}|\theta^{A'})|\theta^{C'}) \quad (12)$$

The DQN algorithm and DDPG algorithm discussed above are the most widely used DRL algorithms in energy management strategies for HEVs. Figure 5 illustrates the framework in which the DRL algorithm is used in a domain related to the energy management of HEVs. On the MDP, the deep neural network corresponds to the agent and the vehicle model corresponds to the environment. As the simulation proceeds, the experience, consisting of the state, action, reward, and the next state are stored in the replay memory, which is a kind of database. The experience stored in the replay memory is then used to train the agent which is constructed by the deep neural networks.

Energy management strategies for HEVs that use DRL mainly uses either the DQN or DDPG algorithms (Zhao *et al*., 2018; Han *et al*., 2019; Zhu *et al*., 2020; Hu *et al*., 2018; Aljohani *et al*., 2021; Wang *et al*., 2019; Liessner *et al*., 2018; Wu *et al*., 2019; Tan *et al*., 2019; Wang *et al*., 2020b; Liessner *et al*., 2019; Lian *et al*., 2020; Guo *et al*., 2020).

Zhao *et al*. (2018) constructed a power distribution strategy for parallel HEVs using the DQN algorithm. The action was designed to control the battery current and gear ratio, the reward was defined as the negative of fuel consumption. The state for mapping the action was composed of the power demand, the vehicle speed, the battery charge, etc. and was given in a discrete value format. Liessner *et al*. (2018) constructed a power distribution strategy for parallel HEVs using the DDPG algorithm. In that study, the state was composed of continuous values of the wheel velocity, wheel torque, battery SOC, battery temperature, and gear ratio. The reward was designed in consideration of the fuel consumption and energy consumed by the battery, actions were defined with the power of the electric motor expressed as a continuous value.

An important criterion for classifying the DRL algorithm is its definition of state, action, and reward, as well as the algorithm used. According to these, we classified the research on DRL-based energy management mentioned above in Table 3.

Wu *et al*. (2019) established an energy management strategy for hybrid electric buses using the DDPG algorithm. In that study, a more efficient methodology was presented by utilizing not only vehicle information, but also considering the number of passengers and traffic information. In their research, Zhu *et al*. (2020) used the DQN algorithm. The reward was configured to minimize fuel consumption, and when excessive SOC deviation occurred, a SOC penalty was imposed on the agent guding the agent to secure SOC-sustainability. Lee and Cha (2020) constructed an energy management system that considers the engine on/off state using the Q-learning algorithm. Hu *et al*. (2018) constructed rewards according to a SOC range and compared DRL applied using an online-learning algorithm with a general offline DRL algorithm. Wang *et al*. (2020b) developed a

Table 3. Summary of EMS based on reinforcement learning framework.

Refs	Algorithm	States	Action	Reward
Hu *et al*. (2018)	DQN	Power demand, battery SOC	Engine torque	Fuel consumption, battery SOC
Liessner *et al*. (2018)	DDPG	Wheel speed, wheel torque, battery SOC, battery temperature, gear ratio	Motor power	Fuel consumption, electric energy
Zhao *et al*. (2018)	DQN	Required power, velocity, battery charge, predicted power demand	Battery discharging current, gear ratio	Fuel consumption
Lee and Cha (2020)	Q-learning	Battery SOC, demanding power, engine on/off state, vehicle speed	Engine torque	Fuel consumption, engine on/off state, battery SOC
Wu *et al*. (2019)	DDPG	Vehicle speed, acceleration, battery SOC, number of passengers, driving distance, speed at links	Engine torque, motor torque, engine rational speed	Cost for fuel, cost for electric energy
Wang *et al*. (2020b)	DDPG	Battery SOC, vehicle speed, vehicle acceleration, camera image	Engine power	Fuel consumption, battery SOC
Zhu *et al*. (2020)	DQN	Battery SOC, vehicle speed, acceleration, gear selection	Motor torque	Fuel consumption, SOC penalty

DDPG-based algorithm by extracting the number of vehicles and traffic light-related information from camera images.

