Article

Carbon-Aware Reinforcement Learning for Cost-Optimal Energy Management of Hydrogen-Electric Heavy-Duty Trucks

Jiawei Li[1,*], Yue Luo[2] and Yunting Ling[3]

1 Henan University, Kaifeng, China
2 University of Southern California, Los Angeles, USA
3 University of California, San Diego, La Jolla, USA
* Correspondence: Jiawei Li, Henan University, Henan, China

Abstract: The decarbonization of heavy-duty road transportation is a critical component of global carbon neutrality strategies, particularly under the increasing adoption of carbon credit pricing mechanisms such as emissions trading systems. Hydrogen-electric hybrid heavy-duty trucks (H2-E HDTs) offer a promising pathway toward low-carbon freight transport; however, their operational cost-effectiveness strongly depends on coordinated energy management decisions that jointly account for energy prices and carbon emission costs. In this paper, we propose a Carbon-aware Reinforcement Learning based Cost-optimal Energy Management System (CRL-CEMS) for hydrogen-electric heavy-duty trucks operating under carbon pricing environments. The proposed framework explicitly integrates dynamic carbon credit prices into the reinforcement learning decision process, enabling adaptive optimization of hydrogen and electricity usage with respect to both energy economics and carbon economics. A real-time carbon emission estimation model is developed to map hydrogen and electricity consumption into monetary carbon costs based on upstream carbon intensities. The energy management problem is formulated as a continuous control task and solved using the Soft Actor-Critic algorithm, where the agent determines the optimal power split between the fuel cell and battery systems. Extensive simulation experiments are conducted under multiple driving cycles and carbon price scenarios. Simulation results show that CRL-CEMS reduces the total operating cost by 7.2% and lowers carbon emissions by 12.3% compared with a carbon-unaware reinforcement learning baseline, while maintaining stable battery operation. These findings demonstrate that carbon-aware reinforcement learning effectively balances economic and environmental objectives, providing a practical approach for low-carbon, cost-optimal hydrogen-electric heavy-duty truck operation and supporting intelligent fleet energy management under dynamic carbon pricing scenarios.

Keywords: Hydrogen-electric heavy-duty truck; Carbon pricing; Reinforcement learning; Energy management system; Soft Actor-Critic; Low-carbon transportation

Received: 30 April 2026
Revised: 14 June 2026
Accepted: 27 June 2026
Published: 03 July 2026

1. Introduction

The heavy-duty freight transportation sector plays a fundamental role in global economic activities, yet it remains one of the largest contributors to greenhouse gas emissions. With increasingly stringent carbon neutrality targets and the rapid expansion of carbon credit pricing mechanisms such as emissions trading systems (ETS), reducing the carbon footprint of heavy-duty trucks has become a critical challenge for both industry and policymakers. Traditional diesel-powered trucks face mounting regulatory and economic pressures, motivating the transition toward alternative powertrain technologies with lower carbon emissions.

Hydrogen-electric hybrid heavy-duty trucks (H2-E HDTs) have emerged as a promising solution due to their zero tailpipe emissions, high energy density, and flexible energy supply configurations. By combining hydrogen fuel cells with onboard battery systems, H2-E HDTs can dynamically allocate power between multiple energy sources to improve efficiency and operational flexibility. However, the economic viability of such vehicles is highly sensitive not only to hydrogen and electricity prices, but also to carbon credit pricing, which introduces an additional layer of financial cost associated with upstream carbon emissions from energy production. Existing energy management strategies for hybrid trucks primarily focus on minimizing fuel consumption or direct energy costs, while the economic impacts of carbon pricing are often overlooked.

Recent advances in reinforcement learning (RL) have enabled data-driven energy management strategies that can adapt to complex and time-varying operating conditions. RL-based energy management systems have demonstrated strong performance in handling nonlinear dynamics and long-term optimization objectives in hybrid vehicles. Nevertheless, most existing RL approaches remain carbon-agnostic, treating energy prices as the sole economic signal and ignoring carbon credit prices that increasingly influence fleet-level operational decisions. This limitation reduces their applicability in carbon-regulated transportation systems.

