Energy 359 (2026) 141442

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

High-altitude adaptive net power optimization of fuel cell systems using deep reinforcement learning

Bieshu Jin [a], Tianfeng Tang [a], Zhongkun Xiao [a], Chaoyi Chen [b], Bo Wang [c], Qing Shi [a], Jin Zhao [a], Guangwei Wang [a,*]

[a] School of Mechanical Engineering, Guizhou University, Guiyang, 550025, China
[b] School of Vehicle and Mobility, Tsinghua University, Beijing, 100084, China
[c] Department of Mechanical Engineering, The City College of New York, NY, 10031, USA

ARTICLE INFO

Keywords:
Fuel cell systems
Net power optimization
Deep reinforcement learning
High-altitude mountainous regions

ABSTRACT

Fuel cells offer zero-emission and high-efficiency power generation, yet their net power capability and durability degrade severely at high altitudes due to reduced oxygen partial pressure. To address the resulting non-stationary operating conditions, this paper proposes an altitude-aware deep reinforcement learning strategy built on Twin Delayed Deep Deterministic Policy Gradient (H-TD3) for real-time air supply optimization. By integrating real-time altitude into the agent state, the control policy adapts continuously to changing atmospheric conditions. The agent is trained to maximize net power while enforcing the oxygen excess ratio (OER) safety window and minimizing hydrogen consumption. Under a varying-altitude driving cycle, H-TD3 raises the peak net power from 38.09 kW (conventional feedforward baseline) to 42.93 kW (+12.7%) and constrains the OER to 1.74–3.19, preventing oxygen starvation. Compared with the ablated TD3 agent without altitude information, H-TD3 reduces cumulative hydrogen consumption by 1.54% at comparable peak power. Data-driven simulations using RTK-GPS altitude profiles collected from real plateau roads show that H-TD3 restores 97.76% of the ideal sea-level net power capability, demonstrating robust adaptability for fuel-cell vehicles operating in mountainous regions.

1. Introduction

Fuel cells are efficient, environmentally friendly energy technologies with promising applications in aerospace, electric vehicles, and ships [1]. Among various fuel cell types, PEMFC is widely adopted in automotive powertrains owing to its high power density and rapid dynamic response, while alternative technologies such as direct ammonia fuel cells [2] and direct methanol fuel cells [3] have also attracted growing interest for heavy-duty and marine applications. With the rapid growth of the FCHEV industry, their performance in high-altitude environments has become a critical concern [4]. In high-altitude environments, FCHEVs encounter challenging operational conditions characterized by reduced atmospheric pressure and diminished oxygen availability [5], thereby necessitating advanced control strategies to unlock their full operational potential.

Research indicates that the maximum net output power of fuel cell engines decreases by about 6% for every 1000-meter rise in altitude [6]. This performance degradation stems from the synergistic effects of altitude-induced environmental pressure attenuation (atmospheric pressure drops ~12% per 1000 m) and temperature gradient decline (−6.5 °C/km), establishing challenging low-pressure, low-temperature operational environments. Under these harsh conditions, cathode oxygen partial pressure diminishes substantially, directly impeding electrochemical reaction kinetics within the fuel cell. This leads to pronounced activation polarization losses and compromised proton migration rates in PEMFC [7], culminating in diminished FCS output power and operational efficiency, with potential consequences including performance deterioration and premature stack degradation due to oxygen starvation. To mitigate high-altitude oxygen deficiency, increasing air flow rates and system pressure can alleviate adverse effects [8]. However, this approach presents inherent trade-offs: excessive OER and compressor pressure ratios induce proton exchange membrane dehydration, generate substantial anode-cathode pressure differentials, and dramatically escalate parasitic power consumption in auxiliary systems, particularly compressors [9]. Consequently, optimizing the delicate balance between the OER, compressor pressure ratio, and net power output has emerged as a critical research frontier [10,11]. The

* Corresponding author.
E-mail address: gwwang@gzu.edu.cn (G. Wang).

https://doi.org/10.1016/j.energy.2026.141442
Received 21 October 2025; Received in revised form 2 March 2026; Accepted 20 May 2026
Available online 26 May 2026

Nomenclature

Abbreviations

DDPG	Deep Deterministic Policy Gradient
DQN	Deep Q-Networks
DRL	Deep Reinforcement Learning
FCHEV	Fuel Cell Hybrid Electric Vehicle
FCS	Fuel Cell System
GIS	Geographic Information System
GPS	Global Positioning System
OER	Oxygen Excess Ratio
PEMFC	Proton Exchange Membrane Fuel Cell
RTK	Real-Time Kinematic
TD3	Twin Delayed Deep Deterministic Policy Gradient

Symbols

A	Active area (cm^2)
F	Faraday constant (C/mol)
$F_{(\cdot)}$	Mass flow rate (kg/s)
h	Altitude (m)
I	Current (A)
J_{cp}	Compressor moment of inertia (kg m^2)
m_{H_2}	Hydrogen consumption (g)
N	Number of cells
P	Power (kW)
p	Pressure (Pa)
Q_{heat}	Reaction heat (kW)
R	Gas constant (J/(mol K))
T	Temperature (K)
V	Voltage (V)
γ	Specific heat ratio
η	Efficiency
λ_{O_2}	Oxygen excess ratio
ρ	Air density (kg/m^3)
τ	Torque (N m)
ω	Angular speed (rad/s)

Subscripts

act	Activation
atm	Atmospheric
ca	Cathode
cm	Compressor motor
con	Concentration
cp	Compressor
fc	Fuel cell
in	Inlet
net	Net
ohm	Ohmic
out	Outlet
sm	Supply manifold
st	Stack
w	Water

development of FCS control strategies specifically engineered for high-altitude operation, with a particular emphasis on precision air supply system management, thus represents a pivotal technological challenge in advancing FCHEV high-altitude adaptability.

Regarding the optimization and control of FCS in high-altitude environments, scholars have conducted extensive research. Xu et al. [12] proposed a comprehensive closed-loop control strategy that integrates virtual calibration and feedforward compensation. This approach enables independent and coordinated regulation of air flow and pressure, significantly enhancing the dynamic response performance of FCS in high-altitude environments. Chen et al. [13] proposed a real-time power optimization strategy based on particle swarm optimization feedforward and perturbation and observation algorithm, finding that FCS power decreased by 9.1% at 3000 m altitude under a 300 A load current demand. Li et al. [14] proposed a power optimization method based on extremum seeking and model-free adaptive control to improve system net power at different altitudes through adaptive adjustment of the OER. Yan et al. [15] proposed an observer-based nonlinear three-level controller that adjusts the OER to the reference values under rapid current transitions, avoiding oxygen starvation and achieving maximum net power. Jian et al. [16] proposed a third-order model of an air supply system with a feedback linearization controller, avoiding oxygen starvation and reducing power consumption by tracking optimal reference OER values. Qi et al. [17] proposed a hierarchical performance enhancement control strategy to handle optimal trajectory planning and nonlinear control problems under multi-objective constraints. Yin et al. [18] proposed an online supervisory control method combining model-guided extremum seeking and active disturbance rejection control, solving the online maximum efficiency trajectory search problem and achieving real-time system efficiency optimization. Xie et al. [19] proposed a high-altitude performance guarantee control strategy, achieving net power optimal trajectory search and optimal trajectory tracking for multi-input multi-output systems, effectively improving the net power of high-altitude PEMFC systems. Wang et al. [20] proposed a coordinated optimal allocation strategy for multi-stack FCS at different altitudes, achieving optimal power allocation among systems, resulting in higher average efficiency and lower hydrogen consumption.

Nevertheless, these control frameworks often exhibit inherent limitations in dynamic response characteristics and robustness, rendering them inadequate for effectively managing the intricate coupling dynamics caused by changes in atmospheric pressure at high altitudes. This challenge is particularly pronounced in FCS characterized by strong multi-variable interdependencies, where the establishment of precise mathematical models remains elusive, thereby presenting significant obstacles to real-time control implementation [21]. To address these challenges, some researchers have applied DRL algorithms for net power optimization in air supply systems. For instance, Jia et al. [22] proposed a multi-agent distributed controller based on DDPG, using multi-agent exploration to optimize PEMFC air supply. Wang et al. [23] developed an adaptive control strategy integrating operating parameter optimization, combining fuzzy Q-learning for OER and pressure reference optimization with adaptive neural network control to coordinate compressors and flow valves. Their results demonstrated superior transient and steady-state performance under varying conditions compared to traditional methods. The TD3 algorithm has recently gained attention in FCS control for its ability to overcome Q-value overestimation and stabilize learning. Deng et al. [24] applied TD3 to optimize energy management in fuel cell hybrid rail vehicles, simultaneously minimizing hydrogen consumption and fuel cell degradation under stochastic operational conditions. Tao et al. [25] utilized a TD3-based strategy combined with terrain prediction to improve both the power allocation and fuel economy of FCHEVs. Furthermore, Li et al. [26] implemented an improved TD3 algorithm as a real-time tuner for PID controllers to achieve precise temperature regulation in PEMFC systems, demonstrating strong robustness against nonlinear uncertainties. While DRL has demonstrated success in optimizing FCS under standard conditions, these works implicitly assume a stationary environment. The continuous variation of altitude during real-world driving introduces a critical non-stationarity to the system dynamics. A DRL agent trained at a fixed altitude would likely fail catastrophically when deployed in high-altitude regions due to a significant state-distribution shift. Therefore, merely applying existing DRL methods is insufficient. The fundamental challenge lies in enabling the agent to perceive and adapt to this environmental non-stationarity in real-time.

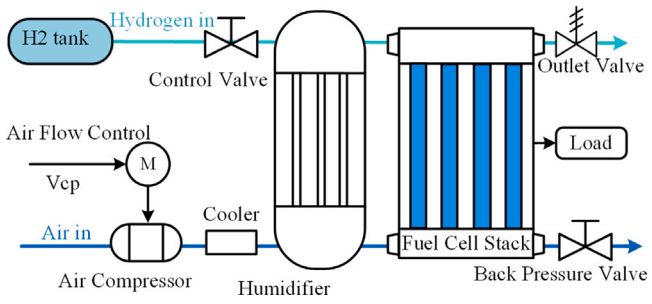

Fig. 1. Topology of the fuel cell system.