DRL has shown that DRL agents can solve a variety of very complex problems beyond a human's ability, this part of the reason that DRL is currently one of the most popular research areas in the field of AI (Silver *et al*., 2017; Vinyals *et al*., 2019). Since DRL algorithms are developing at such a fast pace, power distribution strategies using DRL algorithms is an area that is expected to be very active and see lots of interesting developments.

4.3. Energy Management Strategies Using Unsupervised Learning

Unsupervised learning is not used as much as the supervised learning or reinforcement learning that were discussed above, instead unsupervised learning is mainly used to classify driving profiles through the use of clustering algorithms or used to reduce the dimensions of a set of features through dimensionality reduction algorithms.

Choi (2019) classified driving patterns through the k-means clustering algorithm and constructed an equivalent factor map for the ECMS for each driving pattern. Montazeri-Gh *et al*. (2011) created 21 features for driving cycles, they classified driving cycles through a clustering algorithm that looked at features. The classified driving patterns were then used to improve the efficiency of the energy management strategy. Finesso *et al*. (2016) used a clustering algorithm to group vehicle conditions based on the velocity, battery SOC, and velocity variation. Liu *et al*. (2019) used principal component analysis (PCA) to reduce feature space and classify features of driving conditions.

As of yet, unsupervised learning has not been actively used in establishing energy management strategies for HEVs. However, in the field of unsupervised learning, various generative models such as generative adversarial networks (GANs) have recently been developed (Goodfellow, 2016; Zhang *et al*., 2018). Since these generative models provide a methodology that can compensate for insufficient data, it is thought that generative models could provide a method to reduce the cost and time involved in collecting actual driving data. Due to the development of unsupervised learning, it is expected that unsupervised learning will be actively applied to EMSs for HEVs soon.

5. CONCLUSION

Hybrid electric vehicles are vehicles that have two or more power sources, they have the advantage that each power source can be controlled so it is used under efficient operating conditions. Therefore, HEVs are expected to have higher efficiency than internal combustion engine vehicles. However, in order to achieve such high efficiency, energy management strategies that can guarantee high efficiency must be developed. Motivated by this, many EMSs have been developed through many studies with the goal of increasing the efficiency of HEVs.

In recent years, the use of machine learning when developing these energy management strategies for HEVs has increased greatly due to the rapid developments in machine learning and AI techniques. Many researchers are making continuous efforts to overcome the low generalization performance of optimal control theory-based strategies and the low efficiency of rule-based strategies through the application of machine learning algorithms. In this study, we classified machine learning technology into three areas. We then examined how each machine learning technology is used in the field of EMS development for hybrid electric vehicles. Currently, HEV energy management strategies are being developed based on supervised learning and reinforcement learning, while unsupervised learning approaches have great potential for use in the energy management systems of HEVs.

Since continuous innovation is taking place in the field of machine learning and the AI, it is expected that new innovations in machine learning will be introduced to the development of energy management strategies for HEVs in the future. We hope that this study will be of great help in understanding how machine learning techniques have been used to establish energy management strategies for HEVs.

ACKNOWLEDGEMENT–This work was partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. NRF-2019R1G1A1100393).

This research was supported by the Hydrogen Energy Innovation Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government (Ministry of Science and ICT (MSIT)) (2019M3E6A1064695).

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2019R1A4A1025848).

REFERENCES

Abdelsalam, A. A. and Cui, S. (2012). A fuzzy logic global power management strategy for hybrid electric vehicles based on a permanent magnet electric variable transmission. *Energies* **5**, **4**, 1175–1198.

Aljohani, T. M., Ebrahim, A. and Mohammed, O. (2021). Real-time metadata-driven routing optimization for electric vehicle energy consumption minimization using deep reinforcement learning and Markov chain model. *Electric Power Systems Research*, **192**, 106962.

Asadi, B. and Vahidi, A. (2010). Predictive cruise control: Utilizing upcoming traffic signal information for improving fuel economy and reducing trip time. *IEEE Trans. Control Systems Technology* **19**, **3**, 707–714.

Chambon, P., Curran, S., Huff, S., Love, L., Post, B., Wagner, R., Jackson, R. and Green Jr, J. (2017). Development of a range-extended electric vehicle powertrain for an integrated energy systems research printed utility vehicle. *Applied Energy*, **191**, 99–110.

