doi:10.6041/j.issn.1000-1298.2025.07.065

基于深度强化学习的氢燃料电池电动拖拉机能量管理策略

李利桥[1,2]　陈江春[1,3]　刘　伟[2]　聂　晶[1,3]　高宗余[1,3]

(1.石河子大学机械电气工程学院,石河子 832003;2.兵团石河子国家农业科技园区管委会,石河子 832011;
3.现代农业机械兵团重点实验室,石河子 832003)

摘要:针对氢燃料电池电动拖拉机(Hydrogen fuel cell electric tractor,HFCET)能量管理策略在线运行时对工况适应性差的问题,提出一种基于深度 Q 网络(Deep Q-networks,DQN)学习的混合能量管理策略。将深度强化学习方法用于氢燃料电池(Hydrogen fuel cell,HFC)电动拖拉机对提高燃料经济性和延长燃料电池使用寿命具有重要作用。首先,以燃料电池氢耗量为目标,将 Q-学习算法与 DQN 算法进行对比,并与动态规划(Dynamic programming,DP)方法进行比较。将燃料电池性能退化因子纳入目标函数,通过调整性能退化因子与氢耗量实现氢燃料电池经济性和系统性能退化之间的动态平衡。通过电动拖拉机实际运行工况验证所提策略的有效性。实际运行工况试验结果表明,在训练中纳入氢燃料电池性能退化因子时,能量管理策略(Energy management strategy,EMS)能耗下降 2.46%,达到实际运行工况 DP 方法 EMS 的 87.63%,有效抑制了氢燃料电池性能衰退。同时,与 DP 方法相比,计算效率提高 78% 以上。

关键词:电动拖拉机;氢燃料电池;能量管理策略;深度强化学习;Q-学习;动态规划

中图分类号:S219.4　　**文献标识码**:A　　**文章编号**:1000-1298(2025)07-0691-10　　**OSID**:

Deep Reinforcement Learning Energy Management Strategy for Hydrogen Fuel Cell Electric Tractors

LI Liqiao[1,2]　CHEN Jiangchun[1,3]　LIU Wei[2]　NIE Jing[1,3]　GAO Zongyu[1,3]

(1. *College of Mechanical and Electrical Engineering, Shihezi University, Shihezi* 832003, *China*
2. *Management Committee of Xinjiang Production and Construction Corps Shihezi National Agricultural Science and Technology Park, Shihezi* 832011, *China*
3. *Xinjiang Production and Construction Corps Key Laboratory of Modern Agricultural Machinery, Shihezi* 832003, *China*)

Abstract: In order to solve the problem of poor adaptability of energy management strategy for hydrogen fuel cell electric tractor (HFCET) when running online, a hybrid energy management strategy was proposed based on deep Q-networks (DQN) learning. Applying deep reinforcement learning method to hydrogen fuel cell (HFC) electric tractor played an important role in improving the fuel cell economy and prolonging the service life of fuel cell. Firstly, Q-learning method was compared with deep Q-networks learning method and dynamic programming (DP) method with the fuel cell hydrogen consumption as the target. Secondly, the fuel cell performance degradation factor was incorporated into the objective function, and the dynamic balance between the hydrogen fuel cell economy and system performance degradation was achieved by adjusting the performance degradation factor and hydrogen consumption. Finally, the effectiveness of the proposed strategy was verified by the actual operating condition of the electric tractor. The experimental results of actual operating conditions showed that when hydrogen fuel cell performance degradation factor was included in the training, the energy consumption of energy management strategy (EMS) was decreased by 2.46%, reaching 87.63% of the actual operating

收稿日期:2024-11-13　修回日期:2025-02-28

基金项目:国家自然科学基金项目(52265038)、新疆维吾尔自治区 2023 人才发展基金-天池英才创新领军人才项目(CZ002507)、兵团科技计划项目(2023ZD056)、石河子大学国际科技合作推进计划项目(GJHZ202208)、石河子大学高层次人才科研启动项目(RCZK202309、RCZK2018C33)和石河子大学青年创新人才培育计划项目(CXPY201904)

作者简介:李利桥(1988—),女,副教授,博士后,主要从事农业工程研究,E-mail:liliqiao1108@163.com

通信作者:高宗余(1974—),男,副教授,博士,主要从事机械电子工程研究,E-mail:gzy19750510@163.com

condition of dynamic programming method energy management strategy, which effectively inhibited the decline of hydrogen fuel cell performance. At the same time, compared with dynamic programming method, the computational efficiency was increased more than 78%.

Key words: electric tractor; hydrogen fuel cell; energy management strategy; deep reinforcement learning; Q-learning; dynamic programming

0 引言

农业机械装备是提高农业生产效率、实现资源有效利用、推动农业可持续发展的不可或缺工具[1-2]。鉴于全球温室气体排放量大幅增加及不可再生资源开采量急剧减少[3-4],氢燃料电池(Hydrogen fuel cell,HFC)驱动电动拖拉机为发展绿色农机技术开辟了新途径。能量管理策略(Energy management strategy,EMS)直接影响 HFC 驱动电动拖拉机(Electric tractor,ET)的动力性与经济性,是电驱动拖拉机研究的关键技术之一。因此设计合理有效的 EMS 对提高 HFC 驱动拖拉机性能至关重要。

作为 HFC 和超级电容(Supercapacitor,SC)混合动力 ET 的一项关键技术,混合动力电动拖拉机 EMS 的作用是合理分配 HFC 和 SC 的功率,从而提高 ET 的效率和燃油经济性。根据文献[5],将 EMS 分为基于规则、基于优化和基于学习 3 类。

基于规则的 EMS 具有简单、计算量小、易于实现等优点。然而,规则的设计严重依赖经验,不能保证最优控制。为了实现最优控制,越来越多的研究者开始关注基于优化的 EMS。基于优化的 EMS 通过定义一个带约束的目标函数,并最小化目标函数来优化控制目标。基于优化的 EMS 可以分为全局优化和瞬时优化[6]。动态规划(Dynamic programming,DP)作为典型全局优化算法,将优化问题分解为一系列子问题,在每个采样时间形成一个代价函数,然后通过反向递归求解子问题,最后得到最优控制序列[7]。DP 方法虽然可以实现全局优化,但难以实现实时在线控制,通常作为其他控制策略的一个比较基准。随着实时-在线控制系统的深入研究,一些基于瞬时优化的控制算法获得广泛研究,如庞特里亚金最小化方法(Pontryagin minimum principle,PMP)、等效消耗最小化策略(Equivalent consumption minimization strategy,ECMS)和模型预测控制方法(Model predictive control,MPC)等。PMP 方法通过减少计算量,将全局优化的约束问题简化为一个局部哈密尔顿最小化问题,并实现实时控制[8]。在 PMP 的基础上,通过引入等价因子来辅助能源装置的等效能耗,从而分析和推导出 ECMS,ECMS 可作为一种易于在实际应用中实现的瞬时优化方法[9]。然而,其控制性能严重依赖于等效因子的选择[10]。而 MPC 方法具有滚动优化和反馈修正的特性,可以有效解决约束非线性动态优化问题[11-12]。基于瞬时优化的 EMS 算法计算量小,易于实现。然而,瞬时优化并不等于全局优化,不能保证结果的全局最优。