To address this research gap, we hypothesize that high-altitude fuel cell operation should be modeled as a *structured non-stationary decision process* rather than a stationary control problem with disturbances. Specifically, we posit that explicitly incorporating real-time altitude into the DRL state representation enables the agent to internalize environmental variation and learn altitude-consistent policies. Based on this premise, we develop an altitude-aware TD3 framework (H-TD3), in which altitude is treated as a dynamic state variable governing policy adaptation. This design allows the agent to proactively adjust control actions under continuous elevation changes, leading to coordinated optimization of net power, OER regulation, and hydrogen consumption. The main contributions are summarized as follows:

(i) An altitude-aware control paradigm is introduced, where real-time altitude is embedded as an explicit conditioning variable in the DRL state representation. This formulation converts implicit environmental non-stationarity into an observable structured variation, enabling proactive policy adaptation under continuous elevation changes.
(ii) A multi-objective H-TD3 framework is developed with a safety-prioritized reward decomposition that coordinates net power maximization, OER constraint enforcement, and hydrogen consumption reduction while preserving stack health.
(iii) Comprehensive validation using real plateau altitude profiles demonstrates that the proposed method mitigates performance degradation under elevation variation, recovering up to 97.76% of the nominal sea-level net power and improving peak net power by 12.7% compared with conventional feedforward strategies.

The remainder of this paper is organized as follows. Section 2 introduces the FCS model and analyzes the impact of high-altitude environments on system performance. Section 3 outlines the proposed framework and optimization methodology. Section 4 presents and discusses the results under varying altitude conditions. Finally, Section 5 concludes the paper.

2. System modeling and high-altitude impact analysis

This section details the development of a high-fidelity FCS model, which serves as the simulation testbed for our control strategy design and validation. The model treats altitude as a time-varying boundary condition and integrates sub-models for the atmospheric environment, the fuel cell stack, and the air supply system. The overall system architecture is depicted in Fig. 1.

2.1. Atmospheric environment model

To quantify the impact of altitude on ambient conditions, the International Standard Atmosphere model is employed to determine the temperature, pressure, and air density as functions of altitude [27].

$$\begin{cases} T_h = T_0 - \frac{6.5h}{1000} \\ p_h = p_0\left(1 - \left(0.0065 \cdot \frac{h}{T_0}\right)\right)^{5.2561} \\ \rho_h = \frac{p_h M_{air}}{R \cdot T_h}\left(1 - x_w\left(1 - \frac{M_w}{M_{air}}\right)\right) \end{cases} \tag{1}$$

where T_h is the environmental temperature at altitude h, with T_0 being the sea-level temperature (25 °C); p_h and p_0 are the air pressures at altitude h and sea level (1 atm = 0.101325 MPa), respectively; ρ_h is the air density at altitude h; M_{air} and M_w are the molar masses of air and water vapor, respectively; x_w is the molar fraction of water vapor; and R is the gas constant.

2.2. Fuel cell stack modeling

2.2.1. Voltage output model

The stack model describes the relationship between the stack output voltage and various operating parameters, including system voltage losses that comprise activation overpotential, ohmic resistance losses, and concentration polarization losses. The stack's output voltage is determined by subtracting these voltage losses from the open-circuit voltage. For a single fuel cell, the output voltage [4,28] is calculated as:

$$V_{fc} = E_{Nernst} - V_{act} - V_{ohm} - V_{con} \tag{2}$$

The Nernst voltage and voltage losses [29] are calculated by:

$$\begin{cases} E_{Nernst} = 1.229 - 0.00085(T - 298.15) + \frac{RT[\ln(P_{H_2} \times \sqrt{P_{O_2}})]}{2F} \\ V_{act} = \xi_1 + \xi_2 T + \xi_3 T[\ln(\frac{P_{O_2}}{5.08\times10^6 \exp(\frac{-498}{T})})] + \xi_4 T[\ln(I_{fc})] \\ V_{ohm} = \frac{r_M I}{A}, \quad r_M = \frac{181.6[1+0.03\frac{I_{fc}}{A}+0.062(\frac{T}{303})^2(\frac{I_{fc}}{A})^{2.5}]}{[2-0.634-3\frac{I_{fc}}{A}]\exp^{4.18(\frac{T-303}{T})}} \\ V_{con} = \frac{RT}{2F}\ln(1 - \frac{I_{fc}}{I_{fc,\max}}) \end{cases} \tag{3}$$

where V_{fc} is the fuel cell voltage, E_{Nernst} denotes the Nernst voltage, V_{act} represents the activation voltage loss, V_{ohm} is the ohmic voltage loss, V_{con} is the concentration polarization voltage loss, T is the fuel cell operating temperature in Kelvin (K), P_{H_2} and P_{O_2} represent the hydrogen and oxygen partial pressures in Pascals (Pa), I_{fc} denotes the fuel cell current, $I_{fc,\max}$ denotes the maximum current, R is the gas constant, F is the Faraday constant, ξ_i are constants and the parameter values are shown in Table 1, r_M is the ohmic internal resistance, and A is the effective cross-sectional area of the FCS [30].

For a stack composed of N cells, the total output voltage V_{st} of the FCS can be expressed as:

$$V_{st} = N \cdot V_{fc} \tag{4}$$

Therefore, the stack output power P_{st} and hydrogen consumption m_{H_2} are calculated as [31]:

$$\begin{aligned} P_{st} &= V_{st} I_{st} \\ m_{H_2} &= \frac{1.2 N M_{H_2} I_{fc}}{2F} \end{aligned} \tag{5}$$

where V_{fc} and I_{fc} are the output voltage and current of a single fuel cell, and I_{st} is the stack current. Since the cells are connected in series, $I_{st} = I_{fc}$.

2.2.2. Cathode flow model

The cathode flow dynamics are modeled based on the principles of mass conservation and air thermodynamics with the following assumptions [32]: (1) The stack operating temperature T is regulated by the thermal management system and remains constant at 353.15 K; (2) the cathode inlet gas temperature equals the stack temperature, and the gas temperature within the cathode is spatially uniform; and (3) when the

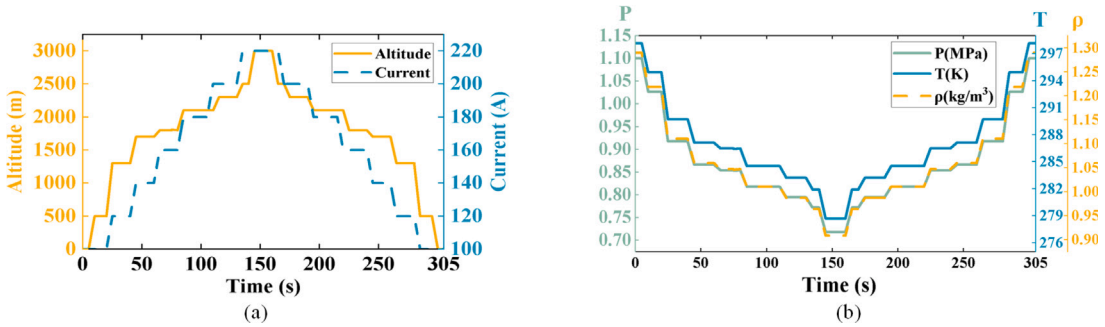

Fig. 2. System input profiles: (a) Altitude and load current; (b) Corresponding atmospheric pressure (P), temperature (T) and air density (ρ).

Table 1
Parameters of the fuel cell stack model.

Parameter	Value	Parameter	Value
Temperature T	353.15 K	ξ_1	−0.948
Gas constant R	8.314 J/(mol K)	ξ_2	0.00354
Faraday constant F	96485 C/mol	ξ_3	0.000078
Number of cells N	281	ξ_4	−0.000196
Active area A	282 cm^2	$I_{fc,\max}$	1.5

relative humidity of the gas is below 100%, liquid water evaporates and accumulates inside the cathode. The mass balance equations for oxygen, nitrogen, and water vapor within the cathode volume are given by [33]:

$$\begin{cases} \frac{dm_{O_2}}{dt} = F_{O_2,ca,in} - F_{O_2,ca,out} - F_{O_2,reacted} \\ \frac{dm_{N_2}}{dt} = F_{N_2,ca,in} - F_{N_2,ca,out} \\ \frac{dm_{w,ca}}{dt} = F_{w,ca,in} - F_{w,ca,out} + F_{w,ca,gen} - F_{w,mem} \end{cases} \tag{6}$$

where m_{O_2}, m_{N_2}, and $m_{w,ca}$ represent the masses of oxygen, nitrogen, and water vapor at the cathode, respectively; $F_{O_2,ca,in}$, $F_{O_2,ca,out}$, $F_{N_2,ca,in}$, $F_{N_2,ca,out}$, $F_{w,ca,in}$, and $F_{w,ca,out}$ represent the mass flow rates of oxygen, nitrogen, and water vapor entering and exiting the cathode, respectively; $F_{O_2,reacted}$ and $F_{w,ca,gen}$ represent the oxygen consumption rate and water vapor generation rate at the cathode, respectively, which can be determined from electrochemical reaction equations; $F_{w,mem}$ represents the mass transfer rate of water through the proton exchange membrane.

The inlet mass flow rates of the constituent gases are determined by the air composition and the total inlet flow from the supply manifold:

$$\begin{cases} F_{O_2,ca,in} = x_{O_2} \times F_{air,ca,in} \\ F_{N_2,ca,in} = (1 - x_{O_2}) \times F_{air,ca,in} \\ F_{w,ca,in} = F_{ca,in} - F_{O_2,ca,in} - F_{N_2,ca,in} \\ x_{O_2} = \frac{y_{O_2} \times M_{O_2}}{y_{O_2} \times M_{O_2} + (1 - y_{O_2}) \times M_{N_2}} \end{cases} \tag{7}$$

where $F_{air,ca,in}$ represents the dry air mass flow rate at cathode, $F_{ca,in}$ denotes the total gas mass flow rate at cathode, which equals the mass flow rate $F_{sm,out}$ from the supply manifold; x_{O_2} is the oxygen mass fraction, y_{O_2} is the oxygen mole fraction, M_{O_2} is the molar mass of oxygen, and M_{N_2} is the molar mass of nitrogen.

The cathode outlet mass flow rate is modeled using a linearized orifice equation, assuming a small pressure drop to the return manifold:

$$F_{ca,out} = k_{ca,out}(P_{ca} - P_{rm}) \tag{8}$$

where $F_{ca,out}$ represents the total gas mass flow rate exiting the cathode, $k_{ca,out}$ is the cathode outlet flow coefficient, P_{ca} denotes the combined partial pressures of dry air and water vapor, and P_{rm} is the return manifold pressure. The dynamics on the anode side are modeled symmetrically and are omitted for brevity.