Chen, Z., Hu, H., Wu, Y., Xiao, R., Shen, J. and Liu, Y. (2018). Energy management for a power-split plug-in hybrid electric vehicle based on reinforcement learning. *Applied Sciences* **8**, **12**, 2494.

Chia, Y. Y., Lee, L. H., Shafiabady, N. and Isa, D. (2015). A load predictive energy management system for supercapacitor-battery hybrid energy storage system in solar application using the Support Vector Machine. *Applied Energy*, **137**, 588–602.

Choi, J. (2019). Energy management strategy for hybrid vehicles based on time series prediction model considering driving characteristics. Ph.D. Dissertation. Seoul National University, Seoul, Korea.

de Souza, R. B. and Dedini, F. G. (2009). Energy management strategy for hybrid electric vehicles. *SAE Technical Paper No*. 2009-36-0328.

Dehghan, S. M., Mohamadian, M., Rajaei, A. H. and Yazdian, A. (2011). Full modeling a series HEV based on BZS-NSI. *2nd Power Electronics, Drive Systems and Technologies Conf. (PEDSTC)*. Tehran, Iran.

Denis, N., Dubois, M. R. and Desrochers, A. (2015). Fuzzy-based blended control for the energy management of a parallel plug-in hybrid electric vehicle. *IET Intelligent Transport Systems* **9**, **1**, 30–37.

Ehsani, M., Gao, Y., Longo, S. and Ebrahimi, K. M. (2018). *Modern electric, hybrid electric, and fuel cell vehicles*. CRC press.

Ehsani, M., Gao, Y. and Miller, J. M. (2007). Hybrid electric vehicles: Architecture and motor drives. *Proc. IEEE* **95**, **4**, 719–728.

Feldkamp, L., Abou-Nasr, M. and Kolmanovsky, I. V. (2009). Recurrent neural network training for energy management of a mild hybrid electric vehicle with an ultra-capacitor. *IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems (CIVVS)*. Nashville, TN, USA.

Finesso, R., Spessa, E. and Venditti, M. (2016). An unsupervised machine-learning technique for the definition of a rule-based control strategy in a complex HEV. *SAE Int. J. Alternative Powertrains* **5**, **2**, 308–327.

Gaikwad, T. D., Asher, Z. D., Liu, K., Huang, M. and Kolmanovsky, I. (2019). Vehicle velocity prediction and energy management strategy part 2: Integration of machine learning vehicle velocity prediction with optimal energy management to improve fuel economy. *SAE Technical Paper No*. 2019-01-1212.

Gao, Y., Rahman, K. M. and Ehsani, M. (1997). Parametric design of the drive train of an electrically peaking hybrid (ELPH) vehicle. *SAE Technical Paper No*. 970294.

Geng, C., Jin, X. and Zhang, X. (2019). Simulation research on a novel control strategy for fuel cell extended-range vehicles. *Int. J. Hydrogen Energy* **44**, **1**, 408–420.

Goodfellow, I. (2016). Nips 2016 tutorial: Generative adversarial networks. *arXiv:* 1701.00160.

Guo, X., Liu, T., Tang, B., Tang, X., Zhang, J., Tan, W. and Jin, S. (2020). Transfer deep reinforcement learning-enabled energy management strategy for hybrid tracked vehicle. *IEEE Access*, **8**, 165837–165848.

Han, L., Jiao, X. and Zhang, Z. (2020). Recurrent neural network-based adaptive energy management control strategy of plug-in hybrid electric vehicles considering battery aging. *Energies* **13**, **1**, 202.

Han, X., He, H., Wu, J., Peng, J. and Li, Y. (2019). Energy management based on reinforcement learning with double deep Q-learning for a hybrid electric tracked vehicle. *Applied Energy*, **254**, 113708.

Harold, C. K., Prakash, S. and Hofman, T. (2020). Powertrain control for hybrid-electric vehicles using supervised machine learning. *Vehicles* **2**, **2**, 267–286.

Hashimoto, T., Yamaguchi, K., Matsubara, T., Yaguchi, H., Takaoka, T. and Jinno, K. (2009). Development of new hybrid system for compact class vehicles. *SAE Technical Paper No*. 2009-01-1332.

