Review

A comparative study of the machine learning-based energy management system for hydrogen fuel cell electric vehicles

A K M Rubaiyat Reza Habib*

Department of Information Technology, Arkansas Tech University, 1811 N Boulder Ave, Russellville, AR, 72801, USA

A R T I C L E   I N F O

*Corresponding author
Email address:
rubaiyat.reza@gmail.com

A B S T R A C T

The demand for better energy technologies has sparked research and development of electric and hybrid vehicles. Due to their clean, sustainable, and high energy density, fuel cell vehicles have begun to stand out above the rest. Therefore, fuel cell hybrid vehicles can compete with internal combustion engine-powered vehicles in the future. However, fuel cells face obstacles, including slow dynamics that necessitate managing their operation together favorably. To reduce an HEV's operational costs, this study analyzes the HEV energy management issue utilizing machine learning approaches, particularly reinforcement learning. This paper aims to comprehensively review the existing work on a couple of machine-learning-based energy management systems for an electric vehicle run by hydrogen fuel cell, it can be concluded that progress was evident when the Q-learning-based algorithm was utilized towards lowering the SOC $_{battery}$ variation of about 0.7 per unit which was the primary task, while a Deep Deterministic Policy Gradient (DDPG) based energy management system (EMS) starts operating the fuel cell at a higher efficiency rate comparatively while using the battery.

1. Introduction

Diverse industrial areas have been acting internationally to lower CO_2 emissions. In response, the transportation and communications sectors concentrated on further electrifying what are referred to as new energy vehicles. This is especially true of on-road transportation, a significant pollution source [1]. In the prior studies [2,3], specifics about the difficulties, technology, and legislation related to these endeavors were covered. The accelerated electrification of light and heavy-duty vehicles is necessary to transition current powertrain technology to climate-neutral new energy automobiles quickly. Overall, electrification in the context of movement, automobiles, and logistics refers to the greater utilization of electric power and the incorporation of a wide range of blended technologies to aid in successful power generation and transmission in addition to leading to a sustainable vehicle movement. The entirety of the electric powered transportation systems used in the automotive industry [3], including battery electric vehicles (BEVs), fuel cell electric vehicles (FCEVs), and fuel cell hybrid electric vehicles (FCHEVs) [4,5], need a form of energy that is guided by certain power conversion, or transformation, first before frictional force arises. The origin of the energy source determines the battery-powered car's lifespan pollutants and its contribution to the reduction of rising temperatures. As a result, the mobility and transport sectors have been directed towards fusing electrification governed by clean energy sources [5]; therefore, a power-to-vehicle pathway was considered in the current study. This kind of global view enables a smooth linking among the energy, climate, and transport sectors. Including hydrogen generation, distribution, and utilization stages is a useful baseline for understanding the complex strategic interactions among sectors and various technologies while performing engineering studies. If the concept is solely applied to a power-to-vehicle chain assuming that the end-user vehicle is a fuel cell electric vehicle (FCEV or FCV), the landscape can be demonstrated as illustrated in Figure 1.

The primary form of energy for hydrogen is thought to come from renewable or alternatives like solar and wind. The electricity produced can be electrolyzed to produce the needed CO_2-neutral green hydrogen after processing using network technology. Additionally, the produced hydrogen can be condensed into a liquid form for storage or transformed chemically into certain fluid states [6]. The transportation of hydrogen is sought using techniques like flowlines or current natural gas connectivity, or it is forced to move in big trailers with maritime transportation before being inevitably conveyed in limited volumes as well as native trailers to the
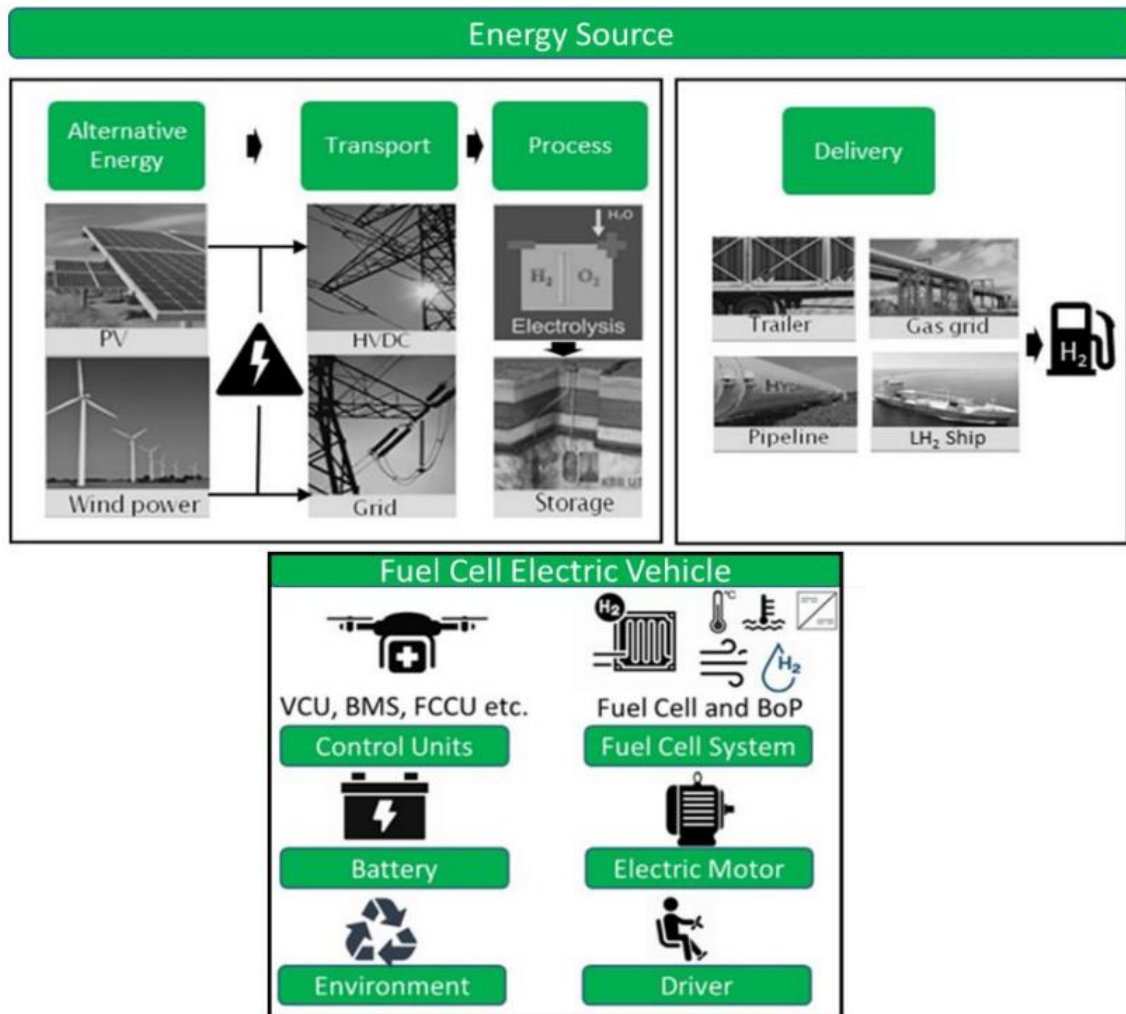
Figure 1. The transition from renewable energy to an electric fuel cell car

geographic recharging stations. A pressure valve lowers the hydrogen's captured pressure from strong to half-court prior to it being launched into the car, and a gas pressure decreases it even further to a level that fuel cells can handle. The complicated structure powering the electric fuel cell car is multi-physics dynamics. The motor, fuel cell system, battery, and the actual automobile are all subject to a variety of energy needs and model parameters because of various driving behaviors and ambient factors. Every control system transmits and gets data to produce a system that performs smoothly overall [7]. In complement to the piled devices, the fuel cell within the car also includes auxiliary modules, such as those shown in Figure 2.
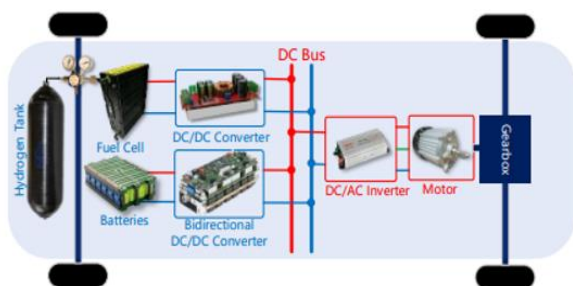