从上述分析可知,现有的 EMS 通常存在计算成本高、优化性能不理想和适应性差等缺点。因此,基于学习的人工智能算法在能量管理中的应用受到了广泛关注,如机器学习、强化学习等。在混合动力驱动电动拖拉机能源管理中,机器学习广泛应用于实时驾驶模式识别,如支持向量机[13]、神经网络[14-16]、马尔可夫[17]等。然而,这些机器学习算法严重依赖于已知数据质量。强化学习作为机器学习的另一个分支,具有自主学习能力,被认为是解决电动拖拉机燃料电池-超级电容混合能量管理问题的有效方法。Q-学习作为强化学习的经典算法,REDDY 等[18]将其应用于电动汽车的 EMS 中,通过最小化电池荷电状态(State of charge,SOC)变化,提高电池寿命和混合系统效率。SUN 等[19]提出一种基于 Q-学习的层次化 EMS,有效提高了计算效率和燃油经济性,并稳定了燃料电池系统的功率变化。为了适应不同的驾驶风格,ZHOU 等[20]引入一种分层 Q-学习框架来实现实时能量管理和近似全局优化,并进一步提出一种基于在线递归 Q-学习的混合动力车辆成本最小化方法。然而,Q-学习算法在高维状态空间中会导致计算负担增加、优化效率降低[21]。为了解决这些限制,将深度学习与强化学习相结合的深度强化学习算法应用于混合动力车辆中。WU 等[22]将深度 Q-学习算法应用于功率分流混合动力车辆中,试验结果表明,与 Q-学习相比,DQN 在训练时间和收敛速度方面表现更优。DU 等[23]为系列履带混合动力车辆开发了一种新的基于 DQN 的 EMS,其中神经网络通过一种新的优化方法进行优化,所提出的控制策略实现了更快的训练速度和更低的燃料消耗。

根据上述分析,深度强化学习在实时性、自主学习和优化性能方面均优于传统的能量管理策略。然而,将深度强化学习应用于电动拖拉机燃料电池-超

级电容混合能量管理系统中的研究较少。此外，现有研究侧重于降低燃料电池的氢耗量、保持超级电容 SOC 稳定性以及提高计算效率，而忽略了能量管理对燃料电池性能退化的影响。

针对以往研究的不足，本文提出一种基于深度强化学习的氢燃料电池电动拖拉机能量管理策略。首先，设计深度强化学习能量管理框架以实现优化氢燃料电池经济性并保持超级电容 SOC 稳定性。之后，采用深度 Q 网络（Deep Q-networks，DQN）算法进行能量管理，采用优先级经验回放方法存储和提取经验；将优化结果与传统的 Q-学习算法进行比较。此外，提出一种基于 DQN 加退化因子（DQN add degradation factors，DQN + deg）的考虑 HFC 系统性能退化的能量管理方法，以实现 HFC 能耗经济性和 HFC 退化之间的平衡。

1　混合动力系统模型

1.1　系统结构

电动拖拉机混合动力系统建模是设计高效 EMS 的关键步骤。HFCET 驱动系统结构如图 1 所示。质子交换膜燃料电池（Proton exchange membrane fuel cell，PEMFC）凭借功率密度高、噪声低、启动快等优点，作为 HFC 的典型代表广泛用于电动农机的主能源供给，超级电容根据犁耕状况及运输路况变化通过充放电缓冲负载功率波动及补偿瞬时功率。PEMFC 与超级电容共同构成分布式混合动力系统，两个能量源共同构成驱动系统。功率传输系统通过 DC/DC 变换器、能量管理系统与混合动力系统母线输出电压进行匹配。行走系统采用 4 个同步电机分布排列，并通过轮边行星传动机构进行加减速[24]。能量管理系统用于提高混合动力系统能量利用率，以提高电动拖拉机工作循环作业经济性。表 1 为 HFCET 动力系统参数。

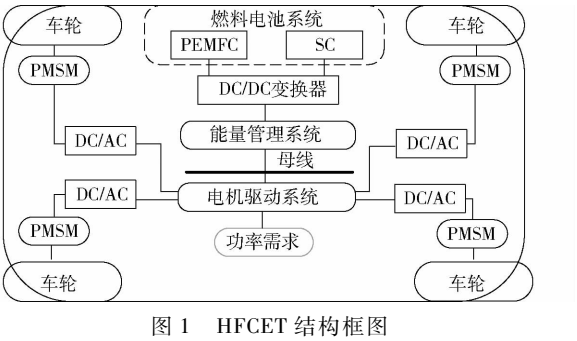

图 1　HFCET 结构框图

Fig. 1　Structure schematic of HFCET

1.2　混合动力系统建模

根据拖拉机纵向动力学方程，动力系统需求功率 P_{req} 为

表 1　氢燃料电池电动拖拉机主要参数

Tab. 1　Main parameters of hydrogen fuel cell electric tractor

参数	数值
拖拉机质量/kg	1 640
牵引功率/kW	40
最大牵引力/kN	10.6
持续工作时间/h	5
PEMFC 功率/kW	43
超级电容电荷量/F	3 200
氢容量/(Pa·L^{-1})	3.5×10^5
驱动电机额定功率/kW	10.2
驱动电机额定转速/(r·min^{-1})	1 200
驱动电机额定扭矩/(N·m)	210
驱动电机最大效率/%	92

$$\begin{cases} P_{req}=\dfrac{P_{drive}}{\eta_{mec}\eta_{DC/AC}\eta_{motor}}=\eta_{DC/DC}(P_{sc}+P_{FC}) \\ P_{drive}=v(ma+F_{aero}+F_{roll}) \end{cases} \tag{1}$$

式中　P_{drive}——电动拖拉机牵引功率

η_{mec}——机械传动系统效率

$\eta_{DC/AC}$——DC/AC 逆变器效率

η_{motor}——电机效率

$\eta_{DC/DC}$——DC/DC 变换器效率

P_{sc}——超级电容输出功率

P_{FC}——PEMFC 系统输出功率

v——电动拖拉机速度

m——电动拖拉机质量

a——电动拖拉机加速度

F_{aero}——电动拖拉机气动阻力

F_{roll}——电动拖拉机滚动阻力

PEMFC 作为 HFCET 的主要能源，通过氢与氧的反应产生电能[25]。文献[26]中，通过准静态的氢消耗模型描述 PEMFC 的氢功率与输出功率的关系，根据燃料电池系统的输出特性建立电压、电流和功率的等效电路模型，如图 2 所示。PEMFC 输出特性表示为

$$\begin{cases} P_{FC}=(U_{0_FC}-U_{ohmic}-U_{act}-U_{con})I_{FC} \\ U_{ohmic}=I_{FC}R_{FC} \\ U_{act}=\dfrac{R_gT}{\alpha_{FC}F}\ln\dfrac{I_{FC}}{I_0} \\ U_{con}=-\dfrac{R_gT}{n_{FC}F}\ln\left(1-\dfrac{I_{FC}}{I_{limit}}\right) \end{cases} \tag{2}$$

