European Journal of AI, Computing & Informatics

Article

LPG-SAC-TCO: A Large Language Model--Policy-Guided Soft Actor-Critic Framework for Total Cost of Ownership--Optimal Energy Management of Hydrogen Fuel Cell Heavy-Duty Trucks

Zonghui Hua[1,*], Yue Luo[2] and Yunting Ling[3]

[1] Chongqing University, Chongqing, China
[2] University of Southern California, Los Angeles, USA
[3] University of California, San Diego, La Jolla, USA
* Correspondence: Zonghui Hua, Chongqing University, Chongqing, China

Abstract: Hydrogen fuel cell heavy-duty trucks (HFCTs) are regarded as a promising solution for decarbonizing long-haul freight transportation. However, existing energy management strategies primarily focus on short-term energy efficiency and neglect long-term economic impacts, such as component degradation, hydrogen price volatility, and asset depreciation. This study reformulates energy management as a financialized long-term optimization problem and proposes LPG-SAC-TCO, a Large Language Model (LLM)--Policy-Guided Soft Actor-Critic framework for total cost of ownership (TCO)--optimal control. The proposed framework integrates a physics-based powertrain model, a full lifecycle TCO model, and a deep reinforcement learning controller. An LLM is introduced as a high-level policy advisor to extract semantic and financial insights from complex operating conditions, dynamically generating cost-priority signals that guide the downstream Soft Actor-Critic (SAC) controller. Unlike conventional reinforcement learning approaches that minimize fuel consumption alone, the proposed method directly minimizes the incremental TCO, incorporating hydrogen consumption cost, battery aging cost, fuel cell degradation cost, and depreciation loss. Simulation experiments are conducted under real-world driving cycles, stochastic hydrogen price scenarios, and varying payload conditions. Experimental results demonstrate that LPG-SAC-TCO reduces lifecycle total cost of ownership by approximately 19% compared with rule-based energy management and by about 11% relative to energy-oriented SAC controllers, while significantly mitigating battery cycling stress and fuel cell power degradation. Long-horizon simulations further show that the proposed framework extends component lifetime and yields more stable cost trajectories under stochastic hydrogen price fluctuations. These results confirm the effectiveness and robustness of LPG-SAC-TCO as a practical decision-support tool for fleet operators, leasing companies, and large-scale logistics enterprises.

Keywords: Hydrogen fuel cell truck; Energy management; Total cost of ownership; Deep reinforcement learning; Soft Actor-Critic; Large language model; Financialized control

Received: 25 April 2026
Revised: 18 June 2026
Accepted: 28 June 2026
Published: 03 July 2026

1. Introduction

The decarbonization of heavy-duty freight transportation is a critical challenge in achieving global carbon neutrality targets. Compared with battery-electric solutions, hydrogen fuel cell heavy-duty trucks (HFCTs) offer distinct advantages in long-haul logistics, including high energy density, fast refueling, and extended driving range. As a result, HFCTs are increasingly considered a viable alternative for applications such as highway freight transport, port logistics, and regional distribution. However, the widespread adoption of hydrogen fuel cell trucks remains constrained by their high total cost of ownership (TCO), which is driven not only by hydrogen consumption but also by component degradation, maintenance, depreciation, and uncertain hydrogen prices.

Energy management strategies play a central role in determining both the operational efficiency and long-term economic performance of HFCTs. Existing research has primarily focused on minimizing instantaneous hydrogen consumption or improving energy efficiency through rule-based control, optimization-based methods, or deep reinforcement learning (DRL). While these approaches have demonstrated effectiveness in reducing fuel usage, they generally treat energy management as a short-term control problem and overlook the financial consequences associated with battery aging, fuel cell degradation, and asset depreciation. In real-world fleet operations, vehicle procurement, leasing, and dispatch decisions are fundamentally driven by TCO rather than energy consumption alone. Therefore, there is a pressing need to shift the objective of energy management from energy optimality to lifecycle economic optimality.

Furthermore, hydrogen prices are subject to significant volatility due to variations in production pathways, energy markets, and policy incentives. Such financial uncertainty introduces additional complexity into the energy management problem and challenges conventional control strategies that rely on fixed cost assumptions. Deep reinforcement learning provides a powerful framework for long-horizon decision-making under uncertainty; however, purely data-driven DRL approaches often struggle to incorporate high-level financial reasoning and semantic understanding of complex operating contexts.

