

文章编号: 1003-1421(2026)09-0022-13

引用格式: 陈宇轩, 彭其渊, 黄平. 基于两阶段多智能体强化学习的列车速度曲线及运行图顺序优化模型[J]. 铁道运输与经济, 2026, 48(9): 22-34.

CHEN Yuxuan, PENG Qiyuan, HUANG Ping. Sequential Optimization Model for Train Speed Curve and Train Working Diagram Based on Two-Stage Multi-Agent Reinforcement Learning[J]. Railway Transport and Economy, 2026, 48(9): 22-34.

基于两阶段多智能体强化学习的列车速度曲线及运行图顺序优化模型

陈宇轩^{1, 2}, 彭其渊^{1, 2}, 黄平^{1, 2}

(1. 西南交通大学 交通运输与物流学院, 四川 成都 611756; 2. 西南交通大学 综合交通运输智能化国家地方联合工程实验室, 四川 成都 611756)

摘 要: 降低列车运行能耗对于实现轨道交通绿色低碳发展、提高铁路运营收益有着重要意义。本研究从列车速度曲线与运行图2个层面入手, 综合考虑列车运行中的限速、时间、舒适度等约束条件, 分别采用DQN单智能体和A2C多智能体强化学习对列车的运行速度曲线和列车整体运行图进行优化, 构建相应的仿真运行环境, 建立两阶段的能耗优化模型。最后以武广高速铁路部分线路为例, 其结果表明该模型可以使原本的列车运行能耗降低16%以上, 同时满足安全与准点运行要求。针对不同的线路与运行情况该模型依然适用, 可以为列车运行能耗优化提供一定的参考。

关键词: 高速铁路; 强化学习; 速度曲线; 列车运行图; 节能优化; 两阶段优化

中图分类号: U238 **文献标识码:** A **DOI:** 10.16668/j.cnki.issn.1003-1421.20251225002

Sequential Optimization Model for Train Speed Curve and Train Working Diagram Based on Two-Stage Multi-Agent Reinforcement Learning

CHEN Yuxuan^{1, 2}, PENG Qiyuan^{1, 2}, HUANG Ping^{1, 2}

(1. School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 2. National and Local Joint Engineering Research Center of Integrated Transportation Intelligence, Southwest Jiaotong University, Chengdu 611756, Sichuan, China)

Abstract: Reducing the train operation energy consumption is significant for achieving low-carbon and green rail transit development and improving the railway operation revenue. This study started with the train speed curve and the train working diagram and comprehensively considered the constraints such as speed limits, travel time, and passenger comfort during train operation. It optimized the train speed curve and train working diagram via deep Q-network (DQN) single-agent and advantage actor-critic (A2C) multi-agent reinforcement learning, respectively, constructed corresponding simulation operation environments, and established a two-stage energy consumption optimization model. A section of the Wuhan-Guangzhou High Speed Railway was taken as the research case, and the results show that the proposed model can reduce the train operation energy consumption by more than 16% while satisfying the safe and on-time operation requirements. The

收稿日期: 2025-12-25 修订日期: 2026-05-29

通信作者: 黄平(1990—), 男, 四川宜宾人, 西南交通大学交通运输与物流学院副教授。

基金项目: 国家重点研发计划项目(2022YFB4300502); 国家自然科学基金项目(72301221)

model is applicable to different lines and operation conditions and provides certain reference for the train operation energy consumption optimization.

Keywords: High Speed Railway; Reinforcement Learning; Speed Curve; Train Working Diagram; Energy-Saving Optimization; Two-Stage Optimization

0 引言

随着我国铁路网络扩张与列车运营规模扩大,铁路能耗过大的问题日益显现。在轨道交通运输系统中,列车能源消耗主要来自于机车或动车组进行的牵引任务。截至2024年,我国铁路能源消耗折算标准煤1 801.2万t,同比增长2.8%^[1]。减少运输过程中的能源消耗不仅符合保护环境、绿色发展的理念,有利于实现“双碳”目标,还能减轻企业的运营成本压力^[2]。

列车运行时刻表是组织列车在网络和线路中运行的文件,它规定了列车的运行速度、到发时刻等参数;优化列车运行图中列车的能耗是减少铁路系统能耗的重要措施之一^[3]。列车运行图能耗优化问题主要是针对既有运行图,通过调节列车的运行速度,在考虑安全运行、停站方案与乘客满意度等基础情况下,优化列车的运行工况(驾驶策略),最终减少运行能耗。

当前的节能优化方法主要沿着2个相对独立的方向展开,一是调整列车的运行工况,通过优化区间速度曲线(即微调列车到发时刻)以减少牵引能耗;二是提高列车再生制动能耗的利用率,通过调整运行图、改变列车制动时刻使运行过程中产生的再生能被充分利用^[4]。针对列车运行速度曲线和列车运行图的优化都存在一定的深入研究空间,速度曲线针对单列车细分工况,运行图基于多列车整体调控,忽略了彼此之间的协同互补。

前者常基于列车动力学模型,考虑限速、再生制动、分相区等因素,采用启发式智能算法优化速度曲线^[5-6]。例如,张可心^[7]从列车目标速度曲线优化和速度跟踪控制两大方面对列车优化运行策略展开研究。陈昱^[8]建立了列车多质点模型,采用遗传算法在列车运行过程中实现坡道速度与惰行工况双重节能优化。侯涛等^[9]提出了一种基于改进PID搜索算法的列车节能优化方法,分析临时限速和运行时间调整情况下的列车节能优化。朱成琨^[10]、黄子凌等^[11]分别采用DQN、SAC算法求解列车新的节能

运行速度曲线。

后者主要关注列车在路网上的时空关系,提升运行图能效。例如,革新^[12]针对高速铁路分别提出了与区间运行时分配、再生制动能利用、动车组运用计划协同优化3种模型。马阳阳^[13]针对列车晚点情况,建立了路网条件下的列车运行调整与节能控制协同优化模型,通过列车工况改变期间运行能耗。李荣喆等^[14]以同一供电分区内各列车牵引与制动工况的重叠时间最大化为目标建立模型对运行图能耗进行优化。赵思哲^[15]研究了单列车节能优化方法,并考虑列车再生制动能量设计了高速铁路多列车递阶节能优化方法。

当前以上2方面的研究较为割裂:在优化列车速度曲线时固定区间运行时间,忽略了可以调整的区间时间,或是在优化运行图时采用简化模型(例如假设能耗与区间运行时间线性相关),不能准确反映操作策略对能耗的影响,导致能耗优化存在一定的限制。将速度曲线优化和运行图调整相结合,组成顺序优化过程,将运行图与精确动力学模型结合,实现铁路线路上能耗的进一步降低。一方面,列车区间规定的运行时间是计算运行能耗的基础,由于线路坡度等条件存在,采用精确动力学模型计算的能耗更为准确。另一方面,节能曲线生成的能耗可以反过来指导运行图调整区间运行时间,采用区间能耗更低的运行方案。

因此,构建一种两阶段优化模型,在给定区间运行时间的条件下,基于列车动力学模型优化区间运行速度曲线,建立列车在不同区间的节能速度曲线备选集合。再将生成的速度曲线备选集作为输入,在满足所有车次总运行时间不变的前提下,以系统总能耗最低为目标,优化所有车次在运行图中的到发时间。通过两阶段优化实现运行图宏观时间资源与列车微观操纵策略的整体最优,为后续编制节能运行图提供一定的参考。