2.3. Key component models of the supply system

2.3.1. Compressor model

The air compressor serves as the interface between the FCS and the atmospheric environment. Its dynamic behavior [14,30] is described by:

$$\begin{cases} \tau_{cm} = \eta_{cm} \frac{k_t}{R_{cm}} \left(V_{cp} - k_i \omega_{cp} \right) \\ \tau_{cp} = \frac{C_p}{\omega_{cp}} \frac{T_{atm}}{\eta_{cp}} \left(\left(\frac{p_{sm}}{p_{atm}} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right) F_{cp} \\ \frac{d\omega_{cp}}{dt} = \frac{1}{J_{cp}} \left(\tau_{cm} - \tau_{cp} \right) \end{cases} \tag{9}$$

where J_{cp} is the moment of inertia of the compressor; τ_{cm} is the compressor motor driving torque calculated from the static motor equation; τ_{cp} is the compressor load torque calculated from the thermodynamic equation; ω_{cp} is the compressor speed; R_{cm} is the motor resistance; η_{cp} and η_{cm} are the mechanical efficiency of the air compressor and its motor, respectively; k_t and k_i are the torque constant and current constant, respectively; V_{cp} is the compressor input voltage; γ is the specific heat ratio; C_p is the specific heat capacity of air at constant pressure; F_{cp} is the compressor outlet mass flow rate; T_{atm} is the atmospheric temperature; p_{atm} is the atmospheric pressure; p_{sm} is the supply line pressure.

The air compressor power demand P_{cp} is the product of input voltage V_{cp} and current I_{cp}. The net power of the FCS, P_{net}, is the fuel cell stack output P_{st} minus the compressor power demand, as shown below:

$$P_{net} = P_{st} - P_{cp} = I_{st}V_{st} - I_{cp}V_{cp} \tag{10}$$

2.3.2. Supply and return manifold model

The dynamics of the supply manifold are modeled based on mass conservation [34], relating the change in gas mass to the inlet and outlet flow rates:

$$\frac{dm_{sm}}{dt} = F_{sm,in} - F_{sm,out}, \quad F_{sm,in} = F_{cp} \tag{11}$$

where m_{sm} is the gas mass in the supply manifold, $F_{sm,out}$ is the outlet mass flow rate of the supply manifold, and $F_{sm,in}$ is the inlet mass flow rate of the supply manifold.

Considering the energy balance and the ideal gas law, the pressure dynamics in the supply manifold are governed by:

$$\begin{cases} \frac{dp_{sm}}{dt} = \frac{\gamma R_a}{V_{sm}} \left(F_{sm,in} T_{cp,out} - F_{sm,out} T_{sm} \right) \\ F_{sm,out} = k_{sm,out}(p_{sm} - p_{ca}) \end{cases} \tag{12}$$

where V_{sm} is the supply manifold volume; R_a is the air gas constant; T_{sm} is the gas temperature in the supply manifold, determined from the ideal gas law; and $T_{cp,out}$ is the compressor outlet temperature. Since the pressure difference between the supply manifold and cathode is relatively small, a linear orifice equation is used to calculate the supply manifold outlet mass flow rate $F_{sm,out}$, where $k_{sm,out}$ is the supply manifold orifice constant and p_{ca} is the cathode pressure.

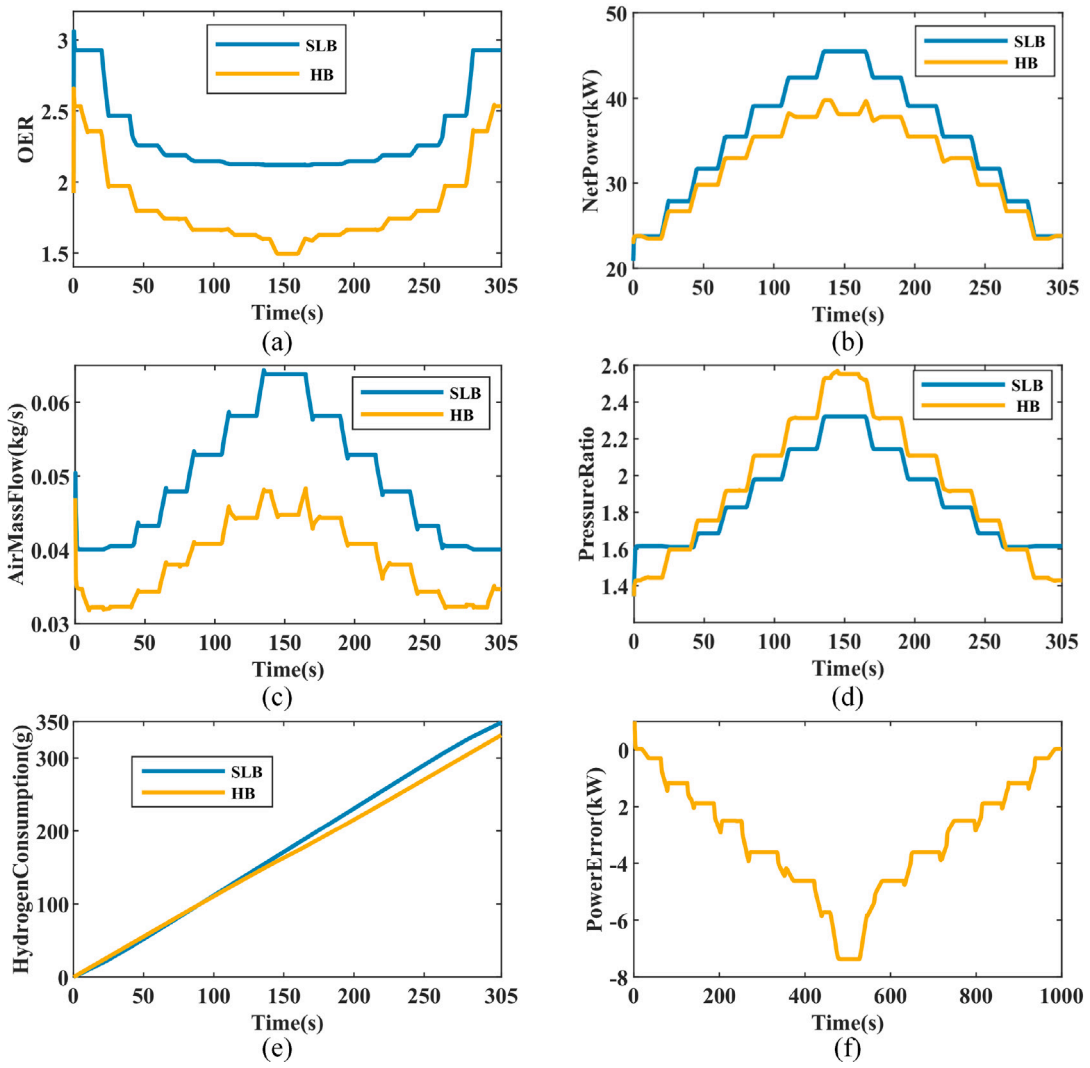

Fig. 3. Simulation results of the FCS with baseline feedforward control: (a) OER; (b) Net power; (c) Air mass flow; (d) Pressure ratio; (e) Hydrogen consumption; (f) Power error.

Table 2
FCS performance under conventional feedforward control.

Method	m_{H_2} (g)	Value	OER	Value	NetPower (kW)	Value
SLB	348.386	0.0	2.12–3.06	0.0/0.0	45.47	0.0
HB	331.395	+4.88%	1.49–2.65	−29.72%/−13.40%	38.09	−16.23%

The pressure in the supply manifold model determines the air compressor pressure ratio, while the pressure in the return manifold model determines the cathode outlet mass flow rate. Since the air temperature leaving the fuel cell stack is lower than that leaving the compressor, temperature variations in the return manifold are neglected. The return manifold pressure p_{rm} is calculated as:

$$\frac{dp_{rm}}{dt} = \frac{R_a T_{rm}}{V_{rm}}(F_{ca,out} - F_{rm,out}) \tag{13}$$

where V_{rm} is the return manifold volume, T_{rm} is the gas temperature, and $F_{ca,out}$ is the mass flow rate entering the return manifold. The return manifold outlet mass flow rate $F_{rm,out}$ is calculated using a nonlinear orifice equation.

2.4. System performance analysis in high-altitude environment

To quantify the performance degradation at high altitudes, we simulated the FCS using a conventional map-based feedforward controller. The analysis was designed as a controlled experiment with two scenarios. Both were subjected to an identical vehicle load and altitude trajectory sourced from a mountainous driving cycle, as depicted in Fig. 2.

The scenarios were: (1) Sea Level Baseline (SLB) which serves as the control group, simulating the system as if it were operating at a constant sea-level altitude; (2) High-altitude Baseline (HB), the test group, which accounts for the effects of altitude variation on system performance.

This comparative framework ensures that any observed performance difference is directly attributable to the effects of altitude. To isolate the air-supply dynamics from thermal effects, we assume an ideal thermal management system that regulates both the stack operating temperature and the cathode inlet temperature at 80 °C; altitude-induced ambient temperature variations only affect the compressor inlet boundary conditions through the International Standard Atmosphere model [35]. This driving profile (Fig. 2) serves as the standard input for all subsequent experiments in this paper.