To address this gap, this paper proposes a Carbon-aware Reinforcement Learning based Cost-optimal Energy Management System (CRL-CEMS) for hydrogen-electric heavy-duty trucks operating under carbon credit pricing environments. The CRL-CEMS framework explicitly integrates dynamic carbon prices into the RL decision-making process, enabling the agent to jointly optimize hydrogen consumption, electricity usage, and carbon emission costs. By incorporating a real-time carbon emission estimation model, CRL-CEMS establishes a closed-loop connection between vehicle energy management and carbon financial mechanisms. The energy management problem is formulated as a continuous control task and solved using the Soft Actor-Critic algorithm, allowing stable and efficient learning under continuous action spaces.

The main contributions of this study are summarized as follows:

1. A carbon-aware energy management framework (CRL-CEMS) is proposed, explicitly embedding carbon credit pricing into reinforcement learning for hydrogen-electric heavy-duty trucks.
2. A real-time carbon emission and cost modeling approach is developed to map energy consumption into monetary carbon costs based on upstream carbon intensities.
3. A continuous-control RL formulation using the Soft Actor-Critic algorithm is designed to achieve cost-optimal power allocation under dynamic energy and carbon price signals.
4. Comprehensive simulation experiments are conducted to demonstrate the effectiveness of the proposed framework in reducing both total operating costs and carbon emissions under multiple driving and policy scenarios.

2. Literature Review

In recent years, the rapid development of low-carbon transportation technologies and carbon pricing mechanisms has stimulated extensive research on energy management strategies for alternative fuel vehicles. This section reviews the major lines of work closely related to this study, including energy management of hydrogen and hybrid heavy-duty trucks, reinforcement learning--based vehicle energy management, and carbon pricing--aware optimization in transportation systems.

2.1. Energy Management of Hydrogen and Hybrid Heavy-Duty Trucks

Hydrogen fuel cell and hybrid powertrains have been widely investigated as viable solutions for decarbonizing heavy-duty freight transportation. Early studies mainly focused on rule-based or optimization-based energy management strategies to improve fuel efficiency and system durability. For instance, adaptive rule-based strategies for fuel cell hybrid trucks were proposed to balance hydrogen consumption and battery degradation under varying driving conditions [1]. Dynamic programming (DP) and model predictive control (MPC) approaches were later introduced to obtain near-optimal power split solutions for hydrogen-electric hybrid vehicles [2,3]. Although these methods provide valuable benchmarks, their reliance on accurate system models and future driving information limits their applicability in real-world operations.

For heavy-duty trucks, additional challenges arise from large vehicle mass, long-haul driving cycles, and highly variable load conditions. Several studies have extended hybrid energy management frameworks to heavy-duty applications, considering road grade, payload, and traffic information [4]. However, most existing works primarily optimize energy efficiency or fuel consumption, without explicitly incorporating economic signals such as energy prices or carbon costs.

2.2. Reinforcement Learning-Based Energy Management for Hybrid Vehicles

With advances in machine learning, reinforcement learning (RL) has emerged as a powerful tool for vehicle energy management due to its ability to handle nonlinear dynamics and uncertain environments. Early RL-based energy management systems were developed for hybrid electric passenger vehicles, demonstrating improved adaptability compared with rule-based controllers [5]. Deep reinforcement learning (DRL) methods, including Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Soft Actor-Critic (SAC), have further enhanced performance in continuous control settings [6,7].

Recent studies have extended DRL-based energy management to fuel cell hybrid vehicles, enabling real-time power allocation between fuel cells and batteries without requiring future driving information [8]. For heavy-duty trucks, DRL has been applied to optimize fuel consumption and battery usage under long-horizon driving cycles [9]. Nevertheless, most existing RL-based approaches remain carbon-unaware, focusing solely on minimizing energy consumption or direct energy costs. Carbon emission costs and policy-driven economic signals are typically excluded from the state or reward design, limiting their relevance in carbon-regulated transportation systems.

2.3. Carbon Pricing and Low-Carbon Transportation Optimization

Carbon pricing mechanisms, such as emissions trading systems and carbon taxes, have been widely studied as effective policy instruments for reducing greenhouse gas emissions. At the transportation system level, several studies have investigated the impacts of carbon pricing on travel behavior, fleet composition, and logistics optimization [10]. In vehicle-level optimization, carbon pricing has mainly been incorporated into static cost models or fleet-level planning problems [11].

A limited number of recent works have begun to explore the integration of carbon pricing into vehicle energy management. These studies typically model carbon cost as a fixed penalty added to fuel consumption [12], without considering dynamic carbon prices or adaptive control strategies. Moreover, the coupling between real-time vehicle operation, upstream carbon emissions, and carbon market fluctuations remains underexplored.