1 问题描述

单列车在区间平直道上运行时,研究人员根据

其能耗特性，提出了四阶段运行策略，该策略指出，列车在理想道路上运行时，运行策略应该包含牵引、巡航、惰行、制动4个阶段，求解列车最优速度曲线的关键在于寻找列车工况的切换时机。该运行方案可以在一定程度上保障运行效率，但由于未考虑线路坡度等因素，当线路坡道较多时，其固定的运行策略限制了能耗进一步优化。多列列车常通过调整列车运行时刻的方法，延长列车运行时间降低速度需求或是重叠不同列车的加速、减速时刻以利用再生能，最终达到降低能耗的目的。在调整运行时刻时需避免列车冲突，并保证准点与乘客舒适度，不能随意延长运行时间。

为了进一步降低列车在区间的能耗，将列车的运行区间按照需求分为若干个更小的区间，列车在每个小区间可以选择当前区间的运行工况，直至终点。列车在区间的运行方案以里程-速度曲线进行表示，该曲线直接显示了列车运行到某一位置时的运行速度，根据同一区间的不同运行时间，可以求解出对应的速度运行曲线与运行能耗，构成备选集，列车运行速度曲线图如图1所示。

基于速度运行曲线备选集，对列车运行图进行

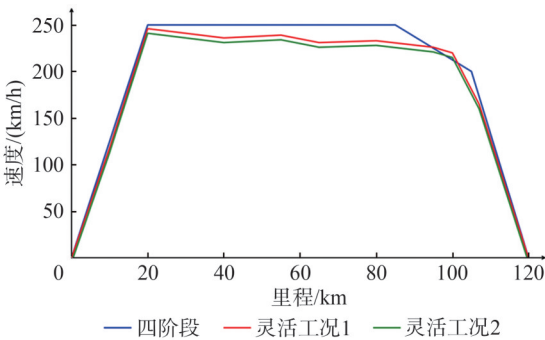


图1 列车运行速度曲线图
Fig.1 Train operation speed curve

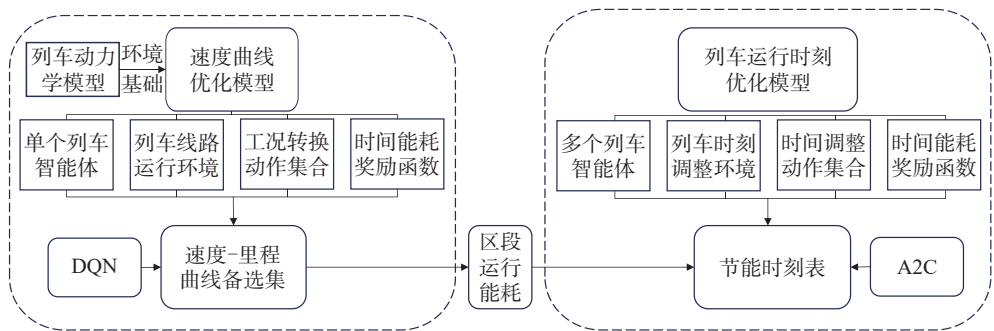


图2 模型结构
Fig.2 Model architecture

微调。模型的核心约束之一为列车总运行时间，即运行图中所有班次列车从起点至终点全程的区间运行时间总和。在不改变列车停站方案与运行图中所有车次运行时间之和的情况下，基于列车安全约束、区间运行时间约束、车站停站时间约束、时间连续约束等重新分配每列车在每个区间的运行时间，通过调整部分车次在车站的出发、到达时间，实现降低列车在区间的运行能耗的目的。

因此，此处的优化问题可以总结为：以总运行能耗最低为目标，满足列车与列车之间、列车运行的约束条件，考虑所有运行车次的总运行时间，优化不同列车在不同车站的进站、出站时间，确定能耗优化情况下的列车运行时刻表。

研究采用单智能体强化学习优化列车速度曲线，多智能体强化学习优化列车时刻表，将速度曲线的优化能耗作为时刻表调整的基本数据，以时刻表调整的时间作为规定时间优化列车的运行工况。

2 列车时刻表优化模型

本研究模型主要由2个部分组成：优化单列车运行能耗的运行速度曲线优化模型和优化多列车运行过程的时刻表优化模型，其模型结构如图2所示。速度曲线优化模型求解出不同时间与区间下的列车运行速度曲线及运行能耗，构成备选集；运行图优化模型修改列车的运行时间，根据前者所求的速度曲线备选集计算运行图的整体能耗，最终求解出能耗优化后的时刻表。

2.1 列车动力学模型

列车在区间运行主要受到自身牵引制动力、基本阻力以及附加阻力影响。本研究将列车进行单质点化，不考虑车厢之间的相互作用力，求解列车在

区间的纵向运行速度曲线,得到符合节能要求的里程-速度曲线^[16-18]。

列车的牵引力与制动力由列车机车主动提供,通常可由列车的牵引制动特性曲线获得。本研究所采用的列车型号为CRH3型高速列车,其牵引力、制动力与运行速度关系函数表现如下^[3]。

$$\begin{cases} F_q = 300 - 0.285v & 0 \leq v \leq 125 \text{ km/h} \\ F_q = 31500/v & 125 \text{ km/h} < v \end{cases} \quad (1)$$

$$\begin{cases} F_b = 59.8v & v \leq 5 \text{ km/h} \\ F_b = -0.285v + 300 & 5 < v \leq 106.7 \text{ km/h} \\ F_b = 28880/v & v > 106.7 \text{ km/h} \end{cases} \quad (2)$$

式中: v 为列车运行速度, km/h; F_q 为列车牵引力, kN; F_b 为列车制动力, kN。

列车运行阻力是列车在运行过程中阻碍列车前行的外力,一般分为基本阻力与附加阻力。

列车运行基本阻力是列车在平直轨道运行时所受的各种阻力之和,其涉及因素众多且情况复杂,一般采用经验公式进行计算。

$$W_0 = a + bV + cV^2 \quad (3)$$

式中: a , b , c 为经验常数。

附加阻力主要与运行时的线路状态相关,通常包含隧道阻力、坡道阻力以及曲线阻力等。

列车的隧道阻力由在隧道中运行的车头与车尾的压力差产生,计算公式为

$$W_t = 1.3 \times 10^{-4} lmg \quad (4)$$

式中: l 为隧道长度, m; m 为列车质量, t; g 为重力加速度。

当列车运行至坡道时,其重力会分解成垂直于坡道与平行于坡道的两个方向的力,垂直力与钢轨的支持力大小相同,但是平行力将会影响列车的运行。坡道阻力主要与列车自身的重力以及坡道的坡度相关,计算公式为

$$W_i = mg \sin(i) \approx mgi \quad (5)$$

式中: i 为线路坡度。

列车在弯道上行驶时,由于离心力的作用会使列车的运行摩擦力变大,因此在弯道上行驶会产生更大的运行阻力,计算公式为

$$W_r = \frac{600mg}{R} \quad (6)$$

式中: R 为曲线半径, m。

列车的附加阻力为

$$W_f = W_0 + W_t + W_i + W_r \quad (7)$$

总体受力为

$$F = u_f F(v) - u_b B(v) - W_f \quad (8)$$

$$\frac{dv}{dx} = \frac{u_f F(v) - u_b B(v) - W_f}{(1 + \rho)vm} \quad (9)$$

式中: ρ 为列车回转质量系数; u_f 和 u_b 分别为列车的牵引、制动级位。

2.2 基于DQN的单列车速度曲线优化模型

Q学习是一种经典的基于值函数的强化学习算法,DQN在Q学习的基础上,采用神经网络去拟合状态转移时产生的Q值。DQN算法采用了Q值网络和目标网络2个网络,分别负责迭代更新和稳定保持、训练速度提高,同时还设计了经验回放机制^[19-21]。

