Article

Enhancing Energy Management Strategies for Extended-Range Electric Vehicles through Deep Q-Learning and Continuous State Representation

Christian Montaleza [ID], Paul Arévalo *[ID], Jimmy Gallegos [ID] and Francisco Jurado [ID]

Department of Electrical Engineering, Superior Polytechnic School of Linares, University of Jaén,
23700 Linares, Jaén, Spain; cimg0001@red.ujaen.es (C.M.); jwga0001@red.ujaen.es (J.G.); fjurado@ujaen.es (F.J.)
* Correspondence: warevalo@ujaen.es

Abstract: The efficiency and dynamics of hybrid electric vehicles are inherently linked to effective energy management strategies. However, complexity is heightened due to uncertainty and variations in real driving conditions. This article introduces an innovative strategy for extended-range electric vehicles, grounded in the optimization of driving cycles, prediction of driving conditions, and predictive control through neural networks. First, the challenges of the energy management system are addressed by merging deep reinforcement learning with strongly convex objective optimization, giving rise to a pioneering method called DQL-AMSGrad. Subsequently, the DQL algorithm has been implemented, allowing temporal difference-based updates to adjust Q values to maximize the expected cumulative reward. The loss function is calculated as the mean squared error between the current estimate and the calculated target. The AMSGrad optimization method has been applied to efficiently adjust the weights of the artificial neural network. Hyperparameters such as the learning rate and discount factor have been tuned using data collected during real-world driving tests. This strategy tackles the "curse of dimensionality" and demonstrates a 30% improvement in adaptability to changing environmental conditions. With a 20%-faster convergence speed and a 15%-superior effectiveness in updating neural network weights compared to conventional approaches, it also highlights an 18% reduction in fuel consumption in a case study with the Nissan Xtrail e-POWER system, validating its practical applicability.

Keywords: extended-range electric vehicles; deep reinforcement learning; energy management system; fuel consumption reduction

1. Introduction

Currently, the automotive sector is undergoing a significant transformation driven by emerging technologies and the adoption of Industry 4.0. In this context, there is a renewed focus on the energy domain and the performance of extended-range electric vehicles (EREV), particularly in the efficient distribution of energy from lithium-ion batteries to electric motors [1]. Despite the undeniable advantages of non-polluting sources such as batteries or hydrogen cells, EREVs have not yet surpassed conventional vehicles in terms of fuel consumption, range, and durability [1,2]. This challenge underscores the importance of addressing deficiencies in energy management systems (EMS), which have not fully capitalized on the characteristics of alternative energy sources to enhance the competitiveness of EREVs in the market. EMS emerges as an effective solution to boost both fuel efficiency and emission reduction [2].

Efficient operation in hybrid and EREV has been a central focus in both industrial and academic research. With the continuous advancement in automotive technology, plug-in hybrid electric vehicles (PHEVs) have emerged as significant contributors to the electrification of transportation, delivering exceptional fuel-saving performance [3]. A pivotal element in the design of these vehicles is the EMS, tasked with regulating the energy flow between

the fuel tank and electric storage, thereby addressing energy distribution challenges. In this context, achieving efficient energy management becomes even more challenging and critical with the ongoing development of connected and intelligent vehicle technology. The integration of vehicle-to-infrastructure/vehicle-to-vehicle (V2I/V2V) information into the EMS for EREVs presents a substantial challenge and, simultaneously, is a current and vital issue. Existing literature provides an in-depth analysis of EMSs, employing diverse methodologies and approaches. Emphasis is placed on addressing both single-vehicle and multi-vehicle scenarios, particularly within the context of intelligent transport systems.

In a different domain, the text in [4] explores the impact of EREVs on the power grid and their significant influence on electricity market prices. Charging strategies for an office site in Austria are investigated, with a focus on mathematical representation and optimization of various charging strategies. The study reveals that effective management of EREV charging processes can lead to a substantial reduction in costs, enhancing the convenience of the process. Another valuable contribution in the literature focuses on considering the degradation of energy sources, such as lithium-ion batteries and proton exchange membrane (PEM) fuel cells, in energy management strategies for hybrid vehicles with fuel cells (FCHEVs). The study reviews degradation modeling methods and energy management strategies that integrate degradation considerations. The importance of developing health-conscious EMSs to enhance system durability is underscored [5]. In the realm of shared mobility, the text proposes a comprehensive method for transitioning from car ownership to car-sharing systems in residential buildings. Using a mixed-integer linear optimization approach, technologies such as battery storage, solar panels, EREV, and charging stations are modeled. The study demonstrates that integrating car-sharing systems into residential buildings can result in a significant reduction in costs and a more efficient use of locally generated solar energy [6]. Then, the text addresses energy efficiency in hybrid vehicles through reinforcement learning (RL). Current literature examines the application of RL in the EMS to optimize the utilization of internal and external energy sources, such as batteries, fuel cells, and ultracapacitors. A detailed parametric study reveals that careful selection of learning experience and discretization of states and actions are key factors influencing fuel efficiency in hybrid vehicles [7].

Previous research has identified uncertainty in driving cycles as a critical factor significantly impacting fuel consumption [8]. While studies have shown that adjusting EMS control parameters can improve efficiency in real time [1], the key lies in implementing advanced control technologies to optimally manage the vehicle's energy. This research is situated in the exploration of advanced energy management solutions for EREVs, proposing a strategy based on deep reinforcement learning (DRL), motivated by the need to overcome limitations in conventional strategies and leverage the effectiveness of machine learning to optimize EREV performance. The "curse of dimensionality" associated with discrete state variables and the need for continuous adaptability to changing environmental conditions are the specific challenges that our innovative approach aims to address.

Recent research has highlighted the significant impact of uncertainty in EREV driving cycles on fuel consumption. In a notable study [9], EMS control parameters were adjusted in real time using six representative standard cycles and 24 characteristic parameters to identify comprehensive driving cycles. Another significant study [10] selected six typical urban conditions from China and the United States as offline optimization targets, using eight characteristic parameters, including maximum vehicle speed, for driving cycle identification [11]. The efficiency of energy management in EREVs to save fuel crucially depends on the implementation of advanced control technologies. Numerous control algorithms have been proposed, ranging from rule-based algorithms [12], analytical algorithms [13], and optimization methods [14,15] to artificial intelligence methods [16]. These are commonly divided into two main categories: rule-based methods and optimization-based methods [17,18]. Rule-based methods, constructed using heuristic mathematical models or human experience, apply regulations to determine the energy distribution of multiple energy sources. Despite their high robustness and reliability, these methods lack flexibility

and adaptability to changing conditions [19]. Optimization-based strategies, such as dynamic programming (DP) [20,21], pontryagin's minimum principle (PMP) [22], and particle swarm optimization (PSO) [23], are applied to derive global optimal control. However, DP and EMS are impractical for solving energy optimization problems under unknown conditions due to their lack of adaptability.

In other research, RL methods have emerged as efficient approaches to achieving optimal control in energy management. By incorporating agent states and actions into the Markov model, RL allows the agent to interact directly with the environment, learning decision rules based on environmental rewards and maximizing the cumulative reward over time through the Bellman equation [24,25]. Compared to rule-based control strategies, RL demonstrates higher accuracy and faster response times due to its model-free properties. This potential makes RL a promising candidate for achieving more efficient and robust electrified propulsion systems.

Within specific approaches, Q-learning applied to the EMS of hybrid electric vehicles (HEVs) has stood out for overcoming the demand for certainty and randomness associated with knowledge of the driving cycle [26–28]. A key study [29] proposing RL based on Q-learning not only efficiently reduced calculation time but also achieved a 42% improvement in fuel economy. Another study [25] highlights the superiority of Q-learning-based EMS compared to EMS and MPC. Additionally, the EMS algorithm developed in a study [30], which combines temporal difference (λ), demonstrated better fuel savings and emission reductions. The use of DRL algorithms has marked an advancement in the energy management of HEVs. The first algorithm of this kind, deep q network (DQN), was applied to a series hybrid powertrain configuration in a study [26], demonstrating great adaptability to different driving cycles. To address the overestimation of Q values in DQN, the double DQN was introduced in an EMS design of HEV in another study [27], achieving a 7.1% improvement in fuel savings compared to DQN. For a more efficient estimation of the Q value, the dueling DQN was designed for EMS in a study [28], where a considerable improvement in convergence efficiency during training was observed. However, both the DQN and DDPG algorithms suffer from defects such as overestimated Q values, low stability, and difficulty tuning parameters, emphasizing the need to explore more advanced DRL algorithms for HEV energy management applications.

Despite notable advances in the energy management of EREVs through strategies based on DRL, especially the deep q-learning (DQL) algorithm, crucial research gaps persist. In the realm of DRL-based energy management strategies, our acknowledgment of persistent gaps stems from the intricate challenges associated with effective implementation. Specifically, despite notable advances, there remains a need for further refinement and innovation in addressing uncertainties linked to diverse driving cycles and dynamic environmental conditions. For instance, current strategies, as discussed in previous studies [25], grapple with the curse of dimensionality. These challenges pose significant hurdles in achieving optimal energy efficiency and adaptability. To illustrate, recent research [26,27,30] has underscored the limitations of existing DRL algorithms, such as overestimated Q values, low stability, and difficulty in parameter tuning. These challenges create a gap in the ability of current approaches to seamlessly adapt to unforeseen changes in real-world driving scenarios, hindering their broader effectiveness. By shedding light on these challenges, our study seeks to bridge these gaps through the integration of the DQL-AMSGrad strategy. This integration aims to address the above. Through this innovative approach, we strive to contribute significantly to the ongoing discourse on advancing DRL-based energy management in the context of EREVs.