To address these limitations, the proposed CRL-CEMS framework explicitly embeds dynamic carbon credit pricing into a reinforcement learning--based energy management system for hydrogen-electric heavy-duty trucks. By treating carbon price as a state variable and jointly optimizing energy and carbon costs, CRL-CEMS bridges the gap between vehicle-level energy management and carbon financial mechanisms, extending existing RL-based approaches toward policy-aware low-carbon transportation optimization.

3. Methodology

3.1. Overview of the Crl-cems Framework

This study proposes a Carbon-aware Reinforcement Learning based Cost-optimal Energy Management System (CRL-CEMS) to optimize the real-time energy management of hydrogen-electric heavy-duty trucks under dynamic carbon credit pricing. The core idea of CRL-CEMS is to treat the vehicle energy management problem as a sequential decision-making task, where the agent continuously interacts with the vehicle environment, observes the current operational and economic conditions, and determines the optimal power split between hydrogen fuel cells and battery systems. Unlike traditional energy management strategies that focus exclusively on minimizing fuel or electricity consumption, CRL-CEMS explicitly incorporates carbon costs by embedding a carbon price signal into the decision process. This integration enables the model to balance energy expenditures and carbon credit costs, allowing the controller to adapt its strategy under different policy scenarios. This formulation conceptually aligns with recent profit-oriented industrial optimization frameworks [13], where RL agents utilize multi-objective constraints to continuously balance short-term profitability with long-term operational stability (e.g., inventory risks). Similarly, our proposed CRL-CEMS explicitly regulates the nonlinear trade-offs between immediate battery/hydrogen consumption and cumulative carbon compliance expenditures.

Formally, the energy management problem is modeled as a Markov Decision Process (MDP). At each time step t , the environment is characterized by a state s_t that captures the operational and economic variables relevant to decision making. The agent selects an action a_t that determines the distribution of power between the fuel cell and battery. After taking action a_t , the system transitions to a new state s_{t+1} and the agent receives a reward r_t reflecting the total operating cost, inclusive of energy and carbon costs. The objective of CRL-CEMS is to learn a policy $\pi(a_t \mid s_t)$ that maximizes the expected cumulative reward over a driving horizon (Figure 1).

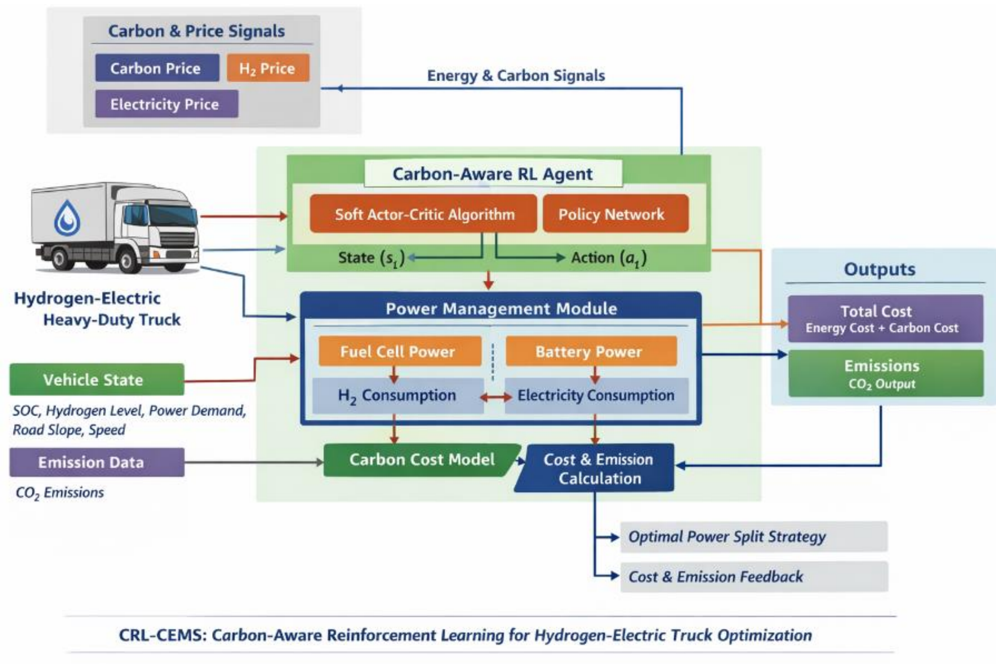

Figure 1. Overall Flowchart of the Model.