列车运行速度曲线优化模型是在规定运行时间内,通过在每个工况转换点选择不同的运行工况,从而改变列车运行速度曲线,最终减少列车的牵引输出以达到列车运行能耗优化的模型^[22-25]。

(1) 状态及状态转移。设列车的初始状态为 $s_0(x_0, v_0, t_0)$,当智能体选择了动作时,状态进行改变。整个运行区间根据需要按照状态转移步长 Δx 进行分割,当进入下一个状态时,总里程增加相应距离。

$$x_{i+1} = x_i + \Delta x \quad (10)$$

式中: x_i 为列车当前所处里程, km; Δx 为区间离散间隔, km。

列车的运行速度按照仿真运行步长计算,变化满足下式。

$$v_{i+1} = \int_{x_i}^{x_{i+1}} \sqrt{v_i^2 + 2adx} \quad (11)$$

式中: a 为列车加速度。

列车的运行时间变化满足

$$t_{i+1} = t_i + \int_{x_i}^{x_{i+1}} \frac{dx}{v} \quad (12)$$

式中: t_i 为列车当前阶段运行时间, s。

在区间运行时,列车应该满足区间各类约束。

列车的运行里程应满足公式(13)。

$$0 \leq x_i \leq x_l \quad (13)$$

式中: x_l 为终点里程, km。

列车在区间的运行速度满足区间限速条件。

$$0 \leq v_i \leq v_{\text{limit}}^i \quad (14)$$

式中: v_{limit}^j 为区间 j 限速, km/h。

列车从起点运行至终点时, 区间运行时间应满足规定时间。

$$0 \leq t_i \leq t_{\text{limit}} \quad (15)$$

式中: t_{limit} 为区间规定运行时间, s。

(2) 动作及动作屏蔽。该模型的智能体动作集合基于列车的最大牵引与制动输出, 主要分为惰行、最大牵引、最大制动、部分牵引、部分制动 5 种工况, 组成列车的离散动作集合, 智能体在每个阶段选择动作与环境进行互动, 环境接受动作后返回被改变的列车状态。

为了提高模型求解的速度, 在模型环境中添加动作屏蔽, 屏蔽特定状态下列车的不合理动作, 提高模型取得可行解的速度, 加速收敛。

为了保证乘客乘车时的舒适度, 列车的加速度变化必须保持在一定的范围之内。

$$\Delta a \leq k \quad (16)$$

式中: k 为舒适度系数。

列车当前的运行时刻与列车从当前位置到达终点的最短时间之和, 应该满足预计的时间限制。

$$t_i + t_{i, \text{end}}^{\min} \leq t_{\text{limit}} \quad (17)$$

式中: $t_{i, \text{end}}^{\min}$ 为列车从当前阶段运行至终点的最短所需时间, s。

当列车抵达终点时, 列车应当稳当停车, 且停车位置应该有一定的误差容量。表现为

$$v_{\text{end}} = 0 \quad (18)$$

$$x_{\text{end}} = x_i \pm \xi \quad (19)$$

式中: v_{end} 为列车终点速度, km/h; x_{end} 为列车终点运行里程, km。

(3) 奖励函数。环境的奖励函数主要包含时间限制奖励和能耗奖励两种。时间奖励鼓励列车准点到达, 能耗奖励尽可能使能耗更低, 如公式(20)、公式(21)所示。

$$R_t = t_{\text{plan}} - t_{\text{end}} \quad (20)$$

$$R_e = a - b \int f(x) dx \quad (21)$$

式中: t_{plan} 为列车预计运行时间, s; t_{end} 为列车实际抵达终点运行时间, s。

为了促使智能体向下一个阶段转移, 需要在智能体成功转移时对其奖励一个较小的固定值。在一轮训练结束, 智能体抵达最后阶段时, 给予一个极

大值作为完成训练的奖励; 在训练中由于违反约束、无可选择动作等情况导致提前中断一轮训练时, 给予一个极大的负值作为训练出错的惩罚。

$$R_0 = r + R_{\text{terminal}} + R_{\text{truncation}} \quad (22)$$

式中: R_{terminal} 为终止奖励; $R_{\text{truncation}}$ 为中断奖励。

在能耗与时间奖励前添加参数确保智能体不会过于追求某一项指标而忽略另一项指标。最终的奖励计算方式为

$$R = \omega_1 \cdot R_t + \omega_2 \cdot R_e + R_0 \quad (23)$$

式中: ω_1 , ω_2 分别为时间参数、能耗参数。

最后, 根据所得到的动作状态值函数, 采用 ϵ -greedy 策略选取列车运行过程中的运行工况, 形成列车在该区间的速度-距离曲线, 求得该曲线下的列车运行能耗。

模型运行流程如图 3 所示。

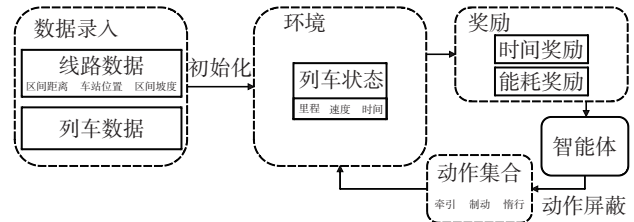


图3 模型运行流程

Fig.3 Model operation process

将线路按照车站划分成若干区间, 对于每个区间设定列车行驶的时间范围, 以 1 min 为间隔, 求得每个区间不同时间的运行速度曲线, 组成速度曲线备选集。

2.3 A2C 多智能体列车运行时刻表优化模型

AC 算法是一种基于策略梯度和价值函数的强化学习算法。该算法结合了 2 种网络, 一个 Actor 网络用于产生动作, 另一个 Critic 网络用于评判动作、估计状态值函数, 指导 Actor 改进策略。

列车运行时刻表优化模型是基于强化学习的优化模型, 该模型通过调整既有线路上的列车运行时刻, 改变列车在任意车站的发车、到达时间, 来做到对时刻表运行时刻的调整, 对列车进行规定时间的能耗优化^[26-30]。由于线路不同位置的坡度、曲线半径不同, 即使相同的运行时间其列车所消耗的能耗也会不同, 该模型主要调整列车在区间的运行时刻, 减少部分车次的运行时间, 将其分配给降低能耗更多的运行区间, 即使部分列车能耗增加, 也能

确保在所有车次总运行时间不变的情况下,降低总运行能耗^[31]。模型的训练智能体采用了多智能体的形式,将每班次列车的运行数据匹配一个智能体,构建一个 Agent Environment Circle(AEC)结构的多智能体交互环境。多智能体优化模型拥有复数智能体,这些智能体能够独自选择动作并与环境交互,每个智能体拥有不同的状态和奖励值。