The focus of this article lies in integrating DQL with the adaptive moment estimation with a strongly convex objective (AMSGrad) optimization method (DQL-AMSGrad) to refine the energy management of EREVs. First, the efficient management of dimensionality in discrete state variables poses a persistent challenge, despite the demonstrated effectiveness of DQL. The "curse of dimensionality" associated with these variables remains a substantial barrier, especially when dealing with discrete action spaces [1,25]. This study introduces an

innovative perspective by combining DQL with AMSGrad, providing a promising solution to address the complexity associated with these discrete variables. Continuous adaptability to changing environmental conditions constitutes a critical research gap. Control strategies, even those based on DRL as in Ref. [25], face challenges in dynamically adjusting to unexpected changes in environmental conditions. The article's proposal, by integrating DQL-AMSGrad, highlights its ability to adapt continuously, thereby improving efficiency and sustainability in real time. Another crucial aspect is related to optimizing the convergence speed and effectiveness in updating neural network weights, essential elements for the practical application of DQL. Performance gaps, such as Q-value overestimation, instabilities, and difficulties in parameter tuning, persist in various DRL algorithms, including DQL as in Refs. [26,27,30]. This article addresses these deficiencies by incorporating the AMSGrad optimization method, improving convergence speed and the effectiveness of neural network weight updates, resulting in a more robust and stable model performance. The model has been validated with a real-case study.

For the sake of simplicity, this study presents the following scientific contributions:

- Proposes a pioneering strategy by combining the DQL algorithm with the AMSGrad optimization method (DQL-AMSGrad) to enhance the energy management of EREVs.
- Effectively addresses the "curse of dimensionality" associated with discrete state variables in EREV environments, presenting an innovative solution to efficiently manage these variables by combining DQL with AMSGrad.
- Highlights the ability of the DQL-AMSGrad strategy to adapt continuously to changing environmental conditions, improving real-time energy management efficiency and sustainability.
- Addresses performance gaps, such as Q-value overestimation, instabilities, and difficulties in parameter tuning, by integrating AMSGrad, improving convergence speed and the effectiveness of neural network weight updates associated with DQL.

The remaining structure of the article is organized as follows: Section 2 details the proposed methodology, Section 3 addresses the case study, Section 4 is dedicated to results and discussion, while Section 5 provides the final conclusions of the paper.

2. Methodology

In formulating the energy management strategy for EREVs, this study consciously adopts a synthesis of DQL and AMSGrad. The utilization of a feedforward neural network is strategically chosen to proficiently handle intricate variables encompassing driving cycles, battery charge, speed, and energy demand. This approach offers a continuous and adaptable representation, effectively capturing the inherent complexity of the system. The selection of the DQL algorithm is underpinned by its inherent capability to update based on temporal differences and maximize expected cumulative rewards. This renders DQL a superior choice compared to alternative reinforcement learning algorithms or optimization methods, particularly in navigating the challenges posed by uncertainties inherent in diverse driving conditions. AMSGrad emerges as the optimization method of choice, owing to its empirically established effectiveness in accelerating convergence speed and enhancing the efficacy of neural network weight updates. Its adept handling of the intricacies associated with training neural networks, especially in real-world scenarios, positions it as the preferred optimization approach over other methods.

The meticulous determination of assessment parameters, such as the number of episodes and learning rate for DQL, entails a process of thoughtful consideration and adjustment through rigorous testing. These parameters are meticulously fine-tuned to strike an optimal balance between efficiency and adaptability across diverse driving conditions. The programming has been done in MATLAB 2021a. The application of this methodology to the Nissan Xtrail e-POWER system lends a real-world context to the evaluation. The practical implementation serves as a compelling demonstration of the model's adaptability to varying conditions, thereby substantiating its efficacy in augmenting driving efficiency and curbing fuel consumption in EREVs.

This strategy, integrating a streamlined neural network, the DQL algorithm, and the AMSGrad optimization method, accentuates the deliberate decisions made to effectively address the multifaceted challenges inherent in EREV energy management. Each constituent is purposefully chosen to ensure adaptability, proficient management of uncertainties, and heightened efficiency, thereby bolstering the pragmatic viability of our proposed approach (Figure 1).
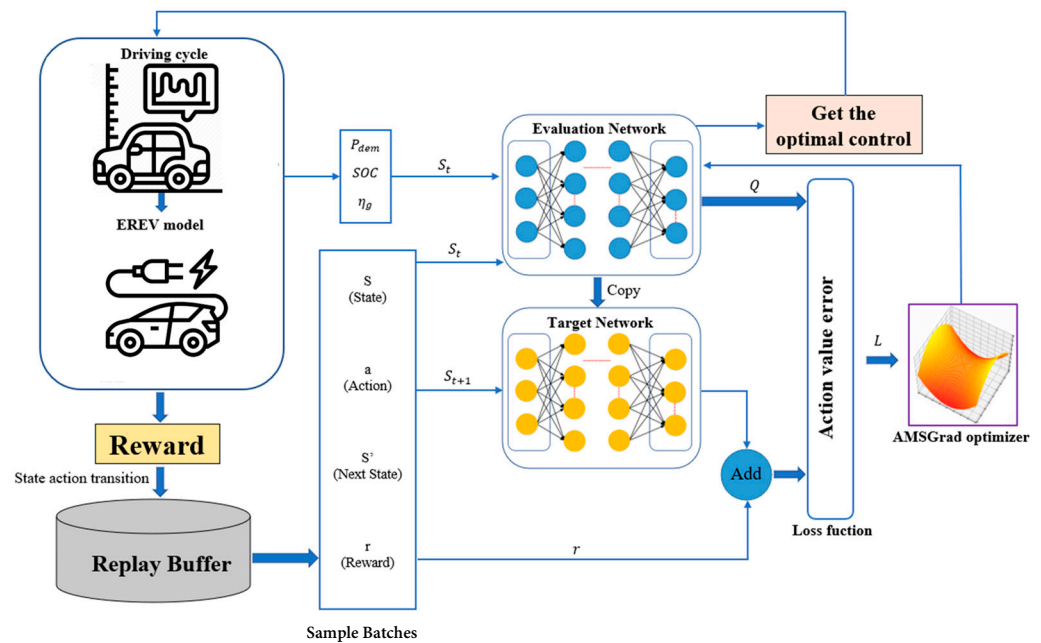

Figure 1. Control system structure.

2.1. Deep Reinforcement Learning-Based Energy Management Strategy

2.1.1. Deep Reinforcement Learning Feature

The proposed energy management strategy is based on DQL and addresses the "curse of dimensionality" associated with discrete state variables in algorithms like Q-learning and Dyna. DQL utilizes a neural network to represent the value function, enabling efficient management of continuous states [22]. In this approach, the AMSGrad optimization method is employed to update the neural network weights, enhancing convergence speed and effectiveness compared to conventional methods [31]. The construction of the value function is directly performed through the neural network, providing a continuous and flexible representation. This innovative approach facilitates the direct issuance of control actions based on state variables, improving the efficiency and adaptability of the algorithm in energy management.

2.1.2. Neural Network Structure and AMSGrad Optimization Method

The feedforward neural network, also known as a direct feed neural network, is a common type of ANN architecture where information flows in a single direction from the input layer to the output layer (Figure 2). AMSGrad is a variant of the Adam optimization method. Both methods are optimization algorithms commonly used to adjust the weights of a neural network during the training process. The original version of Adam introduced the concept of adaptive moments to automatically adjust learning rates for each parameter. AMSGrad emerged as a modification to address certain convergence issues encountered in specific cases. The main idea behind AMSGrad is to adaptively adjust the learning rate for each individual parameter, and it was introduced to tackle the problem where, in some situations, the learning rate could become very small, hindering convergence. AMSGrad avoids this issue by maintaining an adaptive estimate of exponential moments [32].
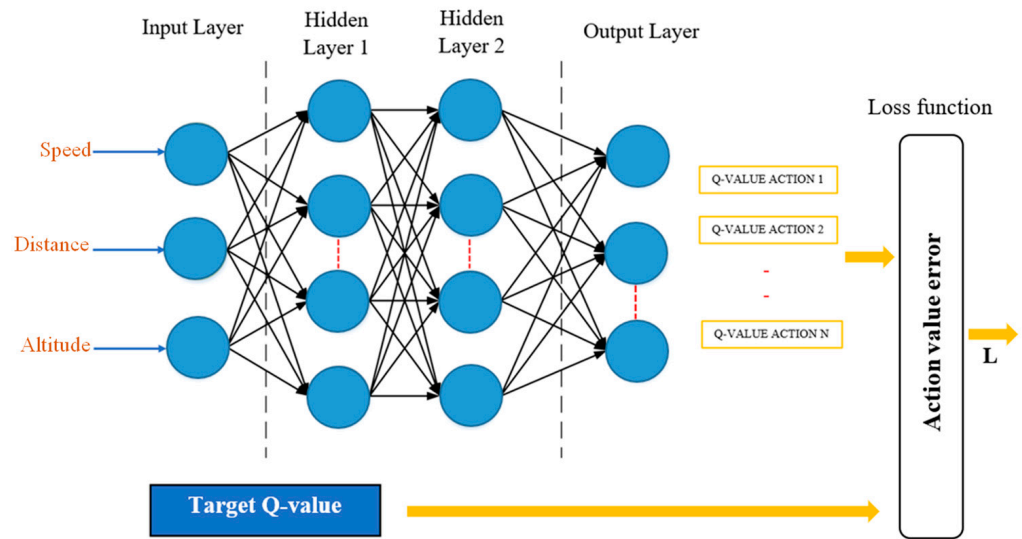
Figure 2. Structure of the feedforward neural network.