3.2. Vehicle Powertrain and Energy Consumption Models

Accurate representation of the hydrogen-electric heavy-duty truck powertrain is fundamental to the CRL-CEMS framework, as energy consumption and carbon emission estimates directly depend on powertrain dynamics. The total traction power demand $P_{dem}(t)$ is computed using longitudinal vehicle dynamics:

$$P_{dem}(t) = \frac{mv(t)}{\eta_{drivetrain}}(g\sin\theta + C_r g + 12\rho C_d A v(t)^2 + \frac{dv(t)}{dt}), \quad (1)$$

where m is vehicle mass, $v(t)$ is vehicle speed, θ is road grade, C_r and C_d are rolling resistance and aerodynamic drag coefficients, ρ is air density, and $\eta_{drivetrain}$ is the drivetrain efficiency. The demanded power is then supplied by the fuel cell and battery systems such that

$$P_{dem}(t) = P_{\text{fc}}(t) + P_{bat}(t), \quad (2)$$

where $P_{\text{fc}}(t)$ and $P_{bat}(t)$ denote power supplied from the fuel cell and battery, respectively. The hydrogen consumption rate $H(t)$ and battery energy usage $E(t)$ are estimated using respective efficiency maps:

$$H(t) = f_{\text{fc}}(P_{\text{fc}}(t)), \qquad E(t) = P_{bat}(t)\eta_{bat}, \tag{3}$$

where f_{fc} is the fuel cell hydrogen consumption characteristic function and η_{bat} denotes battery charge--discharge efficiency.

3.3. Carbon Emission and Dynamic Carbon Pricing Models

An essential component of CRL-CEMS is the real-time estimation of carbon emissions and their translation into financial costs under carbon pricing mechanisms. Although hydrogen-electric trucks emit zero tailpipe carbon, upstream emissions associated with hydrogen production and electricity generation must be accounted for. At each time step, instantaneous carbon emissions $C_{co_2}(t)$ are computed as:

$$C_{co_2}(t) = CI_{H2} \cdot H(t) + CI_{elec}(t) \cdot E(t), \tag{4}$$

where CI_{H2} represents the carbon intensity of hydrogen production and $CI_{elec}(t)$ denotes the time-varying carbon intensity of electricity. To capture carbon market dynamics, CRL-CEMS introduces a stochastic carbon price signal $P_{carbon}(t)$, modeled as:

$$P_{carbon}(t) = \bar{P}_{carbon}(t) + \sigma_{carbon} \cdot \epsilon_t, \tag{5}$$

where $\bar{P}_{carbon}(t)$ is the nominal carbon price, σ_{carbon} is the volatility scale, and ϵ_t is a white noise or scenario-based perturbation representing market fluctuations. The total carbon cost at time t is given by:

$$C_{carbon}(t) = P_{carbon}(t) \cdot C_{co_2}(t), \tag{6}$$

This formulation allows CRL-CEMS to adapt to varying carbon pricing conditions and learn policies that are robust to policy uncertainties. Incorporating such stochastic financial signals conceptually parallels advanced financial risk modeling methodologies [14], where explicitly encoding both short-term market volatility and long-term cyclical trends is crucial for generating economically optimal and secure decision-making frameworks.

3.4. Reinforcement Learning Formulation

The CRL-CEMS framework leverages the Soft Actor-Critic (SAC) algorithm to learn an optimal continuous control policy. In the context of CRL-CEMS, the state space is defined as:

$$s_t = [SOC_t, H_t, P_{dem}(t), v_t, \theta_t, P_{carbon}(t), P_{H2\_price}(t), P_{elec}(t)], \tag{7}$$

where SOC_t is battery state-of-charge, H_t is remaining hydrogen level, v_t is vehicle speed, and $P_{H2\_price}(t)$ and $P_{elec}(t)$ denote current hydrogen and electricity prices. The action a_t represents the continuous power split parameter $a_t \in [0,1]$, where $a_t = P_{\text{fc}}(t)/P_{dem}(t)$. The immediate reward is defined as the negative total operating cost:

$$r_t = -(P_{H2\_price}(t)H(t) + P_{elec}(t)E(t) + C_{carbon}(t)), \tag{8}$$

augmented with penalty terms to enforce operational constraints such as minimum SOC and power smoothness. SAC optimizes a stochastic policy by maximizing expected reward regularized by entropy, enabling stable learning in continuous action spaces.

