



Research on Coordinated Control Strategy of Fuel Cell Anode Pressure and Flow Based on Deep Reinforcement Learning

Wei Lu

School of Automotive and Transportation Engineering, Xihua University, Chengdu, China
Email: 18010541241@163.com

How to cite this paper: Lu, W. (2025) Research on Coordinated Control Strategy of Fuel Cell Anode Pressure and Flow Based on Deep Reinforcement Learning. *Open Access Library Journal*, 12: e13537. <https://doi.org/10.4236/oalib.1113537>

Received: April 28, 2025

Accepted: June 16, 2025

Published: June 19, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In view of the control problem of hydrogen pressure and flow coupling in the anode subsystem of the Proton Exchange Membrane Fuel Cell (PEMFC) under dynamic load conditions, this paper proposes an intelligent control framework based on the deep deterministic policy gradient (DDPG) algorithm. Firstly, a dynamic model of the fuel cell anode system is established to transform the multivariable coupling control problem into a reinforcement learning state space; secondly, a multi-objective reward function integrating pressure tracking accuracy, hydrogen utilization efficiency and actuator smoothness is designed, and a dual-channel Actor-Critic network structure with time series feature extraction capability is constructed to achieve transient complex condition control. Simulation experiments show that compared with the traditional model predictive control (MPC), the DDPG controller reduces the anode pressure fluctuation amplitude by 40.6% and the hydrogen excess coefficient overshoot by 20.2% under step load scenarios.

Subject Areas

Engineering

Keywords

Proton Exchange Membrane Fuel Cell, Deep Learning, Predictive Control, Anode System

1. Introduction

质子交换膜燃料电池(PEMFC)作为清洁能源转换装置,其动态响应性能与寿命可靠性高度依赖于阳极子系统的气体供给控制[1]。在车辆变载工况下,

阳极需实时平衡氢气压力稳定性与流量供给效率[2]。然而，阳极系统的强非线性、强耦合性、执行机构响应延迟以及负载扰动的随机性，使得传统基于模型的控制设计面临严峻挑战。

当前阳极控制普遍采用 PID 与模型预测控制实现阳极压力流量调节。如 Pukrushpan [3] 等人建立的线性化模型，可以实现稳态控制，但难以处理电堆电流剧烈变化时的非高斯噪声干扰；郭爱[4]等人提出的多变量模型预测方案虽改善了动态性能，但存在计算复杂难以满足实时性需求的问题。近年来，强化学习在能源控制场景中展现出优势，但其在 PEMFC 气体供给系统的应用仍存在以下瓶颈：

状态空间维度灾难：阳极动态涉及压力、流量、湿度等多物理场耦合，传统 Q-learning 难以处理高维连续状态；奖励稀疏性问题：系统安全约束(如压力超限)仅在故障时触发惩罚信号，导致策略收敛速度慢[5]。