Figure 2. The standard fuel cell electric vehicle's layout

With the advent of artificial intelligence (AI), the need for quickly running algorithms that convert digital data has also advanced to a critical phase. The revolution of multi-physics modeling is developing because of advanced technological AI [8]. The large customer skill necessary for conventional approaches like the Finite Element Method (FEM) or Computational Fluid Dynamics (CFD) was a limitation, and it was also challenging to arrive at profitable remedies without establishing the knowledge acquired. But since complicated multi-physics investigations' enormous amount of high-end data and the updated AI skills have worked together, a more potent type of expedited forecasting and optimization capacity has been made available. The abstracted digitization process of the real issue turns bulk data sets that may be saved and analyzed relatively easily. Emulating the findings allows for other projection possibilities to be enabled. A significant shift in the simulation and modeling mentality was achieved via understanding the new data and enhancing the techniques built into AI-based devices. Aspects of AI have been employed in a variety of fields, including power, communication, healthcare, and several other scientific knowledge [9].

2.   Summary of fuel cell–battery HEV

In the last five years, there has been an average rise in solar PV-based energy output of 27% and a 13% growth in wind energy [10-11]. Renewable energy sources (RES) are complex, unpredictable, and limited in capacity because of weather conditions. These features contribute to several problems and difficulties in traditional energy systems, including significant active and reactive power losses, voltage profile equalization, and network dependability [12-13]. An effective energy management system (EMS) can be put together for hybrid-electric propulsion systems using rule-based, improvement, and knowledge-based methodologies [14]. A rule-based EMS was promoted by Hofman et al. [15] to restrict similar consumptions for hybrid vehicles. A mixed rule-based EMS for a module hybrid electric vehicle was suggested by Padmarajan et al. [16]. Peng et al. [17] presented an EMS adjustable by robust scripting to implement perfectly for a typical load characteristic. Although developing a rule-based EMS is reasonably straightforward and may be done with ease to operate an electric hybrid network's internet microcontroller, the technique has a limited scope, which makes it challenging to develop the criteria for immensely complicated data sets. Disengaged and virtual techniques are used to uptick EMS; detached streamlining-based EMS typically wants data extrapolated from typical stress periods. They can communicate the appropriate layout inside the load patterns that are utilized to synchronize the EMS in this way. Regardless, detached streamlined EMS can serve as a standard to evaluate the practicality of alternative digital EMS, if knowledge of the overall load profile was known beforehand [18].

A method requiring the internet, such as Equivalent Consumption Minimization Strategies (ECMS) and Model Predictive Control (MPC), in contrast, does not require inferred details on the piled portfolio and is typically scheduled like an instant rate of success for operation with limited computational resources [19]. For a hybrid boat impulsion technology, Kalikatzarakis et al. [20] promoted an ECMS-based EMS with the aid of the web. According to their numerical simulations, ECMS can save an additional 6% on gasoline for a handful of parameters. An MPC-based EMS for a hybrid-electric-powered automobile was suggested by Wang et al. [21] and achieved a 6% increase in fuel efficiency overrule-based EMS for a few given load scenarios. EMS that requires the internet is related to growth, such as ECMS and MPC, which appears to achieve close-to-ideal completion when used on a particular set number of load categories. However, their actual prolonged endurance is still unclear, especially for instances like vessels with extremely high scholastic load profiles [22]. With a large amount of factual evidence, a model can be created for a training EMS that will allow for continuing forecasting of energy pleas in the future. Several additional unrelated methodologies, such as nonlinear optimization, that are computationally expensive for repeated procedures [22-23], can be used to assess complex logic options. Even if a process is acceptable, the suited regulation capacity obtained through vibrant authoring computer software is expensive and constrained by the complexity [24], rendering it inappropriate for entities with highly multilayered or theoretically continuous phase and activity fields. It is noteworthy to observe that a Reinforcement Learning (RL) expert acquires a perfect or almost perfect power conduct approach by persistently interacting in accordance with the conditions. To address the problem of competence beyond evaluations in highly educated environments, Wu et al. [24] developed their RL expert using a broad range of reliable observational data. While the vector representation is still defined, Wu et al. [25] extended the subspace to be endless. The fact that each of the RL-based EMS clearly stated was designed to manage a unique source of power is crucial. It is clear from the analysis above that equally conventional and cutting-edge knowledge-based EMS can deliver optimum or almost optimal energy management performance for hybrid-electric systems with simple load levels. However, some operations, like sends, demand large levels of explicit repetition from inside their power units and use a variety of sources of energy; these applications typically have extremely intricate and variable demand profiles. An EMS still needs to be tested to see if it can provide concurrent control for devices, such as the ones found in ships, that manage multiple forms of energy operating amid extremely demanding workload patterns. Reddy et al. [26] suggested a PEMS for FCHEVs that fosters an aged Q-learning-based related PEMS that can autonomously get acquainted with the best technique continually by collaboration with reproduction designs of the existing hybrid energy system and teach the Q-learning estimation [27].

3.   Machine Learning (ML) and Reinforcement Learning (RL)

AI created by humans intends to connect humans and machines so that the combined level of intellect rises [28]. According to prevailing consensus, ML, an effective approach that uses computers to benefit from meticulously obtained and extant data, is one of the main man-made intelligence strategies currently in use. The study of ML is directly connected to how learners behave; as a result, it aims to understand and analyze data profoundly, equates, and reasons. The goal is to compile the methods used by the core components of human learning. To produce computational assemblages, reasoning, deduction, backpropagation, coding, observations, statistical inference, and building relationships, whereas a human developer is not required to create tasks when studying from a data framework, which is a benefit. Unlike traditional programming with immediate instruction and obtaining a solution, the relationship between the computer programmer and the computer is so strong that they operate separately. This is since the learning procedure of the machine is making decisions given the data presented, i.e., past, and is thereby mimicking rational decisions. By providing models that demonstrate how the system should function, these are implied. The ability to synthesize additional material is increased as the knowledge source becomes more complex, but more customer data is also needed. Thus, careful planning serves as the driving force behind the teaching method; the lifecycle is incredibly dependent on the supplied and shaped facts [29-30]. Reinforcement learning is the final and most advanced ML computation framework. A cluster of events, not a single response, is the outcome of a well-researched foundation. The applied strategy involves arranging clever actions to reach a particular goal. In the unlikely event that it is crucial to a purported goal, an act is proper. This kind of layout is produced by the system of education based on appropriate prior experiences. The RL strategy, in contrast to supervised and unsupervised instructional techniques, continuously improves its theory by gathering feedback from prior focus in a cycle. This prevents it from continuing indefinitely after the model has been created from data collected during preparation and testing, as shown in Figure 3. There are three characteristics that define RL. The first characteristic is how