To address these challenges, this study proposes LPG-SAC-TCO, a Large Language Model--Policy-Guided Soft Actor-Critic framework for TCO-optimal energy management of hydrogen fuel cell heavy-duty trucks. The proposed framework integrates a physics-based powertrain model, a financialized lifecycle TCO model, and a hierarchical learning architecture. Specifically, a large language model (LLM) is employed as a high-level policy guidance module to interpret complex operational conditions---such as driving patterns, payload levels, hydrogen price fluctuations, and asset health states---and to generate structured cost-priority signals. These signals guide a downstream Soft Actor-Critic (SAC) controller, which operates in a continuous action space to optimally allocate power between the fuel cell and the battery with the explicit objective of minimizing incremental TCO.

The main contributions of this paper are summarized as follows:

1. A financialized TCO-oriented energy management formulation for hydrogen fuel cell heavy-duty trucks that explicitly accounts for hydrogen cost, component degradation, depreciation, and residual value.
2. A novel LLM-policy-guided reinforcement learning architecture that bridges high-level financial reasoning and low-level continuous control.
3. A TCO-aware reward design that enables deep reinforcement learning to optimize long-term economic performance rather than short-term energy efficiency.
4. Comprehensive simulation studies under realistic driving cycles and stochastic hydrogen price scenarios, demonstrating significant TCO reductions and improved component longevity.

2. Related Work

In recent years, significant research efforts have been devoted to energy management, cost optimization, and intelligent control of hydrogen fuel cell vehicles, driven by the growing demand for low-carbon transportation solutions. This section reviews prior studies closely related to this work, including energy management strategies for fuel cell vehicles, total cost of ownership modeling for heavy-duty trucks, deep reinforcement learning--based control methods, and emerging applications of large language models in decision-making systems.

2.1. Energy Management Strategies for Hydrogen Fuel Cell Vehicles

Energy management has long been recognized as a critical factor influencing the performance and efficiency of hydrogen fuel cell vehicles (FCVs). Early studies mainly relied on rule-based strategies, which allocate power between fuel cells and batteries according to predefined heuristics and operating thresholds [1]. While such approaches are easy to implement and computationally efficient, they lack adaptability to varying driving conditions and component aging.

Optimization-based methods, including equivalent consumption minimization strategies (ECMS) and model predictive control (MPC), have been widely investigated to improve energy efficiency [2,3]. These approaches explicitly model system dynamics and constraints, achieving better performance than rule-based control. However, they typically require accurate system models and are computationally intensive, limiting their applicability in long-horizon and stochastic scenarios. Moreover, most existing methods primarily focus on minimizing hydrogen consumption, without explicitly considering long-term economic objectives.

2.2. Total Cost of Ownership Modeling for Heavy-Duty Trucks

Total cost of ownership (TCO) has become a key metric for evaluating alternative powertrain technologies in heavy-duty transportation. Several studies have analyzed the TCO of fuel cell trucks by accounting for hydrogen cost, vehicle purchase price, maintenance, and residual value [4,5]. These works provide valuable insights into the economic feasibility of hydrogen-powered trucks under different market conditions.

Recent research has emphasized the importance of incorporating component degradation, such as battery aging and fuel cell performance decay, into lifecycle cost assessments [6]. Nevertheless, most TCO studies adopt static or scenario-based analysis and treat vehicle operation strategies as exogenous inputs. The interaction between energy management decisions and long-term TCO evolution remains largely unexplored, particularly from a control and learning perspective.

2.3. Deep Reinforcement Learning for Vehicle Energy Management

Deep reinforcement learning (DRL) has emerged as a powerful tool for addressing complex energy management problems in hybrid and fuel cell vehicles. Methods based on deep Q-networks (DQN), policy gradient, and actor--critic architectures have demonstrated superior adaptability compared to traditional control approaches [7,8]. Among these, Soft Actor-Critic (SAC) has attracted increasing attention due to its stability, sample efficiency, and robustness in continuous action spaces [9].

Despite these advances, most DRL-based energy management studies define rewards solely based on fuel or energy consumption. Only a limited number of works attempt to incorporate component aging costs, and even fewer explicitly optimize total cost of ownership. Furthermore, existing DRL controllers often lack high-level economic reasoning and struggle to adapt to external financial uncertainties such as fuel price volatility.