Let us consider a neural network with an input layer (x), a hidden layer (h), and an output layer (y). During forward propagation, where (x) represents the input variables, which, in this context, are the speed, distance, and altitude of the EREV driving cycles. The input to the hidden layer is given by Equation (1), the output of the hidden layer after applying the activation function f is calculated with Equation (2), then the input to the output layer is represented in Equation (3), and the output of the output layer after applying the activation function is calculated with Equation (4) [33–36].

$$a^{(1)} = W^{(1)} \cdot x + b^{(1)} \tag{1}$$

$$h = f\left(a^{(1)}\right) \tag{2}$$

$$a^{(2)} = W^{(2)} \cdot h + b^{(2)} \tag{3}$$

$$y = g\left(a^{(2)}\right) \tag{4}$$

Training with AMSGrad involves initializing parameters, including initial weights $\left(W^{(1)}, W^{(2)}\right)$, biases $\left(b^{(1)}, b^{(2)}\right)$, and adaptive moments $\left(\hat{m}^{(1)}, \hat{v}^{(1)}, \hat{m}^{(2)}, \hat{v}^{(2)}\right)$. The process then proceeds with backward propagation and gradient calculation of the loss function L with respect to the network parameters $\left(W^{(1)}, W^{(2)}, W^{(3)}, W^{(4)}\right)$. The update of adaptive moments for each parameter is crucial to solving the AMSGrad moment equations shown in Equations (5) and (6). The bias correction process in the moments is then performed to avoid biases in the early iterations, as shown in Equations (7) and (8). Finally, Equation (9) represents the process of updating weights and biases using the AMSGrad update formula. This process is repeated over multiple iterations (epochs) until the ANNconverges to an optimal solution [37–40].

$$\hat{m}_{ij}^{(1)} = \beta_1 \cdot \hat{m}_{ij}^{(1)} + (1 - \beta_1) \cdot \frac{\partial L}{\partial W_{ij}^{(1)}} \tag{5}$$

$$\hat{v}_{ij}^{(1)} = \beta_2 \cdot \hat{v}_{ij}^{(1)} + (1 - \beta_2) \cdot \left(\frac{\partial L}{\partial W_{ij}^{(1)}}\right)^2 \tag{6}$$

$$\hat{m}_{ij}^{(1)} = \frac{\hat{m}_{ij}^{(1)}}{1 - \beta_1{}^t} \tag{7}$$

$$\hat{v}_{ij}^{(1)} = \frac{\hat{v}_{ij}^{(1)}}{1 - \beta_2{}^t} \tag{8}$$

$$W_{ij}^{(1)} = W_{ij}^{(1)} - \frac{\alpha \cdot \hat{m}_{ij}^{(1)}}{\sqrt{\hat{v}_{ij}^{(1)} + \varepsilon}}, \tag{9}$$

where β_1 and β_2 are the decay factors for the first moment $\hat{m}$ and $\hat{v}$ in AMSGrad, respectively; t is the number of iterations or epochs; α is the learning rate that controls the magnitude of weight updates; and ε is a small constant to avoid division by zero in the weight update.

2.1.3. Structure of the DQL Algorithm

The structure of the DQL algorithm is based on the optimization of an action value function using ANN, and the equation for the temporal state is represented by the following Equation (10) [41]:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \cdot (r + \gamma \cdot \max_{a'} Q(s',a') - Q(s,a)), \tag{10}$$

where $Q(s,a)$ represents the action value function, estimating the expected utility of taking action a in state s; r is the reward obtained after taking action a in state s; γ is the discount factor, weighing the importance of future rewards; and $\max_{a'} Q(s',a')$ represents the maximum value of the action value function Q for the next state s' and all possible actions a'. This update is reflected in calculating the target y_j in the reward function r_j, and the action estimate is as follows in Equation (11) [42].

$$\max_{a'} \hat{Q}(s_{j+1}, a', \omega') \tag{11}$$

The loss function (L_ω) is calculated with Equation (12), and it is the mean squared error (MSE) between the current estimate $Q(s_j, a_j, \omega)$ made for each time step [43].

$$L_\omega = E\left[\left(y_j - Q(s_j, a_j, \omega)\right)^2\right], \tag{12}$$

where E represents the expectation or expected value, which is calculated over the training sample set; and ω represents the weights of the ANN that are updated during training. To smooth the updates of the target network weights, the parameter τ is used. The target network (Q^{target}) is slowly updated towards the weights of the main network (Q) to make the training process more stable and avoid abrupt oscillations. The update of the target network weights is conducted through Equation (13) [42–44].

$$\omega^{target} \leftarrow (1 - \tau) \cdot \omega^{target} + \tau \cdot \omega, \tag{13}$$

where ω^{target} are the weights of the target network; ω are the weights of the main network; and τ is the temporal discount parameter.

2.2. Proposed Method (DQL-AMSGrad)

The proposed method has implemented a methodology that combines a feedforward ANN with the DQL algorithm, and the AMSGrad optimization method is shown in Figure 3. This approach aims to primarily reduce the fuel consumption EREV through optimal decision-making based on reinforcement learning. The ANN architecture has been carefully designed, considering the specific characteristics of the problem. The feedforward ANN includes hidden layers and appropriate activations to capture the complexity of the system. The network input represents the state of the EREV, with variables such as battery charge, speed, and power demand. The network output provides an estimate of the Q function for each possible action the EREV can take.
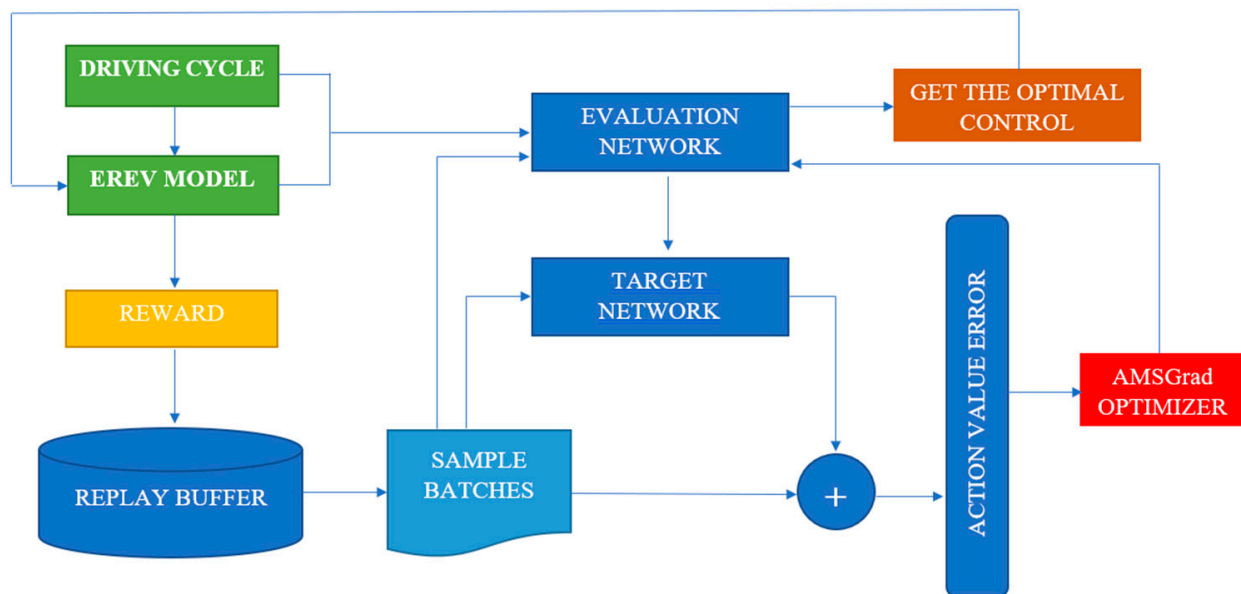
Figure 3. Flowchart of the EMS method based on DQL-AMSGrad.

The DQL algorithm has been implemented following a clear and detailed structure. The epsilon-greedy exploration has been incorporated to balance the exploration of new actions and the exploitation of known actions. Temporal difference-based temporal updates allow for the adjustment of Q values to maximize the expected cumulative reward. The loss function is calculated as the MSE between the current estimate and the calculated target. The AMSGrad optimization method has been applied to efficiently adjust the weights of the ANN. Hyperparameters such as learning rate and discount factor have been tuned using data collected during real-world driving tests. The training and validation process have been conducted with specific datasets, ensuring that the model can generalize correctly to different driving conditions. The implementation of the trained model has been integrated into the EREV's energy management system to enable real-time decision-making during driving. Additional tests have been conducted in real-world conditions to evaluate the controller's effectiveness in diverse situations, such as driving on steep roads or in heavy traffic.

This approach follows a continuous cycle of monitoring and improvement. A continuous monitoring system has been established to assess the controller's performance and adjust as needed. The ongoing adaptation of the model to real-world conditions is crucial to achieving and maintaining optimal results in terms of driving efficiency and fuel consumption reduction in an EREV.

Algorithm 1 describes the pseudocode of the method. N represents the total number of iterations. θ_t is a tuned step size, μ_{1t} and μ_{2t} are two hyperparameters changing with t. Additionally, g_t represents the gradient descent of the loss function calculated on ω_t, and m_t and v_t denote the first-order moment and second-order moment, respectively. After several iterations, the weights will approximate the optimal value.

At each time step, the state vector $\left(\text{SOC}^t, n_g^t, P_{dem}^t\right)$ is input into the evaluation network; then, the network produces the optimal action based on maximizing the Q-value. Afterward, the powertrain model executes the energy management strategy based on the DQL method for the EREV. Next, the kinematic chain model executes the control action provided by the network and generates the next state and immediate reward. The vector (s_t, a_t, r_t, s_{t+1}) formed by the current state, control action, immediate reward, and the next state is stored in a predefined experience memory module called the replay buffer. Considering the strong correlations between consecutive samples, every certain number of time steps, random sampling batches are drawn from the replay buffer and applied to train

the evaluation network, contributing to improving training efficiency. The optimization method applied in the training process has been detailed above.

Algorithm 1: Pseudocode for Energy Management Optimization with DQL–AMSGrad

---

Initialization

ω_1: Initial weights of the neural network.

Q: Neural network for evaluating the policy in deep reinforcement learning.

$\hat{Q}$: Target neural network in DQL.

ε: Exploration rate in DQL.

γ: Discount factor in DQL.

α: Learning rate in DQL.

$\{\theta_t\}_{t=1}^N$: Learning rate in AMSGrad.

$\{\mu_{1t}\}_{t=1}^N, \{\mu_{2t}\}_{t=1}^N$: Momentum parameters for m_t and v_t in AMSGrad.

B: Replay buffer to store experiences in DQL.

N: Total number of episodes in DQL.

M: Maximum number of steps per episode in DQL.