exploration and exploitation are balanced and coordinated. After gathering information about the environment during the exploration phase, the agent moves on to the exploitation phase, which entails initiating control action based on knowledge already at hand. Through its encounters with the environment, the agent may determine the status of the surroundings and act accordingly to modify it if necessary. This capacity to adapt measures without the need for external control is crucial in situations when the environment is hazy or uncertain. The final characteristic is that it is Markovian, which means that the conditional probability distribution of potential environmental states relies only on the present state condition and not on the previous occurrences [31].
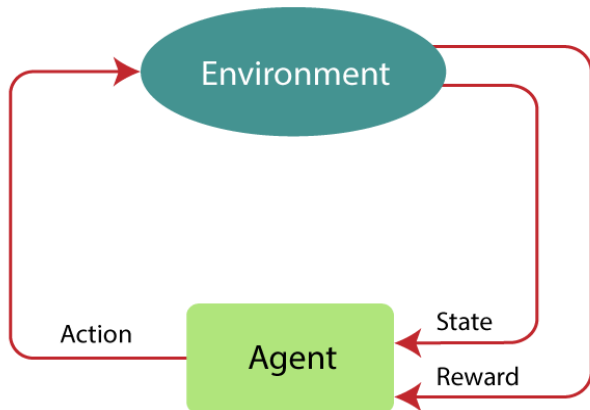

Figure 3. Reinforcement Learning principle

4. Energy Storage and Conversion Systems in FCHEV

As every automobile is devoted to moving for a certain amount of time with essentially no other energy resources, power capacity in transport is undoubtedly an important topic. In the past, people had energy stored as gasoline in gas tanks, which limited on how far they could go. The most prevalent energy converters now are combustion motors, which run on gasoline, diesel, or lamp oil. However, these automobiles use batteries more as a focus for decoration than as a source of power for locomotion. However, as the problem with these vehicles' emissions became more widespread, ignition engines started to use alternative sources of energy instead of the traditional oil-based fuels they previously used. However, the switch from internal combustion engines (ICE) to battery electric vehicles (BEV) has not been seamless because previous batteries' energy and power capacity did not allow for long-range driving and necessitated frequent recharging [32]. Optimizing several cost functions, including fuel consumption, battery life, emissions, and driving control, by balancing the distribution of power among various energy sources and the power source is the primary goal of an energy management system. This problem is typically constructed as a control optimization problem with clear accountability targets and limitations, including exhaust temperature, pollution, gas mileage, and battery state of charge (SOC). The energy management issue for a typical HEV is shown in Figure 4.

4.1 Battery

Battery technology developments have made it feasible for them to be used in both BEV and HEV in addition to PHEV. To replace Nickel Metal Hydride (NiMH), Nickel Cadmium (NiCd), and Lead Corrosive batteries, a type of battery known as lithium-ion battery has started to be used widely in a variety of electric cars. The three key advantages that lithium-ion batteries have over other types of batteries are their increased charge storing competency, ability to store large quantities of energy, and ineffective self-release [33]. As shown in Figure 3, in comparison to their predecessors, lithium-ion batteries encompass an area with a larger power and energy density area and even offer additional likelihood from this perspective due to the choice of different materials for the cathode. There are several cells in lithium-ion batteries. Battery packs are created by joining the cells collectively. Battery packs are made at the facility that is used in automobiles. The cells can then be created to satisfy a converter's requirements for the highest possible voltage and current, hence extending the range of possible applications.

4.2 Fuel Cell

The innovative fuel cell is another alternative energy source. It converts hydrogen-based chemical energy stored in electricity. Fuel cells are an excellent option to replace traditional advancements since the processes that take place inside them produce both energy and water as their final products. Figure 4 shows a variety of fuel cell types that are used in diverse sectors and emphasize different attributes. PEM fuel cells are popular in transportation infrastructure due to their advantages, such as increased efficacy, operating in low temperatures, and reduced susceptibility to rust [34]. Table 1 depicts a few parameters of the two energy sources.

5. Multiple Energy Management Strategies

A large variety of factors must be verified upon every occurrence when using rule-based techniques. Because they were developed abstractly, they cannot guarantee a perfect functioning area. Nevertheless, it is widely used now since it is economical, easy to just use, and works quickly. The controller decides how to divide the car's energy demand across the sources of energy depending on such principles. An FCHEV is also treated using a comparative idea [37], and the modification will be used for comparison in this evaluation. Approaches for real-time optimization (RTO) anticipate the optimal outcome within a procedure and keep analyzing the recent data. Therefore, a few efficiency difficulties are designed to be addressed in every sampling interval rather than trying to identify a global equilibrium position. The technique was designed to combat the susceptibility of the controller's actual climate interface. Due to the existence of a disturbing impact, gold is unable to respond effectively. The method provides knowledge from the history, current, and tomorrow with a framework that can be employed. Model predictive control (MPC) and equal cost minimization strategy (ECMS) are two of the most well-known RTO strategies. Using a variety of data that are most typically generating phases in our circumstance, Global Optimization (GO) procedures try to find the optimum point of a particular objective function. The concept often includes a mix of several drive cycles, and the finest possible options are chosen to enhance the provided performance. On this basis, Dynamic Programming (DP), a tactic that makes use of the Bellman equation's criterion of optimality, is widely used [37]. However, DP remains the most popular strategy since it almost ensures that the global optimum is obtained. Direct coding and enhanced refinement are also used for this rationale. The computational complexity of the control signal, which can be avoided to a significant degree at the expense of the sampling period, is the primary drawback in terms of accuracy [38]. These tactics establish the highest achievable aim for the main functionality, making them unquestionably
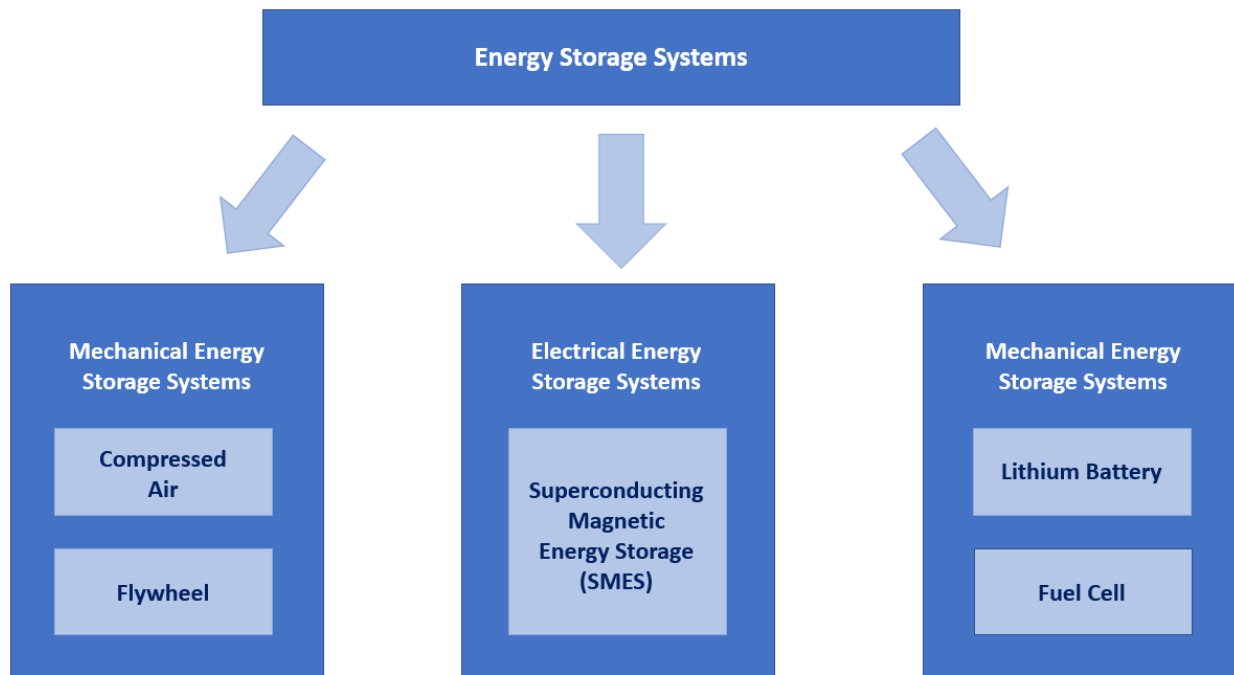
Figure 4. Types of traditional energy storage systems