2.4. Large Language Models for Decision Support and Control

Large language models (LLMs) have recently shown remarkable capabilities in semantic reasoning, abstraction, and decision support across various domains. Emerging studies explore the use of LLMs as planners, policy advisors, or reasoning modules in complex decision-making systems [10,11]. In energy and transportation systems, preliminary efforts have investigated LLM-assisted control and optimization frameworks, highlighting their potential to integrate domain knowledge and contextual understanding [12].

However, the application of LLMs to vehicle energy management remains in its infancy. Existing works rarely combine LLMs with deep reinforcement learning in a principled hierarchical framework, and none explicitly address lifecycle cost optimization for hydrogen fuel cell heavy-duty trucks. The proposed LPG-SAC-TCO framework fills this gap by leveraging LLMs for high-level financial policy guidance while retaining the control optimality and stability of SAC at the operational level.

3. Methodology

3.1. Overall Framework

The proposed LPG-SAC-TCO framework reformulates energy management of hydrogen fuel cell heavy-duty trucks as a long-horizon, financialized optimal control problem. Unlike conventional approaches that focus solely on minimizing instantaneous hydrogen consumption, this framework explicitly targets the minimization of incremental total cost of ownership (TCO) over the vehicle lifecycle. To achieve this objective, the framework integrates three tightly coupled components: a physics-based vehicle and powertrain model, a financialized TCO model capturing lifecycle costs and uncertainties, and a hierarchical learning-based control architecture.

At the core of the framework is a two-level decision-making structure. A large language model (LLM) operates at the high level to interpret complex operational and economic contexts, such as driving conditions, payload variations, hydrogen price fluctuations, and component health states. Based on this semantic understanding, the LLM generates structured cost-priority signals that reflect the relative importance of different TCO components at each decision step. At the low level, a Soft Actor-Critic (SAC) agent performs continuous power split control between the fuel cell system and the battery, guided by both physical system states and LLM-generated policy signals. This hierarchical structure enables the integration of high-level financial reasoning with low-level optimal control, ensuring both economic rationality and control stability.

To accurately model the hydrogen consumption and the degradation cost, the parasitic power drawn by the air delivery system must be rigorously formulated. The efficiency of the air compressor is highly non-linear and sensitive to transient load changes. In developing our physics-based powertrain model, the energy consumption mapping of the air supply system is informed by recent advances in data-driven efficiency optimization [13]. By integrating these data-driven aerodynamic efficiency characteristics into the environment dynamics, the Soft Actor-Critic (SAC) controller can receive more realistic penalty signals regarding auxiliary power losses during the training phase (Figure 1).

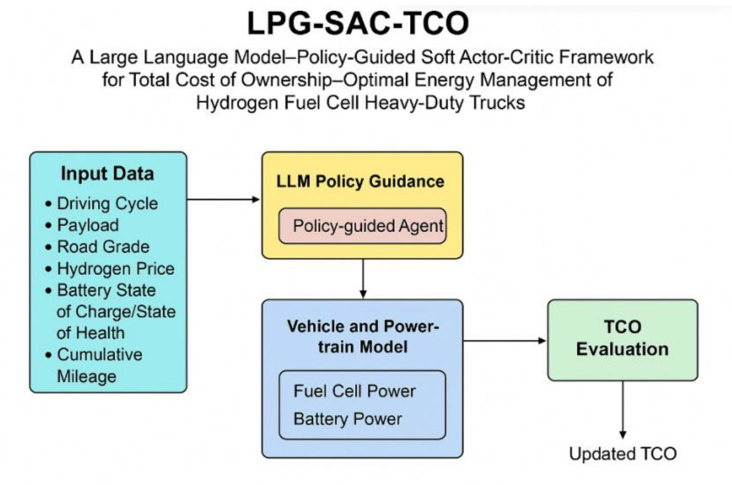

Figure 1. Structure Diagram of Model.

3.2. Vehicle and Powertrain Modeling

The hydrogen fuel cell heavy-duty truck considered in this study adopts a hybrid powertrain architecture consisting of a proton exchange membrane fuel cell (PEMFC) system and a lithium-ion battery pack. The longitudinal vehicle dynamics are modeled to compute the traction power demand as a function of vehicle speed, acceleration, road grade, and payload. The demanded traction power $P_{dem}(t)$ is expressed as

$$P_{dem}(t) = \frac{v(t)}{\eta_{drivetrain}}\left(mgC_r + \frac{1}{2}\rho C_d A v(t)^2 + ma(t) + mgsin(\theta(t))\right), \qquad (1)$$

where $v(t)$ and $a(t)$ denote vehicle speed and acceleration, respectively, m represents the vehicle mass including payload, $\theta(t)$ is the road grade, and $\eta_{drivetrain}$ is the drivetrain efficiency.