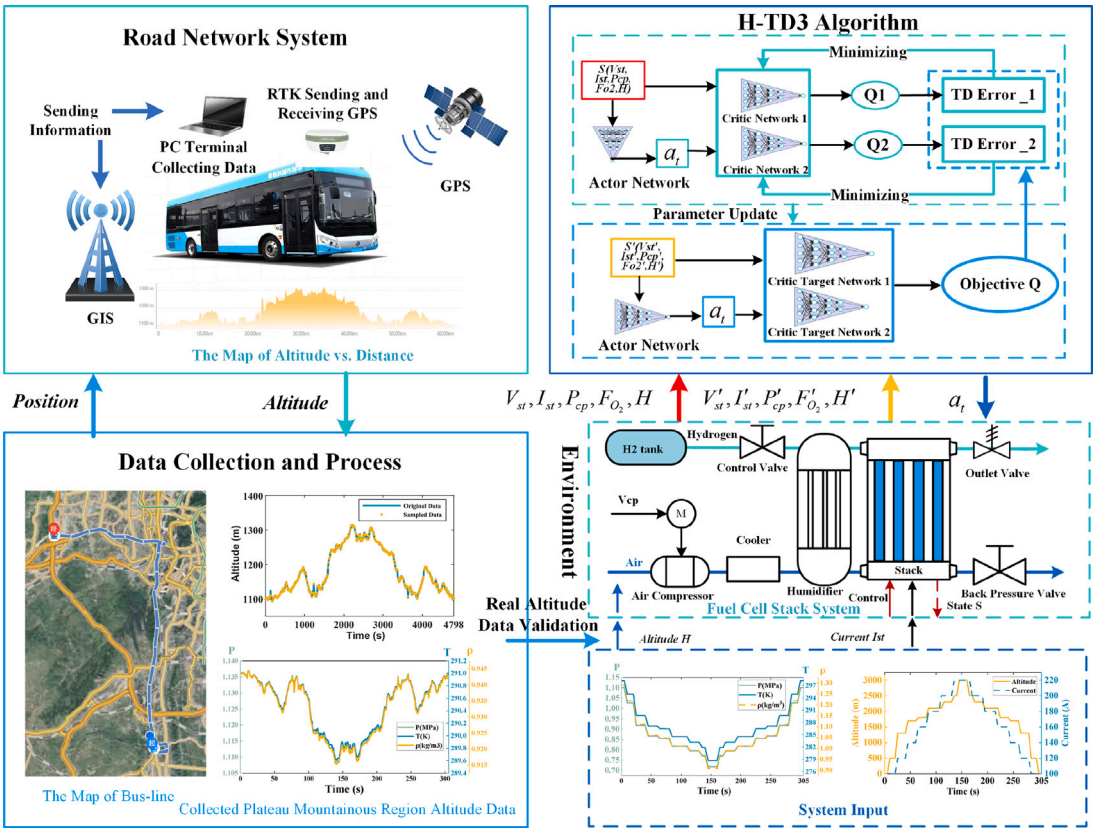

Fig. 4. H-TD3 algorithm framework.

The simulation results, presented in Fig. 3 and Table 2, reveal the severe performance penalty imposed by high-altitude operation. The net power output shows a progressive divergence from the sea-level baseline as altitude increases (Fig. 3(b)). At the cycle's peak elevation of 3000 m, the power deficit reaches 7.38 kW. Overall, the maximum achievable net power plummets from 45.47 kW at sea level to just 38.09 kW, a 16.23% reduction. This loss in net power is directly coupled with a compromised oxygen supply. As shown in Fig. 3(a), the OER value, a critical indicator of stack health, deteriorates significantly with altitude. The minimum OER value drops by 29.72% in the HB scenario, falling to levels that risk oxygen starvation, which can cause irreversible damage to the catalyst and membrane.

The reason for this performance collapse is the inherent limitation of the static feedforward controller, which fails to account for the reduced air density at higher altitudes. To deliver the required mass of oxygen from thinner air, the compressor is forced to spin faster, resulting in a much higher pressure ratio (Fig. 3(d)) and consequently, a massive increase in parasitic power consumption. Despite this increased effort, the air mass flow delivered to the stack is still insufficient (Fig. 3(c)), leading to the dangerously low OER. The controller is trapped in a dilemma: it either starves the stack of oxygen or consumes an excessive amount of power, in both cases crippling the net power output.

This baseline analysis highlights that conventional control strategies are fundamentally inadequate for ensuring both the efficiency and durability of FCS in vehicles operating across diverse topographies. This establishes the compelling need for an intelligent control framework that can dynamically adapt its strategy in real-time to the changing environmental conditions.

3. Net power optimization algorithm design

This study introduces a DRL framework to optimize the net power of FCS under varying altitude conditions. The proposed approach, termed H-TD3, is built upon the TD3 algorithm. As illustrated in Fig. 4, the framework integrates three core components: (1) a real-time environmental data module that uses RTK-GPS to provide the agent with accurate altitude information; (2) the H-TD3 agent, which learns to control the air supply system by processing system states and environmental data; and (3) a high-fidelity FCS model that serves as the simulation environment for agent training and evaluation. This integrated architecture enables the agent to autonomously learn a control policy that adapts to altitude variations, thereby enhancing both system performance and environmental robustness.

3.1. TD3 algorithm

The DDPG algorithm effectively combines DQN and Actor-Critic methodologies for continuous action spaces [36,37]. It employs a dual network architecture comprising a Q-value function (Critic) and a deterministic policy function (Actor) $\pi(s)$ parameterized by θ^{π}, with actions computed as $A_t = \pi(S_t|\theta_t^{\pi})$. To balance the exploration-exploitation trade-off, DDPG incorporates stochastic noise injection during training [38].

The action-value function $Q(s,a|\theta^{Q})$ is updated using the Bellman equation. Under state S_t, the policy π executes action $A_t = \pi(S_t|\theta^{\pi})$, obtaining next state S_{t+1} and reward R_t. The Critic network minimizes loss function L using gradient descent, while the Actor network is updated through policy gradient methods by applying the chain rule to the expected reward function $J = \mathbb{E}_{R_t,S_t\sim\mathcal{E},A_t=\pi}[R_t]$.

$$\begin{cases} \nabla_{\theta^{\pi}} J \approx \mathbb{E}_{S_t\sim\rho^{\beta}}\left[\nabla_{\theta^{\pi}} Q(s,a\mid\theta^{Q})\Big|_{s=S_t,a=\pi(S_t|\theta^{\pi})}\right] \\ \quad = \mathbb{E}_{S_t\sim\rho^{\beta}}\left[\nabla_a Q(s,a\mid\theta^{Q})\Big|_{s=S_t,a=\pi(S_t)} \nabla_{\theta^{\pi}}\pi(s\mid\theta^{\pi})\Big|_{s=S_t}\right] \\ \nabla_{\theta^{\pi}} J \approx \frac{1}{N}\sum_i \nabla_a Q(s,a\mid\theta^{Q})\Big|_{s=S_i,a=\pi(S_i)} \nabla_{\theta^{\pi}}\pi(s\mid\theta^{\pi})\Big|_{s=S_i} \end{cases} \tag{14}$$

Target networks are updated using exponential smoothing: $\theta^{Q'} \leftarrow \tau\theta^{Q}+(1-\tau)\theta^{Q'}$ and $\theta^{\pi'} \leftarrow \tau\theta^{\pi}+(1-\tau)\theta^{\pi'}$, where $\tau \ll 1$ enhances learning stability. However, DDPG suffers from Q-function overestimation bias, which accumulates during training and degrades policy quality.

The TD3 algorithm addresses these limitations through three key enhancements [39,40]:

(i) *Dual Q-network architecture*: Two independent Q-value estimators use the minimum estimate to reduce overestimation bias;
(ii) *Delayed policy updates*: Policy networks update less frequently than value networks, reducing estimation variance;
(iii) *Target policy smoothing*: Clipped Gaussian noise is added to target actions to regularize Q-value computation.

These algorithmic improvements enable TD3 to achieve superior performance compared to DDPG across multiple evaluation metrics, including training stability, convergence efficiency, and asymptotic performance, with particular advantages in high-dimensional continuous control tasks. Consequently, this research employs TD3 as the foundational algorithm for net power control optimization.

3.2. H-TD3 algorithm formulation

The high-altitude control strategy focuses on optimizing OER, net power, and pressure parameters [41] through real-time adjustment of air supply system control. The objective is to maximize net power output in high-altitude environments while maintaining OER within the optimal range of 1.5–3.5 [13], achieving dual improvements in FCS altitude adaptability and hydrogen fuel economy.

Given that air pressure and density vary with altitude and influence FCS compressor power requirements, this study defines the control system state variables as: stack current I_{st}, stack voltage V_{st}, compressor power P_{cp}, oxygen mass flow rate F_{O_2}, and real-time altitude H. To maintain FCS stability while ensuring efficient exploration by reinforcement learning agents, the agent action is defined as the desired change in net power variation dP_{net}, which serves as a setpoint for a low-level controller that modulates the compressor voltage. The state and action spaces are defined as:

$$\begin{cases} S = \left(V_{st}, I_{st}, P_{cp}, F_{\mathrm{O}_2}, H\right) \\ A = \left\{a_t \mid a_t = dP_{net} \in [-3, 3]\right\} \end{cases} \tag{15}$$

For the non-adaptive TD3 baseline, the altitude dimension H is excluded from the state vector.

The reward function is engineered to guide the agent toward multi-objective optimization: maintaining system power variation stability, ensuring fuel consumption economy, maximizing FCS output power, and keeping the OER value within the optimal range. Therefore, the cost function, $J_{P_{net}}$, is formulated to penalize undesirable behaviors. The final reward R_t at each timestep is the negative of this cost, $R_t = -J_{P_{net}}$. The cost function includes four weighted terms: the deviation from the maximum achievable net power at sea level ($P_{0,\max}$), the deviation of the OER λ_{O_2} from its optimal setpoint $\lambda_{\mathrm{O}_2,\mathrm{ref}}$ [13], the rate of hydrogen consumption (m_{H_2}), and the magnitude of the change in action to encourage smooth control.

$$J_{P_{net}} = \begin{cases} a|P_{net} - P_{0,\max}| + b|\lambda_{\mathrm{O}_2} - \lambda_{\mathrm{O}_2,\mathrm{ref}}| \\ \quad + c\, m_{\mathrm{H}_2} + d|a_t - a_{t-1}|, & \text{if } \lambda_{\mathrm{O}_2} \in [1.5, 3.5] \\ P_{penalty}, & \text{otherwise} \end{cases} \tag{16}$$

where a, b, c, d are positive weighting coefficients prioritized as $a > b > c > d$, and P_{penalty} is a large constant penalty (e.g., 1000) for violating OER safety constraints. The pseudocode for the H-TD3 algorithm is presented in Algorithm 1, and its hyperparameters are listed in Table 3.

Table 3
Hyperparameters of the H-TD3.

Parameter	Value
Optimizer	Adam
Actor learning rate	0.001
Critic learning rate	0.001
Discount factor γ	0.99
Delayed update frequency X	2
Update rate τ	0.005
Target smoothing noise	$\mathcal{N}(0, 0.4)$, $C = 0.5$
Exploration noise	$\mathcal{N}(0, 0.3)$
Hidden layers (neurons)	3 (64, 32, 16)
Experience replay capacity	10^6
Batch size	128

Remark 1. The hierarchy $a > b > c > d$ reflects the operational priorities of the FCS: net power maximization (weighted by a) is the primary propulsion objective; OER regulation (weighted by b) enforces the safety constraint against oxygen starvation; hydrogen economy (weighted by c) and control smoothness (weighted by d) are secondary objectives of comparatively lower sensitivity. Empirical tuning confirms that system performance is most sensitive to the ratio between a and b: under-weighting a leads to significant net power loss at elevation, whereas under-weighting b causes the OER to breach the starvation threshold (1.5) during transient load changes. The coefficients c and d exert a minor influence on policy behavior provided they remain sufficiently smaller than a and b.