Table 1. A comparison between a Li-battery and fuel cell [35-36]

Energy storage system	Power consumption rate (MW)	Discharge time	Power density	Energy density	Self discharge rate (%/day)	Response time	Efficiency (%)	Life span (years)	Cycle life (cycles)
Lithium-ion batteries	0–0.1	minutes-h	200–340	130–250	0.1–0.3	ms	65–95	5–8	600–1200
Fuel cell	0–50	s-days	>500 (W/L)	500–3000	0.5–2	ms-minutes	20–66	5–30	$10^3 - 10^4$

fantastic for distinguishing the display of another strategy. To get closer to the best outcome, the effects of various optimization algorithms or maybe even rule-based procedures might be updated [38-39]. Furthermore, stochastic optimization methods are available, some of which are only concerned with the decoupled evaluation and the search for the target function's optimal global threshold. The Bellman equation is combined with those approaches using a Markov Decision Process (MDP). Phase, deed, probability of switching period, reward, and technique are the fundamental concepts to understand an MDP. In a framework that may be used to observe how the components interact, phases are variables. The elements that interact with the model or environment and lead to a switching cycle are actions. There is a possibility for every shift, and the activity is given a reward [39-40]. To sum up, the approach serves as a manual for every state's activity aspect, and the intention is to consider it as a tool for increasing reward chances. Due to advancements in the sans model's predicted values that can be used in any setting classified as an MDP, reinforcement learning calculations have just started to emerge as a viable

approach and are currently being utilized to power-management concerns. In tests focusing on HEV or PHEV, which are like FCHEV, Q learning, and Deep Q learning (DQN) are commonly used without prototype computations. Q learning, also known as Q table learning, involves selecting an erratic activity that results in a switching period and a reward [40]. The Q table, which includes status and action data, collects the total number of upcoming and immediate awards. The table showing the value of the incentive in comparison to that state and action regard is updated for each iteration presuming that the current reward is more substantial than the previous compensation. Applying the Q learning computation to an FCHEV is described in [40-42] as a smart plan, and the results are compared to those of traditional control methods. Studies on adapting a DQN subordinate, a tactic like Q learning but utilizing a deep structure for scheduling, for FCHEV are also being conducted. The focus is on how the initial concentration situation affects performance, and they show how reinforcement learning estimates are beneficial. Besides the energy reduction using DQN, another key feature that is highlighted is fuel cell

deterioration [41-42]. By juxtaposing the display and the outcomes obtained by linear DP while also connecting them to Q Learning-based EMS, [42] further highlights the importance of selecting the proper target function. The converter is typically only depicted as a competency atlas or as a single performance rating in these tests. In comparison to simulation studies that are thought to last lengthier, DQN is suggested as a computational approach that performs better and employs neural activity instead of a database. Barriers within the institution are updated in lieu of the criteria on the table. As a method for maximizing the merits of such strategies, instantaneous and global refinement mergers have also been developed [43]. Only little encouraged research to alter the calculations to enhance the union rate since velocity in the computation pretreatment procedure is crucial; for greater consistency, a different computation dubbed DynaH is suggested [44]. Another RL computation that was created from the DQN calculations is Deep Deterministic Policy Gradient (DDPG). The primary advantage of the DDPG analysis over DQN is the ability to generate a continuous motion field as opposed to DQN, which requires that the activity field be partitioned. A hybrid electric vehicle (HEV) and a hybrid electric transport (HEB) [45] both use the DDPG computation to illustrate that they are able to achieve superior energy usage.

6. RL-Based Energy Management Strategy

By distributing the energy demand among several resources whilst continuing to meet that requirement, energy management systems in HEV aim to maintain SOC within predetermined threshold limits and reduce actual electricity usage. Whenever a battery is used to supply electricity [46], it needs to be recharged, and the only options are a fuel cell or regenerative braking system because distant charging is not an option. Fuel cells could not be used for energy in areas where their efficiency is close to the highest, according to a primary strategy for this problem. The fuel cell runs perhaps low efficiently as the need for power increases, but the battery can still be of service in power production up to a limit that is constrained by its SOC. Because the main requirement is to ensure that the transport follows the speed outline, there is little of choice for having fuel cells operate forward with poor efficiency, supposing the energy requirement significantly increases higher. Maintaining SOC within a period is another issue. It must be ensured that SOC level is maintained within the defined limit by the fuel cell or regenerative decelerating throughout normal driving when outside recharging is unthinkable in an HEV, unlike a PHEV [47]. Rule-based and performance-tuning strategies are frequently used to accomplish these goals. Stochastic and deterministic iterative methods are two further categories under which the improvement methodologies can be divided. Collaboration with AI-based tools for machine learning has also improved the usage and application of conventional simulations [48]. The present part gives a quick overview of the latest machine learning techniques that can be applied to transdisciplinary fuel cell electric vehicle studies to speed up modeled workflows, enhance comprehension, enhance evaluations, and generate current queries and hypotheses. This will build a structure of information sharing and implement an effective human-computer interaction (HCI) strategy [48-49] and must give a clear and succinct explanation of the empirical evidence, their analysis, and any possible design inferences [50]. A rapidly developing science, AI now helps people make forecasts regarding a variety of diverse scientific disciplines. For instance, AI strives to link computers and people in a way

that raises the combined level of intellect. Machine learning (ML), a methodical strategy that enables devices to gain knowledge from properly gathered and pre-processed information, is universally recognized as one of the fundamental AI technologies employed [51]. Since ML is closely related to how students remember, its primary goal is to comprehend and analyze the data, strategies, and philosophy that make up this area. The objective is to group the approaches used by the fundamental components of ML. Hence producing computational answers using reasoning, coding, statistics, probabilistic assessment and backpropagation, and parallelism. The benefit of gaining knowledge from information gathering is that assignments do not need to be established by a skilled developer. In contrast to conventional computer science, where commands are given directly and answers are received, the coder and the machinery act autonomously from one another. This is to ensure that the ML steps could emulate human judgment by creating choices based on the information provided, or familiarity. These are created by giving samples that demonstrate how the technology ought to operate. Additional facts are accessed but ample consumer expertise is needed the further complicated the data stream is [51-52]. As a result, thorough coaching is what propels the educational experience. The provided and molded facts are important to the procedure. At interacting using organized sets of data, comprising tables of qualitative and numerical parameters, it provides an incredible tool. It appears to coincide with the phrase data mining, despite the differences between the two terms. ML is primarily focused on learning by self from data and creating models that serve as the basis for projections for multiple situations, whilst data analysis aims to recognize as well as recognize trends from huge amounts of information to gain knowledge from the previous era [53]. Table 2 indicates the comparison of HESSs's use of RL algorithms.