Hwang, H. Y. and Chen, J. S. (2020). Optimized fuel economy control of power-split hybrid electric vehicle with particle swarm optimization. *Energies* **13**, **9**, 2278.

He, H., Cao, J. and Peng, J. (2017). Online prediction with variable horizon for vehicle's future driving-cycle. *Energy Procedia*, **105**, 2348–2353.

Hoeijimakes, M. J. and Ferreira, J. A. (2006). The electrical variable transmission. *IEEE on Industry Application* **42**, **4**, 1092–1100.

Hofman, T., Steinbuch, M., Van Druten, R. and Serrarens, A. (2007). Rule-based energy management strategies for hybrid vehicles. *Int. J. Electric and Hybrid Vehicles* **1**, **1**, 71–94.

Hou, C., Ouyang, M., Xu, L. and Wang, H. (2014). Approximate Pontryagin's minimum principle applied to the energy management of plug-in hybrid electric vehicles. *Applied Energy*, **115**, 174–189.

Hu, Y., Li, W., Xu, K., Zahid, T., Qin, F. and Li, C. (2018). Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning. *Applied Sciences* **8**, **2**, 187.

Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. *Int. Conf. Neural Networks (ICNN)*. Perth, WA, Australia.

Kim, N., Cha, S. and Peng, H. (2010). Optimal control of hybrid electric vehicles based on Pontryagin's minimum principle. *IEEE Trans. Control Systems Technology* **19**, **5**, 1279–1287.

Kim, T. Y. and Lee, S. H. (2015). Combustion and emission characteristics of wood pyrolysis oil-butanol blended fuels in a DI diesel engine. *Int. J. Automotive Technology* **16**, **6**, 903–912.

Land Transport Guru (2018). http:/landtransportguru.net/European-emission-standards/

Langari, R. and Won, J. S. (2005). Intelligent energy management agent for a parallel hybrid vehicle-part I: system architecture and design of the driving situation identification process. *IEEE Trans. Vehicular Technology* **54**, **3**, 925–934.

Lee, H. and Cha, S. W. (2020). Reinforcement learning based on equivalent consumption minimization strategy for optimal control of hybrid electric vehicles. *IEEE Access*, **9**, 860–871.

Lee, H., Jeong, J., Park, Y. I. and Cha, S. W. (2017). Energy management strategy of hybrid electric vehicle using battery state of charge trajectory information. *Int. J. Precision Engineering and Manufacturing-Green Technology* **4**, **1**, 79–86.

Lee, H., Kang, C., Park, Y. I. and Cha, S. W. (2016). Study on power management strategy of HEV using dynamic programming. *World Electric Vehicle J*. **8**, **1**, 274–280.

Lee, H., Song, C., Kim, N. and Cha, S. W. (2020). Comparative analysis of energy management strategies for HEV: Dynamic programming and reinforcement learning. *IEEE Access*, **8**, 67112–67123.

Leroy, T., Vidal-Naquet, F. and Tona, P. (2014). Stochastic dynamic programming based energy management of HEV's: an experimental validation. *IFAC Proc. Volumes* **47**, **3**, 4813–4818.

Li, H., Ravey, A., N'Diaye, A. and Djerdir, A. (2019a). Online adaptive equivalent consumption minimization strategy for fuel cell hybrid electric vehicle considering power sources degradation. *Energy Conversion and Management*, **192**, 133–149.

Li, Q., Huang, W., Chen, W., Yan, Y., Shang, W. and Li, M. (2019b). Regenerative braking energy recovery strategy based on Pontryagin's minimum principle for fell cell/supercapacitor hybrid locomotive. *Int. J. Hydrogen Energy* **44**, **11**, 5454–5461.

Lin, C. C., Jeon, S., Peng, H. and Lee, J. M. (2004). Driving pattern recognition for control of hybrid electric trucks. *Vehicle System Dynamics* **42**, **1-2**, 41–58.

Lin, C. C., Peng, H., Grizzle, J. W. and Kang, J. M. (2003). Power management strategy for a parallel hybrid electric truck. *IEEE Trans. Control Systems Technology* **11**, **6**, 839–849.