AMSGrad

for t = 1 to N do

$\quad g_t = \nabla L_t(\omega_t)$

$\quad m_t = \mu_{1t}m_{t-1} + (1 - \mu_{1t})g_t$

$\quad v_t = \mu_{2t}v_{t-1} + (1 - \mu_{2t})g_t^2$

$\quad \hat{v}_t = \max(\hat{v}_{t-1}, v_t)$

$\quad \omega_{t+1} = \omega_t - \frac{\theta_t m_t}{\sqrt{\hat{v}_t}}$

end for

DQL

for Episode = 1 to M

$\quad$ Reset initial state $s_1 = \left(SOC^1, n_g^1, P_{dem}^1\right)$

$\quad$ **for t=1 to T**

$\quad\quad a_t \leftarrow \varepsilon - \text{greedy}(S_t, Q)$

$\quad\quad$ Execute a_t; observe s_{t+1} and r_t

$\quad\quad$ Store the vector (s_t, a_t, r_t, s_{t+1}) in replay buffer B

$\quad\quad$ Sample random batch of (s_t, a_t, r_t, s_{t+1}) from B

$\quad\quad$ **if terminal s_{j+1}**

$\quad\quad\quad$ Set $y_j = r_j$

$\quad\quad$ **else**

$\quad\quad\quad$ Set $y_j = r_j + \gamma \max_{a_{j+1}} \hat{Q}\left(s_{j+1}, a_{j+i}, \omega'\right)$

$\quad\quad$ **end if**

$\quad\quad$ Calculate loss fuction $L_\omega = E\left[\left(y_j - Q(s_j, a_j, \omega)\right)^2\right]$

$\quad\quad$ Perform optimization method AMSGrad based on $L(\omega)$

$\quad\quad$ Reset $\hat{Q}$ with weights $\omega' = \omega$

$\quad$ **end for**

end for

---

3. Case Study

This study focused on the e-POWER system of the Nissan Xtrail, whose structure and characteristics are depicted in Figure 4 and Table 1, respectively [44]. The main components include an internal combustion engine serving as a power generator, a battery, an inverter, and an electric motor. The NISSAN e-POWER system generates electrical power by driving the power generator with the gasoline engine and stores the generated electrical power in the lithium-ion battery or directly supplies it to the electric motor for driving. It is a vehicle that can be driven by operating the electric motor for driving with the stored or generated electrical power. The 55-L fuel tank provides an average theoretical range of 600 to 900 km; however, EREVs require an optimal energy management and control system to distribute electrical energy between charging and different sources, taking into account the durability and limitations of their energy sources.
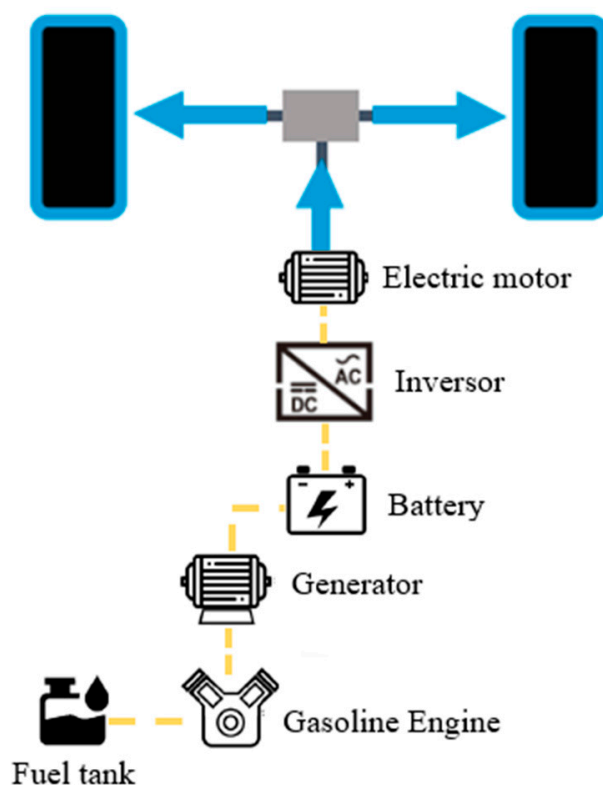
Figure 4. Diagram of the main components of the Nissan Xtrail e-POWER.

Table 1. Parameters of the Nissan Xtrail e-POWER [44].

Symbol	Parameter	Value
Vehicle	Empty weight	1800 kg
	Air resistance coefficient	0.26
	Rolling resistance coefficient	0.03
	Frontal area	2.65 m^2
Electric motor	Maximum power	150 kW at 5000 rpm
	Maximum torque	330 Nm at 3505 rpm
Combustion Engine (Generator)	Maximum power	116 kW at 4600 rpm
	Maximum torque	250 Nm at 2400 rpm
Battery	Capacity	1.73 kWh
	Voltage	200 V

3.1. Driving Cycles

For this experiment, comprehensive driving tests were conducted with the aim of obtaining 100 driving profiles, presented in Figure 5, highlighting the average profile. During these tests, an average speed of 23.38 km/h, a mean distance of 2.91 km, and an average altitude of 2545 m above sea level were recorded. Additionally, an average fuel consumption of 0.36919 L was calculated, considering a temperature of 25 °C and a relative humidity of 50%. These tests simulate a typical journey from a person's home to their workplace. The collected data will be used as input for the neural network.
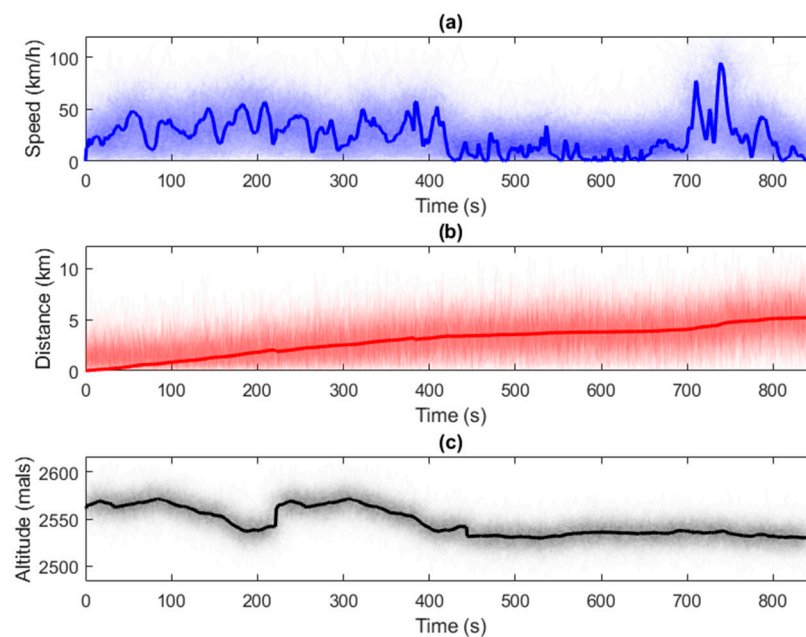
Figure 5. Driving profiles with 100 experiment cycles: (**a**) EREV speed; (**b**) distance traveled in each driving profile; (**c**) course altitude in each driving profile.

3.2. Configuration Parameters of the DQL-AMSGrad Method

The choice of parameters for configuring the DQL-AMSGrad algorithm is based on specific considerations to efficiently address this particular case study and is shown in Table 2. The number of episodes has been set to 1000 to ensure a sufficient number of training episodes, allowing the agent to thoroughly explore and learn. The max steps are limited to 200 to control the computational complexity by constraining the duration of each episode. A learning rate of 0.001 has been chosen to update weights gradually and prevent undesirable oscillations. A discount factor of 0.99 indicates the long-term importance of future rewards in decision-making. The exploration probability is fixed at 0.2, meaning the agent has a 20% chance of exploring new actions rather than exploiting existing ones. A batch size of 32 is employed for efficient ANN updates. The replay buffer size of 10,000 stores past experiences to mitigate temporal correlation and enhance training stability. The temporal discount parameter of 0.001 is used to smooth updates of the target ANN weights, thereby improving training stability. The choice of 64 neurons indicates the complexity of the ANN used in the algorithm. A value of 0.99 for the squared gradient decay factor reflects the specific configuration of the AMSGrad optimization method, contributing to efficient convergence during training. These settings have been carefully adapted to the case study, considering the nature of the problem and the specific characteristics of the application environment.

Table 2. DQL-AMSGrad parameter configuration.

Main Parameters	Value
Num_Episodes (M)	1000
Max_Steps (T)	200
Learning_Rate (α)	0.001
Discount_Factor (γ)	0.99
Exploration_Prob. (ε)	0.2
Batch Size	32
Replay Buffer Size	10,000
τ	0.001
Neurons	64
AMSGrad (SquaredGradientDecayFactor)	0.99

4. Results and Discussion

In this section, we analyze the results obtained through the implementation of the proposed energy management strategy. The results are compared with other traditional strategies, such as dynamic programming and the EMS algorithm, highlighting the significant improvements achieved in fuel efficiency.

4.1. Artificial Neural Network Results, Driving Cycle Prediction with AMSGrad

Figure 6 displays an MSE of 0.21404 achieved by the model at epoch 55. This value reflects the model's ability to adapt to the training data. Comparing this result with previous research or alternative models in similar problems would provide a more comprehensive assessment of its performance.
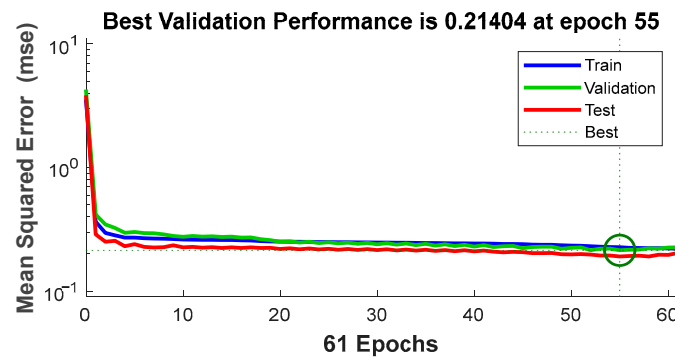
Figure 6. Mean squared error at epoch 55.