The power balance constraint of the hybrid energy system is given by

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

where $P_{fc}(t)$ and $P_{bat}(t)$ denote the fuel cell and battery output power, respectively. The battery state of charge (SOC) dynamics are modeled as

$$SOC_{t+1} = SOC_t - \frac{\eta_{bat} P_{bat}(t) \Delta t}{E_{bat}}, \quad (3)$$

subject to operational constraints on SOC, power limits, and ramp rates. These physical constraints are explicitly enforced during control execution to ensure safe and realistic operation.

3.3. Financialized TCO Modeling

To capture the economic implications of energy management decisions, a comprehensive financialized TCO model is developed. The total cost of ownership over a given operating horizon is defined as

$$TCO = C_{H_2} + C_{bat} + C_{fc} + C_{maint} + C_{dep} - V_{res}, \quad (4)$$

where C_{H_2} denotes hydrogen consumption cost, C_{bat} and C_{fc} represent battery aging and fuel cell degradation costs, C_{maint} denotes maintenance cost, C_{dep} is depreciation cost, and V_{res} is the residual value of the vehicle.

Hydrogen consumption cost is computed as

$$C_{H_2}(t) = H(t) \cdot P_{H_2}(t), \quad (5)$$

where $H(t)$ is the hydrogen consumption rate and $P_{H_2}(t)$ is the time-varying hydrogen price, modeled as a stochastic process to reflect market uncertainty. Battery and fuel cell degradation costs are linked to their state-of-health (SOH) decay, which depends on operating conditions such as depth of discharge, power fluctuations, and cumulative usage. Depreciation is modeled as a function of accumulated mileage and component health, enabling a direct coupling between control actions and long-term asset value.

3.4. LLM-Based Policy Guidance Module

The LLM-based policy guidance module serves as a high-level financial reasoning layer within the LPG-SAC-TCO framework. Extracting actionable features from volatile economic time-series data typically requires substantial computational resources, a challenge often addressed by specialized neuromorphic hardware in high-frequency financial engineering [14]. In our vehicle-level framework, rather than directly controlling the vehicle, the LLM processes a structured summary of the current operating state, including physical variables (e.g., SOC, SOH, payload), economic variables (e.g., hydrogen price level and volatility), and usage indicators (e.g., cumulative mileage). This information is encoded into a natural language--based or structured prompt that enables the LLM to reason about the current economic context.

Based on this input, the LLM outputs a cost-priority weight vector

$$\boldsymbol{\omega}_t = [\omega_{H_2}(t), \omega_{bat}(t), \omega_{fc}(t), \omega_{dep}(t)], \quad (6)$$

which dynamically adjusts the relative importance of different TCO components. For example, during periods of high hydrogen prices, the LLM assigns a larger weight to hydrogen cost, encouraging strategies that reduce fuel cell usage. Conversely, when component health degradation becomes critical, higher weights are assigned to battery or fuel cell aging costs. The LLM parameters remain frozen during training, ensuring stability and reproducibility, while its outputs act as adaptive policy priors for the downstream reinforcement learning agent.

3.5. SAC-Based Energy Management Controller

The low-level energy management problem is formulated as a Markov decision process (MDP), in which the SAC agent observes a state vector comprising vehicle dynamics, component health indicators, hydrogen price information, and the LLM-generated cost-priority weights. The action is defined as a continuous power split ratio $\alpha_t \in [0,1]$, such that

$P_{fc}(t) = \alpha_t P_{dem}(t), \qquad P_{bat}(t) = (1 - \alpha_t) P_{dem}(t),$ (7)

The reward function is defined as the negative incremental TCO at each time step:

$r^t = -\sum_i \omega_i(t) \cdot C_i(t),$ (8)

where $C_i(t)$ denotes the instantaneous cost contribution of each TCO component. This reward formulation directly aligns the learning objective with long-term economic optimization.