针对上述挑战，本文提出基于深度确定性策略梯度(DDPG)的阳极智能控制框架，主要贡献包括：

建立燃料电池系统模型，构建包含时序关联特征的状态空间，解决传统 RL 状态表征不完整问题；设计压力流量短期指标与膜电极损伤长期指标相结合的分层奖励函数。

2. 系统建模与问题描述

2.1. 系统模型建立

PEMFC 氢气供气系统主要包括氢气减压阀、氢气循环泵和引射器等部分，如图 1 所示。氢气供给系统主要功能是为电堆阳极提供合适的氢气流量和压力。

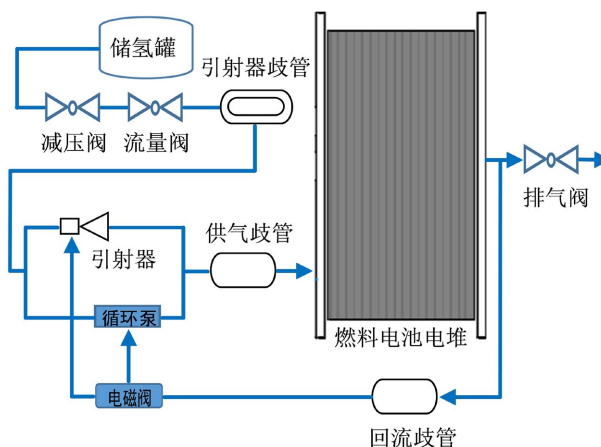


图 1. 燃料电池供氢系统结构

本文参考文献[6]和[7]建立氢气循环结构的氢气供给系统，主要动态模型描述如下：

1) 氢气循环泵模型

$$\begin{cases} \frac{d\omega_{HCP}}{dt} = \frac{\tau_1 - \tau_2}{J_{HCP}} \\ \tau_1 = \eta_{HCP} \frac{k_t}{R_{cm}} (u_{HCP} - k_v \omega_{HCP}) \\ \tau_2 = \frac{C_p}{\omega_{cp}} \frac{T_{rm}}{\eta_{cp}} \left[\left(\frac{P_{sm}}{P_{rm}} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right] W_{HCP} \end{cases} \quad (1)$$

其中, J_{HCP} 是转动惯量, u_{HCP} 是控制电压, W_{HCP} 是泵输出质量流量, ω_{HCP} 是泵转速, k_t 、 k_v 、 R_{cm} 为电机参数。

2) 引射器模型

混合室喉部马赫数 Ma_t 与面积比关系:

$$\frac{A_t}{A_p^*} = \frac{1}{Ma_t} \left(\frac{2}{\gamma+1} \left(1 + \frac{\gamma-1}{2} Ma_t^2 \right) \right)^{\frac{\gamma+1}{2(\gamma-1)}} \quad (2)$$

其中: A_t 喉部截面积, A_p^* 一次流临界截面积, $\gamma = 1.4$ 氢气比热比。

定义引射比 $\omega = \frac{\dot{m}_s}{\dot{m}_p}$, 其简化表达式为:

$$\omega = \frac{\sqrt{\frac{P_s}{P_p}} \sqrt{\frac{T_p}{T_s}}}{\sqrt{\left(\frac{\gamma+1}{2} \right)^{\frac{\gamma+1}{\gamma-1}} - \left(\frac{P_s}{P_p} \right)^{\frac{\gamma-1}{\gamma}}}} - 1 \quad (3)$$

式中: P_p , P_s 为一次/二次流入口压力, \dot{m}_p , \dot{m}_s 为一次流二次流质量流量

3) 阳极流道模型

通过理想气体状态方程, 燃料电池阳极氢气压力计算公式如下:

$$P_{an} = \frac{m_{H_2} R_{H_2} T_{an}}{V_{an} M_{H_2}} + \min \left\{ 1, \frac{m_v V_{an} R_v T_{st}}{m_v V_{an} P_{sat}^{st}} \right\} P_{sat}^{st} + \frac{m_{N_2} R_{N_2} T_{an}}{V_{an} M_{N_2}} \quad (4)$$

式中: T_{an} 为电池阳极温度, P_{H_2} 氢气的分压力, V_{an} 阳极和阴极的激活电压降。

2.2. 模型验证

引射器模型验证使用电堆实验数据, 建立的模型验证结果如图 2 所示。

3. DDPG控制算法设计

3.1. 算法原理

深度确定性策略梯度算法(DDPG)是一种基于行动者-评论家(Actor-Critic)的深度强化学习算法, 可用于处理连续状态和动作的问题, 含有策略网络和价值网络, 策略网络根据观测的环境输出确定的动作, 价值网络对状态-动作进行评价, 指导策略网络更新[8]。DDPG 的策略网络表示为 $\mu(s; \theta)$, 价值网络表示为 $Q(s, a; \omega)$ 。策略网络目标函数的梯度如式所示:

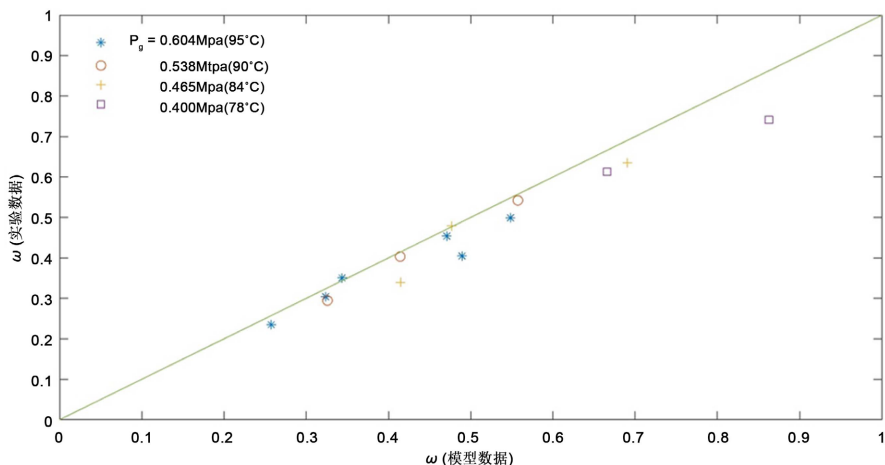


图 2. 映射器模型验证

$$\nabla_{\theta} J(\theta) = \sum_{i=1}^M \nabla_{\theta} Q(s, a; w) \Big|_{s=s_i, a=\mu(s_i)} \mu(a_i | s_i; \theta) \quad (5)$$

价值网络表示为 $Q(s, a; \omega)$ ，用它来近似真实的动作值函数。定义 DDPG 的目标函数为：

$$y_i = r_i + \gamma Q(s_{i+1}, \mu(s_{i+1}; \theta); \omega) \quad (6)$$

定义损失函数如下：

$$L_{\omega} = \frac{1}{n} \sum_{i=1}^n (y_i - Q(s_i, a_i; \omega))^2 \quad (7)$$

DDPG 通过目标函数梯度上升更新策略参数 θ ，通过最小化损失函数来更新价值网络参数 ω ，使其愈加接近动作值函数 $Q_{\pi}(s, a)$ 。

3.2. 控制目标设计

氢气作为燃料电池的燃料，其成本十分昂贵，需要保证氢气浓度在一个合理的范围，过氢比(Hydrogen Excess Ratio, HER)作为关键控制参数，定义为阳极输入的氢气流量与化学反应消耗的氢气流量之比：

$$\text{HER} = \frac{W_{\text{H}_2, \text{in}}}{W_{\text{H}_2, \text{react}}} \quad (8)$$

式中： $W_{\text{H}_2, \text{in}}$ 为进入的氢气流量， $W_{\text{H}_2, \text{react}}$ 为消耗的氢气流量。

HER 过低易引发氢饥饿，导致催化剂活性下降和电压波动；HER 过高则加剧氢气浪费，显著提升燃料成本，同时增加氢气渗透风险，加速质子交换膜老化。为了实现氢气供给效率与系统寿命的全局最优，本文设计过氢比控制目标为 1.8。

阳极压力对燃料电池有着至关重要的影响。当阳极压力过低时，参与电化学反应的氢气不足，导致发生“氢饥饿”电堆峰值功率。但过高的阳极压力会导致氢气向阴极渗透，降低氢气利用率，并且会导致“氧饥饿”，增大动态负载下的电压下冲量。因此本文参考某电堆手册进行阳极压力控制。