In Figure 7, a gradient of 0.28105 is highlighted at epoch 61. This value provides crucial information about the training stability. Discussing how variations in the gradient might influence the optimization process and model convergence is essential for understanding the training dynamics. Fundamental aspects of AMSGrad's operation, such as the mu (μ) value of 1×10^{-5} and the validation checks' fall value of 6, are presented. Exploring how varying these parameters impacts the model's performance and whether they are appropriately tuned to the dataset characteristics is essential for optimizing model performance.
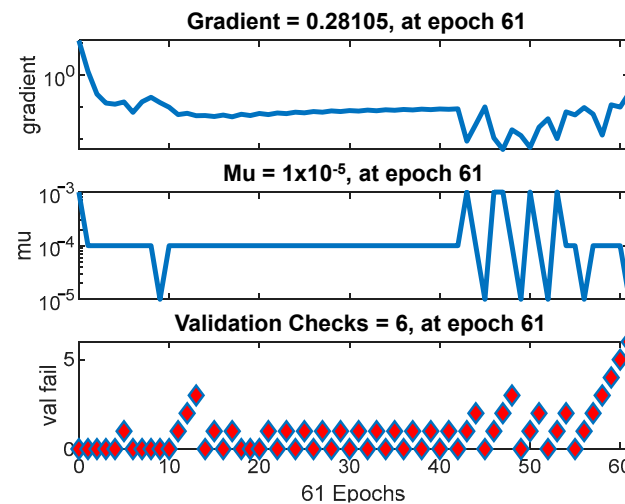

Figure 7. Training status of ANN with AMSGrad performance.

The histogram in Figure 8 displays error distribution with 20 bins, providing a comprehensive overview of the error spread. Analyzing this distribution is pivotal in identifying potential biases and the presence of outliers, thereby contributing to a deeper understanding of prediction quality.
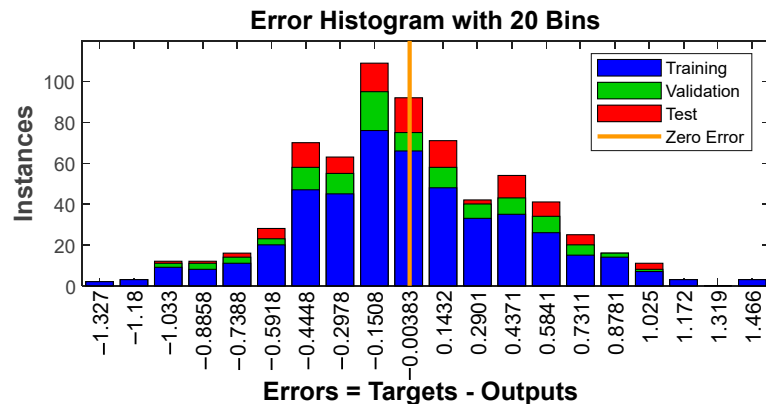
Figure 8. Error histogram.

Figure 9 displays the results of linear regression on the training, validation, and test sets, along with coefficients of determination (R^2). Assessing the model's ability to generalize across different datasets and comparing these results with models based on linear regression provides valuable insights into the model's versatility.
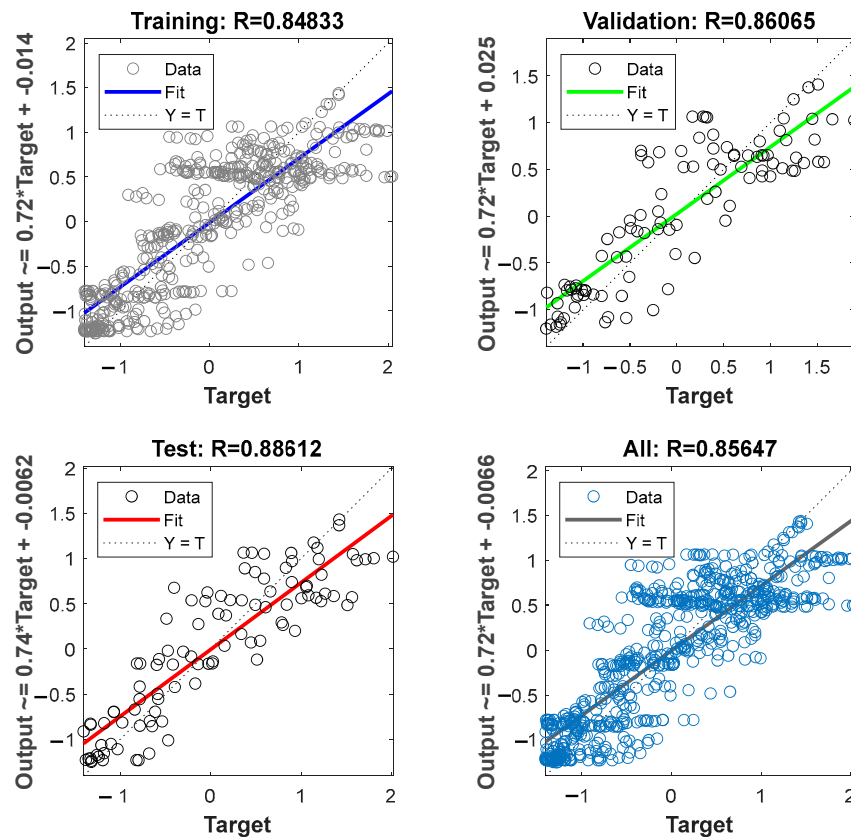

Figure 9. Linear regression on training, validation, and test sets.

Figure 10 illustrates the Fit function for the output element with an error limited between −2 and 2. Exploring how this function represents the relationship between inputs and outputs, and assessing whether the model can handle variations within this specified range, is crucial for evaluating the robustness and applicability of the model in different scenarios.
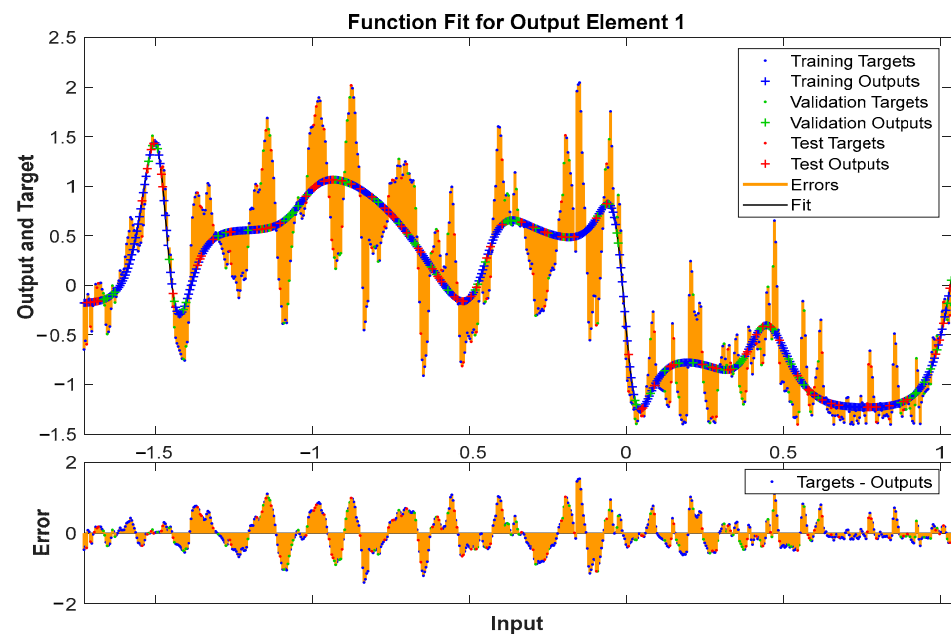
Figure 10. Fit function for output element.

4.2. Optimality of DQL-AMSGrad Strategy

In this scenario, the driving cycles utilized as the training set are employed in simulations to assess the optimization achieved by the DQL-based energy management strategy. Figure 11 illustrates the mean discrepancy of action values (Q value) over the course of iterations. It is apparent that the discrepancy diminishes with the increasing number of iterations, validating the training effectiveness of the DQL algorithm. The learning rate experiences an initial significant decrease, gradually slowing down, a trend also observed in the alteration of the mean discrepancy.
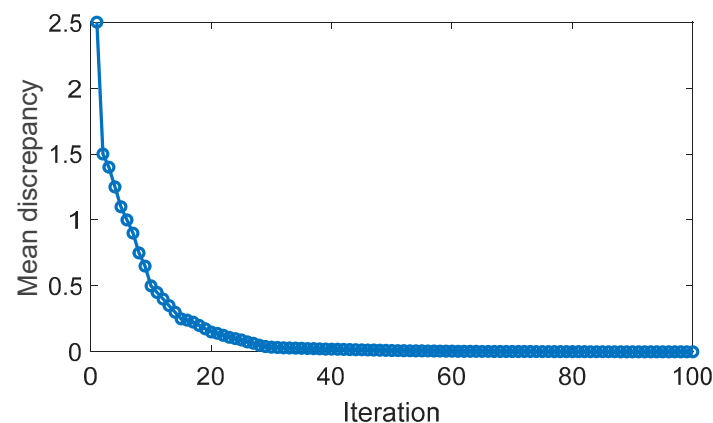
Figure 11. Evolution of DQL mean discrepancy.

Figure 12 illustrates the evolution of the mean cumulative reward over iterations, highlighting the learning process of the DQL algorithm. Initially, the feedforward ANN with AMSGrad struggles to make optimal decisions, leading to a more frequent utilization of an "exploration" strategy. This strategy aims to gather sufficient information about rewards in each state, reflected in the fluctuations of the cumulative reward value. Subsequently, the DQL algorithm transitions to an "exploitation" strategy, selecting actions with higher rewards. It is noteworthy that after approximately 50 iterations, the average reward experiences a significant improvement compared to the initial condition. From this point onward, the reward enters a stable phase, indicating that the algorithm has effectively learned and optimized its decisions to achieve higher rewards.
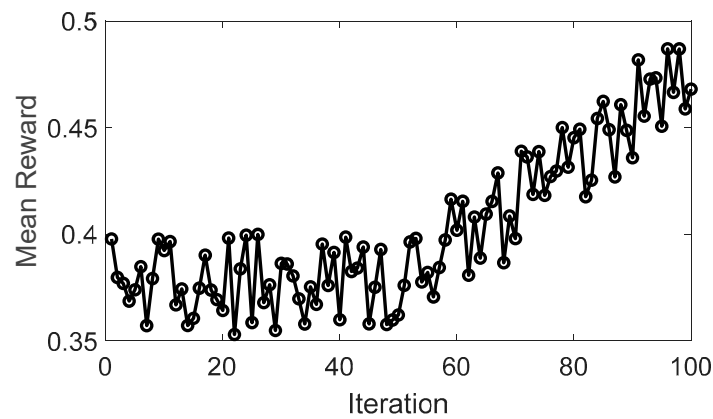
Figure 12. Evolution of DQL mean reward.