(1) 状态及状态转移。该模型环境导入了运行图的基础运行时间数据与区间的车站线路数据,记录了所有列车在不同阶段的发车、到达时间,反映列车的运行情况。将以上所有列车的到发阶段按照顺序设定为环境的不同步骤阶段。

设单个列车的初始状态为 $s_0^{\text{train}} = (t_0, o_0)$, 其中, t 表示列车在当前状态的修改时间, o 表示运行方向。那么与环境交互的列车状态集合、动作集合、列车的时刻集合即为

$$\begin{aligned} s_0 &= \{\text{train}_0: (t_0^0, o_0^0), \text{train}_1: (t_0^1, o_0^1), \dots\} \\ \text{actions} &= \{\text{train}_0: \text{action}_0, \text{train}_1: \text{action}_1, \dots\} \\ \text{schedule} &= \{\text{train}_0: (\text{time}_i^0, \text{time}_{i+1}^0, \dots), \text{train}_1: \\ &(\text{time}_i^1, \text{time}_{i+1}^1, \dots), \dots\} \end{aligned}$$

列车的运行方向在抵达终点前保持不变,每个智能体在当前阶段 s 选择需要执行的动作 a , 即改变时间 Δt , 修改列车的运行时刻。

$$t_{i+1}^{\text{train}} = t_i^{\text{train}} + \Delta t \quad (24)$$

式中: t_i^{train} 为列车当前运行时刻, $\min$ 。

所有的智能体动作采用集合的方式与环境进行互动,环境以相同的方式返回各智能体的状态,即

$$\text{schedule}_{i+1} = \text{schedule}_i + s_i \quad (25)$$

(2) 动作及动作屏蔽。模型的动作集合为时间调整数值,针对每班次列车的出发与到达时间进行提前或者延后调整,其出发-到达时间间隔构成列车的区间运行时间,到达-出发时间构成停站时间。每个阶段,智能体从离散动作集合中选择一个时间用于修改当前的时间节点,列车 j 在当前阶段 s 的时间节点为 t_s^j , 所选动作为 a_s , 改变后的时间即为 $t_s^j + a_s$ 。

针对每个智能体,预设动作屏蔽掩码,在此基础上,智能体选择动作更新状态,并影响下一个智能体。

列车的站间运行时间应满足连续性约束。

$$t_x^f < t_{x+1}^d \quad (26)$$

$$t_x^d \leq t_{x+1}^f \quad (27)$$

式中: t_x^f 为列车在 x 车站的出发时间, $\min$; t_x^d 为列车在 x 车站的到达时间, $\min$; x 表示任意车站。

当列车时间进行调整时,该列车后续所有车站的出发、到达时间进行同样幅度的平移调整。

$$\forall y > x, t_y' = t_y + \Delta t \quad (28)$$

式中: x, y 表示任意车站; t_y' 表示调整后的时间。

任意一列列车应该满足区间运行时间约束与规定的停站时间约束。

$$t_{\min} \leq t_{\text{run}} \leq t_{\max} \quad (29)$$

$$t_{\min} \leq t_{\text{dwell}} \leq t_{\max} \quad (30)$$

式中: t_{run} 为列车区间运行时间, $\min$; t_{dwell} 为列车车站停站时间, $\min$ 。

同方向任意相邻列车应该满足追踪到达时间间隔^[32]。

$$t_j^d - t_k^d \geq I_d \quad (31)$$

同方向任意相邻列车应该满足追踪出发时间间隔。

$$t_j^f - t_k^f \geq I_f \quad (32)$$

上述公式中, j 与 k 分别表示不同列车; f 与 d 分别表示出发与到达时间; I 表示追踪时间间隔。

任意车站同一时间内所停留的列车不应超过车站的最大容量。

$$n_x \leq n_x^{\text{limit}} \quad (33)$$

式中: n_x^{limit} 为该车站的最大可容纳列车数。

运行图中共计 n 列列车, m 个运行区间,图中所有车次的区间运行时间之和应该保持不变。

$$\sum_{p=1}^m \sum_{j=1}^n t_{j,p}^{\text{run}} = \sum_{p=1}^m \sum_{j=1}^n t_{j,p}^{\text{run}'} \quad (34)$$

式中: $t_{j,p}^{\text{run}}$ 表示 j 次列车在区间 p 的运行时间; $t_{j,p}^{\text{run}'}$ 为调整过后的运行时间。

AEC 环境运行逻辑及模型运行流程如图 4 所示。智能体与环境互动后,返回当前智能体的调整时间与运行方向,更新列车运行时刻表数据,同时更新动作掩码,由下一个智能体与环境交互, AEC 环境具体流程如图 4a 所示。当所有的智能体完成了当前步骤的交互之后,返回相应的奖励值。

(3) 奖励函数。模型的奖励函数主要以能耗为目标,同时考虑运行时间、停站时间、运行图基本时间间隔等约束,调节运行图中的列车运行时间。在不改变列车原本的停站方案的情况下,首先调整

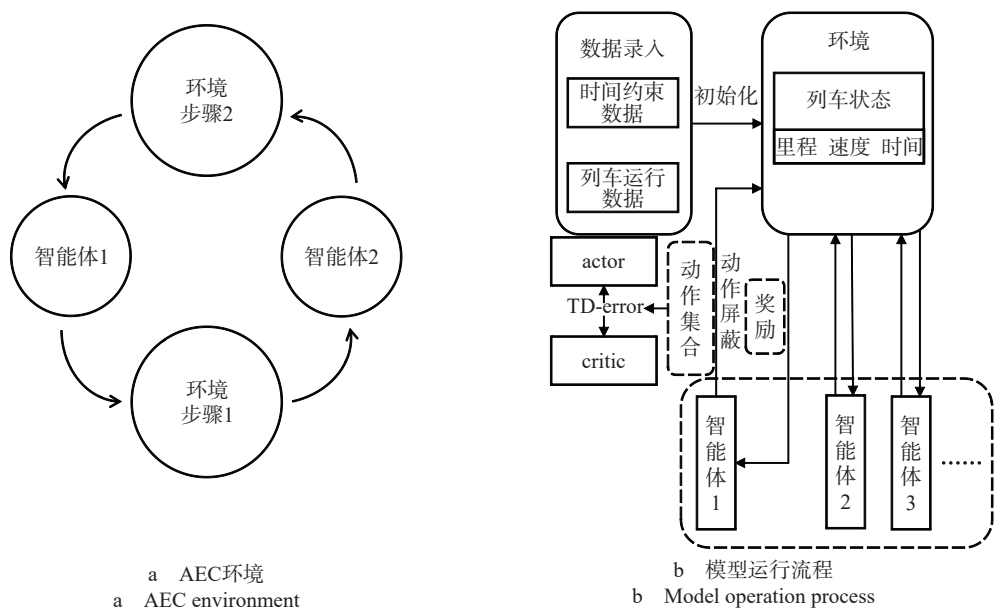


图4 AEC环境运行逻辑及模型运行流程
Fig.4 AEC environment operation logic and model operation process

列车的区间运行时间，根据列车的起点、终点与运行时间，从速度曲线备选集中选取相应的运行曲线，获得该区间的运行能耗，所有停站区间组成总区间的运行能耗。模型运行流程如图4b所示。

3 实例

以武广高速铁路(武汉—广州南)为例，对某一时时间段的列车运行图进行优化，确保运行过程中无运行时间冲突。所选择区间为广州南站至乐昌东站，共计6个车站，时间段为6:00—11:00，共计开行29列列车。