3.3. 控制框架设计

DDPG 智能体通过传感器实时获取燃料电池系统多物理场状态观测向量, 其中包含阳极压力动态、氢气过量比、电堆负载特性及历史控制动作序列等。Actor 网络基于当前策略输出执行器动作向量(氢气阀开度、氢气循环泵电压)。燃料电池在动作向量驱动下产生动态响应, 并计算目标奖励信号, 此时再根据奖励信号对 critic 网络更新, 最终达成控制策略的渐进式优化。控制框架如图 3 所示。

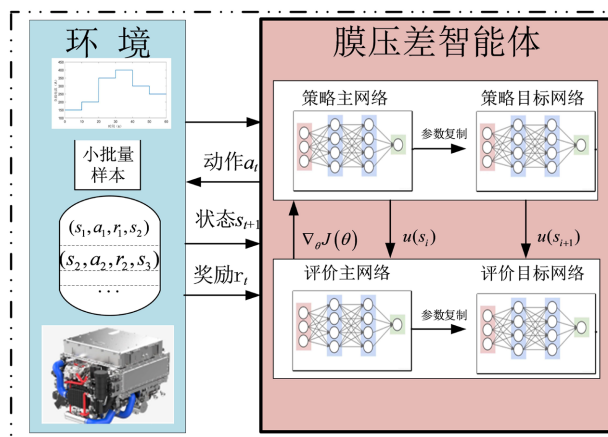


图 3. DDPG 控制框架

3.4. 智能体设计

氢气供给系统智能体设置主要包括状态变量、动作变量、奖励函数和网络结构设计。

1) 状态变量

对于氢气供给系统控制智能体的状态变量 s_{pan} 选取, 应充分体现阳极的工作状态。

$$s_{pan} = \left[e_{pan}(t) \quad e_{HER}(t) \quad I \quad \frac{de_{HER}(t)}{dt} \quad \frac{de_{pan}(t)}{dt} \quad \frac{\theta_{an}(t-1)}{\theta_{an}^{max}} \quad \frac{V_{HCP}(t-1)}{V_{HCP}^{max}} \right] \quad (9)$$

式中: $e_{HER}(t)$ 、 $e_{pan}(t)$ 、 I 分别为过氢比偏差与阴极压力偏差和电流。

2) 动作变量

PEMFC 氢气供给系统过氢比控制的执行器是流量阀开度, 用来控制氢气压力设置背压阀开度变化量作为动作变量:

$$a_{pan} = [\Delta\theta_{rm} \quad \Delta V_{HCP}] \quad (10)$$

式中: $\Delta\theta_{rm}(t)$ 满足 $-\Delta\theta_{rm}^{max} < \Delta\theta_{rm}(t) < \Delta\theta_{rm}^{max}$, $\Delta V_{HCP}(t)$ 满足 $-\Delta V_{HCP}^{max} < \Delta V_{HCP}(t) < \Delta V_{HCP}^{max}$ 。

t 时刻的氢气循环泵的电磁阀开度:

$$\theta_{rm}(t) = \theta_{rm}(t-1) + \Delta\theta_{rm}(t) \quad (11)$$

t 时刻的氢气循环泵的电压:

$$V_{HCP}(t) = V_{HCP}(t-1) + \Delta V_{HCP} \quad (12)$$

3) 奖励函数

阳极系统控制智能体主要的控制目标是保证氢气压力与流量满足系统需求值，设计分层奖励函，短期指标与膜电极损伤风险长期约束相结合的奖励函数：

$$r_{Pan} = \begin{cases} -[\omega_1 e_{pan}^2(t) + \omega_2 e_{HER}^2(t)] & e_{pan} < 2000 \text{ pa} \\ -10 & e_{pan} > 2000 \text{ pa} \end{cases} \quad (13)$$

式中： ω_1 为阳极压力偏差值系数， ω_2 过氢比偏差系数，由于阴极压力与过氢比之间存在耦合作用，所以通常 ω_1 与 ω_2 设置为 0.5，但通过多次设置定值实验发现 ω_1 取 0.6， ω_2 取 0.4 训练效果更佳。