Algorithm 1: H-TD3 Algorithm

Initialize critic networks $Q_{\theta_1}, Q_{\theta_2}$ and actor network π_ϕ with random parameters Q_1, Q_2, π ;
Initialize target networks $\theta_1' \leftarrow \theta_1, \theta_2' \leftarrow \theta_2, \phi' \leftarrow \phi$;
Initialize experience buffer $\mathcal{B}$;
for $t = 1$ **to** T **do**
  Select action with exploration noise: $A \sim \pi_\phi(s) + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma)$, execute action A and observe reward R and next state S';
  Store experience (S, A, R, S') in buffer $\mathcal{B}$;
  Sample mini-batch of N experiences (S, A, R, S') from $\mathcal{B}$;
  $\tilde{a} \leftarrow \pi_{\phi'}(s') + \epsilon$, $\epsilon \sim \mathrm{clip}(\mathcal{N}(0, \sigma), -c, c)$;
  Calculate target Q-values: $y \leftarrow R + \gamma \min_{i=1,2} Q_{\theta_i'}(S', \tilde{a})$;
  Update critic networks: $\theta_i \leftarrow \arg\min_{\theta_i} N^{-1} \sum (y - Q_{\theta_i}(s, a))^2$;
  if t mod X **then**
    Update actor network ϕ by the deterministic policy gradient: $\nabla_\phi J = N^{-1} \sum \nabla_a Q_{\theta_1}(s, a)\big|_{a=\pi_\phi(s)} \nabla_\phi \pi_\phi(s)$;
    Update target networks: $\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i'$;
    $\phi' \leftarrow \tau\phi + (1-\tau)\phi'$;
  end
end

4. Simulation and data-driven validation

To evaluate the performance of our proposed control strategy, we used a dynamic driving cycle that features simultaneous variations in load current and altitude, as depicted in Fig. 2. This profile serves as the standard test condition for all subsequent simulation experiments to ensure a fair comparison.

4.1. Baseline algorithm selection

To establish the effective foundation for the altitude-adaptive controller, we first compared three baseline DRL algorithms: PG, DDPG, and TD3. The agents were trained and evaluated on the dynamic

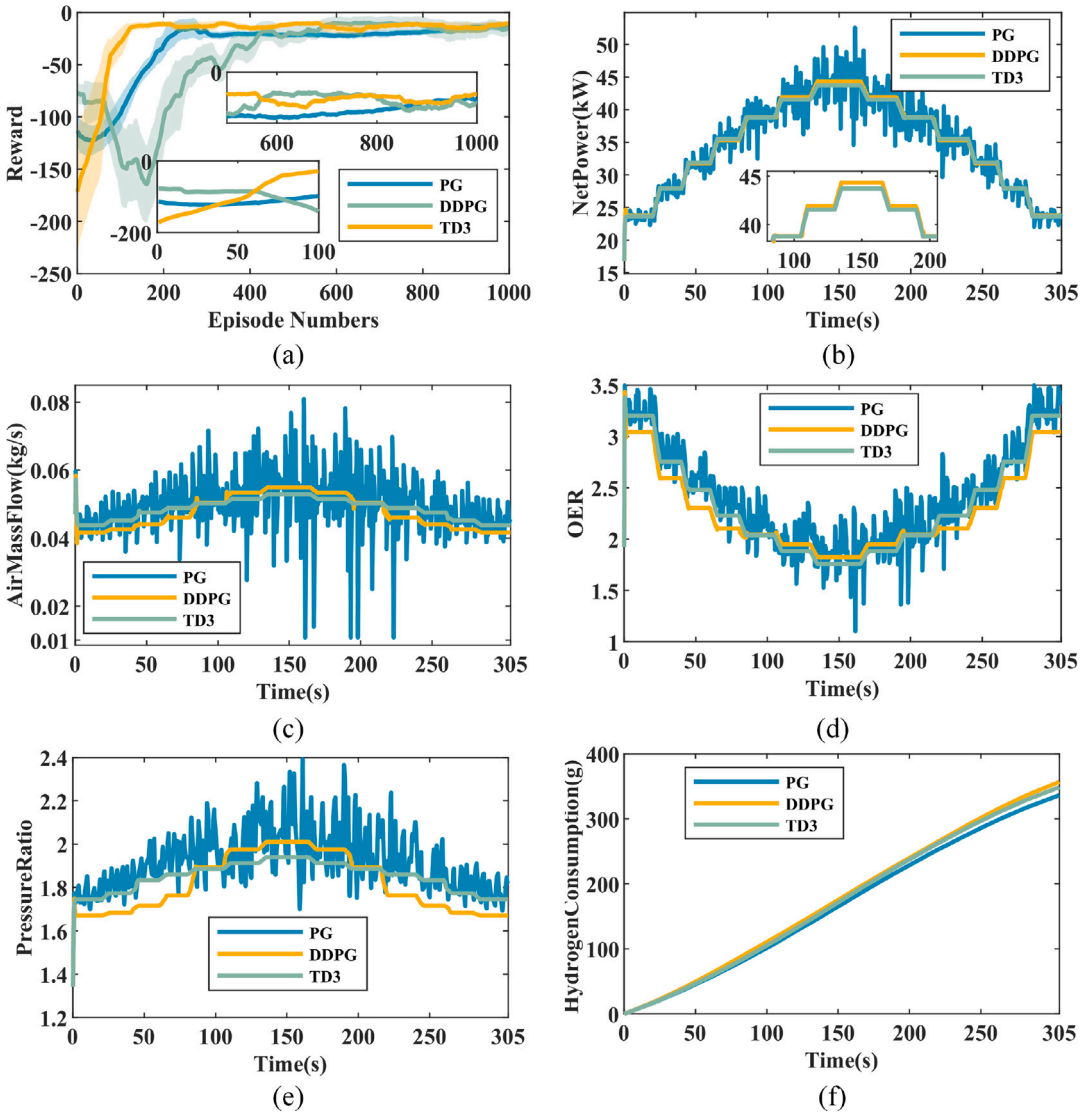

Fig. 5. Performance comparison of PG, DDPG, and TD3 under standard conditions.

driving cycle, as shown in Fig. 2. In this initial comparison, the agents were not provided with altitude information in their state space.

The results, presented in Fig. 5 and Table 4, highlight the superior performance of TD3. The convergence curves in Fig. 5(a) show that TD3, despite a slower initial start, converges rapidly to the highest and most stable reward plateau. In contrast, DDPG exhibits substantially slower convergence, while PG suffers from convergence to a suboptimal policy. Although DDPG appears to yield a slightly higher net power output during steady-state segments of the cycle (Fig. 5(b)), this marginal power advantage comes at the expense of control stability. DDPG exhibits aggressive and oscillatory control over the air mass flow and pressure ratio, particularly during load changes (Fig. 5(c, e)). These actions lead to larger OER fluctuations (Fig. 5(d)), which are heavily penalized by the reward function, resulting in a lower overall reward. Conversely, the TD3 agent learns a more balanced control strategy. It prioritizes system stability and operational safety, resulting in exceptionally smooth control over the air supply. As shown in Fig. 5(c, d, e), the TD3 policy effectively dampens the sharp transients and oscillations characteristic of DDPG and PG. It maintains the OER within a tighter and safer operational band, preventing both oxygen starvation and excessive parasitic losses. This balanced approach, which is rewarded more highly by the proposed multi-objective reward function, confirms its alignment with the objectives of long-term system durability and reliability.

Table 4
Training efficiency comparison of DRL algorithms.

Method	Episodes to convergence		Average time per episode	
	Value	Change	Value (s)	Change
PG	315	−136.84%	10.379	+42.93%
DDPG	435	−227.07%	19.075	−4.89%
TD3	133	0.0	18.186	0.0
H-PG	343	−157.89%	15.638	+14.01%
H-DDPG	131	+1.5%	33.879	−91.79%
H-TD3	246	−83.96%	18.170	+0.88%

Based on its superior convergence, stability, and control performance, TD3 was selected as the foundational algorithm for developing our altitude-aware strategy.

4.2. Performance of altitude-aware agents

Having established TD3 as the optimal algorithmic core, we next incorporated real-time altitude information into the state vector of

Table 5
FCS performance under DRL-based control strategies.

Method	m_{H_2} (g)	Value	OER	Value	NetPower (kW)	Value
SLB	348.386	0.0	2.12–3.06	0.0/0.0	45.47	0.0
PG	393.987	−13.09%	1.10–3.49	−48.11%/+14.05%	44.67	−1.76%
DDPG	356.663	−2.38%	1.82–3.42	−14.15%/+11.76%	44.33	−2.51%
TD3	348.662	−0.08%	1.76–3.17	−16.98%/+3.59%	43.71	−3.87%
H-PG	336.331	+3.46%	2.34–3.99	+10.38%/+30.39%	42.9	−5.65%
H-DDPG	367.817	−5.58%	1.43–3.09	−32.54%/+0.98%	43.13	−5.15%
H-TD3	343.31	+1.46%	1.74–3.19	−17.92%/+4.25%	42.93	−5.59%