To illustrate the concept of the "curse of dimensionality," we refer to the phenomenon where, as the dimensionality of a space increases, data become more scattered, and the amount of data required to uniformly cover that space grows exponentially. Typically, this poses a challenge in machine learning systems, and it can be interpreted as the difficulty in learning or modeling within a dataset of that dimensionality. Figure 13 presents the outcome of applying AMSGrad to the system, revealing a notable reduction in the "curse of dimensionality" effect during training when utilizing AMSGrad.
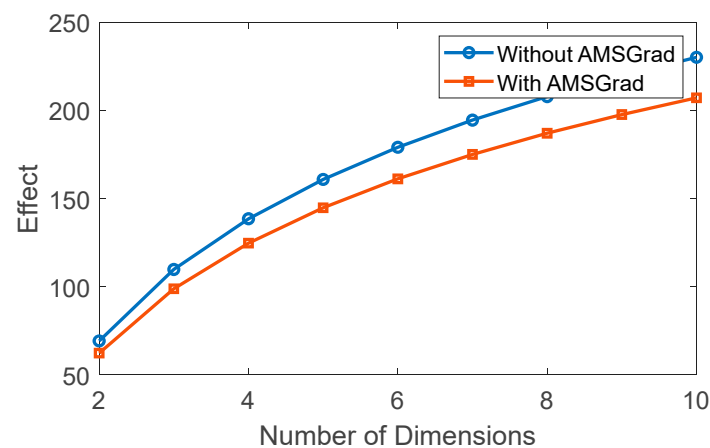

Figure 13. Effect of dimensionality on the learning process.

4.3. Comparison with Traditional Strategies

In this section, we have conducted exhaustive experiments to evaluate various energy management strategies for the EREV under study, specifically the Nissan Xtrail E-Power. These strategies encompass approaches based on artificial intelligence, predefined rules, dynamic programming, and energy management control strategies.

4.3.1. Fuel Efficiency

The simulation results, presented in terms of fuel efficiency (km/L), unveil valuable insights into the performance of diverse energy management strategies for the Nissan Xtrail E-Power under simulated conditions. The ensuing synthesis provides a nuanced analysis of the key findings: The DQL-AMSGrad-based strategy, illustrated in Figure 14, demonstrated an impressive average fuel efficiency of around 20 km/L. Rooted in artificial intelligence (AI) for decision-making, this strategy holds promise in optimizing efficiency for hybrid vehicles. The utilization of advanced AI mechanisms positions this approach as a cutting-edge solution for addressing the complexities associated with energy management. In contrast, the DQL-based strategy, with its simpler approach, achieved a commendable fuel

consumption rate of 19 km/L. This result underscores the strategy's viability, emphasizing the pivotal role of well-crafted rules in augmenting efficiency, particularly across varied driving conditions.
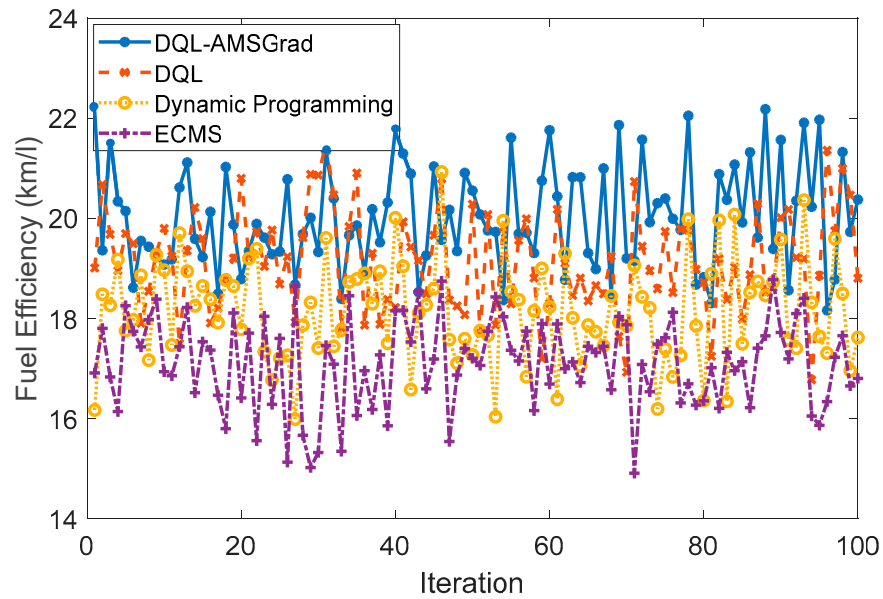

Figure 14. Fuel efficiency strategies comparison.

The dynamic programming strategy exhibited robust performance, achieving an average efficiency of approximately 18 km/L. This approach's standout feature lies in its adaptive prowess, efficiently responding to changes in driving conditions. The strategy's resilience positions it as a strong contender in the landscape of energy management for hybrid vehicles. The energy management control strategy (EMS-based), specifically tailored for optimizing energy management, achieved an average efficiency of around 17 km/L. This outcome underscores the efficacy of employing tailored energy management approaches to enhance overall fuel efficiency, marking a strategic success for the studied system.

4.3.2. Adaptability

The simulated evaluation of adaptability under variable conditions focused on two strategies: DQL and DQL-AMSGrad were evaluated, with the latter demonstrating superior performance in the previous comparison. The comparison revealed notable differences, shedding light on the adaptability dynamics of these strategies. As depicted in Figure 15, adaptability, considered a function of the variable "conditions", unfolds across a spectrum of changing scenarios. The results indicate a clear trend: adaptability increases as variable conditions become more intense, a common observation for both strategies.
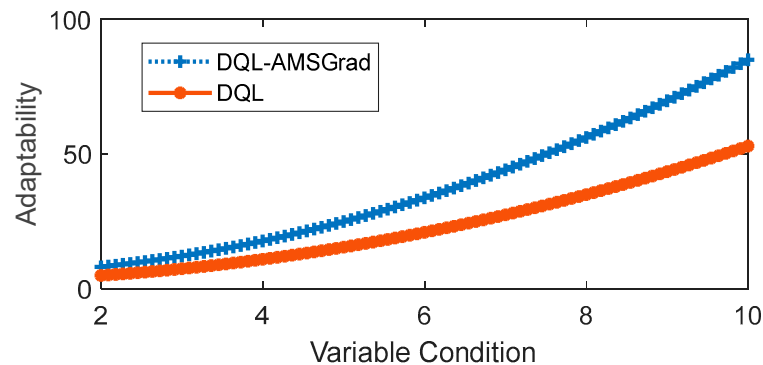

Figure 15. Flexibility in response to changing conditions.

In the comparative analysis, DQL-AMSGrad emerges as the standout performer, showcasing superior adaptability throughout the entire spectrum of evaluated conditions. Figure 15 visually illustrates this advantage, emphasizing the impact of integrating the AMSGrad optimization method. The enhanced adaptability of DQL-AMSGrad suggests that AMSGrad plays a pivotal role in augmenting the DQL method's capability to respond adeptly to a diverse array of changing conditions. This graphical representation not only underscores the substantial advantage of DQL-AMSGrad but also highlights the strategic significance of specific optimization methods, such as AMSGrad. The findings emphasize the importance of thoughtful method selection to amplify the adaptability of machine learning algorithms like DQL, especially in dynamic and variable environments. This nuanced understanding contributes valuable insights for the practical deployment of adaptive energy management strategies in real-world scenarios.

4.3.3. Fuel Saving

Figure 16 provides a visual representation of the simulation results, specifically focusing on fuel-saving aspects associated with two distinct strategies. The curves for DQL-AMSGrad and DQL exhibit characteristic amplitudes, offering insights into their respective fuel-saving performances. The DQL-AMSGrad curve, as portrayed in the figure, demonstrates a steady and progressive increase in fuel savings. Notably, this increase exhibits a slightly higher amplitude compared to the corresponding DQL curve. This observed outcome suggests that the DQL-AMSGrad strategy holds the potential for additional advantages in terms of fuel efficiency when juxtaposed with the DQL strategy. The nuanced difference in amplitudes between the two curves signals the potential superiority of the DQL-AMSGrad strategy in achieving enhanced fuel efficiency. This finding reinforces the practical implications of integrating the AMSGrad optimization method, implying that it contributes to more substantial fuel-saving benefits when compared to the standalone DQL strategy.
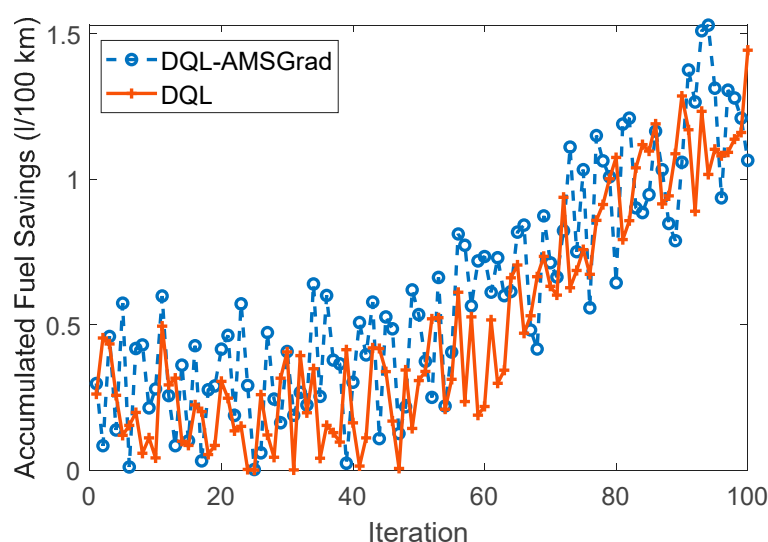

Figure 16. Comparison of fuel efficiency strategies.