7. Simulation Results

The simulations, therefore in research are concentrated on adjusting various Q learning control variables to reach a resolution with the aim of reducing SOC $_{battery}$ variance. To obtain this, the methodology training is broken up into 1000 events, each of which lasts for 2000 seconds. Every episode is broken down into 500 iterations, with a sampling interval of 4 seconds in every cycle [63]. The outcomes of simulations from both the beginning and the end of training are shown in Figure 5. The findings are shown for three various SOC $_{battery}$ starting values, including a minimum value of 0.3 per unit, a desirable value of 0.7 per unit, and the highest value that can be achieved, which is 1 per unit. The three sections (a), (b), and (c) of the simulation results for each episode exhibit distinct onboard hybrid power system and controller parameters (Figure 6) [64]. The NEDC load profile, the power provided by the fuel cell (PFC), and the power provided or absorbed by the battery are all displayed in part (a) of the episodic parameters "P" _"battery"). The battery's state of charge (SOC) and whether the controller activity is exploitation or exploration are both displayed in part (b) of the episodic parameters. When an orange bar is present, it means that the controller is engaging in exploitation, while when it is absent, it means that the controller is engaging in exploration. The SOC values of 0.6 and 0.8 per unit, which are closer to the intended value of 0.7 per unit, are also shown by two lines in section (b). These lines are incorporated into the graph to make the confluence more obvious.

Table 2. Comparison of HESSs's use of RL algorithms [31]

| Reference	Drawbacks	Advantages	Power transmission system	Algorithm
Hsu et al. [54]	Simplification of complex models	Adaptive to riding conditions	Electric bike	Q-learning
Qi and Wu [55], Liu and Murphey [56]	Dependence on driving data	High accuracy	Hybrid electric vehicle	Temporal-difference (TD) learning
Liu et al. [57]	Sporadic local optimization	Ability to run online	Plug-in hybrid electric vehicle	Q-learning
Hu et al. [58]	Computational load	Multiple control objectives	Hybrid electric vehicle	Q-learning
Kamet et al. [59]	Design complexity	Robust against variability	Hydraulic hybrid vehicle	Deep reinforcement learning and dynamic neural programming
Zhao et al. [60], Xiong et al. [34]	Complex mathematics	Real-time control	Hybrid truck	Dynamic learning
Kamet et al. [59] Zhao et al. [60]	Needs specific training	Data-driven model	Plug-in hybrid electric vehicle	Deep learning
Liu et al. [57], Hu et al. [58]	Sensitivity to driving cycle	Fast computation	Electric vehicle	Online reinforcement learning
Hay et al. [61], Lin et al. [62]	Data requirements	Improved battery life	Hybrid electric vehicle and plug-in hybrid electric vehicle	Reinforcement learning and Markov decision system
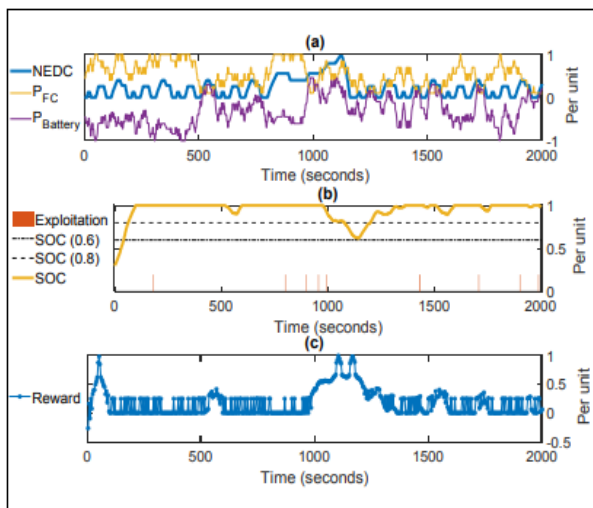
Figure 5. Simulation parameters during episode-1 for SOC $_{initial}$ = 0.3 per unit

Part (c) displays the per-unit value of the incentive received following every series repetition. The simulation findings demonstrate that the controller's activities at the start of training are driven by curiosity and are chosen at random [63-64]. According to the numerical simulations, the controller's basic training actions are driven by discovery and are random in nature. This means that the SOC battery and incentive are not compatible with the object's goal. In comparison, as the course progresses, the controller's activities are governed by exploitative and are selected to enhance compensation. As a result, the SOC $_{battery}$ and compensation are inching relative to the individual's goal. Despitef the SOC $_{battery}$ for show's base value, similar are observed. Nevertheless, SOC $_{battery}$'s value is frequently greater than the desired level of 0.7 per unit. This happens as the methodology may regulate how the battery fills, whilst the load demand affects how the battery empties.

Total energy consumption, average fuel cell efficiency, and SOC deviation will all be compared [64-65]. The agents programmed with DDPG and DQN methods can keep the SOC level within certain constraints during the Urban Dynamometer Driving Schedule (UDDS) cycle, and the difference from the target, which is set at 50%, is not noteworthy [66]. Total energy consumption, average fuel cell efficiency, and SOC deviation will all be compared. The agents programmed with DDPG and DQN methods can keep the SOC level within certain constraints during the Urban Dynamometer Driving Schedule (UDDS) cycle, and the difference from the target, which is set at 50%, is not noteworthy. As illustrated in Figure 7, it still has a suitable result at the end of the cycle, and restarting the loop from that point onward will not dramatically change SOC performance [66-67]. Considering that it is evident that there is an exchange between such two effects, Figure 8 illustrate correlation of the reduced energy area showings of the controller established with Rule-Based EMS and RL-Based EMS educated with DDP SOC levels and performance of the fuel cell under distinct EMS should be mentioned together. To assess how well the EMS is performing, multiple cases are evaluated. The battery is used in lieu of the fuel cell in DDPG-based EMS in the area where the SOC drops below 50%, as has been mentioned previously [59-60]. The scenario where the SOC level is just a little bit above the target is depicted in Figure 8. While rule-based EMS underperforms and starts the fuel cell slightly later, causing the SOC to fall more sharply, DDPG-based EMS starts operating the fuel cell almost where it is most efficient and attempts to keep running the fuel cell on higher efficiency region while using the battery [67].
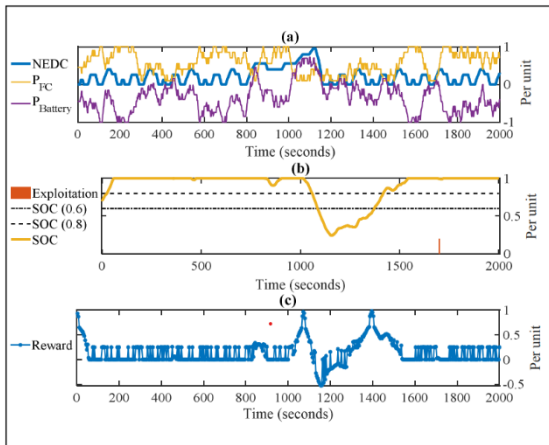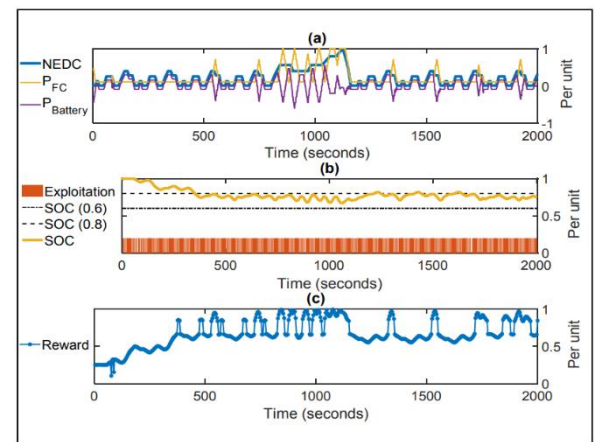
(a)
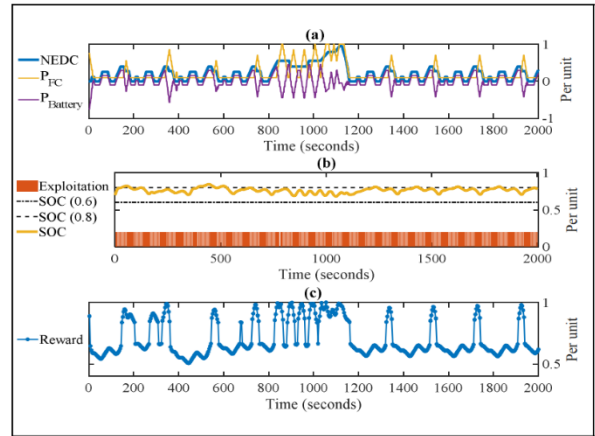
(b)
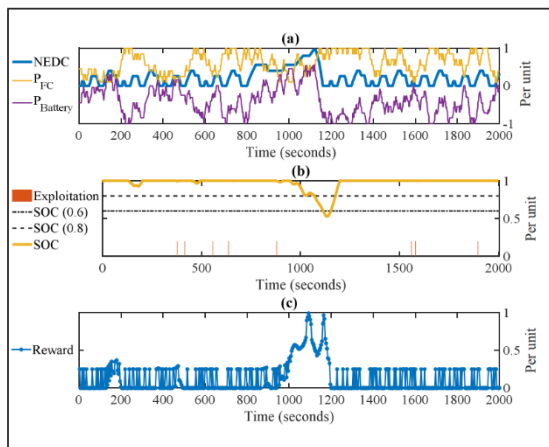
(c)