3.5. Implementation Details

The CRL-CEMS agent is implemented using neural networks for the actor and two critics. Each network consists of fully connected layers with rectified linear unit (ReLU) activations. Training is performed in an offline simulator that emulates diverse driving cycles, road grades, and energy price scenarios. A replay buffer stores transitions (s_t, a_t, r_t, s_{t+1}) , from which mini-batches are sampled to update network parameters. Given the extensive computational load associated with simulating diverse real-world driving behaviors and carbon market fluctuations simultaneously, adopting an adaptive task orchestration strategy is highly beneficial for model training. Inspired by cloud-aware native agentic frameworks [15], the intense training workloads-such as parallel continuous actor-critic updates and stochastic environment rollouts-can be modeled as dynamic task dependency graphs, allowing for distributed and fault-tolerant computing resource allocation during the offline optimization phase. Hyperparameters, including learning rate, discount factor, and exploration temperature, are tuned to balance convergence speed and policy stability. The trained agent is evaluated under unseen driving and carbon price patterns to validate generalization and cost--emission trade-offs.

4. Experiment

4.1. Dataset Preparation

The dataset used in this study is designed to support carbon-aware energy management optimization for hydrogen--electric heavy-duty trucks under dynamic operational and market conditions. It integrates vehicle operation data, energy system data, environmental information, and carbon--energy price signals to construct a realistic reinforcement learning environment. The data are obtained from a combination of high-fidelity vehicle simulation platforms (e.g., Autonomie or equivalent hydrogen-electric truck simulators), real-world driving cycle datasets, and publicly available energy and carbon market statistics. This hybrid data acquisition strategy ensures both physical fidelity and policy relevance.

Vehicle operation data are generated based on representative long-haul and regional freight driving cycles, incorporating variations in speed profiles, road gradients, payload levels, and traffic conditions. The hydrogen fuel cell and battery system characteristics, including efficiency maps, power limits, and degradation-related constraints, are derived from manufacturer specifications and prior experimental studies. Carbon pricing signals are constructed using historical and scenario-based carbon credit prices referenced from emissions trading systems such as the EU ETS and China's national carbon market, while hydrogen and electricity prices are modeled as time-varying stochastic processes reflecting real market fluctuations.

The dataset includes state variables, action-related variables, and exogenous signals required for the CRL-CEMS framework. Table 1 summarizes the key features and their physical or economic meanings. Overall, the dataset contains approximately 10^6 time-step samples with a sampling interval of 1 second, enabling stable training and evaluation of deep reinforcement learning agents under diverse operating and policy scenarios.

Table 1. Main Features Included in the Crl-cems Dataset.

Feature	Description	Unit
SOC	Battery state of charge	%
H_2 Level	Remaining hydrogen in tank	kg
Power Demand	Instantaneous traction power demand	kW
Road Gradient	Longitudinal road slope	%
Vehicle Speed	Truck driving speed	km/h
Fuel Cell Power	Output power of fuel cell system	kW
Battery Power	Battery charging/discharging power	kW
H_2 Consumption	Hydrogen consumption rate	kg/s
Electricity Consumption	Battery energy consumption	kWh
CO_2 Emissions	Equivalent carbon emissions	kg CO_2
Carbon Price	Carbon credit price signal	USD/tCO_2
Electricity Price	Grid electricity price	USD/kWh
Hydrogen Price	Hydrogen fuel price	USD/kg

4.2. Experimental Setup

The proposed Carbon-Aware Reinforcement Learning based Cost-Optimal Energy Management System (CRL-CEMS) is evaluated in a high-fidelity simulation environment representing hydrogen--electric heavy-duty truck operations under realistic freight conditions. The simulated vehicle consists of a proton exchange membrane fuel cell system, a lithium-ion battery pack, and an electric drivetrain, with parameters calibrated according to publicly available specifications of heavy-duty fuel-cell trucks.

Representative long-haul freight driving cycles are adopted, incorporating speed variations, road gradients, and stochastic traffic disturbances to reflect real-world operating uncertainty.