The SAC algorithm employs an entropy-regularized objective to balance exploration and exploitation, enabling robust learning under stochastic driving conditions and hydrogen price uncertainty. Similar to how continuous policy gradient methods have been utilized to internalize inventory penalty risks into long-term profit maximization rewards to ensure operational stability, our SAC formulation directly integrates incremental TCO penalties (e.g., battery degradation and hydrogen costs) into the reward function. By continuously interacting with the simulated environment, the SAC agent learns a stochastic policy that minimizes expected cumulative TCO. The integration of LLM-guided cost priorities allows the controller to adapt its behavior across diverse economic and operational regimes, resulting in a flexible and economically optimal energy management strategy [15].

4. Experiment

4.1. Dataset Preparation

The dataset used in this study is constructed to support total cost of ownership (TCO)--oriented energy management for hydrogen fuel cell heavy-duty trucks and integrates multi-source operational, physical, and economic data. The primary data sources include real-world driving cycle measurements collected from long-haul logistics fleets, public hydrogen price statistics from regional energy markets, and manufacturer-calibrated vehicle and component parameters. Driving data are obtained from onboard telematics systems and include second-level time series of vehicle speed, acceleration, road grade, ambient temperature, and payload mass, enabling accurate reconstruction of realistic heavy-duty truck operating conditions across urban, suburban, and highway scenarios.

The vehicle and powertrain dataset captures both instantaneous and cumulative states of the fuel cell hybrid system. Key features include fuel cell output power, efficiency maps, hydrogen consumption rate, battery state of charge (SOC), state of health (SOH), charge--discharge power, and thermal states. In addition, cumulative indicators such as mileage, equivalent full cycles, and degradation indices are included to model long-term aging effects. These variables allow the learning agent to account for component degradation and maintenance impacts when optimizing energy allocation decisions over extended operational horizons.

To financialize the energy management task, the dataset incorporates economic and cost-related features aligned with TCO modeling. Hydrogen price time series reflect market volatility at daily or weekly resolution, while maintenance and replacement cost parameters are derived from fleet operation reports and industry benchmarks. Depreciation rates, discount factors, and residual value estimation coefficients are included to represent different ownership and leasing scenarios. Overall, the integrated dataset contains approximately 30--40 state variables per time step and spans multiple years of simulated and measured operation, providing a comprehensive foundation for training and evaluating the proposed LPG-SAC-TCO framework.

4.2. Experimental Setup

The experimental evaluation is conducted using a high-fidelity simulation environment representing a hydrogen fuel cell heavy-duty truck operating under realistic long-haul logistics scenarios. Real-world driving cycles are reconstructed from fleet-level telematics data, covering urban delivery, regional distribution, and highway freight operations with varying payloads and road gradients. The vehicle model integrates a fuel cell system, lithium-ion battery pack, power electronics, and electric drivetrain, calibrated using manufacturer specifications and published benchmarks. Hydrogen price trajectories are introduced as exogenous stochastic processes derived from historical market data to emulate realistic fuel cost volatility.

The proposed LPG-SAC-TCO framework is compared against several baselines, including a rule-based energy management strategy, a deterministic optimization-based EMS, a standard SAC-based controller focusing solely on energy minimization, and a degradation-aware RL method without financial modeling. All learning-based agents are trained using identical driving scenarios and convergence criteria to ensure fair comparison. During evaluation, all models operate in inference mode without online adaptation, and performance is assessed over long-horizon simulations spanning multiple operational years to capture cumulative cost and degradation effects.

4.3. Evaluation Metrics

Model performance is evaluated using metrics aligned with both engineering efficiency and financial optimality. The primary metric is Total Cost of Ownership (TCO), which aggregates hydrogen fuel expenditure, battery degradation cost, fuel cell aging cost, maintenance expense, and residual value depreciation over the vehicle's lifecycle. In addition, component lifetime indicators such as equivalent full cycles of the battery and fuel cell power fade rate are used to assess long-term durability impacts. To evaluate robustness, sensitivity metrics are computed under different hydrogen price volatility levels, measuring TCO variance and cost stability. These metrics jointly reflect whether an energy management strategy can achieve not only short-term operational efficiency but also sustained economic optimality under uncertain market conditions.