4) 网络结构设计

网络结构设计如图 4 所示，策略网络包括 1 层输入层、4 层隐藏层和 1 层输出层。它的输入层设有 7 个神经元，与状态变量维度一致，输出层对应动作变量维度为 2 个神经元，4 层隐藏层中，每层均包含 128 个神经元。评价网络用于近似替代动作值函数，以此评估当前状态动作对的优劣。策略网络与评价网络采用全连接方式。

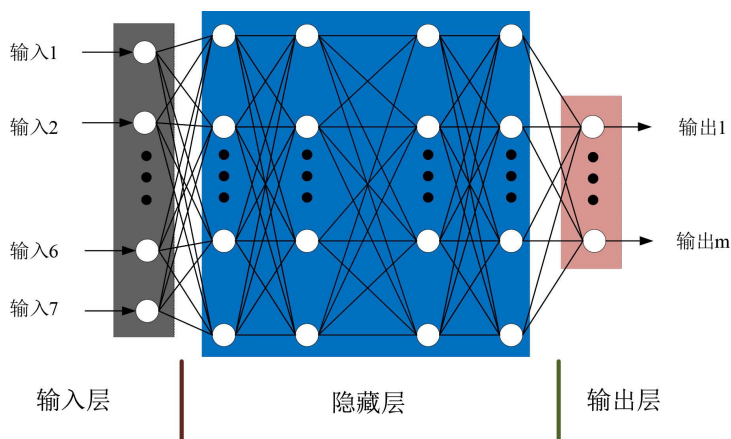


图 4. 网络结构设计

4. 实验与结果分析

基于前文搭建的燃料电池模型与控制策略，使用 100 - 300 A 随机阶跃工况训练智能体，收敛后在 Matlab/Simulink 中进行离线仿真验证，其结果如图 5 所示：

仿真实验表明，本文提出的 DDPG 控制框架在动态工况下展现出显著优势。在 100 到 200 A 阶跃负载场景中，阳极压力最大超调量仅为 7.8% (较 MPC 降低 40.6%)，同时氢气阀的动作更加缓慢有助于提升部件寿命；经济性提升：过氧比维持在 1.8 左右，在工况变化时过氢比最大超调量仅为 52.8% (较 MPC 降低 20.2%)，同时氢气循环泵的电电压变化更平缓，有助于提升氢气循环泵寿命。