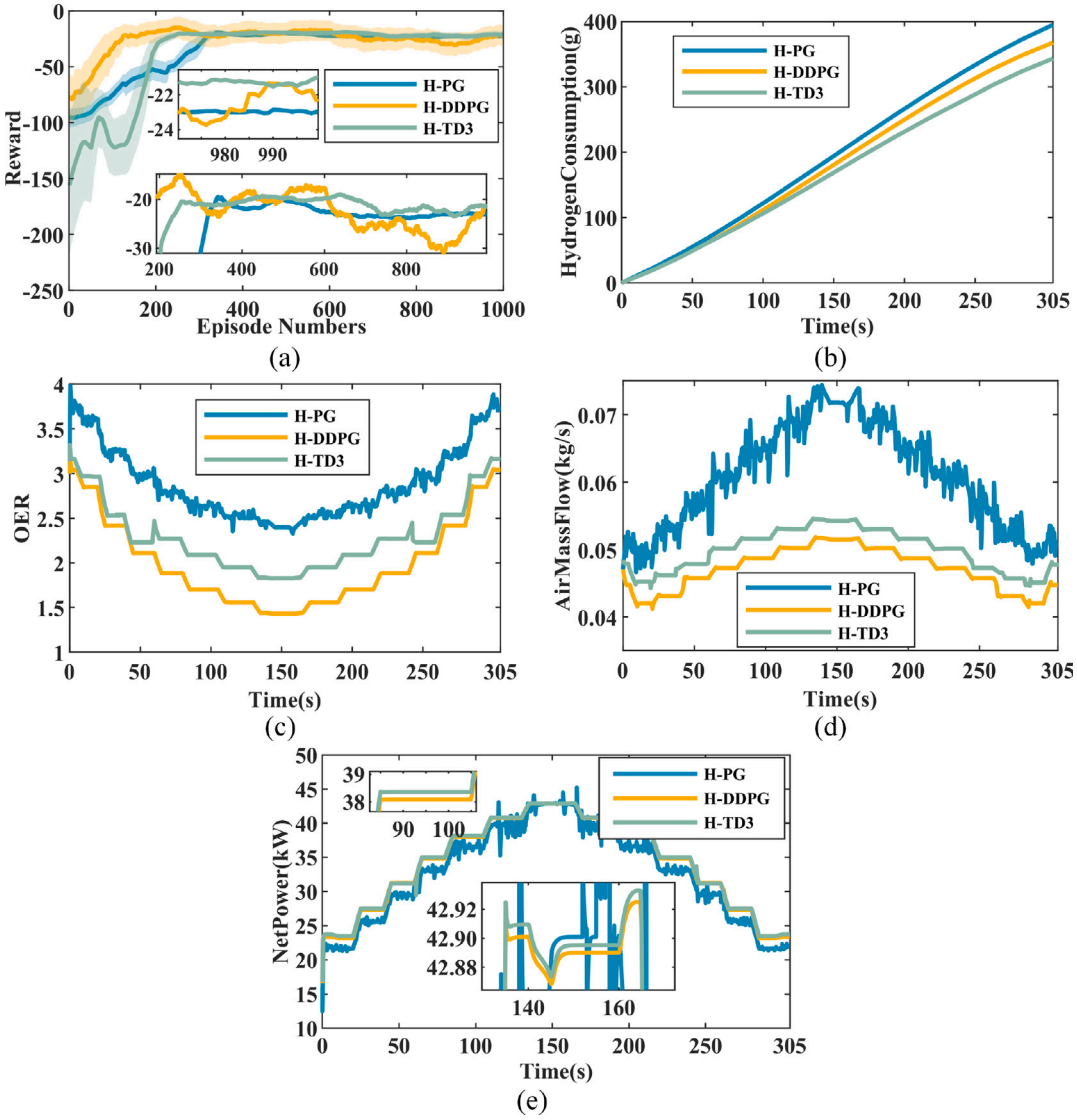

Fig. 6. Performance analysis of H-PG, H-DDPG, and H-TD3 under high-altitude variations.

each agent, yielding altitude-aware variants: H-TD3, H-DDPG, and H-PG. Their performance was then evaluated on the same dynamic, high-altitude driving cycle.

The convergence analysis presented in Fig. 6(a) and the efficiency metrics detailed in Table 4 reveal that H-TD3 converges at episode 246 and maintains stable reward performance thereafter. Although H-DDPG converges earlier, it exhibits persistent reward volatility, suggesting inadequate policy robustness that could compromise FCS reliability through repeated stress cycles. In terms of computational efficiency, H-DDPG requires approximately 86% more time per episode than H-TD3, representing a significant overhead penalty. Conversely, while H-PG demonstrates strong computational efficiency, its substantially lower reward performance indicates convergence to suboptimal solutions that fail to capture effective control policies.

The results in Fig. 6 and Table 5 demonstrate the superiority of the H-TD3 strategy. As shown in Fig. 6(b, c), the net power and OER curves under H-TD3 control are markedly smoother and more stable than those of H-DDPG. This validates the theoretical foundations of H-TD3: through double Q-network clipping and delayed policy updates, H-TD3 effectively addresses Q-value overestimation while eliminating the policy oscillations characteristic of DDPG, resulting in enhanced learning stability and control robustness.

Table 5 summarizes the cycle-level performance of the evaluated controllers. Under varying-altitude conditions, all altitude-aware agents

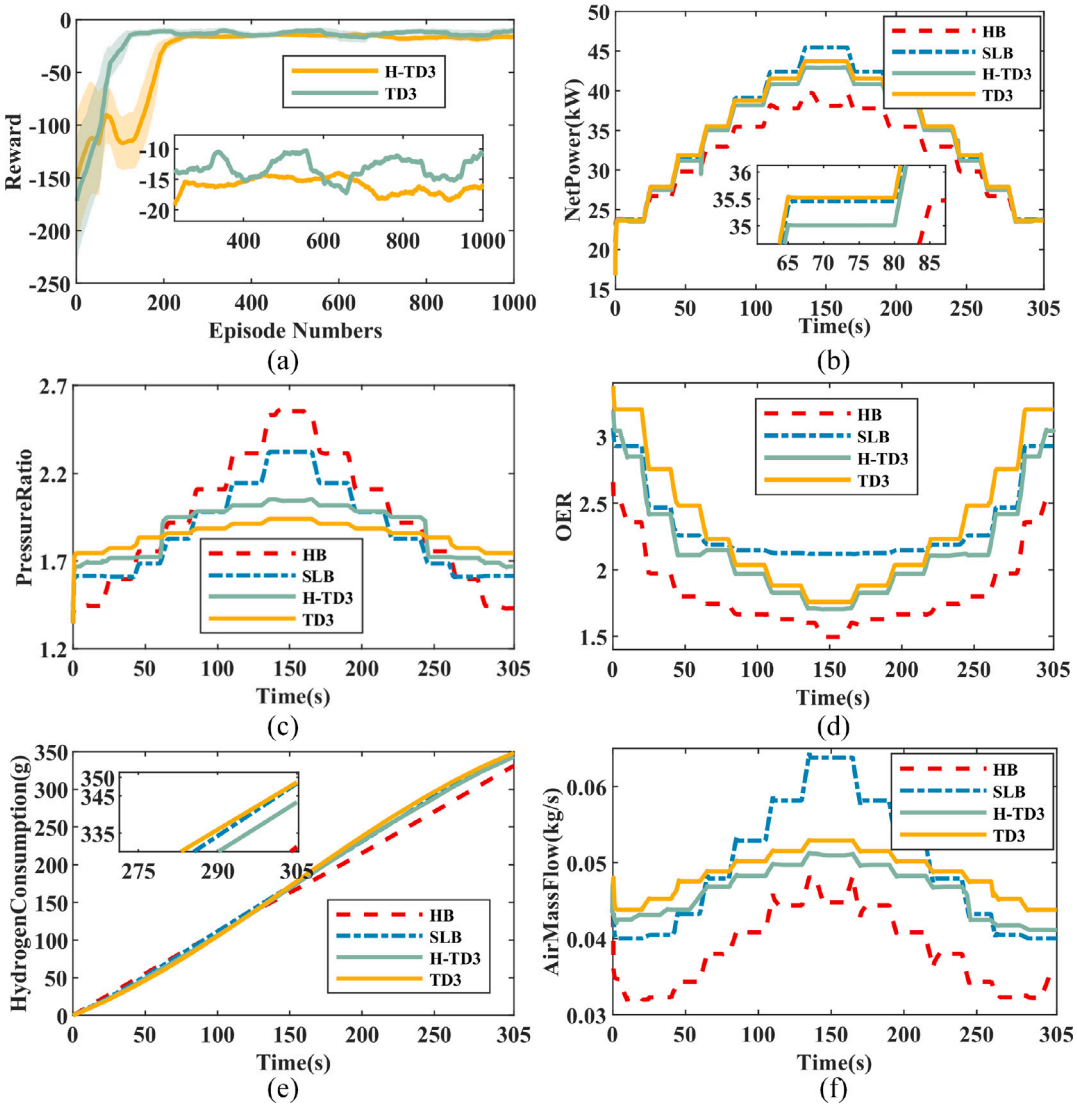

Fig. 7. Performance analysis of TD3 and H-TD3 under high-altitude variations.

substantially outperform the high-altitude baseline (HB), which employs a conventional feedforward controller and yields a peak net power of only 38.09 kW. At 3000 m elevation, H-TD3 achieves a peak net power of 42.93 kW while constraining the OER within 1.74–3.19, entirely preventing oxygen starvation. Compared with H-DDPG, H-TD3 delivers 6.67% lower cumulative hydrogen consumption and a substantially higher minimum OER (1.74 vs. 1.43), indicating improved fuel economy and enhanced oxygen-starvation robustness at comparable peak power levels. Although H-PG maintains operation within safe OER margins, its convergence to a suboptimal policy limits practical applicability. Notably, while HB exhibits the lowest hydrogen consumption of 331.395 g, this apparent efficiency stems from degraded chemical reactions at altitude rather than intelligent control compensation.

These results demonstrate that altitude-adaptive power optimization via H-TD3 yields substantial performance improvements for fuel cell systems operating in high-altitude environments. The proposed strategy effectively enhances stack output and energy utilization efficiency while ensuring robust operation under adverse atmospheric conditions, which are essential for maintaining system durability and reliability in plateau mountainous regions.

4.3. Ablation study: impact of altitude awareness

To quantify the contribution of altitude information, an ablation study was conducted comparing the proposed altitude-aware agent (H-TD3) against its ablated counterpart (TD3) that operates without real-time altitude data. Both agents share identical network architecture, hyperparameters, and reward structure, differing only in the presence of the altitude variable H in the state vector. This isolates the direct impact of altitude awareness on control performance.

The convergence curves (Fig. 7(a)) show that H-TD3 requires more training episodes than TD3. This is expected because the expanded state space demands more extensive exploration to learn a complex, generalized policy. However, H-TD3 ultimately achieves a higher and more stable final reward, indicating that the additional training overhead yields superior control performance.

The ablation comparison reveals that altitude awareness enables a more balanced control trade-off. As shown in Table 5, H-TD3 achieves comparable peak net power to the ablated TD3 agent (42.93 kW vs. 43.71 kW) while reducing cumulative hydrogen consumption by 1.54% (343.31 g vs. 348.66 g). Crucially, H-TD3 maintains a tighter and safer OER band: its minimum OER is only 0.94% below TD3, while

its maximum OER exceeds TD3 by 0.66%, demonstrating consistent performance across altitude gradients. This advantage stems from H-TD3's integration of altitude as a state variable, which allows the agent to anticipate altitude-induced air density changes and proactively adjust the oxygen supply strategy, rather than reactively compensating for performance degradation.

This trade-off also translates into a substantial advantage over the conventional feedforward baseline (HB): H-TD3 delivers a 4.84 kW net power improvement (12.7%) over HB's 38.09 kW, whereas the ablated TD3 achieves a marginally higher peak (43.71 kW) at the cost of 1.54% greater hydrogen consumption. The comparison indicates that altitude awareness does not merely shift performance metrics but enables a fundamentally more efficient allocation of compressor power across varying elevations.