(d)

Figure 6. Simulation results with variable for $SOC_{initial}$ values during different episodes: (a) Episode-2 for $SOC_{initial}$ = 0.7 per unit, (b) Episode-999 for $SOC_{initial}$ = 0.7 per unit, (c) Episode-3 for $SOC_{initial}$ = 1 per unit, (d) Episode-1000 for $SOC_{initial}$ = 1 per unit
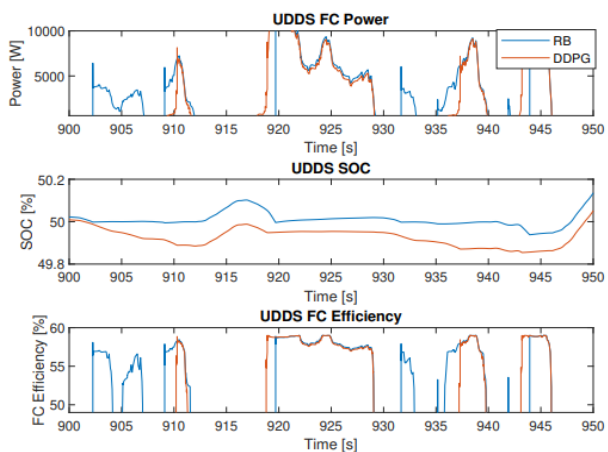

Figure 7. Performance comparison of the controllers created using Rule-Based EMS and RL-Based EMS trained with DDPG in the low power zone
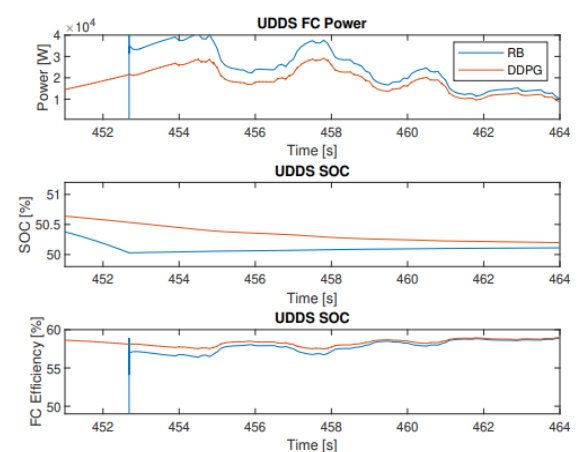

Figure 8. Performances of the controllers generated with Rule-Based EMS and RL-Based EMS trained with DDPG are compared in the mid power area

8.   Conclusion

This study's main driving factor was to review the implemented works of reinforcement learning calculations in an energy management strategy. For hybrid energy storage systems (HESSs) of hybrid electric vehicles, this research offered an overview of reinforcement learning-based energy management techniques (HEVs). The study began by introducing the issue of energy management in this industry and the approaches that can be taken to address it. The current energy management plans with various control goals were then discussed. Finally, as indicated, the future of reinforcement learning-based energy management systems looks promising. The development of more effective AI energy management methods represents a potentially fruitful research topic in this field. Through real-world experiments, applications, and simulations, it is also crucial to assess the suggested method's theoretical and practical viability. For FCHEV supported by fuel cells and batteries, an intelligent PEMS based on conventional reinforcement learning (Q-learning) has been created. The simulation model of the onboard hybrid power system was used to evaluate the Q learning based PEMS algorithm, which was implemented as a MATLAB script. Consequently, a comparison of the existing energy management systems employing machine learning has been made. According to simulated data, the algorithm can progress towards the primary goal of lowering $SOC_{battery}$ variation of about 0.7 per unit. Consequently, the algorithm can be changed beyond to achieve various goals like minimizing fuel usage and prolonging the life of elements. The DQN algorithm, which was later replaced by the DDPG method, emerged as the widely used model-free RL algorithm after Q-learning. The SOC of the battery, fuel cell efficacy, and overall energy usage were the determining factors. The UDDS cycles were chosen as they consider multiple driver behavior types. Energy management solutions based on DDPG and DQN were discovered to have the ability to utilize lower energy than the rule-based technique in all these drive cycles whilst attaining a comparable SOC behavior and minor SOC variance. Even while DQN, is extremely like DDPG and only slightly poorer than DDPG in the UDDS cycle. The best solution, whose active learning can persist in real-time applications, may be reached using the DDPG algorithm, it can be inferred.

Ethical issue

The author is aware of and comply with best practices in publication ethics, specifically with regard to authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The author adheres to publication requirements that the submitted work is original and has not been published elsewhere.

Data availability statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflict of interest

The author declares no potential conflict of interest.

References

[1]   T. Smith, C. Raucci, S. H. Hosseinloo, I. Rojon, J. Calleya, S. S.De La Fuente, P. Wu, and K. Palmer, "CO2 emissions from internationalshipping. Possible reduction targets and their associated pathways,"UMAS: London, UK, Tech. Rep., 2016.

[2]   Wu, P. (2020). Decarbonising coastal shipping using fuel cells and batteries (Doctoral dissertation, UCL (University College London)), uri: https://discovery.ucl.ac.uk/id/eprint/10101024.

[3]   N. Gray, S. McDonagh, R. O'Shea, B. Smyth, and J. D. Murphy,"Decarbonising ships, planes and trucks: An analysis of suitable low-carbon fuels for the maritime, aviation and haulage sectors,"Advancesin Applied Energy, p. 100008, 2021.

[4]   N. Sulaiman, M. Hannan, A. Mohamed, P. J. Ker, E. Majlan, and W. W.Daud, "Optimization of energy management system for fuel-cell hybridelectric vehicles: Issues and recommendations,"Applied energy, vol.228, pp. 2061–2079, 2018.

[5]   H. Peng, J. Li, L. L̈owenstein, and K. Hameyer, "A scalable, causal,adaptive en ergy management strategy based on optimal control theoryfor a fuel cell hybrid railw ay vehicle,"Applied Energy, vol. 267, p.114987, 2020.