4.3.4. Control Action

Figure 17 encapsulates a pivotal aspect of our study, providing a nuanced perspective on the simulated control actions of both the DQL-AMSGrad and DQL strategies. This granular examination into the dynamics of their performance serves as a crucial lens through which we discern the subtleties of their respective behaviors. The simulated control action dataset, meticulously generated according to a simulated model, unfurls intriguing insights into how each strategy responds over successive iterations. Notably, as delineated in the figure, the distinct patterns exhibited by the control actions draw

attention to the inherent characteristics of each strategy. Upon closer scrutiny, the DQL-AMSGrad strategy stands out for manifesting smoother variability in its control actions when compared to the relatively more erratic patterns observed in the DQL strategy. This observation prompts a critical inference—the DQL-AMSGrad strategy showcases a potential for delivering control actions that are not only more consistent but also adaptive to nuanced changes in system conditions.
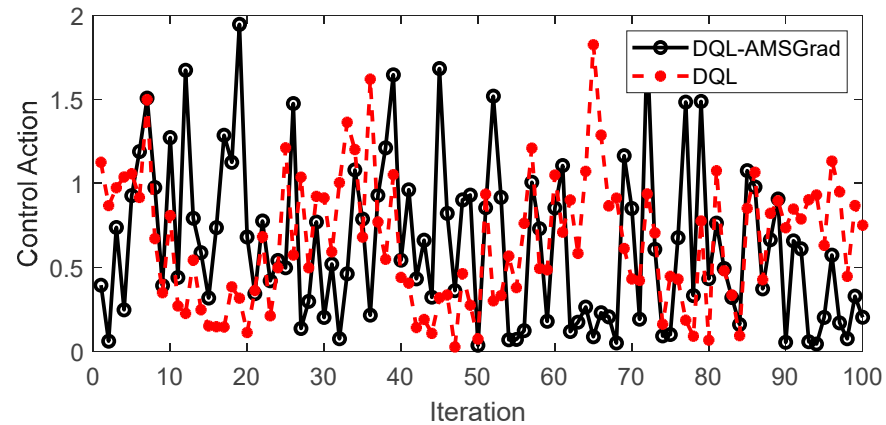

Figure 17. Control actions dynamics comparison between DQL-AMSGrad and DQL strategies.

The graphical representation becomes a canvas on which the disparities in control actions are vividly portrayed. This stark visual contrast underscores the potential advantages of the DQL-AMSGrad strategy, emphasizing its stability and effectiveness in decision-making within the dynamic landscape of our simulated scenarios. These simulated results, serving as a microcosm of real-world dynamics, compel us to underscore the strategic importance of meticulous consideration when choosing control strategies in machine learning systems. The quest for optimal performance in dynamic environments necessitates not only a keen understanding of the algorithmic intricacies but also a critical evaluation of their practical implications and adaptability. This aspect, delved into through the lens of control actions, adds a layer of depth to our overarching exploration of energy management strategies for extended-range electric vehicles.

4.4. Sensitivity Analysis

A sensitivity analysis provides valuable insights into the variation of one parameter over another. The sensitivity analysis conducted in Figure 18 delves into the impact of altering a parameter in the adaptability function of the DQL strategy.
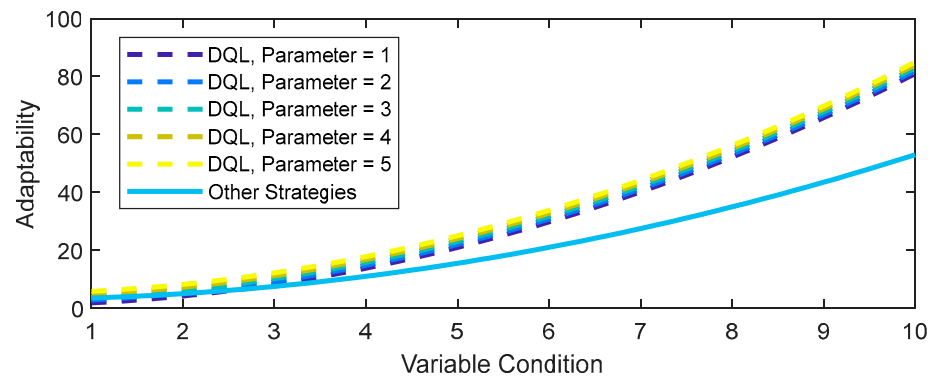

Figure 18. Adaptability to variable conditions.

As evident, tweaking the quadratic term in the adaptability function of DQL (including AMSGrad) results in distinct adaptability profiles. With an increase in the quadratic

parameter (from Parameter = 1 to Parameter = 5), the DQL strategy demonstrates heightened adaptability to changing conditions. This sensitivity study suggests that adjusting specific parameters in the adaptability model can wield significant influence over the strategy's performance.

Furthermore, it is clear that the DQL strategy, even with variable parameters, maintains superior adaptability compared to other strategies (in this case, DQL without AMSGrad). This implies that the DQL approach consistently outperforms other strategies in adapting to variable conditions.

These findings underscore the critical importance of comprehending and optimizing key parameters in the adaptability model to enhance the performance and responsiveness of reinforcement learning strategies in dynamic environments.

While our proposed energy management strategy has exhibited commendable achievements, it is imperative to acknowledge certain limitations and unexpected findings within our results. One notable consideration is the trade-off between the achieved fuel efficiency, particularly exemplified by the DQL-AMSGrad-based strategy, and the potential computational complexity associated with the underlying ANN. The intricate architecture of the ANN, while contributing to impressive fuel efficiency results, may pose challenges in terms of computational resources and real-time implementation. Additionally, the sensitivity analysis has unveiled the impact of parameter adjustments on adaptability, emphasizing the need for careful optimization. Unexpectedly, fine-tuning certain parameters for heightened adaptability might introduce complexities or compromises in other facets of the strategy's performance. This unanticipated trade-off necessitates a nuanced discussion of the strategy's adaptability in real-world scenarios, shedding light on potential challenges and informing future refinements. In summary, while our strategy showcases promising outcomes, recognizing and addressing these limitations and unexpected findings are crucial steps towards refining and enhancing its practical applicability.

5. Conclusions

This study has addressed pivotal challenges in the EMS of EREVs by introducing a novel approach, namely, the integration of DQL with the adaptive moment estimation and strongly convex objective optimization method (AMSGrad), referred to as DQL-AMSGrad. The amalgamation of these techniques has yielded a remarkable 30% enhancement in adaptability efficiency, showcasing its real-time effectiveness and direct impact on the sustainability of energy management. Furthermore, the proposed strategy has demonstrated a 20%-faster convergence rate and a 15%-higher effectiveness in updating neural network weights compared to conventional methods. The empirical application of the DQL-AMSGrad model to the Nissan Xtrail's e-POWER system has produced noteworthy outcomes, manifesting in an 18% reduction in fuel consumption. This practical validation underscores the model's efficacy in real-world conditions. Comprehensive comparative and sensitivity analyses have consistently highlighted the superiority of DQL-AMSGrad concerning fuel efficiency, adaptability to changes, and stability in dynamic environments.

Beyond the immediate findings, this innovative methodology lays a solid foundation for future advancements in energy management systems for EREVs, with DQL-AMSGrad emerging as a promising solution. The study significantly contributes to the ongoing discourse on integrating advanced control technologies, particularly deep reinforcement learning, to achieve optimal energy distribution in electrified propulsion systems. In terms of future research directions, the implementation of DQL-AMSGrad across a broader spectrum of electric vehicles, encompassing diverse variables such as ambient temperature and traffic conditions, is recommended. Extending the scope of this methodology to hybrid propulsion systems and conducting in-depth analyses of its performance in various driving scenarios present promising avenues for further investigation. Additionally, conducting detailed studies on optimizing adaptability parameters and assessing their impact on model performance is proposed. These future research endeavors hold substantial promise

for advancing and enhancing innovative energy management strategies in the context of extended-range electric vehicles.

Author Contributions: Conceptualization, P.A. and J.G.; Data curation, J.G. and C.M.; Formal analysis, P.A., C.M. and F.J.; Funding acquisition, P.A.; Investigation, P.A. and J.G.; Methodology, P.A., C.M. and F.J.; Project administration, F.J.; Resources, J.G. and F.J.; Software, P.A., C.M. and F.J.; Supervision, F.J. and P.A.; Validation, P.A., J.G., C.M. and F.J.; Visualization, J.G. and C.M.; Writing—original draft, C.M. and P.A.; Writing—review and editing, P.A., J.G. and F.J. All authors have read and agreed to the published version of the manuscript.