式中　I_{FC}——PEMFC 工作电流

U_{0_FC}——PEMFC 堆开路电压

U_{ohmic}——PEMFC 欧姆极化压损

U_{act}——PEMFC 活化极化压损

U_{con}——PEMFC 浓度极化压损

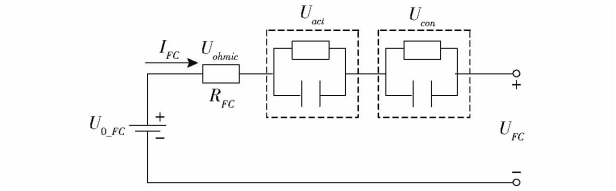

图 2 质子交换膜燃料电池等效电路模型

Fig. 2 Equivalent circuit model of proton exchange membrane fuel cell

R_{FC}——PEMFC 内部阻抗

I_0——电化学反应交换电流

I_{limit}——燃料电池极限工作电流

α_{FC}——电荷转移系数

F——法拉第常数

n_{FC}——单位摩尔浓度反应物转移质子数量

R_g——理想气体常数

T——热力学温度

PEMFC 在 t 时刻工作的氢耗量为

$$m_{H_2} = \int_0^t \frac{M_{H_2} N_{cell}}{2L} i(t)\,\mathrm{d}t \tag{3}$$

式中 M_{H_2}——氢气摩尔质量

N_{cell}——单体 PEMFC 数量

$i(t)$——t 时刻 PEMFC 电流

L——氢气分子电子数

能量管理策略通过控制 PEMFC 的输出,将影响 PEMFC 的使用寿命[11]。PEMFC 系统的降解主要来自于膜电极组件,即催化剂降解、膜降解和气体扩散层降解等[25]。膜电极组件退化的详细描述参考文献[25-27]。CHEN 等[28]提出了 PEMFC 系统退化与负载变化、启停和大功率负载之间关系的定量描述。由于负载变化对 PEMFC 系统退化有重要影响[25,29],为充分考虑其他负载条件[28]的影响,PEMFC 系统始终保持常开状态。PEMFC 系统降解公式为

$$D_{FC} = k_1 t_1 + k_2 t_2 + k_3 \frac{|\Delta P|}{P_{high} - P_{low}} \tag{4}$$

式中 D_{FC}——PEMFC 输出电压衰减度

k_1——PEMFC 系统输出功率大于额定功率 80% 时 PEMFC 电压退化率,取 0.001 48%/h

k_2——PEMFC 系统输出功率小于额定功率 10% 时 PEMFC 电压退化率,取 0.001 27%/h

k_3——瞬时负荷变化时 PEMFC 电压输出退化率,取 0.000 059 6%/h

P_{high}——PEMFC 系统输出的高功率阈值,取 45 kW

P_{low}——PEMFC 系统输出的低功率阈值,取 3 kW

ΔP——负载功率变化量,kW

当 PEMFC 电压输出性能衰减到 10% 时,PEMFC 系统输出功率迅速下降。因此,将电压从额定功率状态下降至 10% 的时间定义为 PEMFC 系统寿命结束时间。

作为辅助能源,SC 系统可辅助 PEMFC 提供峰值功率并回收再生制动能量。在能量管理系统中,SC 动态特性为

$$I_{sc} = \frac{U_{oc} - \sqrt{U_{oc}^2 - 4R_{int}P_{sc}}}{2R_{int}} \tag{5}$$

$$SOC(k+1) = SOC(k) - \frac{I_{sc}(k)\Delta t}{C} \tag{6}$$

式中 R_{int}——SC 内阻 I_{sc}——SC 电流

U_{oc}——SC 开路电压

SOC——SC 荷电状态 C——SC 容量

Δt——离散化时间间隔

P_{sc}——SC 输出功率

为保证系统可靠运行,混合动力系统须满足约束条件

$$\begin{cases} SOC_{\min} \leqslant SOC(t) \leqslant SOC_{\max} \\ P_{sc,\min} \leqslant P_{sc}(t) \leqslant P_{sc,\max} \\ P_{FC,\min} \leqslant P_{FC}(t) \leqslant P_{FC,\max} \end{cases} \tag{7}$$

式中 $SOC_{\min}$、$SOC_{\max}$——SC 最小、最大荷电状态

$P_{sc,\min}$、$P_{sc,\max}$——SC 最小、最大输出功率

$P_{FC,\min}$、$P_{FC,\max}$——PEMFC 最小、最大输出功率

2 能量管理策略

DQN 算法是一种智能算法,融合了深度学习的感知能力和强化学习的决策能力,以实现累积奖励值最大化为目标。在解决分布式电动拖拉机能量管理问题时,将系统模型视为环境,能量管理控制策略视为智能体,则电动拖拉机行驶过程能量变化就是智能体与环境的不断交互过程。

2.1 深度强化学习算法

Q-学习是一种通过智能体与环境交互获得奖励进而引导其行为的强化学习方法,可以有效解决马尔可夫决策问题。基本原理如图 3 所示,智能体观察当前状态 s_t,基于策略 π 选择并执行动作 a_t,获得奖励 r_t,之后根据环境变化获得下一时刻状态 s_{t+1},以此类推,直到最后一个状态。Q-学习的目标是找到马尔可夫决策过程的最优策略 π^*,最优策略可以通过最大化动作值函数获得,表示为

$$\pi^* = \max Q(s_t, a_t) \tag{8}$$

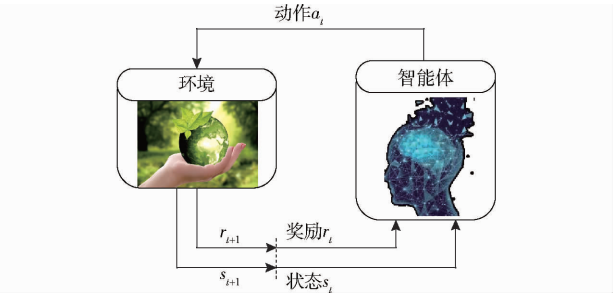

图 3　强化学习基本框架

Fig. 3　Basic framework of reinforcement learning

式中　$Q(s_t,a_t)$——动作值函数，即 Q 值，用于评估智能体在当前状态 s_t 中所采取的动作 a_t

在 Q-学习中，通过将 Q 值存储在 Q 表中来选择动作，Q 值更新公式为

$$Q(s_t,a_t)\leftarrow Q(s_t,a_t)+\eta(r+\gamma\max Q(s_{t+1},a_{t+1})-Q(s_t,a_t)) \tag{9}$$