[6]   P. Wu, J. Partridge, and R. Bucknall, "Costeffective reinforcementlearning energy management for plugin hybrid fuel cell and batteryships,"Applied Energy, vol. 275, p. 115258, 2020.

[7]   W. Wu, J. Partridge, and R. Bucknall, "Stabilised control strategy forpem fuel cell and supercapacitor propulsion system for a city bus,"International Journal of Hydrogen Energy, vol. 43, no. 27, pp. 12 302–12 313, 2018.

[8]   S. Eriksen, M. L̈utzen, J. B. Jensen, and J. C. Sørensen, "Improving theenergy efficiency of ferries by optimizing the operational practices," inProceedings of the Full Scale Ship Performance Conference 2018: TheRoyal Institution of Naval Architects.The Royal Institution of NavalArchitects, 2018, pp. 101–111.

[9]   S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approxi-mation error in actor-critic methods,"arXiv preprint arXiv:1802.09477,2018.

[10]   W. He, L. Tao, L. Han, Y. Sun, P. E. Campana, and J. Yan, "Optimal analysis of a hybrid renewable power system for a remote island," Renewable Energy, vol. 179, pp. 96–104, 2021.

[11]   A. Boretti and S. Castelletto, "Cost and performance of CSP and PV plants of capacity above 100 MW operating in the United States of America," Renewable Energy Focus, vol. 39, pp. 90–98, 2021.

[12]   O. Alves, L. Calado, R. M. Panizio, M. Gonçalves, E. Monteiro, and P. Brito, "Techno-economic study for a gasification plant processing residues of sewage sludge and solid recovered fuels," Waste Management, vol. 131, pp. 148–162, 2021.

[13]   A. Alsharif, C. W. Tan, R. Ayop, A. Dobi, and K. Y. Lau, "A comprehensive review of energy management strategy in vehicle-to-grid technology integrated with renewable energy sources," Sustainable Energy

Technologies and Assessments, vol. 47, article 101439, 2021.

[14] D.-D. Tran, M. Vafaeipour, M. El Baghdadi, R. Barrero, J. Van Mierlo,and O. Hegazy, "Thorough state-of-the-art analysis of electric and hybridvehicle powertrains: Topologies and integrated energy managementstrategies,"R enewable and Sustainable Energy Reviews, vol. 119, p.109596,

[15] J. Wu, H. He, J. Peng, Y. Li, and Z. Li, "Continuous reinforcementlearning of energy management with deep q network for a power splithybrid electric bus,"Applied energy, vol. 222, pp. 799–811, 2018.

[16] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G.Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovskiet al., "Humanlevel control through deep reinforcement learning,"Nature, vol. 518, no. 7540, p. 529, 2015.

[17] P. Wu and R. Bucknall, "Hybrid fuel cell and battery propulsionsystem modelling and multi-objective optimisation for a coastal ferry,"International Journal of Hydrogen Energy, vol. 45, no. 4, pp. 3193–3208, 2020.

[18] "On the design of plugin hybrid fuel cell and lithium batterypropulsion systems for coastal ships," in13th Int ernational MarineDesign Conference (IMDC 2018), P. Kujala and L. Lu, Eds., vol. 2. London: CRC Press/Balkema, 2018, Conference Proceedings, pp. 941–951.

[19] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller,"Deterministic policy gradient algorithms," inProceedings of the 31stInternati onal Conference on Machine Learning, 2014, pp. 387–395.

[20] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa,D. Silver, and D. Wierstra, "Continuous control with deep reinforcementlearning,"arXiv preprint, 2015.

[21] H. van Hasselt, "Double q-learning," inAdvances in Neural InformationProcessing Systems 23, J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor,R. S. Zemel, and A. Culotta, Eds.Curran Associates, Inc., 2010, pp.2613–2621.

[22] Henderson, Peter & Islam, Riashat & Bachman, Philip & Pineau, Joelle & Precup, Doina & Meger, David. (2017). Deep Reinforcement Learning That Matters. Proceedings of the AAAI Conference on Artificial Intelligence. 32. 10.1609/aaai.v32i1.11694.

[23] N. Sulaiman, M. A. Hannan, A. Mohamed, P. J. Ker, E. H. Majlan, and W. R. Wan Daud, "Optimization of energy management system for fuel-cell hybrid electric vehicles: Issues and recommendations," Appl. Energy, vol. 228, pp. 2061–2079, Oct. 2018.

[24] H. Li, A. Ravey, A. N'Diaye, and A. Djerdir, "A Review of Energy Management Strategy for Fuel Cell Hybrid Electric Vehicle," in 2017 IEEE Vehicle Power and Propulsion Conference (VPPC), 2017, pp. 1–6.

[25] J. Han, J.-F. Charpentier, and T. Tang, "An Energy Management System of a Fuel Cell/Battery Hybrid Boat," Energies, vol. 7, no. 5, pp. 2799–2820, May 2014.

[26] M. K. Zadeh, L. Saublet, R. Gavagsaz-Ghoachani, B. Nahid Mobarakeh, S. Pierfederici, and M. Molinas, "Energy management and stabilization of a hybrid DC

microgrid for transportation applications," in 2016 IEEE Applied Power Electronics Conference and Exposition (APEC), 2016, pp. 3397–3402.

[27] M. K. Zadeh, Stability Analysis Methods and Tools for Power Electronics- Based DC Distribution Systems, Applicable to On-Board Electric Power Systems and Smart Microgrids. NTNU, 2016.

[28] S. N. Motapon, L. Dessaint, and K. Al-Haddad, "A Comparative Study of Energy Management Scemes for a Fuel-Cell Hybrid Emergency Power System of More-Electric Aircraft," IEEE Trans. Ind. Electron., vol. 61, no. 3, pp. 1320–1334, Mar. 2014.

[29] S. J. Moura, D. S. Callaway, H. K. Fathy, and J. L. Stein, "Tradeoffs between battery energy capacity and stochastic optimal power management in plug-in hybrid electric vehicles," J. Power Sources, vol. 195, no. 9, pp. 2979–2988, May 2010.

[30] X. Zhang, C. C. Mi, A. Masrur, and D. Daniszewski, "Wavelet transform-based power management of hybrid vehicles with multiple on-board energy sources including fuel cell, battery and ultracapacitor," J. Power Sources, vol. 185, no. 2, pp. 1533–1543, Dec. 2008.

[31] Chen, Tzu-Chia & Al-Azzawi, Fouad & Grimaldo, John & Chetthamrongchai, Paitoon & Dorofeev, Aleksei & masood, aras & Ahmed, Alim Al Ayub & Akhmadeev, Ravil & Latipah, Asslia & Abu Al-Rejal, Hussein. (2022). Development of Machine Learning Methods in Hybrid Energy Storage Systems in Electric Vehicles. Mathematical Problems in Engineering. 2022. 1-8. 10.1155/2022/3693263.

[32] F. Segura and J. M. Andújar, "Power management based on sliding control applied to fuel cell systems: A further step towards the hybrid control concept," Appl. Energy, vol. 99, pp. 213–225, Nov. 2012.

[33] G. Paganelli, S. Delprat, T. M. Guerra, J. Rimaux, and J. J. Santin, "Equivalent consumption minimization strategy for parallel hybrid powertrains," in Vehicular Technology Conference. IEEE 55th Vehicular Technology Conference. VTC Spring 2002 (Cat. No.02CH37367), 2002, vol. 4, pp. 2076–2081 vol.4.

[34] J. Han, Y. Park, and Y. Park, "A novel updating method of equivalent factor in ECMS for prolonging the lifetime of battery in fuel cell hybrid electric vehicle," IFAC Proc. Vol., vol. 45, no. 30, pp. 227–232, Jan. 2012.