This ablation study demonstrates that including altitude as a state variable is essential for robust, high-performance control in varying-altitude environments. The results confirm our hypothesis: by enabling the agent to perceive its operational environment, we transform it from a reactive controller into an adaptive system capable of handling real-world non-stationarity.

4.4. Validation and analysis at real high-altitude conditions

4.4.1. Data collection in plateau terrain

Plateau mountainous terrain presents unique driving challenges due to significant altitude variations. To accurately assess fuel cell vehicle energy consumption under these conditions, we conducted data collection in this challenging environment. The collected data serves to validate the applicability of existing models. This study utilizes a fixed bus route in plateau mountainous region for data collection. As shown in Fig. 8(a), this route traverses suburban, rural, and urban areas with substantial elevation changes, making it highly representative of plateau topography. Given the significant impact of steep terrain on hydrogen consumption, we employed RTK real-time navigation equipment (BDMC-02S) to collect altitude data along the entire route. As shown in Fig. 8(b), the navigation equipment mounted on the bus receives GPS signals and transmits them to an onboard computer, which processes the data to determine vehicle speed and position. At the same time, the GIS acquires GPS and RTK signals, sending them to a server for processing to provide real-time feedback and positioning corrections. The processed data are sent to PC software for real-time reception and storage. As shown in Fig. 9(a), the altitude profile from the collection route demonstrates the complex topographical features of this plateau mountainous region.

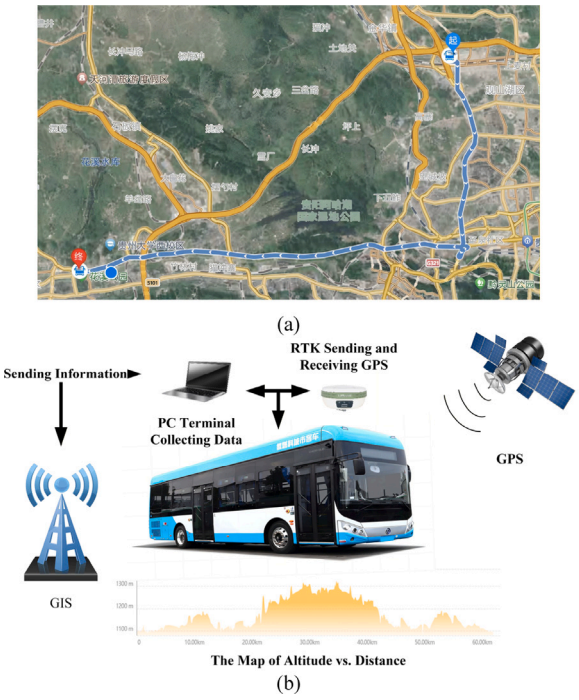

Fig. 8. Plateau mountainous data collection: (a) Map of bus route and stations; (b) Data acquisition equipment and workflow.

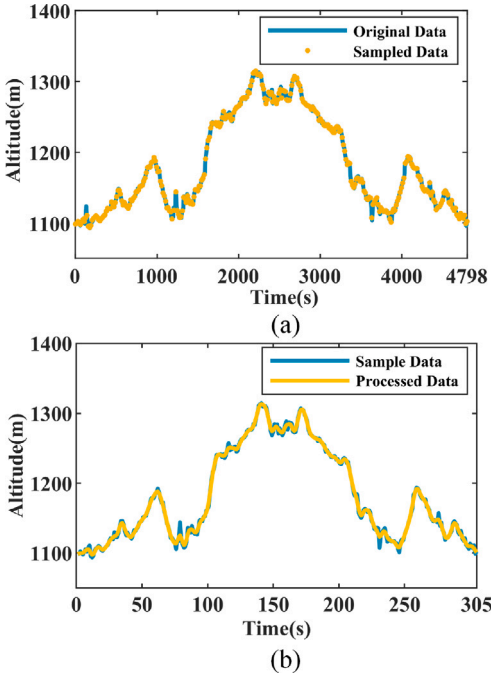

Fig. 9. Altitude data processing: (a) Raw data and sampled points; (b) Smoothed altitude profile.

4.4.2. Validation of adaptability

To validate the adaptability of the H-TD3 framework without the ambiguity of physical hardware testing, this study employed a data-driven simulation approach. Due to the extreme safety constraints and explosion risks associated with operating pressurized hydrogen systems under emulated high-altitude low-pressure environments, full-scale physical bench testing was not conducted. Instead, the actual scope of our validation comprises high-fidelity numerical evaluations driven by the collected real-world plateau RTK-GPS data.

The sampled data were then smoothed to yield the results shown in Fig. 9(b). Based on the processed real-world plateau altitude data, comparative simulations were designed to evaluate the effectiveness of the proposed H-TD3 control strategy. Fig. 10 and Table 6 present the validation results for both controlled and uncontrolled conditions using actual altitude data.

Analysis reveals that when accounting for altitude variations, H-TD3 control achieved a maximum net power of 44.45 kW, representing a 2.73% improvement over the HB condition and reaching 97.76% of the ideal SLB condition. As shown in Fig. 10, the net power curve under H-TD3 control consistently outperforms the HB condition, demonstrating significant performance enhancement. Compared with the conventional HB controller, H-TD3 achieves a substantially wider and safer OER operating band (1.78–2.83 vs. 1.68–2.15), indicating that the agent learns to supply adequate oxygen across diverse altitude–load combinations rather than relying on a fixed feedforward map. In particular, H-TD3 avoids the excessive parasitic losses incurred by the HB controller at high altitudes, where the conventional strategy over-drives the compressor in an attempt to compensate for thin air. This intelligent reallocation of compressor effort yields both higher net power

Table 6
FCS performance under real plateau altitude conditions.

Method	m_{H_2} (g)	Value	OER	Value	NetPower (kW)	Value
SLB	348.386	0	2.12–3.06	0/0	45.47	0
HB	342.814	+1.60%	1.68–2.15	−20.75%/−29.74%	43.21	−4.97%
H-TD3	343.911	+1.28%	1.78–2.83	−16.04%/−7.52%	44.45	−2.24%

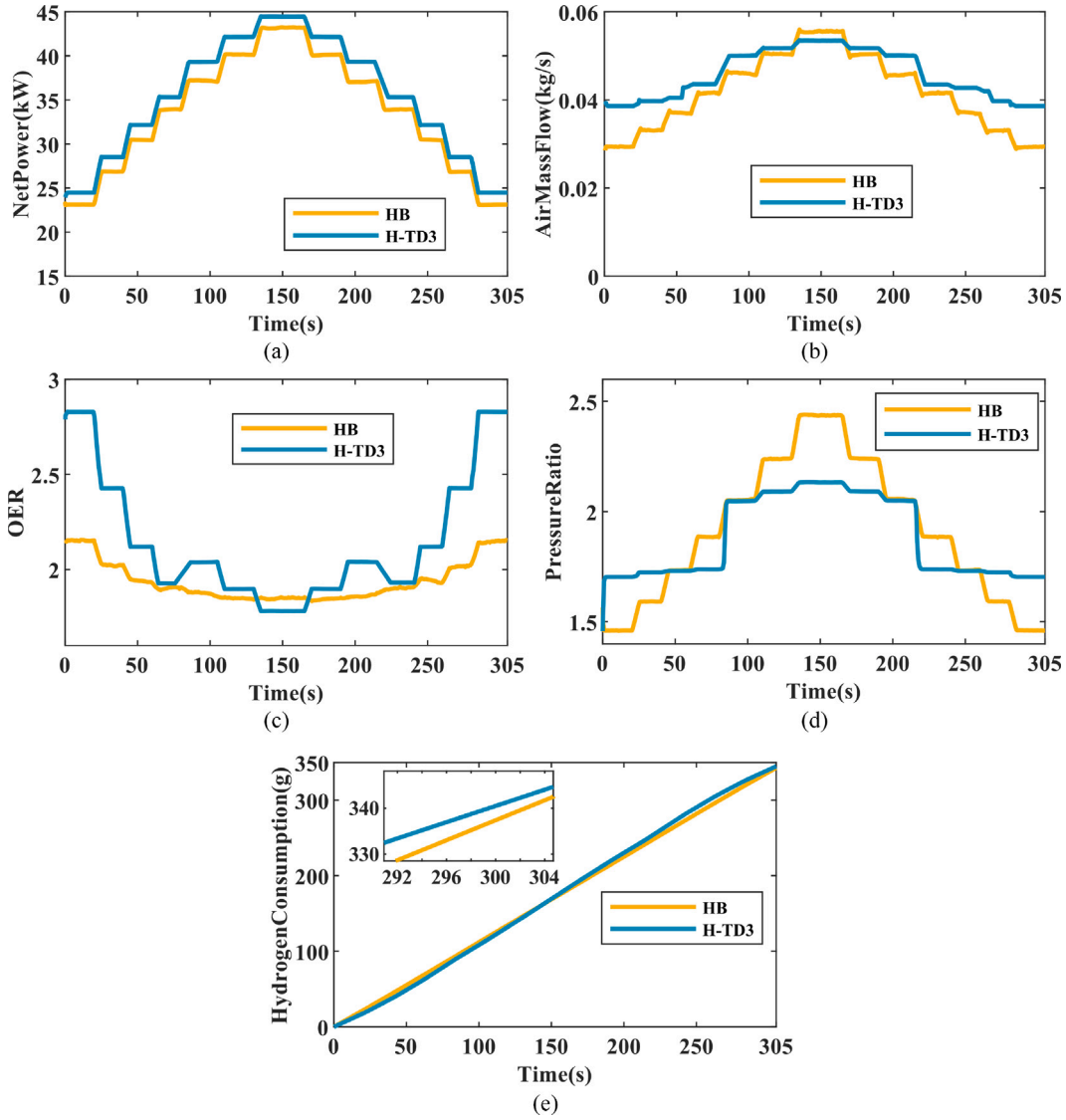

Fig. 10. Data-driven simulation validation and comparison under real plateau mountain altitude conditions.

and improved OER margins across the altitude profile. Although total hydrogen consumption under H-TD3 control increases slightly by 1.1 g compared to the HB condition, the higher net power output translates to lower hydrogen consumption per unit power and greater energy utilization efficiency, confirming the excellent adaptability and robustness of the H-TD3 optimization strategy.