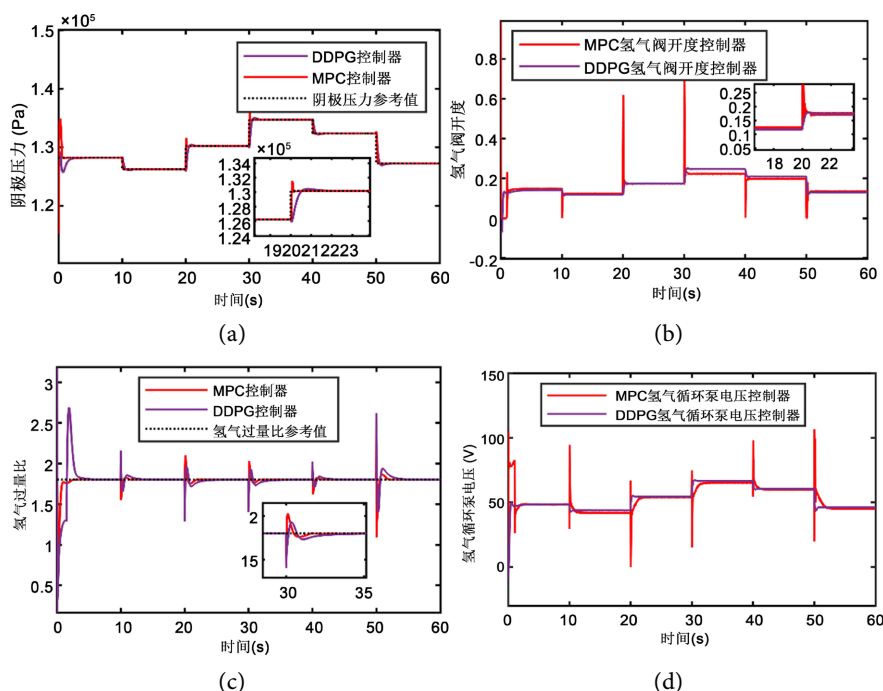


图 5. 控制结果对比

5. 结论

本文阐述了当前燃料电池阳极系统控制存在的问题，并设计基于 DDPG 的控制算法，最后通过仿真验证，证明所提出的控算法相较于 MPC 控制在动态响应和超调量控制方面展现更优的性能。

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] 张思龙, 梁满志, 孙珩凯, 等. 燃料电池氢气供应系统多功能测试平台设计[J]. 汽车工程, 2024, 46(7): 1147-1156.
- [2] 张家明, 马天才, 丛铭, 等. 大功率燃料电池氢气系统建模与控制[J]. 汽车技术, 2021(2): 23-27.
- [3] Pukrushpan, J.T., Stefanopoulou, A.G. and Peng, H. (2004) Control of Fuel Cell Breathing. *IEEE Control Systems Magazine*, **24**, 30-46. <https://doi.org/10.1109/MCS.2004.1275430>
- [4] 郭爱, 陈维荣, 李奇, 等. 车用燃料电池氢气供应系统的预测控制[J]. 太阳能学报, 2013, 34(8): 1484-1491.
- [5] Garcia, M. and Dong, W. (2024) Reinforcement Learning-Based Optimal Control of Uncertain Nonlinear Systems. *International Journal of Control*, **97**, 2839-2850. <https://doi.org/10.1080/00207179.2024.2305727>
- [6] He, J., Ahn, J. and Choe, S. (2011) Analysis and Control of a Fuel Delivery System Considering a Two-Phase Anode Model of the Polymer Electrolyte Membrane Fuel Cell Stack. *Journal of Power Sources*, **196**, 4655-4670. <https://doi.org/10.1016/j.jpowsour.2011.01.019>

- [7] Tri, D.T.L., Vu, H.N., Nguyen, H.L., *et al.* (2023) A Comparative Study of Single and Dual Ejector Concepts for Anodic Recirculation System in High-Performance Vehicular Proton Exchange Membrane Fuel Cells. *International Journal of Hydrogen Energy*, **48**, 27344-27360. <https://doi.org/10.1016/j.ijhydene.2023.03.234>
- [8] Gao, Q., Liu, Y., Zhao, J., *et al.* (2021) Hybrid Deep Learning for Dynamic Total Transfer Capability Control. *IEEE Transactions on Power Systems*, **36**, 2733-2736. <https://doi.org/10.1109/TPWRS.2021.3057523>

Appendix 1. Abstract and Keywords in Chinese

基于深度强化学习的燃料电池阳极压力与流量协同控制策略研究

摘要：针对质子交换膜燃料电池(Proton Exchange Membrane Fuel Cell, PEMFC)阳极子系统在动态负载工况下，存在氢气压力与流量耦合的控制难题，本文提出一种基于深度确定性策略梯度(DDPG)算法的智能控制框架。首先，建立燃料电池阳极系统动态模型，将多变量耦合控制问题转化为强化学习状态空间；其次，设计融合压力跟踪精度、氢气利用效率与执行器平滑性的多目标奖励函数，并构建具有时序特征提取能力的双通道 Actor-Critic 网络结构，实现瞬态复杂工况控制。仿真实验表明：相较于传统模型预测控制(MPC)，DDPG 控制器在阶跃负载场景下将阳极压力波动幅度降低 40.6%，氢气过量系数超调量降低 20.2%。

关键词：质子交换膜燃料电池，深度学习，预测控制，阳极系统