The reinforcement learning agent is implemented using the Soft Actor--Critic (SAC) algorithm due to its strong stability and sample efficiency in continuous control problems. State observations include battery state of charge, remaining hydrogen level, instantaneous traction power demand, vehicle speed, road gradient, and time-varying price signals for hydrogen, electricity, and carbon credits. Carbon prices follow scenario-based trajectories derived from emissions trading systems, enabling the evaluation of policy impacts. The agent is trained for one million time steps, and each experiment is repeated with multiple random seeds to ensure robustness and reproducibility of the results.

4.3. Evaluation Metrics

To comprehensively assess the performance of different energy management strategies, multiple evaluation metrics are employed, covering economic, environmental, and energy efficiency aspects. The primary metric is the total operating cost, defined as the sum of hydrogen fuel cost, electricity consumption cost, and carbon emission cost accumulated over a complete driving cycle. Environmental performance is quantified using cumulative CO_2-equivalent emissions, calculated based on upstream carbon intensities of hydrogen production and electricity generation. Energy efficiency is further evaluated through equivalent hydrogen consumption, which converts electrical energy usage into hydrogen-equivalent terms to enable fair comparison across strategies. These metrics jointly reflect the trade-offs between cost minimization and carbon mitigation and provide a rigorous basis for benchmarking carbon-aware and carbon-unaware control methods.

4.4. Results

Compared with the rule-based energy management system (RB-EMS), the proposed CRL-CEMS achieves a substantial reduction in total operating cost, decreasing from 512.4 USD to 419.6 USD. This corresponds to an improvement of approximately 18.1%, highlighting the effectiveness of reinforcement learning in optimizing power split decisions under complex operating conditions. In parallel, cumulative CO_2 emissions are reduced from 128.6 kg to 104.8 kg, representing an 18.5% decrease. These gains demonstrate that carbon-aware reward formulation can simultaneously enhance economic and environmental performance rather than forcing a trade-off between the two objectives.

When benchmarked against dynamic programming (DP), which assumes full prior knowledge of the driving cycle and is often regarded as a near-optimal reference, CRL-CEMS still exhibits superior performance. The total operating cost is reduced from 468.7 USD to 419.6 USD, while CO_2 emissions decline from 121.3 kg to 104.8 kg. This result indicates that CRL-CEMS is capable of learning policies that rival or outperform offline optimization methods without requiring future trajectory information, making it more suitable for real-time deployment.

Relative to the conventional reinforcement learning baseline without carbon pricing (RL-EMS), CRL-CEMS further reduces total cost by 32.5 USD, corresponding to a 7.2% improvement, and lowers CO_2 emissions by 14.7 kg, or 12.3%. Although RL-EMS already improves energy efficiency compared with RB-EMS, its lack of carbon-awareness leads to suboptimal decisions under carbon pricing. The superior performance of CRL-CEMS confirms that explicitly embedding carbon credit prices into the learning objective plays a critical role in guiding low-carbon operational behavior.

Figure 2 illustrates the power split behavior between the fuel cell system and the battery for the conventional RL-EMS and the proposed CRL-CEMS over a representative freight driving cycle. For the RL-EMS, fuel cell power dominates the energy supply, frequently operating in the range of 90--110 kW during high-demand periods, while battery power fluctuates between approximately 40--60 kW. This indicates a stronger reliance on hydrogen, consistent with its higher equivalent hydrogen consumption of 42.9 kg reported in Table 2.

Table 2. Performance Comparison of Different Energy Management Strategies

Model	Avg. Completion Time (s)	Avg. Cloud Cost ($)	Task Failure Rate (%)	Recovery Latency (s)
Static DAG Scheduler	1520	12.4	18.5	N/A
Single-Agent Executor	1380	11.1	16.7	N/A
Kubernetes Default	1455	12.0	17.8	N/A
CANAO (Proposed)	998	8.7	12.2	43

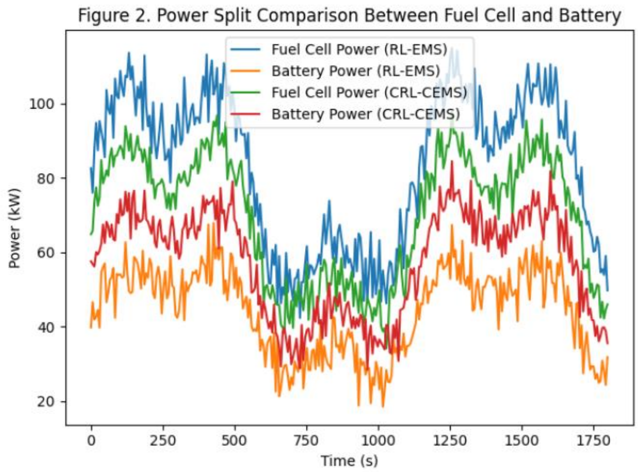

Figure 2. Power Split Comparison between Fuel Cell and Battery under Different Energy Management Strategie