式中　a_{t+1}——$t+1$ 时刻动作

η——学习率　　r——奖励

γ——反映未来奖励重要性的折现因子，γ 越大表示智能体越关注未来的奖励，$\gamma\in[0,1]$

然而，由于 Q-学习算法在处理高维状态空间或者连续的状态变量时，Q 表规模较大，导致计算量大、收敛性差，会产生“维数灾难”。为解决这一问题，提出 DQN 算法。DQN 算法利用 ε-贪婪策略进行最优动作选择，并以 $1-\varepsilon$ 概率随机选择动作。随着迭代次数增加，ε 初始值 ε_{ini} 随着一定的增量 ε_{inc} 逐渐增大，直到达到最大值 $\varepsilon_{\max}$。此外，DQN 算法不仅仅是深度学习和强化学习的简单结合。由于深度学习要求用于训练神经网络的样本之间具有高度独立性，强化学习要求训练样本之间具有高度相关性。因此，训练样本通过经验回放机制以 (s,a,r,s_{t+1}) 的向量模式存储在经验池中。之后利用梯度下降法随机从经验池中选取小批量经验值用于计算损失函数并更新网络参数，损失函数表示为

$$L(\theta)=E[(r+\max Q(s_{t+1},a_{t+1},\theta^-)-Q(s_t,a_t,\theta))^2] \tag{10}$$

式中　$Q(s_t,a_t,\theta)$——评价网络在当前状态下 Q 值

$Q(s_{t+1},a_{t+1},\theta^-)$——目标网络在下一状态下生成的目标 Q 值

θ、θ^-——评价网络和目标网络参数

其中 $r+\max Q(s_{t+1},a_{t+1},\theta^-)$ 为目标。

通过梯度下降法对评价网络进行更新。为了进一步提高算法稳定性，在 DQN 算法更新中使用单独的目标网络生成目标 Q 值。每隔一定步骤将评价网络参数复制到目标网络。

2.2　电动拖拉机深度强化学习能量管理策略

状态变量、动作空间和奖励函数的定义对于分布式电动拖拉机 EMS 中深度强化学习的实现至关重要。将超级电容 SOC、PEMFC 系统输出功率 P_{FC} 和电动拖拉机犁耕作业速度 v 作为 DQN 的状态向量，即

$$\boldsymbol{S}=(SOC,P_{FC},v) \tag{11}$$

选取 PEMFC 系统功率变化量 ΔP_{FC} 作为动作空间 A 的变化量，在 $[-1,1]$ 范围内将动作变量离散化为 10 个动作值，即

$$A=\{\Delta P_{FC}\}=\{-1,-0.8,-0.6,-0.4,-0.2,0.2,0.4,0.6,0.8,1\} \tag{12}$$

奖励函数直接决定深度强化学习的结果。选取氢耗量、电池 SOC 作为奖励函数 R_t，终点 SOC 维持在 SOC 初值的 0.7 附近，表达式为

$$R_t=-(\alpha FH(t)+\beta(SOC(t)-0.7)^2) \tag{13}$$

式中　α——氢耗因子

β——SOC 消耗因子

$FH(t)$——t 时刻氢耗量

基于 DQN 的能量管理策略框架如图 4 所示，状态向量 (SOC,P_{FC},v) 作为评价网络输入，根据评价网络输出 Q 值采用 ε-贪婪策略选择动作。随后，能量管理系统执行动作，生成奖励并移至下一状态。通过以上策略，向量 (s,a,r,s_{t+1}) 由当前状态、动作、奖励和下一状态组成向量集并存储在预定义的经验池中。经过训练后，从经验池中随机抽取样本作为评价网络和目标网络输入，以降低连续样本间的强相关性，提高神经网络训练效率。损失函数分别由评价网络输出 $Q(s_t,a_t,\theta)$、目标网络输出 $Q(s_{t+1},$

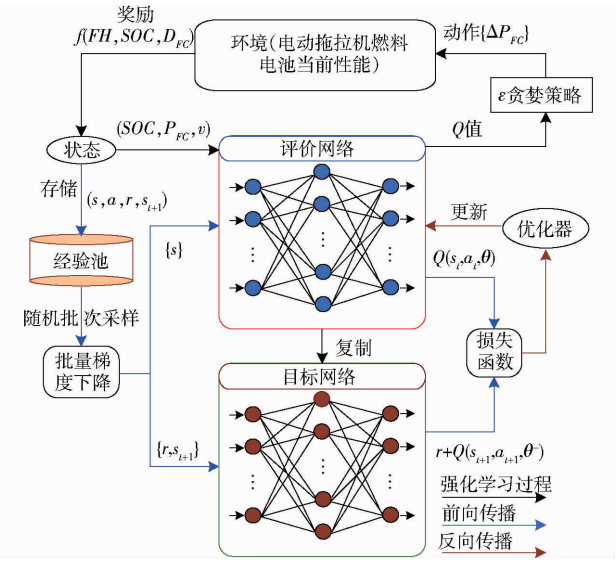

图 4　基于 DQN 的能量管理框架

Fig. 4　DQN-based energy management framework

a_{t+1},θ^-)和采样样本奖励计算获得。采用梯度下降法通过最小化损失函数来更新优化评价网络参数。每隔一定训练步骤,将评估网络参数复制到目标网络中。

深度强化学习参数对所提策略的收敛性和敏感性有重要影响,需在训练前人为调整设定,参数设置如表 2 所示。

表 2 基于 DQN 的能量管理系统参数

Tab. 2 Parameters of DQN-based energy management system

参数	数值
学习率	0.001
折现因子	0.95
经验池容量	10 000
样本批次大小	64
神经网络层数	4
ε 初始值 ε_{ini}	0.2
ε 增量 ε_{inc}	0.000 3
ε 最大值 ε_{max}	0.99

电动拖拉机 DQN 能量管理算法如图 5 所示。

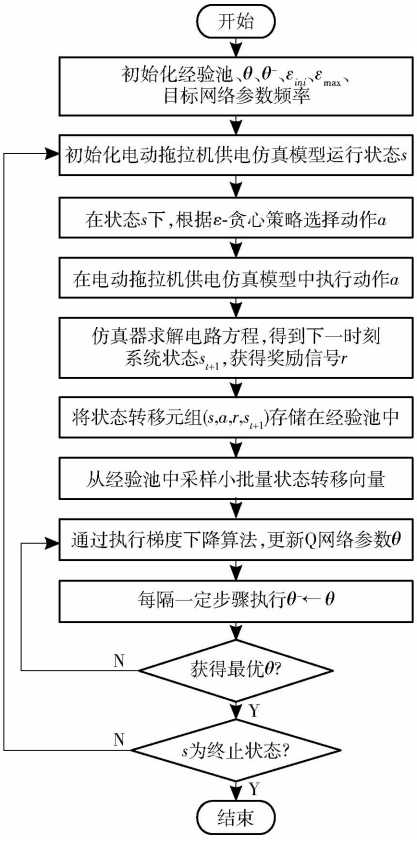

图 5 基于 DQN 的能量管理算法流程图

Fig. 5 Flowchart of DQN-based energy management algorithm

3 仿真试验

电动拖拉机的一个正常犁耕作业循环包括加速、农具入土、行驶作业、农具出土、停车怠速、转向调整 6 个阶段。由于犁耕作业从电动机输出转矩或行驶速度变化曲线上不容易精确找到各个阶段之间的分割点,本文从电动拖拉机行驶工况的构建思路出发,将拖拉机加速、农具入土、行驶作业、停车怠速、转向调整 5 个阶段定义为一个有效作业片段。在一个完整的有效作业片段中,一般包含电动拖拉机充电模式、加载模式、能量回收模式。图 6 为电动拖拉机输出速度和功率随时间变化曲线,共进行 100 组循环深耕试验(100 个有效作业片段),排除在试验过程中数据间断等意外情况产生的明显不完整循环数据外,共记录 10 000 组采样点构成有效数据,以此作为训练集,对基于 Q-学习和 DQN 的 EMS 进行训练并对比差异。