Lian, R., Peng, J., Wu, Y., Tan, H. and Zhang, H. (2020). Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle. *Energy*, **197**, 117297.

Liessner, R., Schmitt, J., Dietermann, A. and Bäker, B. (2019). Hyperparameter optimization for deep reinforcement learning in vehicle energy management. *11th Int. Conf. Agents and Artificial Intelligence (ICAART)*. Prague, Czech Repulic.

Liessner, R., Schroer, C., Dietermann, A. M. and Bäker, B. (2018). Deep reinforcement learning for advanced energy management of hybrid electric vehicles. *10th Int. Conf. Agents and Artificial Intelligence (ICAART)*. Funchal, Portugal.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D. (2016). Continuous control with deep reinforcement learning. *4th Int. Conf. Learning Representations (ICLR)*. San Juan, Puerto Rico.

Lin, X., Bogdan, P., Chang, N. and Pedram, M. (2015). Machine learning-based energy management in a hybrid electric vehicle to minimize total operating cost. *IEEE/ACM Int. Conf. Computer-Aided Design (ICCAD)*. Austin, TX, USA.

Lin, X., Wang, Y., Bogdan, P., Chang, N. and Pedram, M. (2014). Reinforcement learning based power management for hybrid electric vehicles. *IEEE/ACM Int. Conf. Computer-Aided Design (ICCAD)*. San Jose, CA, USA.

Liu, J. and Peng, H. (2008). Modeling and control of a power-split hybrid vehicle. *IEEE Trans. Control Systems Technology* **16**, **6**, 1242–1251.

Liu, K., Asher, Z., Gong, X., Huang, M. and Kolmanovsky, I. (2019). Vehicle velocity prediction and energy management strategy part 1: Deterministic and stochastic vehicle velocity prediction using machine learning. *SAE Technical Paper No*. 2019-01-1051.

Liu, T., Hu, X., Li, S. E. and Cao, D. (2017). Reinforcement learning optimized look-ahead energy management of a parallel hybrid electric vehicle. *IEEE/ASME Trans. Mechatronics* **22**, **4**, 1497–1507.

Liu, T., Yu, H., Guo, H., Qin, Y. and Zou, Y. (2018). Online energy management for multimode plug-in hybrid electric vehicles. *IEEE Trans. Industrial Informatics* **15**, **7**, 4352–4361.

Matsubara, T., Yaguchi, H., Takaoka, T. and Jinno, K. (2009). Development of new hybrid system for compact class vehicles. *SAE Technical Paper No*. 2009-01-1332.

Mayur, M., Gerard, M., Schott, P. and Bessler, W. G. (2018). Lifetime prediction of a polymer electrolyte membrane fuel cell under automotive load cycling using a physically-based catalyst degradation model. *Energies* **11**, **8**, 2054.

Miller, J. M. (2006). Hybrid electric vehicle propulsion system architectures of the e-CVT type. *IEEE Trans. Power Electronics*, **21**, **3**, 756–767.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature* **518**, **7540**, 529–533.

Montazeri-Gh, M., Fotouhi, A. and Naderpour, A. (2011). Driving patterns clustering based on driving features analysis. *Proc. Institution of Mechanical Engineers, Part C: J. Mechanical Engineering Science* **225**, **6**, 1301–1317.

Montazeri-Gh, M., Poursamad, A. and Ghalichi, B. (2006). Application of genetic algorithm for optimization of control strategy in parallel hybrid electric vehicles. *J. Franklin Institute* **343**, **4-5**, 420–435.

Murphey, Y. L., Park, J., Kiliaris, L., Kuang, M. L., Masrur, M. A., Phillips, A. M. and Wang, Q. (2012). Intelligent hybrid vehicle power control—Part II: Online intelligent energy management. *IEEE Trans. Vehicular Technology* **62**, **1**, 69–79.

Musardo, C., Rizzoni, G., Guezennec, Y. and Staccia, B. (2005). A-ECMS: An adaptive algorithm for hybrid electric vehicle energy management. *European J. Control* **11**, **4-5**, 509–524.