These validation results demonstrate that when deployed in real-world conditions, the H-TD3 control strategy effectively mitigates altitude-induced performance degradation, enabling the FCS to maintain near sea-level performance at high elevations. This research provides a practical approach and valuable technical guidance for addressing fuel cell vehicle performance challenges in high-altitude environments.

5. Conclusions

This study proposed a high-altitude adaptive control strategy based on deep reinforcement learning (H-TD3) to address the severe performance degradation of fuel cell systems in plateau environments. By incorporating real-time altitude as a state variable, the H-TD3 framework transforms the control paradigm from static feedforward to dynamic environmental adaptation, effectively resolving the non-stationarity inherent in high-altitude driving. Under dynamic 3000-meter altitude driving cycles, the H-TD3 strategy raised the peak net power by 4.84 kW, an improvement of 12.7% over conventional feedforward controllers, while constraining the OER within a safe range of 1.74–3.19 to prevent oxygen starvation. Data-driven simulations using

real plateau altitude profiles further confirmed that H-TD3 restores 97.76% of the ideal sea-level net power capability, demonstrating robust adaptability for fuel cell vehicles in mountainous regions. The proposed altitude-aware architecture can be generalized to other fuel cell platforms by retraining on platform-specific data.

The current framework assumes an idealized thermal management system and relies on numerical simulation rather than physical hardware testing; moreover, only the air supply subsystem is addressed. Future work will focus on advancing thermal–air coupled control, conducting hardware-in-the-loop validations, and expanding the framework to integrated hydrogen–air collaborative control and vehicle–road–cloud energy management.

CRediT authorship contribution statement

Bieshu Jin: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Tianfeng Tang:** Writing – original draft, Visualization, Software, Formal analysis, Data curation. **Zhongkun Xiao:** Writing – original draft, Validation, Software, Data curation. **Chaoyi Chen:** Writing – review & editing, Visualization, Methodology. **Bo Wang:** Writing – review & editing, Visualization, Investigation. **Qing Shi:** Writing – review & editing, Visualization, Investigation, Formal analysis. **Jin Zhao:** Writing – review & editing, Visualization, Supervision. **Guangwei Wang:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 52502501), and the Guizhou Provincial Science and Technology Projects (Grant No. JSZX(2025)003 and 2024100).

Data availability

Data will be made available on request.

References

[1] Mitra U, Arya A, Gupta S. A comprehensive and comparative review on parameter estimation methods for modelling proton exchange membrane fuel cell. Fuel 2023;335:127080.
[2] Akgun I, Dincer I. Development of a smart powering system with ammonia fuel cells and internal combustion engine for submarines. Energy 2024;294:130747.
[3] Chi X, Chen F, Mo T, Li Y, Wei W. Improve methanol efficiency for direct methanol fuel cell system via investigation and control of optimal operating methanol concentration. Energy 2024;290:130147.
[4] Tang T, Peng Q, Shi Q, Peng Q, Zhao J, Chen C, Wang G. Energy management of fuel cell hybrid electric bus in mountainous regions: A deep reinforcement learning approach considering terrain characteristics. Energy 2024;311:133313.
[5] Zhao D, Hua Z, Dou M, Huangfu Y. Control oriented modeling and analysis of centrifugal compressor working characteristic at variable altitude. Aerosp Sci Technol 2018;72:174–82.
[6] Ma T, Xu Z, Liu W, Qi J, Gu Z, Li R, Lin W, Li C. Energy recovery and efficiency improvement method of proton exchange membrane fuel cell system by applying the expander in plateau environment. J Power Sources 2024;613:234903.
[7] Sodré JR, Soares S. Comparison of engine power correction factors for varying atmospheric conditions. J Braz Soc Mech Sci Eng 2003;25:279–84.
[8] Chen J, He H, Zhang Z, Wu J, Wang Y-X. Optimization and matching of the air loop system in a fuel cell for high-altitude application. Int J Hydrog Energy 2025;141:523–35.
[9] Ou K, Wang Y-X, Li Z-Z, Shen Y-D, Xuan D-J. Feedforward fuzzy-PID control for air flow regulation of PEM fuel cell system. Int J Hydrog Energy 2015;40:11686–95.
[10] Yuan H, Dai H, Wei X, Ming P. A novel model-based internal state observer of a fuel cell system for electric vehicles using improved Kalman filter approach. Appl Energy 2020;268:115009.
[11] Li Y, Pei P, Ma Z, Ren P, Huang H. Analysis of air compression, progress of compressor and control for optimal energy efficiency in proton exchange membrane fuel cell. Renew Sustain Energy Rev 2020;133:110304.
[12] Xu Y, Zhang F, Pei Y, Sun Y, Chen G. Research on the adaptability and control strategies of fuel cell systems in plateau environments. Int J Hydrog Energy 2025;171:151205.
[13] Chen J, He H, Quan S, Wei Z, Zhang Z, Wang Y-X. Real-time power optimization based on PSO feedforward and perturbation & observation of fuel cell system for high altitude. Fuel 2024;356:129551.
[14] Li S, Qiu Y, Yin L, Li R, Gan R, Li Q, Huangfu Y. Net power optimization based on extremum search and model-free adaptive control of PEMFC power generation system for high altitude. IEEE Trans Transp Electrification 2022;9(4):5151–64.
[15] Ma Y, Zhang F, Gao J, Chen H, Shen T. Oxygen excess ratio control of PEM fuel cells using observer-based nonlinear triple-step controller. Int J Hydrog Energy 2020;45(54):29705–17.
[16] Chen J, Liu Z, Wang F, Ouyang Q, Su H. Optimal oxygen excess ratio control for PEM fuel cells. IEEE Trans Control Syst Technol 2017;26(5):1711–21.
[17] Li Q, Yin L, Yang H, Wang T, Qiu Y, Chen W. Multiobjective optimization and data-driven constraint adaptive predictive control for efficient and stable operation of PEMFC system. IEEE Trans Ind Electron 2020;68(12):12418–29.
[18] Yin L, Li Q, Breaz E, Chen W, Gao F. Model guided extremum seeking and active disturbance rejection control for efficiency real-time optimization of PEMFC system. IEEE Trans Ind Electron 2023;71(6):5905–19.
[19] Xie S, Li Q, Yin L, Huo S, Wang T, Chen W. Multivariable cooperative control for performance guarantee of PEMFC system in high-altitude environment. IEEE Trans Ind Electron 2024;71(12):15846–57.
[20] Wang Y, Chen W, Li Q, Han Y, Guo A, Wang T. Coordinated optimal power distribution strategy based on maximum efficiency range of multi-stack fuel cell system for high altitude. Int J Hydrog Energy 2024;50:374–87.
[21] Yang D, Pan R, Wang Y, Chen Z. Modeling and control of PEMFC air supply system based on TS fuzzy theory and predictive control. Energy 2019;188:116078.
[22] Li J, Yu T. Intelligent controller based on distributed deep reinforcement learning for PEMFC air supply system. IEEE Access 2021;9:7496–507.
[23] Wang Y, Wang Y. Adaptive power optimization and control for automobile PEMFC air management system based on fuzzy Q-learning and prescribed tracking performance. IEEE Trans Intell Transp Syst 2024.
[24] Deng K, Liu Y, Hai D, Peng H, Löwenstein L, Pischinger S, Hameyer K. Deep reinforcement learning based energy management strategy of fuel cell hybrid railway vehicles considering fuel cell aging. Energy Convers Manage 2022;251:115030.
[25] Tao F, Fu Z, Gong H, Ji B, Zhou Y. Twin delayed deep deterministic policy gradient based energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles considering predicted terrain information. Energy 2023;283:129173.
[26] Jiawen L, Yaping L, Tao Y. Temperature control of proton exchange membrane fuel cell based on machine learning. Front Energy Res 2021;9.
[27] Lee S-H, Aldredge RC. Analytic approach to determine optimal conditions for maximizing altitude of sounding rocket: Flight in standard atmosphere. Aerosp Sci Technol 2015;46:374–85.
[28] Yan M, Xu H, Jin L, He H, Li M, Liu H. Co-optimization for fuel cell buses integrated with power system and air conditioning via multi-dimensional prediction of driving conditions. Energy Convers Manage 2022;271:116339.
[29] Chen H, Liu Z, Ye X, Yi L, Xu S, Zhang T. Air flow and pressure optimization for air supply in proton exchange membrane fuel cell system. Energy 2022;238:121949.
[30] Mann RF, Amphlett JC, Hooper MA, Jensen HM, Peppley BA, Roberge PR. Development and application of a generalised steady-state electrochemical model for a PEM fuel cell. J Power Sources 2000;86(1–2):173–80.
[31] Li J, Zhou T. Active fault-tolerant coordination energy management for a proton exchange membrane fuel cell using curriculum-based multiagent deep meta-reinforcement learning. Renew Sustain Energy Rev 2023;185:113581.
[32] Pukrushpan JT. Modeling and control of fuel cell systems and fuel processors. University of Michigan; 2003.
[33] Kunusch C, Puleston PF, Mayosky MA, Husar AP. Control-oriented modeling and experimental validation of a pemfc generation system. IEEE Trans Energy Convers 2011;26(3):851–61.
[34] Kang S, Min K, Yu S. Two dimensional dynamic modeling of a shell-and-tube water-to-gas membrane humidifier for proton exchange membrane fuel cell. Int J Hydrog Energy 2010;35(4):1727–41.
[35] Chen J, He H, Yue H. A review of plateau environmental adaptation for proton exchange membrane fuel cells. Int J Hydrog Energy 2024;50:744–64.
[36] Silver D, Lever G, Heess N, Degris T, Wierstra D, Riedmiller M. Deterministic policy gradient algorithms. In: International conference on machine learning. Pmlr; 2014, p. 387–95.
[37] Tan H. Reinforcement learning with deep deterministic policy gradient. In: 2021 international conference on artificial intelligence, big data and algorithms. IEEE; 2021, p. 82–5.

[38] Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, Silver D, Wierstra D. Continuous control with deep reinforcement learning. 2015, arXiv preprint arXiv:1509.02971.

[39] Fujimoto S, Hoof H, Meger D. Addressing function approximation error in actor-critic methods. In: International conference on machine learning. PMLR; 2018, p. 1587–96.

[40] Zhou Y, Huang Y, Mao X, Kang Z, Huang X, Xuan D. Research on energy management strategy of fuel cell hybrid power via an improved TD3 deep reinforcement learning. Energy 2024;293:130564.

[41] Khan SS, Shareef H, Wahyudie A, Khalid S, Sirjani R. Influences of ambient conditions on the performance of proton exchange membrane fuel cell using various models. Energy Environ 2019;30(6):1087–110.