设广州南站作为起始车站，里程为0，其余各车站里程分布如表1所示。

首先对途经车站进行总停站方案求解，任意两

表1 车站里程分布				km
Tab.1 Distribution of station mileage				
车站	里程	车站	里程	
广州南	0	英德西	140	
广州北	47	韶关	227	
清远	83	乐昌东	271	

个车站组成一种列车停站方案，共计15种停站方案。针对每一种停站方案，给出预计的运行时间范围，通过速度曲线优化模型求解不同运行时间要求下该停站方案的速度曲线。不同停站方案规定运行时间如表2所示。

用于强化学习仿真训练基本参数如表3所示。

基于以上设定，编写列车运行仿真环境，搭建DQN强化学习框架，求解不同停站方案下指定运行时间的列车运行速度曲线。首先确定列车的运行区

表2 不同停站方案规定运行时间
Tab.2 Specified operation time for different stop plans

起点站	终点站	里程/km	预计运行时间/min	起点站	终点站	里程/km	预计运行时间/ min
广州南	广州北	47	13~21	广州北	乐昌东	224	45~55
广州南	清远	83	20~30	清远	英德西	57	14~19
广州南	英德西	140	30~39	清远	韶关	144	30~39
广州南	韶关	227	48~56	清远	乐昌东	188	38~48
广州南	乐昌东	271	55~65	英德西	韶关	87	17~26
广州北	清远	36	9~15	英德西	乐昌东	131	26~35
广州北	英德西	93	19~27	韶关	乐昌东	44	9~16
广州北	韶关	180	36~45				

表3 训练基本参数
Tab.3 Basic training parameters

名称	参数	名称	参数
列车型号	CRH3	激活函数	ReLU
列车质量/t	380	折扣系数	0.95
列车初始位置/km	0	算法初始探索率	1
列车初始速度/(km/h)	0	a, b, c	0.66, 0.002 45, 0.000 132
状态转移步长/km	0.1	时间参数 ω_1	0.4
仿真运行步长/km	0.01	能耗参数 ω_2	1
列车停车位置偏差允许值/km	0.3	完成训练正奖励值	+1 000
隐藏层数量	1	中断训练负奖励值	-1 000
各隐藏层神经元数量	64		

间里程与规定运行时间，将列车的运行区间从起点至终点按照0.1 km分为若干状态转移区间，0.01 km作为仿真运行步长，基于速度、时间等约束，使智能体从0开始按照状态转移区间向前运行，每个新的区间选择当前区间的运行工况，直到抵达终点，经过训练后输出符合要求的节能速度曲线。以广州南站至广州北站为例，列车运行距离为47 km，4种