Nguyen, B. H., German, R., Trovão, J. P. F. and Bouscayrol, A. (2018). Real-time energy management of battery/supercapacitor electric vehicles based on an adaptation of Pontryagin's minimum principle. *IEEE Trans. Vehicular Technology* **68**, **1**, 203–212.

Niu, L. (2015). Intelligent HEV fuzzy logic control strategy based on identification and prediction of drive cycle and driving trend. *World J. Engineering and Technology* **3**, **03**, 215.

Onori, S. and Tribioli, L. (2015). Adaptive Pontryagin's Minimum Principle supervisory controller design for the plug-in hybrid GM Chevrolet Volt. *Applied Energy*, **147**, 224–234.

Ou, K., Yuan, W. W., Choi, M., Yang, S., Jung, S. and Kim, Y. B. (2018). Optimized power management based on adaptive-PMP algorithm for a stationary PEM fuel cell/battery hybrid system. *Int. J. Hydrogen Energy* **43**, **32**, 15433–15444.

Paganelli, G., Delprat, S., Guerra, T. M., Rimaux, J. and Santin, J. J. (2002). Equivalent consumption minimization strategy for parallel hybrid powertrains. *IEEE 55th Vehicular Technology Conf. (VTC Spring 2002)*. Birmingham, AL, USA.

Paganelli, G., Tateno, M., Brahma, A., Rizzoni, G. and Guezennec, Y. (2001). Control development for a hybrid-electric sport-utility vehicle: Strategy, implementation and field test results. *Proc. American Control Conf. (ACC 2001)*. Arlington, VA, USA.

Park, J., Chen, Z., Kiliaris, L., Kuang, M. L., Masrur, M. A., Phillips, A. M. and Murphey Y. L. (2009). Intelligent vehicle power control based on machine learning of optimal control parameters and prediction of road type and traffic congestion. *IEEE Trans. Vehicular Technology* **58**, **9**, 4741–4756.

Pei, H., Hu, X., Yang, Y., Tang, X., Hou, C. and Cao, D. (2018). Configuration optimization for improving fuel efficiency of power split hybrid powertrains with a single planetary gear. *Applied Energy*, **214**, 103–116.

Pérez, L. V., Bossio, G. R., Moitre, D. and García, G. O. (2006). Optimization of power management in a hybrid electric vehicle using dynamic programming. *Mathematics and Computers in Simulation* **73**, **1-4**. 244–254.

Plötz, P., Funke, S. A., Jochem, P. and Wietschel M. (2017) CO_2 mitigation potential of plug-in hybrid electric vehicles larger than expected. *Scientific Reports* **7**, **1**, 1–6.

Pourabdollah, M., Murgovski, N., Grauers, A. and Egardt, B. (2014). An iterative dynamic programming/convex optimization procedure for optimal sizing and energy management of PHEVs. *IFAC Proc. Volumes* **47**, **3**, 6606–6611.

Rezaei, A., Burl, J. B. and Zhou, B. (2017). Estimation of the ECMS equivalent factor bounds for hybrid electric vehicles. *IEEE Trans. Control Systems Technology* **26**, **6**, 2198–2205.

Shabbir, W. and Evangelou, S. A (2019). Threshold-changing control strategy for series hybrid electric vehicles. *Applied Energy*, **235**, 761–775.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T. and Hassabis, D. (2017) Mastering the game of Go without human knowledge. *Nature* **550**, **7676**, 354–359.

Son, H. and Kim, H. (2016). Development of near optimal rule-based control for plug-in hybrid electric vehicles taking into account drivetrain component losses. *Energies* **9**, **6**, 420.

Son, H., Kim, H., Hwang, S. and Kim, H. (2018). Development of an advanced rule-based control strategy for a PHEV using machine learning. *Energies* **11**, **1**, 89.

Song, W., Zhang, X., Tian, Y., Zhang, X. and Song J. (2014). The research on vehicle driving pattern characteristic parameters search algorithm based on parallel computing. *Proc. 33rd Chinese Control Conf. (CCC 2014)*. Nanjing, China.

Sun, C., Hu, X., Moura, S. J. and Sun, F. (2014a). Velocity predictors for predictive energy management in hybrid electric vehicles. *IEEE Transactions on Control Systems Technology* **23**, **3**, 1197–1204.