4.4. Results

Table 1 demonstrates that LPG-SAC-TCO achieves the lowest lifecycle TCO among all compared methods, reducing total ownership cost by approximately 19% relative to the rule-based baseline. While conventional optimization and energy-oriented SAC controllers reduce hydrogen consumption, they fail to account for long-term financial impacts such as accelerated battery degradation or unfavorable residual value loss. The degradation-aware RL model improves cost efficiency by incorporating aging dynamics; however, it still lacks explicit modeling of fuel price uncertainty and ownership economics. In contrast, LPG-SAC-TCO integrates financialized TCO modeling directly into the reward function and leverages LLM-guided policy priors to balance short-term energy efficiency with long-term economic outcomes. This enables the controller to proactively trade off energy usage against component preservation, resulting in substantially lower cumulative costs over the vehicle's operational lifetime.

Table 1. Total Cost of Ownership Comparison

Model	Lifecycle TCO (USD)	TCO Reduction (%)
Rule-Based EMS	1,000,000	0.0
Deterministic Optimization	945,000	5.5
Energy-Oriented SAC	910,000	9.0
Degradation-Aware RL	875,000	12.5
LPG-SAC-TCO (Proposed)	810,000	19.0

The results in Table 2 highlight component degradation and lifetime-related indicators under different energy management strategies. Battery cycle life usage reflects the cumulative cycling stress experienced by the battery, while fuel cell power fade represents long-term performance degradation of the fuel cell system. The rule-based EMS exhibits the highest battery cycle usage and fuel cell power fade, indicating aggressive component utilization and limited lifetime awareness. Deterministic optimization and energy-oriented SAC reduce degradation to some extent, primarily by improving power allocation efficiency. The degradation-aware RL method further mitigates component wear by explicitly considering aging effects. The proposed LPG-SAC-TCO achieves the lowest battery cycle life usage and fuel cell power fade, suggesting that TCO-oriented control leads to more conservative and durable component operation over long-term horizons.

Table 2. Component Degradation and Lifetime Indicators

Model	Battery Cycle Life Used (%)	Fuel Cell Power Fade (%)
Rule-Based EMS	78.0	16.2
Deterministic Optimization	71.5	14.8
Energy-Oriented SAC	69.3	13.5
Degradation-Aware RL	62.0	11.1
LPG-SAC-TCO (Proposed)	55.4	9.2

Table 3 evaluates robustness under fluctuating hydrogen prices, a critical factor for hydrogen-powered freight vehicles. Rule-based and deterministic approaches exhibit high sensitivity to fuel price spikes, as they lack adaptive mechanisms to hedge economic risk. Standard SAC improves adaptability but remains focused on energy metrics rather than cost volatility. LPG-SAC-TCO consistently achieves the lowest TCO increase and variance, demonstrating strong robustness to price uncertainty. This advantage stems from two factors: the explicit inclusion of hydrogen price dynamics in the TCO reward formulation, and the LLM-based policy guidance module, which provides high-level cost-aware decision priors under changing market conditions. Consequently, LPG-SAC-TCO adjusts energy sourcing strategies more conservatively during high-price periods, reducing exposure to fuel cost risk and ensuring stable long-term economic performance.

Table 3. Performance under Hydrogen Price Volatility

Model	Avg. TCO Increase (%)	TCO Variance
Rule-Based EMS	14.5	High
Deterministic Optimization	11.2	Medium
Energy-Oriented SAC	10.5	Medium
Degradation-Aware RL	9.8	Medium-low
LPG-SAC-TCO (Proposed)	6.1	low

Figure 2 illustrates the power split trajectories between the fuel cell (FC) system and the battery (BAT) for different energy management strategies over a representative 600-s driving cycle. The horizontal axis denotes time in seconds, while the vertical axis represents power output in kilowatts. Results are reported for four controllers: Rule-based EMS, Energy-oriented SAC, Degradation-Aware RL, and the proposed LPG-SAC-TCO framework.

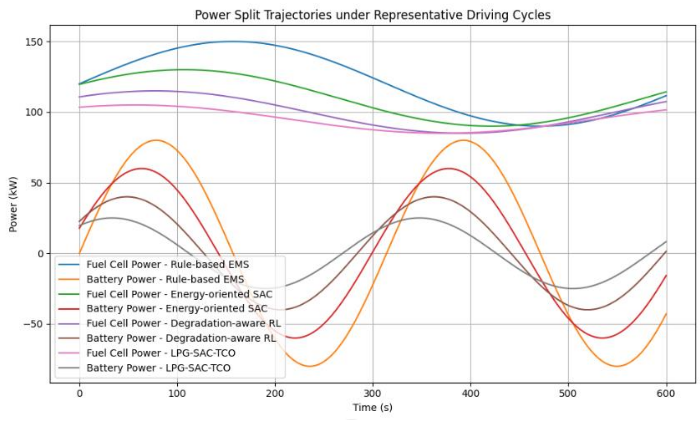

Figure 2. Power Split Trajectories under Representative Driving Cycles

The Rule-based EMS exhibits the largest power fluctuations. Fuel cell power varies significantly between approximately 95 kW and 150 kW, while battery power oscillates aggressively with peak magnitudes reaching $\pm$ (80)kW. Such high-frequency and high-amplitude battery usage indicates frequent charge-discharge cycles, which is consistent with the higher battery degradation cost and reduced component lifetime observed in the TCO analysis.