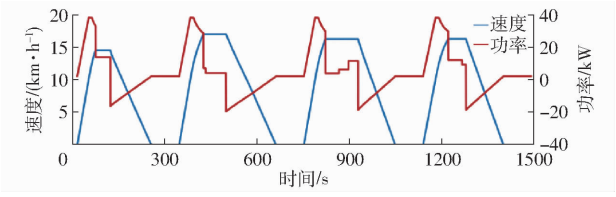

图 6 电动拖拉机输出速度与功率变化曲线

Fig. 6 Electric tractor output speed and power

图 7 为基于 Q-学习和 DQN 的能量管理策略累积奖励随训练次数的变化。在训练初始阶段,需充分探索未知状态以获得更好的训练结果,故累积奖励较低。随着不断训练,智能体逐渐倾向于选择奖励较大的动作,累积奖励增加并趋于稳定。最终,策略收敛到稳定状态,获得的累积奖励较初始条件有显著提升。通过对比可以看出,基于 DQN 的策略收敛后的稳定性以及获得的累积奖励均高于 Q-学习,说明 DQN 可以获得更好的奖励效果。

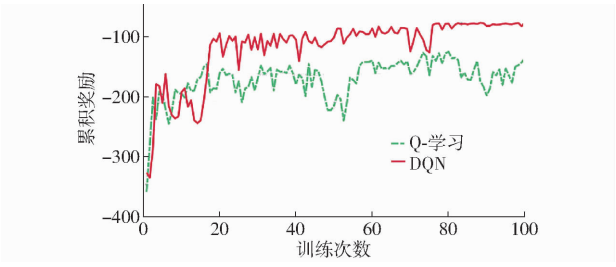

图 7 Q-学习与 DQN 能量管理策略累积奖励变化曲线

Fig. 7 Cumulative reward of Q-learning and DQN energy management strategy

超级电容 SOC 和 PEMFC 的氢耗量稳定性是评价能量管理策略的有效方法。为降低超级电容功耗对氢耗量的影响,超级电容 SOC 初值和终值始终保持在 0.7 左右。从图 8a 可看出,随着智能体的不断学习,超级电容的 SOC 从 1 逐渐减小,最终保持在 0.7 左右。从图 8a 可看出基于 Q-学习的能量管理策略系统的 SOC 终值稳定性略差。能量管理策略氢耗量随训练次数的变化曲线如图 8b 所示。在策

略收敛后，基于 DQN 的能量管理策略的氢耗量保持在 42.23 g 左右，且稳定性优于基于 Q-学习的能量管理策略。因此，本文选择性能更好的 DQN 用于能量管理策略。

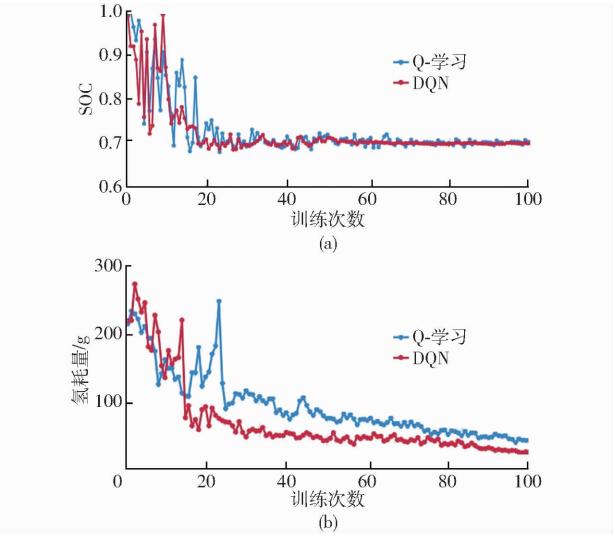

图 8　不同策略下超级电容 SOC 和氢耗量变化曲线

Fig. 8　SOC and hydrogen consumption under different strategies

4　试验验证

为了验证所提出的基于 DQN 的能量管理策略有效性，以图 9 所示循环驾驶工况为研究场景，整个工况运行时间为 1 400 s，总里程为 5.83 km。以图 9 所示驾驶工况作为周期性输入，将所得结果与 DP 进行比较。图 10a 为基于 DP 的 EMS 功率分配性能。从图 10a 可看出，电动拖拉机最初主要由超级电容提供动力，PEMFC 系统无法起到主电源作用。460 s 后，PEMFC 系统开始作为主电源提供动力。尽管基于 DP 的功率分配使 EMS 可以获得全局最优的结果，但其功率分配与实际情况并不匹配。主要原因在于 DP 采用离线数据并在循环驾驶工况下使其表现较好。图 10b 为 DQN 控制策略性能，与 DP 功率分配相比，所提策略的 PEMFC 系统功率随需求功率变化而变化。

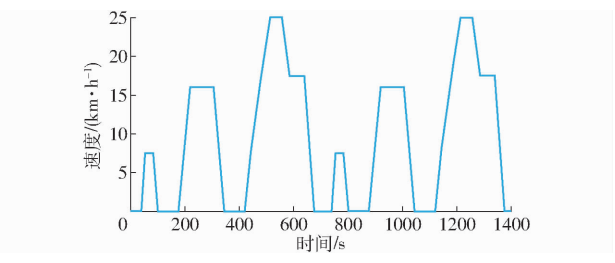

图 9　电动拖拉机运行速度变化曲线

Fig. 9　Electric tractor operating speed

图 11 为基于 DP 和 DQN 的 EMS 超级电容 SOC 变化对比。两种 SOC 曲线都在合理运行区间 0.3 ~ 0.9 内波动，然而，基于 DQN 的 EMS 具有更好的 SOC 维持能力，将波动维持在 0.65 ~ 0.70 范围内。基于 DP 的 EMS 由于在早期消耗超级电容功率，导致 SOC 快速下降，而基于 DQN 的 EMS 由于合理的功率分配使得 SOC 变化更加稳定。