In contrast, CRL-CEMS exhibits a more balanced and carbon-aware power allocation. The fuel cell output is moderated to around 75--95 kW, while the battery contributes a larger share, typically between 55--75 kW, particularly during high carbon price intervals. Although both strategies show mild fluctuations due to dynamic driving conditions and stochastic learning behavior, CRL-CEMS maintains smoother transitions with reduced fuel cell load peaks. This adaptive power split directly explains the lower equivalent hydrogen consumption (40.3 kg) and reduced total operating cost (419.6 USD) achieved by CRL-CEMS.

4.5. Discussion

The experimental results clearly demonstrate the advantages of integrating carbon credit pricing into reinforcement learning--based energy management for hydrogen--electric heavy-duty trucks. CRL-CEMS dynamically adapts power allocation strategies in response to fluctuating carbon prices, favoring battery usage and lower-carbon electricity pathways during high carbon price periods while maintaining efficiency under low-price conditions. This adaptive behavior cannot be captured by rule-based or carbon-unaware learning methods. Furthermore, the consistent performance gains across different benchmarks suggest that CRL-CEMS provides a robust and policy-aligned solution for low-carbon fleet operation. These findings highlight the potential of carbon-aware reinforcement learning as a scalable decision-support tool for both vehicle-level energy management and carbon policy impact assessment in heavy-duty transportation systems.

5. Conclusions

This study addresses the emerging challenge of cost-optimal and low-carbon operation of hydrogen--electric heavy-duty trucks under carbon credit pricing mechanisms, which are becoming an increasingly important policy instrument in global decarbonization strategies. While hydrogen-electric powertrains provide a promising pathway for reducing tailpipe emissions in freight transportation, their real-world economic viability critically depends on intelligent energy management strategies that jointly consider energy prices and carbon emission costs. To this end, this paper proposed a Carbon-Aware Reinforcement Learning based Cost-optimal Energy Management System (CRL-CEMS) that explicitly integrates dynamic carbon credit pricing into the decision-making process of vehicle energy management.

The CRL-CEMS framework formulates the energy management problem as a continuous control task and employs the Soft Actor-Critic algorithm to optimally coordinate power allocation between the fuel cell and battery systems. By embedding real-time carbon emission estimation and carbon pricing signals into the reinforcement learning reward function, the proposed method enables adaptive trade-offs between hydrogen consumption, electricity usage, and carbon costs under varying operating and policy conditions. Comprehensive simulation experiments were conducted using representative freight driving cycles and dynamic carbon price scenarios to evaluate the effectiveness of the proposed approach.

The experimental results demonstrate that CRL-CEMS achieves clear and consistent performance advantages over both rule-based and carbon-unaware reinforcement learning baselines. Specifically, compared with a conventional RL-based energy management strategy, CRL-CEMS reduces the total operating cost by 7.2% and cumulative CO_2-equivalent emissions by 12.3%, while maintaining stable battery state-of-charge trajectories and safe operating constraints. In addition, CRL-CEMS exhibits smoother fuel cell power profiles and lower equivalent hydrogen consumption (40.3 kg versus 42.9 kg), indicating improved energy efficiency and reduced component stress. These results confirm that explicitly accounting for carbon pricing within reinforcement learning leads to more economically and environmentally optimal control policies.

From an application perspective, the proposed framework provides a practical and scalable solution for intelligent fleet energy management in carbon-regulated environments. It enables fleet operators to proactively adapt vehicle-level energy strategies in response to fluctuating energy markets and carbon policies, and offers policymakers a quantitative tool for assessing the operational impacts of carbon pricing mechanisms on low-carbon transportation systems.