不同时间的列车运行速度曲线如图5所示，图中蓝线表示运行速度曲线，红线表示区间限速。

对其他不同里程的停站方案同样求解出各时间运行速度曲线，计算运行能耗，共同构成能耗备选集用于后续列车时刻表优化。不同时间的区间运行能耗如图6所示。

将所有列车的运行时间与能耗数据进行存储。

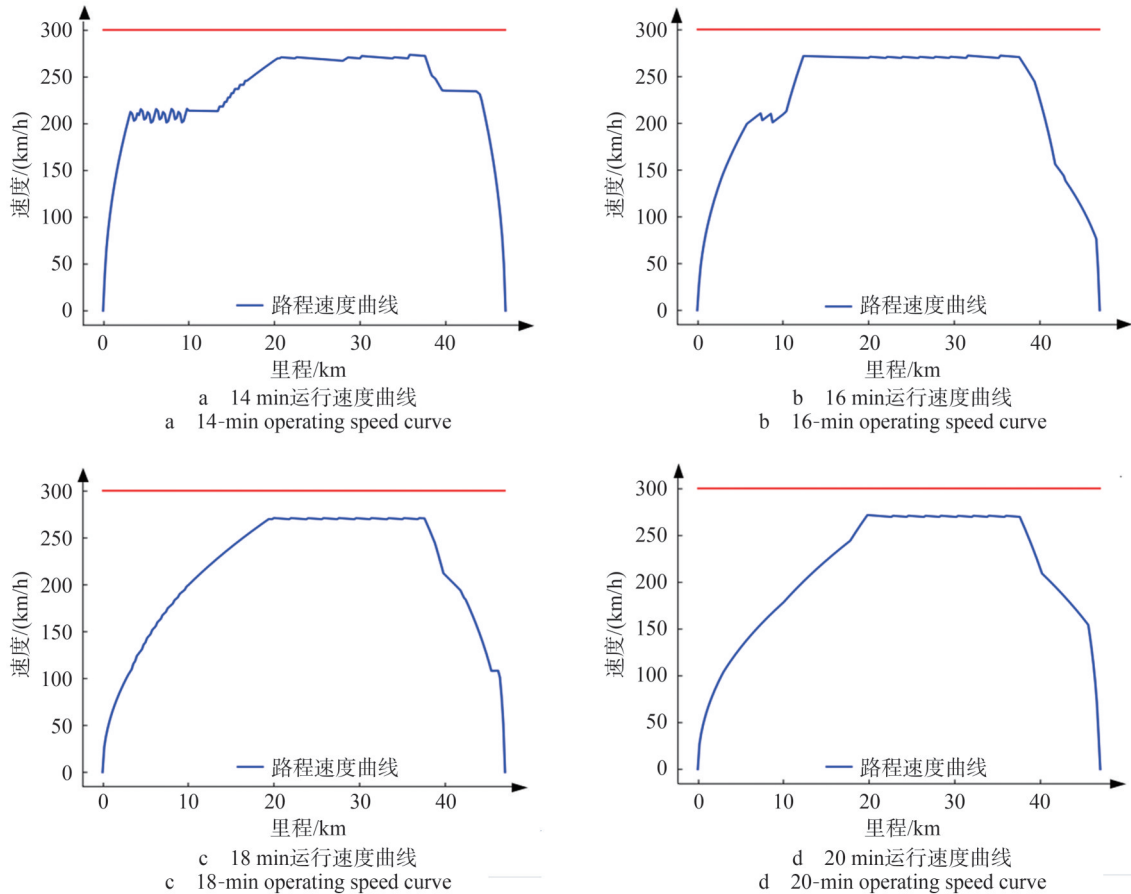


图5 列车运行速度曲线

Fig.5 Train operation speed curve

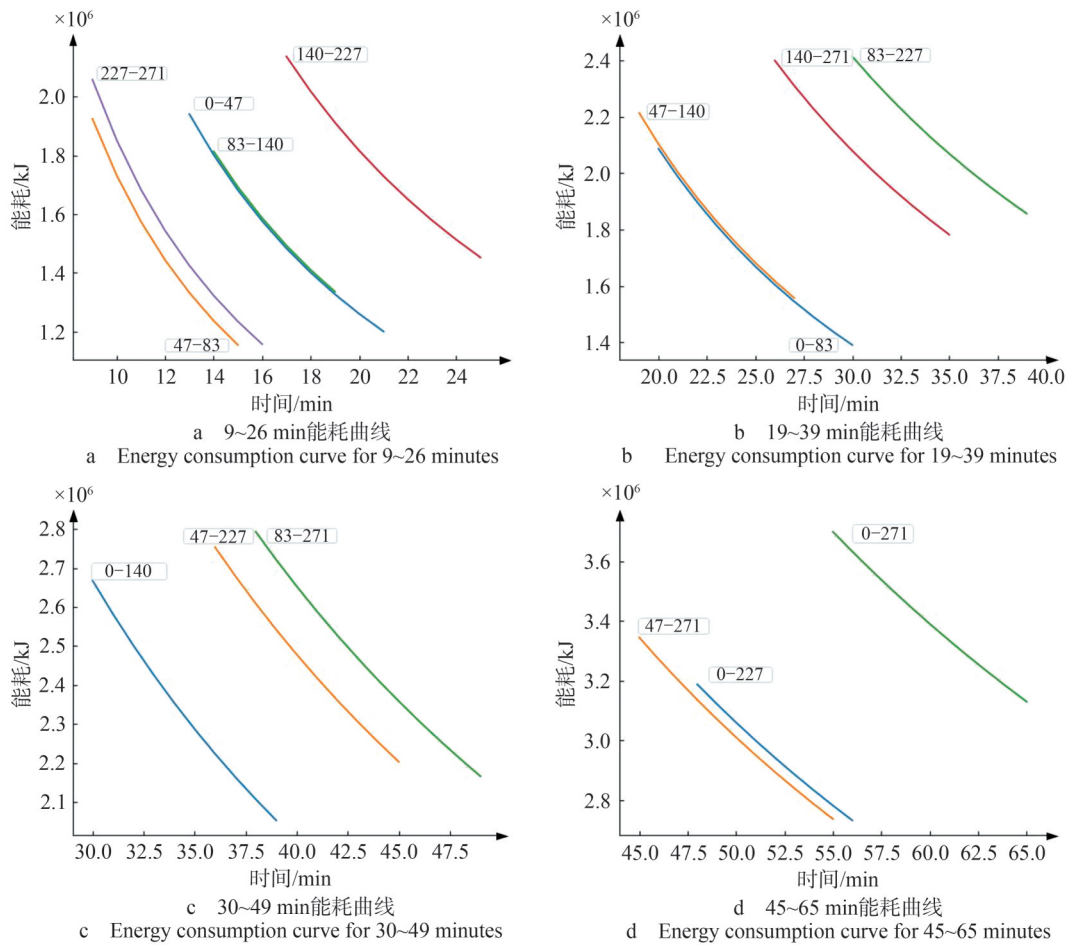


图6 区间运行能耗
Fig.6 Energy consumption of sections

编写多智能体交互环境，设置与开行列车数量相同的智能体匹配列车运行数据，单个智能体只负责改变其绑定列车的到发时间。环境动作设置为离散的时间变量，智能体所选动作会直接改变当前列车所处阶段的时间。从起始车站开始，将每个车站的出发和到达时间设置为一个阶段，列车从起始车站的出发时间开始向前模拟，经历出发、到达、出发等阶段变化，在每个区间选择一个动作改变当前区间的出发或到达时间。在每一个阶段，按照顺序让智能体选择该阶段执行的动作，遇到冲突通过动作裁剪消除冲突，每个智能体都做完选择后，更新状态与动作掩码向下一个阶段推进。抵达终点后，对当前结果进行奖励，输出符合要求的运行图。

采用传统四阶段运行模式对列车能耗进行优化，总体能耗约为 39 374 kW·h。两阶段优化先对列车运行区间的速度曲线进行优化，得到节能曲线后再调整列车区间运行时间，经过训练后输出运行

时间不变的新运行时刻，求解运行图如图 7 所示，该运行图运行能耗共计 32 760.29 kW·h。相比于传统四阶段运行模式其消耗能耗更低，有 16% 以上的优化，能耗优化如表 4 所示。模型奖励值如图 8 所示。列车优化时刻表如表 5 所示。各列车能耗变化如图 9 所示。

相较于传统启发式算法，强化学习由于动作屏蔽的设置求解速度较快，同时保持了良好的优化效果，算法在曲线优化方面有着近 13% 的能耗优化，针对宏观的运行图优化可以将能耗降低 5% 左右。采用粒子群算法优化列车运行能耗时，同样对速度曲线与运行图进行优化，相较于强化学习，其曲线求解时间稍长，优化效果略低，只有运行图能耗优化差距较小。两种算法在求解运行图的时间上都花费较长，总体上对原本四阶段运行模式有 16% 以上的能耗缩减，方法对比如表 6 所示。