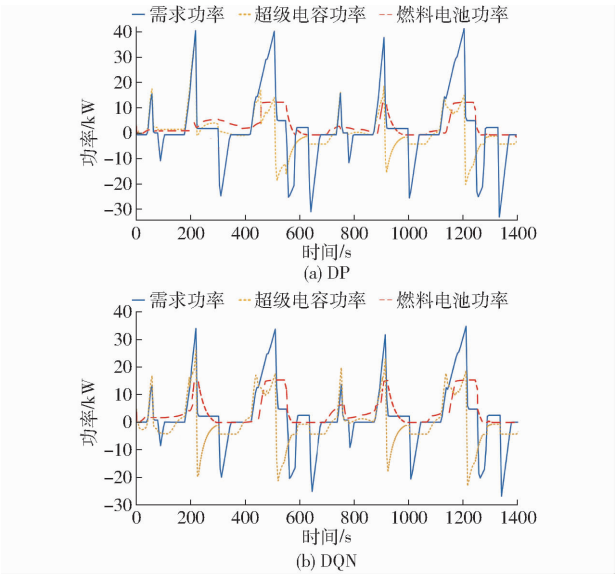

图 10　基于 DP 和 DQN 的功率分配曲线

Fig. 10　Power distribution curves based on DP and DQN

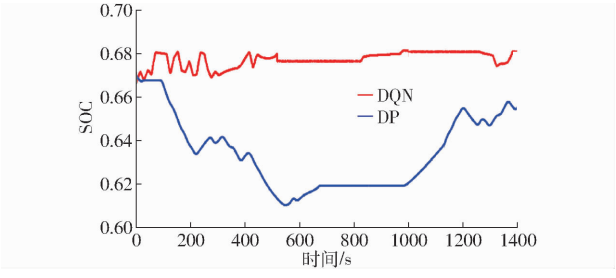

图 11　SOC 变化曲线

Fig. 11　Change curves of SOC

图 12 为两种策略下燃料电池效率分布。与 DQN 策略相比，DP 的 EMS 使燃料电池系统在高氢消耗点上的工作效率更高。所提 DQN 的 EMS 使得燃料电池在低效率区的工作点较为稀疏而集中分布在氢消耗较高效率区。燃料电池平均经济性达到 DP 的 89.84%。

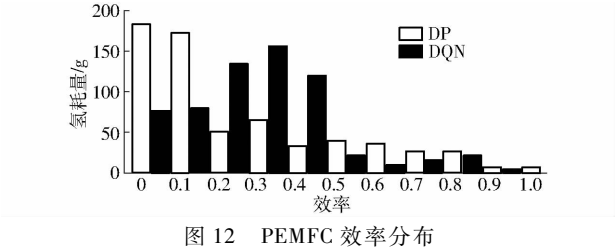

图 12　PEMFC 效率分布

Fig. 12　Efficiency distribution of hydrogen fuel cell

虽然提出的 DQN 可以实现良好的混合动力系统经济性，但尚未考虑到 PEMFC 系统的寿命问题。PEMFC 系统的耐用性指标是阻碍其在实际电动拖

拉机中应用的主要技术障碍之一。因此,在奖励函数中加入 PEMFC 系统退化指标以评价其耐用性,为此提出了包含 PEMFC 退化权重因子 λ 的 DQN 电动拖拉机使用寿命的能量管理设计方法,其奖励函数为

$$R_t = -(\alpha FH(t) + \beta(SOC(t) - 0.7)^2 + \lambda D_{FC,\deg}(t)) \tag{14}$$

式中 $D_{FC,\deg}(t)$——不同退化权重因子下 PEMFC 输出电压衰减度

不同退化权重因子的 PEMFC 系统退化情况如图 13a 所示。氢耗量固定在 0.64 g,PEMFC 退化权重因子从 0 逐渐增加。影响 PEMFC 系统性能下降的主要因素为负载变化、空转和大功率负载。随着 PEMFC 退化权重因子增加,PEMFC 系统退化率在奖励函数中被赋予越来越大的权重(此时奖励函数为负);为了获得更多奖励,智能体会避免选择低-高功率区域运行且功率变化较大的动作,使 PEMFC 系统退化程度降低。该策略在 PEMFC 退化权重因子为 1 时达到稳定,训练稳定后 PEMFC 系统退化率维持在 0.000 7% 左右。这表明 PEMFC 系统降解权重的最优估计值为 1,有效降低了 PEMFC 系统降解。如图 13b 所示,虽然退化权重因子设置与氢耗量无正相关关系,但在退化权重因子的影响下 PEMFC 系统避免在低-高功率区域运行,从而避免氢耗量增加。

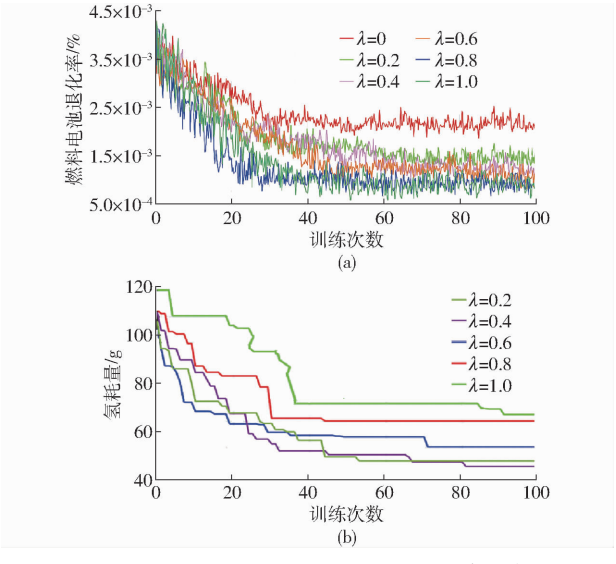

图 13 不同退化权重因子下 PEMFC 退化率和氢耗量变化曲线

Fig. 13 PEMFC degradation rate and hydrogen consumption under different degradation weighting factors

通过以上分析可知,选择 1 作为 PEMFC 系统降解退化权重因子。但在降解权重因子的作用下,PEMFC 系统会加剧老化。因此,氢消耗权重因子的选择需进一步讨论。增加氢消耗老化权重因子对 PEMFC 系统降解和氢耗量的影响如图 14 所示,其奖励函数为

$$R_t = -(\alpha FH(t) + \beta(SOC(t) - 0.7)^2 + \lambda D_{FC,\deg}(t) + \delta Ah_{eff}) \tag{15}$$

式中 δ——PEMFC 老化权重因子

Ah_{eff}——有效安时通量,通过减少 Ah_{eff} 间接实现优化 PEMFC 系统容量损失[30]