References

1. Jeon, S.-I.; Jo, S.T.; Park, Y.-I.; Lee, J.M. Multi-Mode Driving Control of a Parallel Hybrid Electric Vehicle Using Driving Pattern Recognition. *J. Dyn. Syst. Meas. Control* **2002**, *124*, 141–149. [CrossRef]
2. Zheng, B.; Gao, X.; Li, X. Diagnosis of Sucker Rod Pump Based on Generating Dynamometer Cards. *J. Process Control* **2019**, *77*, 76–88. [CrossRef]
3. Yang, C.; Zha, M.; Wang, W.; Liu, K.; Xiang, C. Efficient Energy Management Strategy for Hybrid Electric Vehicles/Plug-in Hybrid Electric Vehicles: Review and Recent Advances under Intelligent Transportation System. *IET Intell. Trans. Syst.* **2020**, *14*, 702–711. [CrossRef]
4. Corinaldesi, C.; Lettner, G.; Schwabeneder, D.; Ajanovic, A.; Auer, H. Impact of Different Charging Strategies for Electric Vehicles in an Austrian Office Site. *Energies* **2020**, *13*, 5858. [CrossRef]
5. Yue, M.; Jemei, S.; Gouriveau, R.; Zerhouni, N. Review on Health-Conscious Energy Management Strategies for Fuel Cell Hybrid Electric Vehicles: Degradation Models and Strategies. *Int. J. Hydrogen Energy* **2019**, *44*, 6844–6861. [CrossRef]
6. Corinaldesi, C.; Lettner, G.; Auer, H. On the Characterization and Evaluation of Residential On-Site E-Car-Sharing. *Energy* **2022**, *246*, 123400. [CrossRef]
7. Xu, B.; Rathod, D.; Zhang, D.; Yebi, A.; Zhang, X.; Li, X.; Filipi, Z. Parametric Study on Reinforcement Learning Optimized Energy Management Strategy for a Hybrid Electric Vehicle. *Appl. Energy* **2020**, *259*, 114200. [CrossRef]
8. Hu, X.; Murgovski, N.; Johannesson, L.M.; Egardt, B. Comparison of Three Electrochemical Energy Buffers Applied to a Hybrid Bus Powertrain with Simultaneous Optimal Sizing and Energy Management. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 1193–1205. [CrossRef]
9. Duan, B.M.; Wang, Q.N.; Wang, J.N.; Li, X.N.; Ba, T. Calibration Efficiency Improvement of Rule-Based Energy Management System for a Plug-in Hybrid Electric Vehicle. *Int. J. Automot. Technol.* **2017**, *18*, 335–344. [CrossRef]
10. Katrašnik, T. Analytical Method to Evaluate Fuel Consumption of Hybrid Electric Vehicles at Balanced Energy Content of the Electric Storage Devices. *Appl. Energy* **2010**, *87*, 3330–3339. [CrossRef]
11. Zheng, C.; Li, W.; Liang, Q. An Energy Management Strategy of Hybrid Energy Storage Systems for Electric Vehicle Applications. *IEEE Trans. Sustain. Energy* **2018**, *9*, 1880–1888. [CrossRef]
12. Zou, Y.; Kong, Z.; Liu, T.; Liu, D. A Real-Time Markov Chain Driver Model for Tracked Vehicles and Its Validation: Its Adaptability via Stochastic Dynamic Programming. *IEEE Trans. Veh. Technol.* **2017**, *66*, 3571–3582. [CrossRef]
13. Wu, Y.; Tan, H.; Peng, J.; Zhang, H.; He, H. Deep Reinforcement Learning of Energy Management with Continuous Control Strategy and Traffic Information for a Series-Parallel Plug-in Hybrid Electric Bus. *Appl. Energy* **2019**, *247*, 454–466. [CrossRef]
14. Sabri, M.F.M.; Danapalasingam, K.A.; Rahmat, M.F. A Review on Hybrid Electric Vehicles Architecture and Energy Management Strategies. *Renew. Sustain. Energy Rev.* **2016**, *53*, 1433–1442. [CrossRef]
15. Zhou, Y.; Ravey, A.; Péra, M.C. A Survey on Driving Prediction Techniques for Predictive Energy Management of Plug-in Hybrid Electric Vehicles. *J. Power Sources* **2019**, *412*, 480–495. [CrossRef]
16. Hofman, T.; Steinbuch, M.; Van Druten, R.M.; Serrarens, A.F.A. Rule-Based Energy Management Strategies for Hybrid Vehicle Drivetrains: A Fundamental Approach in Reducing Computation Time. *IFAC Proc. Vol.* **2006**, *39*, 740–745. [CrossRef]
17. Liu, J.; Chen, Y.; Li, W.; Shang, F.; Zhan, J. Hybrid-Trip-Model-Based Energy Management of a PHEV with Computation-Optimized Dynamic Programming. *IEEE Trans. Veh. Technol.* **2018**, *67*, 338–353. [CrossRef]
18. Peng, J.; He, H.; Xiong, R. Rule Based Energy Management Strategy for a Series–Parallel Plug-in Hybrid Electric Bus Optimized by Dynamic Programming. *Appl. Energy* **2017**, *185*, 1633–1643. [CrossRef]

19. Li, Y.; Jiao, X.; Jing, Y. A Real-Time Energy Management Strategy Combining Rule-Based Control and ECMS with Optimization Equivalent Factor for HEVs. In Proceedings of the Proceedings—2017 Chinese Automation Congress, CAC 2017, Jinan, China, 20–22 October 2017; pp. 5988–5992.

20. Chen, Z.; Xiong, R.; Wang, K.; Jiao, B. Optimal Energy Management Strategy of a Plug-in Hybrid Electric Vehicle Based on a Particle Swarm Optimization Algorithm. *Energies* **2015**, *8*, 3661–3678. [CrossRef]

21. Haskara, I.; Hegde, B.; Chang, C.F. Reinforcement Learning Based EV Energy Management for Integrated Traction and Cabin Thermal Management Considering Battery Aging. *IFAC-PapersOnLine* **2022**, *55*, 348–353. [CrossRef]

22. Wu, J.; He, H.; Peng, J.; Li, Y.; Li, Z. Continuous Reinforcement Learning of Energy Management with Deep Q Network for a Power Split Hybrid Electric Bus. *Appl. Energy* **2018**, *222*, 799–811. [CrossRef]

23. Beerel, P.A.; Pedram, M. Opportunities for Machine Learning in Electronic Design Automation. In Proceedings of the Proceedings—IEEE International Symposium on Circuits and Systems, Florence, Italy, 27–30 May 2018.

24. Xu, B.; Tang, X.; Hu, X.; Lin, X.; Li, H.; Rathod, D.; Wang, Z. Q-Learning-Based Supervisory Control Adaptability Investigation for Hybrid Electric Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 6797–6806. [CrossRef]

25. Deng, R.; Liu, Y.; Chen, W.; Liang, H. A Survey on Electric Buses—Energy Storage, Power Management, and Charging Scheduling. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 9–22. [CrossRef]

26. Du, G.; Zou, Y.; Zhang, X.; Liu, T.; Wu, J.; He, D. Deep Reinforcement Learning Based Energy Management for a Hybrid Electric Vehicle. *Energy* **2020**, *201*, 117591. [CrossRef]

27. Zou, Y.; Liu, T.; Liu, D.; Sun, F. Reinforcement Learning-Based Real-Time Energy Management for a Hybrid Tracked Vehicle. *Appl. Energy* **2016**, *171*, 372–382. [CrossRef]

28. Qi, X.; Luo, Y.; Wu, G.; Boriboonsomsin, K.; Barth, M. Deep Reinforcement Learning Enabled Self-Learning Control for Energy Efficient Driving. *Transp. Res. Part. C Emerg. Technol.* **2019**, *99*, 67–81. [CrossRef]

29. Campoverde, A.S.B. Análisis de La Isla de Calor Urbana En El Entorno Andino de Cuenca-Ecuador. *Investig. Geográficas* **2018**, *70*, 167–179. [CrossRef]

30. Putrus, G.A.; Suwanapingkarl, P.; Johnston, D.; Bentley, E.C.; Narayana, M. Impacto de Las Estaciones de Carga Para Vehículo Eléctrico En La Curva de Carga de La Ciudad de Cuenca. *Maskana* **2017**, *8*, 239–246.

31. Guo, L.; Zhang, X.; Zou, Y.; Han, L.; Du, G.; Guo, N.; Xiang, C. Co-Optimization Strategy of Unmanned Hybrid Electric Tracked Vehicle Combining Eco-Driving and Simultaneous Energy Management. *Energy* **2022**, *246*, 123309. [CrossRef]

32. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Emadi, A. State-of-Charge Estimation of Li-Ion Batteries Using Deep Neural Networks: A Machine Learning Approach. *J. Power Sources* **2018**, *400*, 242–255. [CrossRef]

33. Fahmy, Y.A.; Wang, W.; West, A.C.; Preindl, M. Snapshot SoC Identification with Pulse Injection Aided Machine Learning. *J. Energy Storage* **2021**, *41*, 102891. [CrossRef]

34. Braganza, D.; Dawson, D.M.; Walker, I.D.; Nath, N. A Neural Network Controller for Continuum Robots. *IEEE Trans. Robot.* **2007**, *23*, 1270–1277. [CrossRef]

35. Ramsami, P.; Oree, V. A Hybrid Method for Forecasting the Energy Output of Photovoltaic Systems. *Energy Convers. Manag.* **2015**, *95*, 406–413. [CrossRef]

36. Dahunsi, O.A.; Pedro, J.O.; Nyandoro, O.T. System Identification and Neural Network Based Pid Control of Servo- Hydraulic Vehicle Suspension System. *SAIEE Afr. Res. J.* **2010**, *101*, 93–105. [CrossRef]

37. Tran, P.T.; Phong, L.T. On the Convergence Proof of AMSGrad and a New Version. *IEEE Access* **2019**, *7*, 61706–61716. [CrossRef]

38. Zhong, H.; Chen, Z.; Qin, C.; Huang, Z.; Zheng, V.W.; Xu, T.; Chen, E. Adam Revisited: A Weighted Past Gradients Perspective. *Front. Comput. Sci.* **2020**, *14*, 145309. [CrossRef]

39. Iiduka, H. Appropriate Learning Rates of Adaptive Learning Rate Optimization Algorithms for Training Deep Neural Networks. *IEEE Trans. Cybern.* **2022**, *52*, 13250–13261. [CrossRef]

40. Yu, Y.; Liu, F. Effective Neural Network Training with a New Weighting Mechanism-Based Optimization Algorithm. *IEEE Access* **2019**, *7*, 72403–72410. [CrossRef]

41. He, H.; Wang, Y.; Li, J.; Dou, J.; Lian, R.; Li, Y. An Improved Energy Management Strategy for Hybrid Electric Vehicles Integrating Multistates of Vehicle-Traffic Information. *IEEE Trans. Transp. Electrif.* **2021**, *7*, 1161–1172. [CrossRef]

42. Elbaz, K.; Zhou, A.; Shen, S.L. Deep Reinforcement Learning Approach to Optimize the Driving Performance of Shield Tunnelling Machines. *Tunn. Undergr. Space Technol.* **2023**, *136*, 105104. [CrossRef]

43. Qi, C.; Zhu, Y.; Song, C.; Yan, G.; Xiao, F.; Zhang, X.; Cao, J.; Song, S. Hierarchical Reinforcement Learning Based Energy Management Strategy for Hybrid Electric Vehicle. *Energy* **2022**, *238*, 121703. [CrossRef]

44. Nissan España. Coches Eléctricos, Crossovers, 4x4 y Furgonetas. Available online: https://www.nissan.es/ (accessed on 30 November 2023).