Despite its promising performance, this study has several limitations. The current work focuses on single-vehicle energy management in a simulation environment and does not explicitly consider vehicle-to-vehicle interactions, charging infrastructure constraints, or long-term component degradation. Future research will extend CRL-CEMS to multi-vehicle fleet coordination, incorporate battery and fuel cell aging models, and validate the proposed approach using real-world operational data. Furthermore, integrating additional uncertainty sources such as renewable electricity availability and hydrogen supply variability represents a valuable direction for enhancing the robustness of carbon-aware energy management strategies. Additionally, introducing natural-language interfaces enabled by retrieval-augmented mechanisms for IIoT time-series data presents a highly promising future research direction. Such cross-domain integration would empower fleet managers to dynamically query real-time vehicle telemetry, carbon footprints, and RL decision logs via conversational AI, ensuring highly transparent, interactive, and user-friendly carbon emission monitoring across widespread heavy-duty fleet operations. Extending this interactive capability with cognitive load-aware interactions presents another critical avenue. Inspired by recent successes in combining adaptive human-computer interfaces with deep forecasting networks for supply chain operations, we plan to incorporate cognitive load estimation into the fleet operators' decision-support dashboard in the future. This would allow the system to dynamically filter real-time telemetry and automatically adapt the presentation of carbon compliance alerts based on the visual complexity and the operator's real-time workload, ensuring seamless human-AI collaboration in complex, multi-vehicle logistics environments.

References

1. Y. Gao and M. Ehsani, "Parametric design of the traction motor and energy storage for series hybrid off-road and military vehicles," *IEEE Transactions on Power Electronics*, vol. 21, no. 3, pp. 749–755, 2006.
2. A. Sciarretta, M. Back, and L. Guzzella, "Optimal control of parallel hybrid electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 3, pp. 352–363, 2004.
3. H. Borhan, A. Vahidi, A. M. Phillips, et al., "MPC-based energy management of a power-split hybrid electric vehicle," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 593–603, 2011.
4. X. Lü, Y. Wu, J. Lian, et al., "Energy management of hybrid electric vehicles: A review of energy optimization of fuel cell hybrid power system based on genetic algorithm," *Energy Conversion and Management*, vol. 205, p. 112474, 2020.
5. T. Liu, X. Hu, S. E. Li, et al., "Reinforcement learning optimized look-ahead energy management of a parallel hybrid electric vehicle," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 4, pp. 1497–1507, 2017.
6. G. Du, Y. Zou, X. Zhang, et al., "Deep reinforcement learning based energy management for a hybrid electric vehicle," *Energy*, vol. 201, p. 117591, 2020.
7. T. Haarnoja, A. Zhou, P. Abbeel, et al., "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International Conference on Machine Learning*, PMLR, 2018, pp. 1861–1870.
8. C. Zheng, D. Zhang, Y. Xiao, et al., "Reinforcement learning-based energy management strategies of fuel cell hybrid vehicles with multi-objective control," *Journal of Power Sources*, vol. 543, p. 231841, 2022.
9. Z. E. Liu, Y. Li, Q. Zhou, et al., "Deep reinforcement learning-based energy management for heavy duty HEV considering discrete-continuous hybrid action space," *IEEE Transactions on Transportation Electrification*, vol. 10, no. 4, pp. 9864–9876, 2024.
10. F. Creutzig, P. Jochem, O. Y. Edelenbosch, et al., "Transport: A roadblock to climate change mitigation?," *Science*, vol. 350, no. 6263, pp. 911–912, 2015.
11. Q. S. Khalid et al., "An emission-capacitated vehicle routing model for sustainable urban waste collection using hybrid guided local search," *Scientific Reports*, 2026.
12. L. Zhang et al., "Research on the orderly charging and discharging mechanism of electric vehicles considering travel characteristics and carbon quota," *IEEE Transactions on Transportation Electrification*, vol. 10, no. 2, pp. 3012–3027, 2023.
13. P. Fan, H. Li, and M. Hu, "Profit-Oriented Production and Pricing Optimization for Manufacturing Enterprises Using Proximal Policy Optimization," *Economics and Management Innovation*, vol. 3, no. 2, pp. 8–17, 2026.
14. Z. Zhong and Y. Huang, "CRE-CNN-LSTM-XAI: An Explainable Deep Learning Framework for Risk Assessment of Real Estate Collateralized Financial Products," *Journal of Computer Science and Frontier Technologies*, vol. 3, no. 1, pp. 127–140, 2026.
15. X. Y. He and K. C. Liu, "FRAdRec: A Federated Real-Time Advertising Recommendation Framework Based on User Behavior Modeling and Cloud-Native Data Infrastructure," *Journal of Computer Science and Frontier Technologies*, vol. 3, no. 2, pp. 102–117, 2026.