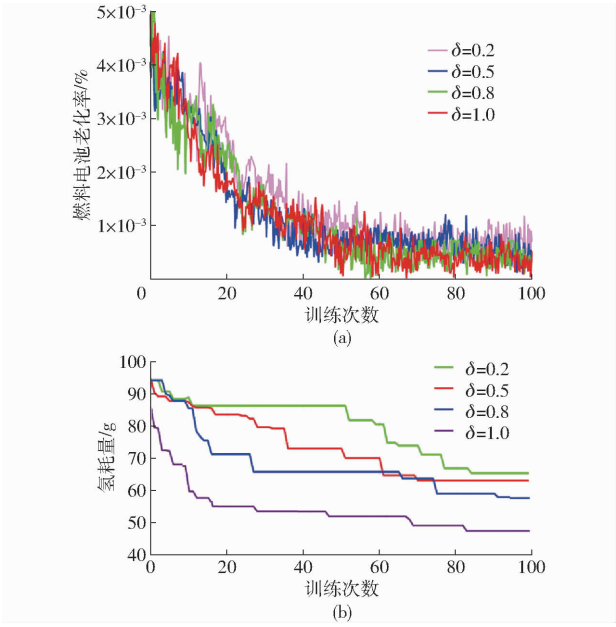

图 14 不同老化权重因子下 PEMFC 老化率与氢耗量变化曲线

Fig. 14 PEMFC aging rate and hydrogen consumption under different aging weighting factors

从图 14a 可看出,随着训练次数增加,PEMFC 老化率总体呈降低趋势。在老化权重因子为 1 且策略达到稳定时 PEMFC 系统老化得到改善。当老化权重因子为 1 时 PEMFC 系统的性能退化维持在 0.000 9% 左右。从图 14b 可看出,随着老化权重因子增加,氢耗量逐渐减少,在老化权重因子为 1 时,氢耗量与未施加老化权重因子时的氢耗量非常接近。氢耗量在策略稳定时维持在 43.28 g 左右,与未考虑能量源寿命问题的能量管理氢耗量(42.23 g)基本相同。考虑到能量源寿命及经济性问题,将 PEMFC 老化权重因子设置为 1。

在确定 PEMFC 退化权重因子和老化权重因子后,所提 DQN + deg 的能量管理策略累积奖励收敛性如图 15 所示。该策略获得的累积奖励通过不断学习逐渐增加,最终在第 120 次训练达到收敛。与基于 DQN 的 EMS 相比,该策略收敛速度有所降低,主要原因是优化目标的增加。

图 16 对比了基于 DQN + deg 的 EMS 和本文提出的 EMS 的功率分布。与未退化的 PEMFC 系统功率曲线相比,退化 PEMFC 系统在低功率区域的工

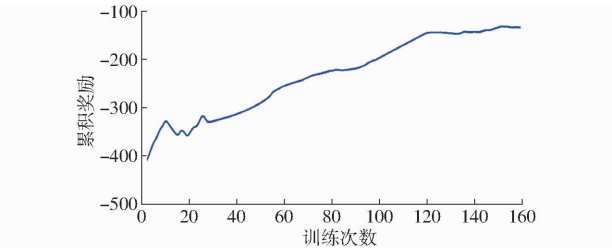

图 15　DQN + deg 能量管理策略累积奖励变化曲线

Fig. 15　Cumulative reward DQN + deg energy management strategy

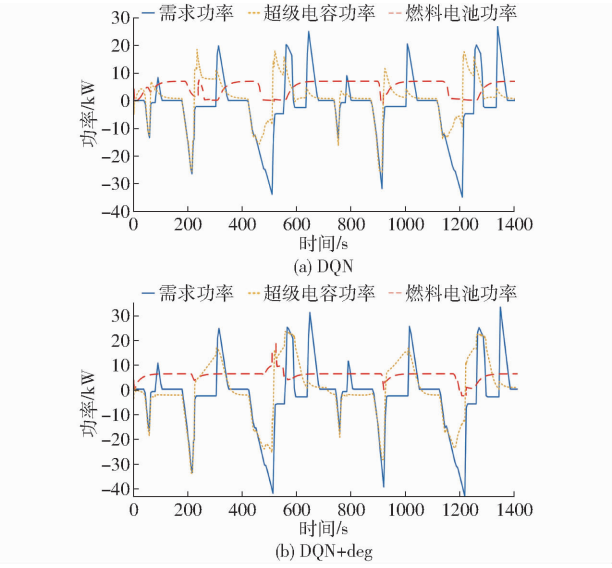

图 16　基于 DQN 和 DQN + deg 的功率分配曲线

Fig. 16　Power distribution curves based on DQN and DQN + deg

作点明显更少，功率变化明显更低。

图 17 对比了 2 种能量管理策略的超级电容 SOC 曲线。在 PEMFC 退化权重因子的影响下，在 200 ~ 400 s 区间需求功率较高，而 PEMFC 系统功率变化较低，导致超级电容功耗增加。因此，与基于 DQN 的 EMS 相比，DQN + deg 策略的超级电容 SOC 变化相对较大。

在图 9 所示循环驾驶工况下不同策略的氢耗量和计算时间对比如表 3 所示。在考虑 PEMFC 系统降解时，EMS 氢耗量降低 2.48%，达到基于 DP 的

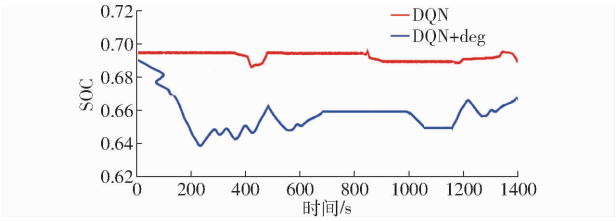

图 17　SOC 曲线对比

Fig. 17　Comparison of SOC

表 3　不同策略结果对比

Tab. 3　Comparison of results of different strategies

算法	氢耗量/g	燃油经济性/%	计算时间/s	计算时间减少率/%
DP	39.28	100	20.67	
DQN	42.23	89.84	4.38	78.81
DQN + deg	43.28	87.63	4.42	78.62

EMS 的 87.63%。此外，所提策略在图 9 所示循环驾驶工况下在 838 s 时的计算时间小于 5 s，执行每个动作耗时 3.41 ms，在实际应用中具有实时性。

5　结论

（1）提出了一种基于 DQN 的 PEMFC 混合动力电动拖拉机能量管理策略。首先，与传统的基于 Q-学习的策略相比，基于 DQN 的 EMS 在训练稳定性方面具有更好的性能。其次，以 DP 为基准，基于 DQN 的 EMS 功率分配更加合理，有效地将超级电容 SOC 波动维持在 0.7 左右，并达到 DP 方法 89.84% 的燃油经济性。最后，提出基于 DQN + deg 的 PEMFC 系统寿命意识的能量管理策略，可以有效降低 PEMFC 系统退化，燃油经济性比基于 DQN 的能量管理策略降低 2.46%。

（2）在循环驾驶工况下进一步验证了所提策略的有效性。仿真结果表明，所提策略能够在燃油经济性略有提升的情况下有效提高 PEMFC 系统使用寿命，在循环驾驶工况中实现了基于 DP 基准的 87.63% 燃油经济性；此外，所提策略计算效率提高 78% 以上，每个动作执行时间为 3.41 ms，保证了实时性。

参　考　文　献

[1]　徐立友，张俊江，闫祥海，等. 农业装备电动化技术研究综述[J]. 农业机械学报，2023，54(9)：1 - 12.
XU Liyou, ZHANG Junjiang, YAN Xianghai, et al. Review of research for agricultural equipment electrification technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 1 - 12. (in Chinese)

[2]　刘孟楠，雷生辉，赵静慧，等. 电动拖拉机发展历程与研究现状综述[J]. 农业机械学报，2022，53(增刊 1)：348 - 364.
LIU Mengnan, LEI Shenghui, ZHAO Jinghui, et al. Review of development process and research status of electric tractors[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 1): 348 - 364. (in Chinese)

[3]　LI X Z, ZHANG M Z, YAN X H, et al. Power allocation strategy for fuel cell distributed drive electric tractor based on adaptive multi-resolution analysis theory[J]. Energy, 2023, 284: 129350.