Sun, C., Moura, S. J., Hu, X., Hedrick, J. K. and Sun F. (2014b). Dynamic traffic feedback data enabled energy management in plug-in hybrid electric vehicles. *IEEE Trans. Control Systems Technology* **23**, **3**, 1075–1086.

Sutton, R. S. and Barto, A. G. (1992). *Reinforcement learning: An introduction*. 2nd edn. The MIT Press. Cambridge, MA, USA; London, UK.

Tan, H., Zhang, H., Peng, J., Jiang, Z. and Wu Y. (2019). Energy management of hybrid electric bus based on deep reinforcement learning in continuous state and action space. *Energy Conversion and Management*, **195**, 548–560.

Tang, L., Rizzoni, G. and Onori, S. (2015). Energy management strategy for HEVs including battery life optimization. *IEEE Trans. Transportation Electrification* **1**, **3**, 211–222.

Tian, H., Wang, X., Lu, Z., Huang, Y. and Tian, G (2017). Adaptive fuzzy logic energy management strategy based on reasonable SOC reference curve for online control of plug-in hybrid electric city bus. *IEEE Trans. Intelligent Transportation Systems* **19**, **5**, 1607–1617.

Vagg, C., Akehurst, S., Brace, C. J. and Ash, L. (2015). Stochastic dynamic programming in the real-world control of hybrid electric vehicles. *IEEE Trans. Control Systems Technology* **24**, **3**, 853–866.

Venditti, M. (2016). Analysis of the performance of different machine learning techniques for the definition of rule-based control strategies in a parallel HEV. *Energy Procedia*, **101**, 685–692.

Vinot, E., Trigui, R., Jeanneret, B., Scordia, J. and Badin F. (2007). HEVs comparison and components sizing using dynamic programming. *Proc. IEEE Vehicle Power and Propulsion Conference (VPPC)*. Arlington, TX, USA.

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., Choi, D. H., Powell, R., Ewalds, T., Georgiev, P., Oh, J., Horgan, D., Kroiss, M., Danihelka, I., Huang, A., Sifre, L., Cai, T., Agapiou, J. P., Jaderberg, M., Vezhnevets, A. S., Leblond, R., Pohlen, T., Dalibard, V., Budden, D., Sulsky, Y., Molloy, J., Paine, T. L., Gulcehre, C., Wang, Z., Pfaff, T., Wu, Y., Ring, R., Yogatama, D., Wünsch, D., McKinney, K., Smith, O., Schaul, T., Lillicrap, T., Kavukcuoglu, K., Hassabis, D., Apps, C. and Silver, D. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* **575**, **7782**, 350–354.

Wang, J., Wang, J., Wang, Q. and Zeng, X (2018). Control rules extraction and parameters optimization of energy management for bus series-parallel AMT hybrid powertrain. *J. Franklin Institute* **355**, **5**, 2283–2312.

Wang, P., Li, Y., Shekhar, S. and Northrop, W. F. (2019). Actor-critic based deep reinforcement learning framework for energy management of extended range electric delivery vehicles. *IEEE/ASME Int. Conf. Advanced Intelligent Mechatronics (AIM)*. Hong Kong, China.

Wang, R. and Lukic, S. M. (2012). Dynamic programming technique in hybrid electric vehicle optimization. *IEEE Int. Electric Vehicle Conf. (IEVC)*. Greenville, SC, USA.

Wang, W., Fang, M., Qu, F. and Wang, X. (2020a). Energy management research of the fourth generation Toyota Prius Prime. *IEEE 4th Information Technology, Networking, Electronic and Automation Control Conf. (ITNEC)*. 2012–2016.

Wang, Y., Tan, H., Wu, Y. and Peng, J. (2020b). Hybrid electric vehicle energy management with computer vision and deep reinforcement learning. *IEEE Trans. Industrial Informatics* **17**, **6**, 3857–3868.

Wikipedia - Electric vehicle (2020). https://en.wikipedia.org/wiki/Electric_vehicle

Wikipedia - Hybrid electric vehicle (2020). https://en.wikipedia.org/wiki/Hybrid_electric_vehicle#cite_note-Sales2012US-114

Wu, J., He, H., Peng, J., Li, Y. and Li, Z. (2018). Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Applied Energy*, **222**, 799–811.