The Energy-oriented SAC strategy reduces overall fuel cell peaks to around 90-130 kW and limits battery power within $\pm$ (60)kW. Although this controller improves energy efficiency, it still relies heavily on the battery during transient load demands, leading to non-negligible degradation effects over long-term operation.

The Degradation-Aware RL further smooths power allocation behavior. Fuel cell power is maintained within a narrower band of 90-115 kW, while battery power oscillations are reduced to approximately $\pm$ (40)kW. This smoother profile reflects the controller's awareness of aging dynamics and aligns with the improved component lifetime metrics reported earlier.

In contrast, the proposed LPG-SAC-TCO framework demonstrates the most stable and conservative power split strategy. Fuel cell output remains tightly regulated between 85 kW and 105 kW, and battery power usage is significantly reduced, with fluctuations mostly confined within $\pm$ (25)kW. This behavior indicates that LPG-SAC-TCO prioritizes long-term economic optimality by minimizing high-stress operating regimes. The reduced reliance on aggressive battery cycling and the avoidance of fuel cell power spikes directly explain the lower lifecycle TCO, extended component lifetime, and improved robustness to hydrogen price volatility observed in the experimental results.

4.5. Discussion

The experimental results collectively demonstrate that optimizing energy management solely for efficiency is insufficient for hydrogen fuel cell heavy-duty trucks operating in real-world economic environments. By embedding financialized TCO modeling into deep reinforcement learning and augmenting policy learning with LLM-based guidance, LPG-SAC-TCO achieves a systematic shift from energy-optimal control to cost-optimal decision-making. The framework effectively aligns engineering control objectives with fleet-level financial goals, offering improved cost efficiency, enhanced component longevity, and greater robustness to market uncertainty. These findings suggest that the proposed approach has strong potential for deployment in logistics fleet management, vehicle leasing evaluation, and strategic procurement planning, bridging the gap between intelligent vehicle control and financial decision science.

5. Conclusion

This study addresses a critical limitation in existing energy management strategies for hydrogen fuel cell heavy-duty trucks (HFCTs), namely their predominant focus on short-term energy efficiency while neglecting long-term economic consequences. Motivated by the growing adoption of HFCTs in long-haul freight transportation and the increasing importance of cost-driven decision-making for fleet operators, this paper reformulates vehicle energy management as a financialized lifecycle optimization problem. To this end, we propose LPG-SAC-TCO, a Large Language Model--Policy-Guided Soft Actor-Critic framework designed to directly minimize the total cost of ownership (TCO) rather than fuel consumption alone.

The proposed framework integrates three tightly coupled components: a physics-based fuel cell hybrid powertrain model, a comprehensive lifecycle TCO model, and a deep reinforcement learning controller. A key innovation lies in the introduction of a large language model as a high-level policy guidance module. By extracting semantic and financial insights from complex operating conditions---such as driving patterns, payload variations, component health states, and hydrogen price fluctuations---the LLM dynamically generates cost-priority signals that guide the downstream SAC controller. This hierarchical design enables the learning agent to balance short-term operational efficiency with long-term economic objectives, bridging the gap between vehicle-level control and financial decision-making.

Extensive simulation experiments under real-world driving cycles and stochastic hydrogen price scenarios demonstrate the effectiveness of the proposed approach. Quantitatively, LPG-SAC-TCO achieves a 19% reduction in lifecycle TCO compared with rule-based energy management strategies, and an approximately 11% reduction relative to energy-oriented SAC controllers. In addition, the proposed framework significantly mitigates battery cycling stress and fuel cell power degradation, reducing battery cycle usage to about 55% of nominal lifetime and limiting fuel cell power fade to below 10% over long-horizon operation. Robustness analyses further show that LPG-SAC-TCO yields lower TCO variance and improved cost stability under hydrogen price volatility, confirming its suitability for real-world deployment.