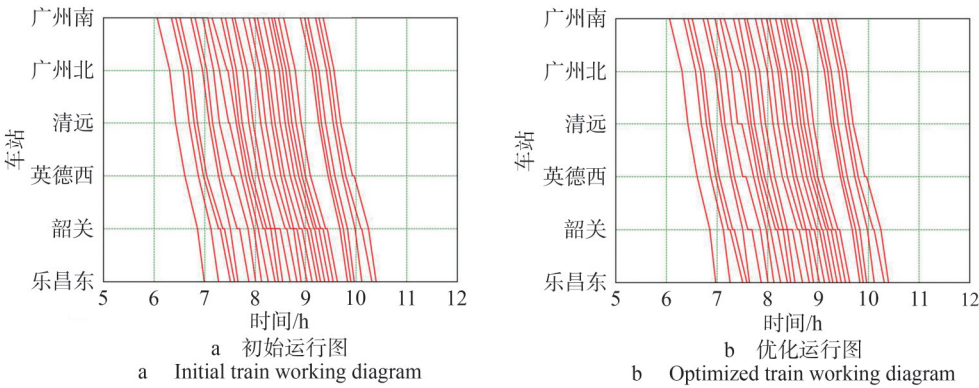


图7 求解运行图
Fig.7 Solution diagram

表4 能耗优化

Tab.4 Energy consumption optimization

运行模式	能耗/(kW·h)	优化率/%
传统四阶段运行	39 374.31	—
速度曲线-运行图优化运行	32 760.29	16.80

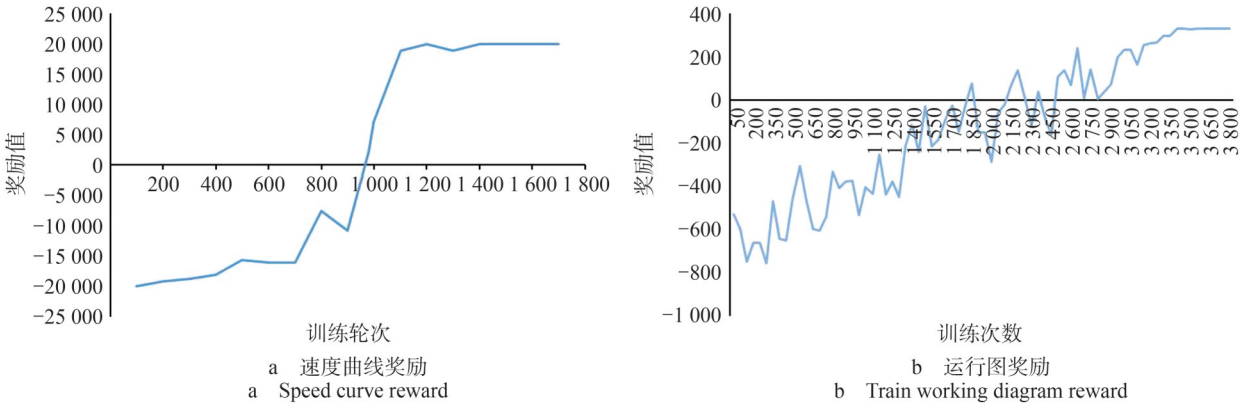


图8 模型奖励值
Fig.8 Model reward

表5 优化时刻表

Tab.5 Optimized timetable

列车编号	运行时刻									
	广州南		广州北		清远		英德西		韶关	
	出发	到达	出发	到达	出发	到达	出发	到达	出发	到达
G9672	364	379	379	386	386	397	397	412	412	420(-1)
G1408	381	395	395	402	402	413	413	428	428	437(-1)
G296	386	401	401	408	408	419	419	437(-3)	439(-3)	451
G1744	391	405	405	412	412	423	423	442(-3)	444(-3)	456
G276	404	418	418	426	426	437	437	451	451	460(-1)
G1316	409	423	423	430	430	441	441	459(-2)	462(-2)	473(-2)
G1102	418	431	431	438	438	453(-2)	455(-2)	472(-2)	472(-2)	481(-2)
G6132	424	438	438	449(-3)	450(-3)	464(-3)	464(-3)	479(-3)	479(-3)	488(-3)
G832	429	447(-1)	448(-1)	458(-1)	458(-1)	469(-1)	469(-1)	484(-1)	484(-1)	496(-4)
G1748	440	454	454	461	461	473	473	490(-2)	492(-2)	504(+1)
G9662	446	458	458	465	465	477	477	494(-1)	510(-1)	522(+2)

列车编号	续表									
	运行时刻									
	广州南	广州北		清远		英德西		韶关		乐昌东
	出发	到达	出发	到达	出发	到达	出发	到达	出发	到达
G1002	453	466	466	473	473	484	484	499	499	508
G94	456	471	471	478	478	489	489	504	504	513(-2)
G72	466	480	480	487	487	498	498	516(-2)	518(-2)	530(+1)
G6012	473	486	486	493	493	504	504	522(-1)	530(-1)	542(+2)
G86	479	493	493	500	500	511	511	526	526	535
G550	483	496	496	504	504	515	515	533(-2)	536(-2)	548(+1)
G1312	490	503	503	510	510	522	522	539(-1)	542(-1)	554(+2)
G280	496	509	509	516	516	527	527	545(-1)	546(-1)	558(+2)
G6164	500	513	513	520	520	532	532	550(-2)	551(-2)	563(+1)
G636	503	517	517	524	524	536	536	553(-2)	556(-2)	568(+1)
G9696	508	521	521	529	529	540	540	558(-2)	561(-2)	573(+1)
G4556	514	528	528	535	535	546	546	564(-2)	566(-2)	578(+1)
G96	534	548	548	556	556	567	567	581	581	590(+2)
G80	539	552	552	559	559	570	570	585	585	594
G6142	543	556	556	564	564	575	575	590	590	598
G1004	552	565	565	573	573	584	584	599	599	607
G1108	556	569	569	577	577	588	588	606(-2)	607(-2)	619(+1)
G74	561	574	574	582	582	596(-2)	598(-2)	615(-2)	615(-2)	624(-2)

注：()内数据表示时刻调整时间，-表示提前，+表示延后。

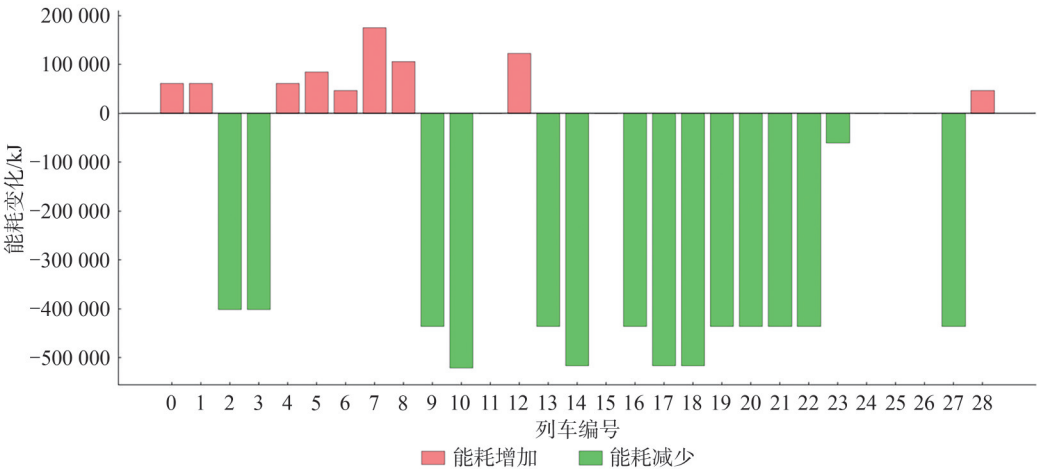


图9 列车能耗变化
Fig.9 Change of energy consumption

表6 方法对比 Tab.6 Method comparison				
算法	单曲线求解时间/s (47 km 为例)	运行图优化时间/min	曲线能耗优化率/%	运行图能耗优化率/%
强化学习	362+	30+	12.80	4.58
粒子群算法	437+	30+	11.50	4.18

4 结论

研究主要从运行图角度进行列车能耗优化研究，从2个方面展开，一是通过优化区段列车的运

行速度曲线，减少单个列车运行能耗；二是通过运行图的整体调控，调整列车运行时间，进一步优化列车运行图总运行能耗。模型使用DQN求解速度曲线，A2C求解运行图，在未改变列车原有次序与

停站方案的基础上,使列车运行总能耗最小,并求得能耗优化过后的新列车运行图;结果表明,相比于当前运行图,优化后运行图能耗平均可节约16%。针对不同情况的线路,该模型仍可适用且无需进行过多修改,通过录入线路数据与列车数据,对模型进行重新训练,能够在短时间内求得符合要求的节能运行图。

该研究仍有较大的提升空间。首先,列车动力学模型采用了单质点模型,未来研究可以采用多质点模型进行研究,能够更准确地描述出列车在运行过程中的受力情况,更精确地求出运行能耗;其次,本研究暂未考虑列车的再生制动,作为列车运行过程中的另一节能手段,如果能够同时考虑速度曲线与再生制动,可进一步提高列车运行效率。最后,模型假设所有列车按计划正常运行,实际场景可能遇到特殊情况导致当前求解计划无法完整执行,需要进行适当调整,因此针对干扰场景和特殊情况的研究仍有必要。

参考文献:

- [1] 国家铁路局. 2024年铁道统计公报[Z]. 北京: 国家铁路局, 2025.
- [2] 刘杰, 彭其渊, 殷勇. 低碳背景下的多式联运路径规划[J]. 交通运输系统工程与信息, 2018, 18(6): 243-249.
LIU Jie, PENG Qiyuan, YIN Yong. Multimodal Transportation Route Planning under Low Carbon Emissions Background[J]. Journal of Transportation Systems Engineering and Information Technology, 2018, 18(6): 243-249.
- [3] 李志远. 面向高速列车的节能优化研究[D]. 南昌: 华东交通大学, 2023.
- [4] 高坤盈泽, 张玉召, 冀璇. 高速列车运行图节能优化研究现状及展望[J]. 铁道运输与经济, 2024, 46(8): 24-36.
GAO Shenyongze, ZHANG Yuzhao, JI Xuan. Research Status and Prospects of Energy-Saving Optimization for High Speed Train Working Diagrams[J]. Railway Transport and Economy, 2024, 46(8): 24-36.
- [5] 田志强, 刘磊, 孙国锋, 等. 铁路长大坡道线路考虑维修天窗设置的列车运行图优化[J]. 交通运输系统工程与信息, 2025, 25(3): 288-298.
TIAN Zhiqiang, LIU Lei, SUN Guofeng, et al. Optimization of Train Operation Diagram Considering Maintenance Windows on Railway Lines with Long and Steep Gradients[J]. Journal of Transportation Systems Engineering and Information Technology, 2025, 25(3): 288-298.
- [6] 盛昭. 考虑电分相的高速列车节能优化方法研究[D]. 北京: 北京交通大学, 2022.
- [7] 张可心. 基于节能规划的高速列车优化运行策略研究[D]. 北京: 北京交通大学, 2023.
- [8] 陈昱. 基于双重优化的高速列车节能运行研究[D]. 兰州: 兰州交通大学, 2021.
- [9] 侯涛, 兰小斌. 基于改进PID搜索算法的列车节能运行优化研究[J]. 重庆交通大学学报(自然科学版), 2025, 44(5): 112-120.
HOU Tao, LAN Xiaobin. Energy-Saving Operation Optimization of Train Operation Based on Improved PID Search Algorithm[J]. Journal of Chongqing Jiaotong University (Natural Science), 2025, 44(5): 112-120.
- [10] 朱成琨. 基于网-车-线耦合的高速列车节能运行优化研究[D]. 兰州: 兰州交通大学, 2024.
- [11] 黄子凌, 石红国. 基于深度强化学习的高速列车运行节能在线优化[J]. 铁道标准设计, 2026, 70(5): 10-19, 34.
HUANG Ziling, SHI Hongguo. Energy-Saving Online Optimization of High Speed Train Operation Based on Deep Reinforcement Learning[J]. Railway Standard Design, 2026, 70(5): 10-19, 34.
- [12] 革新. 考虑节能的高速列车运行图优化研究[D]. 兰州: 兰州交通大学, 2023.
- [13] 马阳阳. 路网条件下高速铁路列车运行调整与节能控制协同优化[D]. 兰州: 兰州交通大学, 2022.
- [14] 李荣喆, 马睿杰. 基于再生制动的朔黄铁路节能运行图优化方法[J]. 科技与创新, 2025(1): 198-201.
- [15] 赵思哲. 高速列车递阶节能牵引优化运行驾驶策略研究[D]. 长沙: 中南大学, 2023.
- [16] 周学影. 基于强化学习的高速列车目标速度曲线与运行时分协同优化[D]. 北京: 北京交通大学, 2021.
- [17] 孟祥麟. 基于Q学习和深度Q学习的列车自动驾驶节能运行控制[D]. 北京: 北京交通大学, 2021.
- [18] ALI SANDIDZADEH M, HAVAEI P. A Comprehensive Study on Reinforcement Learning Application for Train Speed Profile Optimization[J]. Multimedia Tools and Applications, 2023, 82(24): 37351-37386.
- [19] 陈昂扬, 范家铭, 徐辉章, 等. 基于强化学习的铁路列车运行图编制与优化研究综述[J]. 铁道运输与经济, 2025, 47(11): 36-51.

- CHEN Angyang, FAN Jiaming, XU Huizhang, et al. Review on Reinforcement Learning-Based Railway Train Working Diagram and Optimization[J]. Railway Transport and Economy, 2025, 47(11): 36–51.
- [20] 韩忻辰, 俞胜平, 袁志明, 等. 基于 Q-Learning 的高速铁路列车动态调度方法[J]. 控制理论与应用, 2021, 38(10): 1511–1521.
- HAN Xincheng, YU Shengping, YUAN Zhiming, et al. High Speed Railway Dynamic Scheduling Based on Q-Learning Method[J]. Control Theory & Applications, 2021, 38(10): 1511–1521.
- [21] 朱擎阳. 基于强化学习的城轨列车节能驾驶控制方法[D]. 北京: 北京交通大学, 2022.
- [22] 黄 畅, 姜辰宇, 邢昕铨. 基于强化学习的列车驾驶曲线节能优化算法[J]. 数字通信世界, 2022(1): 19–21.
- HUANG Chang, JIANG Chenyu, XING Xinquan. An Energy-Efficient Train Control Approach Based on Deep Q-Network Methodology[J]. Digital Communication World, 2022(1): 19–21.
- [23] 朱宇清, 白宝雪, 陈鸿辉, 等. 基于神经网络的货运列车速度曲线实时规划算法[J]. 控制与信息技术, 2019(3): 7–12.
- ZHU Yuqing, BAI Baoxue, CHEN Honghui, et al. A Real-Time Speed Profile Planning Algorithm for Freight Trains Based on Neural Network[J]. Control and Information Technology, 2019(3): 7–12.
- [24] 张 森. 基于强化学习的列车自动驾驶方法研究[D]. 北京: 中国铁道科学研究院, 2020.
- [25] 张 森, 张 琦, 张梓轩. 基于 Q 学习算法的高速铁路列车节能优化研究[J]. 铁道运输与经济, 2019, 41(12): 111–117.
- ZHANG Miao, ZHANG Qi, ZHANG Zixuan. A Study on Energy-Saving Optimization for High Speed Railways Train Based on Q-Learning Algorithm[J]. Railway Transport and Economy, 2019, 41(12): 111–117.
- [26] 王荣笙, 张 琦, 张 涛, 等. 基于蒙特卡罗树搜索-强化学习的列车运行智能调整方法[J]. 中国铁道科学, 2022, 43(5): 146–156.
- WANG Rongsheng, ZHANG Qi, ZHANG Tao, et al. Intelligent Adjustment Approach for Train Operation Based on Monte Carlo Tree Search-Reinforcement Learning[J]. China Railway Science, 2022, 43(5): 146–156.
- [27] 杜心怡, 邵长虹. 基于强化学习的高速铁路列车运行调整方法研究[J]. 铁道技术标准(中英文), 2024, 6(6): 35–43.
- DU Xinyi, SHAO Changhong. Research on High Speed Railway Timetable Rescheduling Based on Reinforcement Learning[J]. Railway Technical Standard (Chinese & English), 2024, 6(6): 35–43.
- [28] DONG H R, NING L B, ZHOU M, et al. Deep Reinforcement Learning for Integration of Train Trajectory Optimization and Timetable Rescheduling under Disturbances[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36(8): 14009–14021.
- [29] 周靖哲. 基于深度强化学习的高速铁路列车运行图优化方法研究[D]. 北京: 北京交通大学, 2023.
- [30] REYNOLDS E, MAHER S J. A Data-Driven, Variable-Speed Model for the Train Timetable Rescheduling Problem[J]. Computers & Operations Research, 2022, 142: 105719.
- [31] LI W Q, NI S Q. Train Timetabling with the General Learning Environment and Multi-Agent Deep Reinforcement Learning[J]. Transportation Research Part B: Methodological, 2022, 157: 230–251.
- [32] 吕苗苗. 新旧交替列车运行图编制优化理论与方法研究[D]. 成都: 西南交通大学, 2013.

(责任编辑 吴文娟)