Wu, J., Zhang, C. H. and Cui, N. X. (2008a). PSO algorithm-based parameter optimization for HEV powertrain and its control strategy. *Int. J. Automotive Technology* **9**, **1**, 53–69.

Wu, X., Cao, B., Wen J., and Bian Y. (2008b). Particle swarm optimization for plug-in hybrid electric vehicle control strategy parameter. *IEEE Vehicle Power and Propulsion Conf. (VPPC)*. Harbin, China.

Wu, Y., Tan, H., Peng, J., Zhang, H. and He, H. (2019). Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. *Applied Energy*, **247**, 454–466.

Xie, S., He, H. and Peng, J. (2017). An energy management strategy based on stochastic model predictive control for plug-in hybrid electric buses. *Applied Energy*, **196**, 279–288.

Xie, S., Hu, X., Qi, S. and Lang, K. (2018). An artificial neural network-enhanced energy management strategy for plug-in hybrid electric vehicles. *Energy*, **163**, 837–848.

Xie, S., Hu, X., Xin, Z. and Brighton, J. (2019). Pontryagin's minimum principle based model predictive control of energy management for a plug-in hybrid electric bus. *Applied Energy* **236**, 893–905.

Xiong, R., Cao, J. and Yu, Q. (2018). Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. *Applied Energy*, **211**, 538–548.

Xu, J., Zhang, C., Fan, R., Bao, H., Wang, Y., Huang, S., Chin, C. S. and Li, C. (2020). Modelling and control of vehicle integrated thermal management system of PEM fuel cell vehicle. *Energy*, **199**, 117495.

Yang, J. and Zhu, G. (2016). Stochastic predictive boundary management for a hybrid powertrain. *IEEE Trans. Vehicular Technology* **65**, **6**, 4700–4713.

Zhang, H., Goodfellow, I., Metaxas, D. and Odena, A. (2018). Self-attention generative adversarial networks. *arXiv*. 1805.08318.

Zhang, P., Yan, F. and Du, C. (2015). A comprehensive analysis of energy management strategies for hybrid electric vehicles based on bibliometrics. *Renewable and Sustainable Energy Reviews*, **48**, 88–104.

Zhang, Q. and Deng, W. (2016). An adaptive energy management system for electric vehicles based on driving cycle identification and wavelet transform. *Energies* **9**, **5**, 341.

Zhang, R. and Tao, J. (2017). GA-based fuzzy energy management system for FC/SC-powered HEV considering H_2 consumption and load variation. *IEEE Trans. Fuzzy Systems* **26**, **4**, 1833–1843.

Zhao, P., Wang, Y., Chang, N., Zhu, Q. and Lin, X. (2018). A deep reinforcement learning framework for optimizing fuel economy of hybrid electric vehicles. *Proc. Asia and South Pacific Design Automation Conf. (ASP-DAC)*. Jeju, Korea.

Zheng, C. H., Xu, G. Q., Park, Y. I., Lim, W. S. and Cha, S. W. (2014). Prolonging fuel cell stack lifetime based on Pontryagin's Minimum Principle in fuel cell hybrid vehicles and its economic influence evaluation. *J. Power Sources*, **248**, 533–544.

Zheng, Y., He, F., Shen, X. and Jiang, X. (2020). Energy control strategy of fuel cell hybrid electric vehicle based on working conditions identification by least square support vector machine. *Energies* **13**, **2**, 426.

Zhuang, W., Zhang, X., Li, D., Wang, L. and Yin, G. (2017). Mode shift map design and integrated energy management control of a multi-mode hybrid electric vehicle. *Applied Energy*, **204**, 476–488.

Zhu, Z., Liu, Y. and Canova, M. (2020). Energy management of hybrid electric vehicles via deep Q-networks. *Proc. American Control Conf. (ACC)*. Denver, CO, USA.

Zeng, Y., Cai, Y., Kou, G., Gao, W. and Qin, D. (2018). Energy management for plug-in hybrid electric vehicle based on adaptive simplified-ECMS. *Sustainability* **10**, **6**, 2060.