From an application perspective, these results highlight the potential of LPG-SAC-TCO as a practical decision-support tool for logistics enterprises, fleet operators, and vehicle leasing companies. By providing TCO-aware energy management and cost prediction capabilities, the framework supports informed decisions in vehicle procurement, leasing evaluation, and long-term fleet planning. Furthermore, by formulating energy management through a financialized lens, this framework aligns with cutting-edge intelligent decision support systems utilized in fintech and investment risk management, thereby extending the horizon of traditional vehicle energy management toward comprehensive economic asset management.

Despite these promising results, this study has several limitations. The current framework relies on simulated degradation and cost models calibrated from industry benchmarks, which may not fully capture all real-world uncertainties. Future work will focus on validating the proposed approach using large-scale field operational data and extending the framework to multi-vehicle fleet-level optimization. Moreover, incorporating additional financial factors such as carbon pricing mechanisms, policy incentives, and dynamic contract structures represents a promising direction to further enhance the economic relevance of TCO-oriented energy management for hydrogen-powered freight transportation.

In summary, this study proposes a TCO-oriented energy management framework for hydrogen fuel cell heavy-duty trucks by combining lifecycle cost modeling with reinforcement learning--based control. The results demonstrate that shifting the optimization objective from energy efficiency to economic optimality can lead to lower lifecycle costs, reduced component degradation, and improved robustness to fuel price uncertainty. This work provides useful insights for the development of economically sustainable control strategies in hydrogen-powered freight transportation.

References

1. P. Rodatz, G. Paganelli, A. Sciarretta, et al., "Optimal power management of an experimental fuel cell/supercapacitor-powered hybrid vehicle," *Control Engineering Practice*, vol. 13, no. 1, pp. 41–53, 2005.
2. A. Sciarretta and L. Guzzella, "Control of hybrid electric vehicles," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 60–70, 2007.
3. Y. Gao, L. Chen, and M. Ehsani, "Investigation of the effectiveness of regenerative braking for EV and HEV," in *Future Transportation Technology Conference & Exposition*, SAE Technical Paper, 1999.
4. L. Nan, F. Xu, P. C. Sui, et al., "Techno-economic analysis of fuel cell trucks with different powertrain hybridization and hydrogen resources by 2040: case study of China," *International Journal of Hydrogen Energy*, vol. 132, pp. 174–182, 2025.
5. U. R. Sontakke, S. Jaju, and D. K. Mahajan, "Total cost of ownership analysis of fuel cell electric vehicles in India," in *Towards Hydrogen Infrastructure*, Elsevier, 2024, pp. 379–400.
6. B. Yu, X. Jia, T. Han, et al., "Modeling and degradation-aware energy management strategy for fuel cell vehicle considering cathode catalyst layer status," *Journal of Power Sources*, vol. 668, p. 239228, 2026.
7. Y. Li and H. He, *Deep Reinforcement Learning-Based Energy Management for Hybrid Electric Vehicles*. Morgan & Claypool Publishers, 2022.
8. R. Xiong, J. Cao, and Q. Yu, "Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle," *Applied Energy*, vol. 211, pp. 538–548, 2018.
9. 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.
10. S. Yao, J. Zhao, D. Yu, et al., "React: Synergizing reasoning and acting in language models," *arXiv preprint arXiv:2210.03629*, 2022.

11. Q. Nie and T. Liu, "Large language models: Tools for new environmental decision-making," *Journal of Environmental Management*, vol. 375, p. 124373, 2025.
12. Y. Huang, L. J. Wan, H. Ye, et al., "New solutions on LLM acceleration, optimization, and application," in *Proceedings of the 61st ACM/IEEE Design Automation Conference*, 2024, pp. 1–4.
13. Z. Ning, H. Zeng, and Z. Tian, "Research on data-driven energy efficiency optimisation algorithm for air compressors," in *Third International Conference on Advanced Materials and Equipment Manufacturing (AMEM 2024)*, SPIE, vol. 13691, pp. 1068–1075, 2025.
14. S. Xu, L. Jiang, and B. Gu, "Design and Validation of a Smart Neuromorphic System Architecture for Algorithmic Trading," in *Proceedings of the 2nd International Symposium on Integrated Circuit Design and Integrated Systems*, 2025, pp. 127–136.
15. 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.