[4]　李利桥，高宗余，聂晶. 电动拖拉机混合电源负载自适应能量管理策略研究[J]. 农业机械学报，2024，55(9)：459 - 469.
LI Liqiao, GAO Zongyu, NIE Jing. Load-adaptive energy management strategy for electric tractor based on hybrid power source [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(9): 459 - 469. (in Chinese)

[5] LIU T, TANG X L, WANG H, et al. Adaptive hierarchical energy management design for a plug-in hybrid electric vehicle[J]. IEEE Transactions on Vehicular Technology, 2019, 68(12): 11513 - 11522.
[6] HU X S, HAN J, TANG X L, et al. Powertrain design and control in electrified vehicles: a critical review[J]. IEEE Transactions on Transportation Electrification, 2021, 7(3): 1990 - 2009.
[7] ZHANG P, YAN F W, DU C Q, et al. A comprehensive analysis of energy management strategies for hybrid electric vehicles based on bibliometrics[J]. Renewable and Sustainable Energy Reviews, 2015, 48: 88 - 104.
[8] LIU T, TANG X L, HU X S, et al. Human-like energy management based on deep reinforcement learning and historical driving experiences[J/OL]. arXiv, 2020. https://doi.org/10.48550/arXiv.2007.10126
[9] TENG T, ZHANG X, DONG H, et al. A comprehensive review of energy management optimization strategies for fuel cell passenger vehicle[J]. International Journal of Hydrogen Energy, 2020, 45(39): 20293 - 20303.
[10] MUSARDO C, RIZZONI G, GUEZENNEC Y, et al. A - ECMS: an adaptive algorithm for hybrid electric vehicle energy management[J]. European Journal of Control, 2005, 11(4 - 5): 509 - 524.
[11] HUANG Y J, WANG H, KHAJEPOUR A, et al. Model predictive control power management strategies for HEVs: a review [J]. Journal of Power Sources, 2017, 341: 91 - 106.
[12] XIE S B, HU X S, QI S W, et al. Model predictive energy management for plug-in hybrid electric vehicles considering optimal battery depth of discharge[J]. Energy, 2019, 173: 667 - 678.
[13] GUO J Q, HE H W, LI J W, et al. Real-time energy management of fuel cell hybrid electric buses: fuel cell engines friendly intersection speed planning[J]. Energy, 2021, 226: 120440.
[14] ZHENG Y L, HE F, SHEN X Z, et al. Energy control strategy of fuel cell hybrid electric vehicle based on working conditions identification by least square support vector machine[J]. Energies, 2020, 13(2): 426.
[15] ZHANG R, TAO J L, ZHOU H Y. Fuzzy optimal energy management for fuel cell and supercapacitor systems using neural network based driving pattern recognition[J]. IEEE Transactions on Fuzzy Systems, 2019, 27(1): 45 - 57.
[16] SONG K, LI F Q, HU X, et al. Multi-mode energy management strategy for fuel cell electric vehicles based on driving pattern identification using learning vector quantization neural network algorithm[J]. Journal of Power Sources, 2018, 389: 230 - 239.
[17] KANDIDAYENI M, FERNANDEZ A O M, KHALATBARISOLTANI A, et al. An online energy management strategy for a fuel cell/battery vehicle considering the driving pattern and performance drift impacts[J]. IEEE Transactions on Vehicular Technology, 2019, 68(12): 11427 - 11438.
[18] REDDY N P, PASDELOUP D, ZADEH M K, et al. An intelligent power and energy management system for fuel cell/battery hybrid electric vehicle using reinforcement learning[C]// IEEE Transportation Electrification Conference and Expo, Detroit, MI, USA, 2019.
[19] SUN H C, FU Z M, TAO F Z, et al. Data-driven reinforcement-learning-based hierarchical energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles[J]. Journal of Power Sources, 2020, 455: 227964.
[20] ZHOU Y, RAVEY A, PÉRA M C. Multi-mode predictive energy management for fuel cell hybrid electric vehicles using Markov driving pattern recognizer[J]. Applied Energy, 2020, 258: 114057.
[21] LIN X Y, ZHOU B H, XIA Y T. Online recursive power management strategy based on the reinforcement learning algorithm with cosine similarity and a forgetting factor[J]. IEEE Transactions on Industrial Electronics, 2021, 68(6): 5013 - 5023.
[22] WU J D, HE H W, PENG J K, et al. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus[J]. Applied Energy, 2018, 222: 799 - 811.
[23] DU G D, ZOU Y, ZHANG X D, et al. Deep reinforcement learning based energy management for a hybrid electric vehicle [J]. Energy, 2020, 201: 117591.
[24] 李贤哲,张明柱,刘孟楠,等. 基于自适应多分辨率分析的电动拖拉机驱动功率分配策略[J]. 农业工程学报,2023, 39(23):55 - 66.
LI Xianzhe, ZHANG Mingzhu, LIU Mengnan, et al. Drive power allocation strategy for electric tractor based on adaptive multi-resolution analysis[J]. Transactions of the CSAE, 2023, 39(23): 55 - 66. (in Chinese)
[25] LI H, RAVEY A, N'DIAYE A, et al. Equivalent consumption minimization strategy for fuel cell hybrid electric vehicle considering fuel cell degradation[C]// IEEE Transportation Electrification Conference and Expo, Chicago, IL, USA, 2017.
[26] WIPKE K B, CUDDY M R, BURCH S D. Advisor 2.1: a user-friendly advanced powertrain simulation using a combined backward/forward approach[J]. IEEE Transactions on Vehicular Technology, 1999, 48(6):1751 - 1761.
[27] FLETCHER T, THRING R, WATKINSON M. An energy management strategy to concurrently optimise fuel consumption & PEM fuel cell lifetime in a hybrid vehicle[J]. International Journal of Hydrogen Energy, 2016, 41(46): 21503 - 21515.
[28] CHEN H C, PEI P C, SONG M C. Lifetime prediction and the economic lifetime of proton exchange membrane fuel cells[J]. Applied Energy, 2015, 142: 154 - 163.
[29] ZHAO J, LI X G. A review of polymer electrolyte membrane fuel cell durability for vehicular applications: degradation modes and experimental techniques[J]. Energy Conversion and Management, 2019, 199: 112022.
[30] FILIPPI A D, STOCKAR S, ONORI S, et al. Model-based life estimation of Li-ion batteries in PHEVs using large scale vehicle simulations: an introductory study[C]//2010 IEEE Vehicle Power and Propulsion Conference, 2010.