[35] A. Chatzivasileiadi, E. Ampatzi, and I. Knight, "Characteristics of electrical energy storage technologies and their applications in buildings," Renewable and Sustainable Energy Reviews, vol. 25, pp. 814–830, 2013.

[36] O. Z. Sharaf and M. F. Orhan, "An overview of fuel cell technology: fundamentals and applications," Renewable and Sustainable Energy Reviews, vol. 32, pp. 810–853, 2014.

[37] J. Han, Y. Park, and D. Kum, "Optimal adaptation of equivalent factor of equivalent consumption minimization strategy for fuel cell hybrid electric vehicles under active state inequality constraints," J. Power Sources, vol. 267, pp. 491–502, Dec. 2014.

[38] P. García, J. P. Torreglosa, L. M. Fernández, and F. Jurado, "Viability study of a FCbattery-SC tramway controlled by equivalent consumption minimization

strategy," Int. J. Hydrog. Energy, vol. 37, no. 11, pp. 9368–9382, Jun. 2012.

[39] C. Zheng, S. W. Cha, Y. Park, W. S. Lim, and G. Xu, "PMP-based power management strategy of fuel cell hybrid vehicles considering multi-objective optimization," Int. J. Precis. Eng. Manuf., vol. 14, no. 5, pp. 845–853, May 2013.

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

[41] T. Li, H. Liu, and D. Ding, "Predictive energy management of fuel cell supercapacitor hybrid construction equipment," Energy, vol. 149, pp. 718–729, Apr. 2018.

[42] J. P. Torreglosa, P. García, L. M. Fernández, and F. Jurado, "Predictive Control for the Energy Management of a Fuel-Cell–Battery–Supercapacitor Tramway," IEEE Trans. Ind. Inform., vol. 10, no. 1, pp. 276–285, Feb. 2014.

[43] M.-J. Kim and H. Peng, "Power management and design optimization of fuel cell/battery hybrid vehicles," J. Power Sources, vol. 165, no. 2, pp. 819–832, Mar. 2007.

[44] W.-S. Lin and C.-H. Zheng, "Energy management of a fuel cell/ultracapacitor hybrid power system using an adaptive optimal control method," J. Power Sources, vol. 196, no. 6, pp. 3280– 3289, Mar. 2011.

[45] S. Mohan, Y. Kim, and A. G. Stefanopoulou, "Estimating the Power Capability of Li-ion Batteries Using Informationally Partitioned Estimators," IEEE Trans. Control Syst. Technol., vol. 24, no. 5, pp. 1643–1654, Sep. 2016.

[46] R. C. Hsu, S. Chen, W. Chen, and C. Liu, "A Reinforcement Learning Based Dynamic Power Management for Fuel Cell Hybrid Electric Vehicle," in 2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS), 2016, pp. 460–464.

[47] F.-D. Li, M. Wu, Y. He, and X. Chen, "Optimal control in microgrid using multi-agent reinforcement learning," ISA Trans., vol. 51, no. 6, pp. 743–751, Nov. 2012.

[48] P. Kofinas, S. Doltsinis, A. I. Dounis, and G. A. Vouros, "A reinforcement learning approach for MPPT control method of photovoltaic sources," Renew. Energy, vol. 108, pp. 461– 473, Aug.2017.

[49] R. Liessner, C. Schroer, A. Dietermann, and B. Bäker, "Deep Reinforcement Learning for Advanced Energy Management of Hybrid Electric Vehicles," presented at the International Conference on Agents and Artificial Intelligence, 2018, vol. 2, pp. 61–72.

[50] M. Prauzek, N. R. A. Mourcet, J. Hlavica, and P. Musilek, "Q Learning Algorithm for Energy Management in Solar Powered Embedded Monitoring Systems," in 2018 IEEE Congress on Evolutionary Computation (CEC), 2018, pp. 1–7.

[51] S. Kim and H. Lim, "Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings," Energies, vol. 11, no. 8, p. 2010, Aug. 2018.

[52] S. Etemad and N. Mozayani, "Using reinforcement learning to make smart energy storage sources in microgrid," in 2015 30th International Power System Conference (PSC), 2015, pp. 345–350.

[53] C. Liu, "Optimal Power Management Based on Q-Learning and Neuro-Dynamic Programming for Plug-in Hybrid Electric Vehicles," PhD Dissertation, University of Michigan, Dearborn, Michigan, USA, 2018.

[54] R. C. Hsu, C. T. Liu, and D. Y. Chan, "A reinforcementlearning-based assisted power management with QoR provisioning for human–electric hybrid bicycle," IEEE Transactions on Industrial Electronics, vol. 59, no. 8, pp. 3350–3359, 2011.

[55] X. Qi, G. Wu, K. Boriboonsomsin, M. J. Barth, and J. Gonder "Data-driven reinforcement learning-based real-time energy management system for plug-in hybrid electric vehicles," Transportation Research Record: Journal of the Transportation Research Board, vol. 2572, no. 1, pp. 1–8, 2016.

[56] C. Liu and Y. L. Murphey, "Power management for plug-in hybrid electric vehicles using reinforcement learning with trip information," in Proceedings of the 2014 IEEE Transportation Electrification Conference and Expo (ITEC), pp. 1–6, IEEE, Dearborn, MI, USA, June 2014.

[57] T. Liu, Y. Zou, D. Liu, and F. Sun, "Reinforcement learning of adaptive energy management with transition probability for a hybrid electric tracked vehicle," IEEE Transactions on Industrial Electronics, vol. 62, no. 12, pp. 7837–7846, 2015.

[58] J. Hu, D. Liu, C. Du, F. Yan, and C. Lv, "Intelligent energy management strategy of hybrid energy storage system for electric vehicle based on driving pattern recognition," Energy, vol. 198, p. 117298, 2020.

[59] A. A. Kamel, H. Rezk, and M. A. Abdelkareem, "Enhancing the operation of fuel cell-photovoltaic-battery-supercapacitor renewable system through a hybrid energy management strategy," International Journal of Hydrogen Energy, vol. 46, no. 8, pp. 6061–6075, 2021.

[60] P. Zhao, Y. Wang, N. Chang, Q. Zhu, and X. Lin, "A deep reinforcement learning framework for optimizing fuel economy of hybrid electric vehicles," in Proceedings of the 2018 23rd Asia and South Pacific design automation conference (ASP-DAC), pp. 196–202, IEEE, Jeju, South Korea, January 2018.

[61] R. Xiong, J. Cao, and Q. Yu, "Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle," Applied Energy, vol. 211, pp. 538–548, 2018.

[62] H. Yu, M. Kuang, and R. McGee, "Trip-oriented energy management control strategy for plug-in hybrid electric vehicles," in Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference, pp. 1323–1336, Orlando, FL, USA, January 2013.

[63] X. Lin, Z. Wang, and J. Wu, "Energy management strategy based on velocity prediction using back propagation neural network for a plug-in fuel cell

electric vehicle," International Journal of Energy Research, vol. 45, no. 2, pp. 2629–2643, 2021.

[64] N. P. Reddy, D. Pasdeloup, M. K. Zadeh and R. Skjetne, "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, pp. 1-6, doi: 10.1109/ITEC.2019.8790451.

[65] S. Kim and H. Lim, "Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings," Energies, vol. 11, no. 8, p. 2010, Aug. 2018.

[66] Y. Wu, H. Tan, J. Peng, H. Zhang, and H. He, "Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus," Applied energy, vol. 247, pp. 454–466, 2019.

[67] Eğer, Z. E. (2022). Reinforcement learning based energy management strategy for fuel cell hybrid vehicles (Doctoral dissertation), uri: https://research.sabanciuniv.edu/id/eprint/42949.