DOI：10.13890/j.issn.1000-128X.2023.06.017　　文章编号：1000-128X(2023)06-0139-08

基于强化学习的重载组合列车长大下坡操纵优化研究

王建华[1,2]，王春毅[1]，曾　周[1,2]，王　聪[2]，王青元[2]，杨　航[2]
（1.国能朔黄铁路发展有限责任公司，河北 肃宁　062350；
2.西南交通大学 电气工程学院，四川 成都　611756）

摘　要：针对2万t重载组合列车在长大下坡区段列车纵向冲动大，以及连续循环空气制动操纵困难的问题，文章基于数据驱动算法，设计了一种适用于长编组重载组合列车的长大下坡区段操纵优化方法。该方法考虑了不同列车间、同一列车不同制动系统状态间空气制动特性的差异，利用神经网络对列车在不同操作状态下的空气制动性能变化规律进行学习，得到空气制动力预测模型；而后基于强化学习设计重载组合列车操纵优化算法，以速度跟随性为目标建立奖励函数，考虑列车牵引/电制动特性、列车管制动及缓解时间、限速、运行平稳性等约束条件，应用强化学习方法开展重载组合列车长大下坡操纵策略优化。基于实车数据验证空气制动仿真模型与操纵优化算法的可行性与合理性。结果表明，文章建立的空气制动力预测模型对列车运行中空气制动系统性能具有良好的预测作用，优化操纵控制策略相较于司机驾驶能够有效地减小列车运行中的纵向冲动与最大车钩力，保障列车运行安全。

关键词：重载组合列车；空气制动；长大下坡；神经网络；强化学习；重载铁路
中图分类号：U292.92^{+}1　　**文献标识码：**A

Research on operation optimization of heavy-haul combined trains in long and steep downhill sections based on reinforcement learning

WANG Jianhua[1,2], WANG Chunyi[1], ZENG Zhou[1,2], WANG Cong[2], WANG Qingyuan[2], YANG Hang[2]

(1. CHN ENERGY Shuohuang Railway Development Co., Ltd., Suning, Hebei 062350, China;
2. School of Electric Engineering, Southwest Jiaotong University, Chengdu, Sichuan 611756, China)

Abstract: To mitigate longitudinal impulse and address challenge posed by continuous air braking operations in long and steep downhill sections for 20 000-ton heavy haul combined trains, this paper proposes an approach for operation optimization of such trains featuring a long formation in such sections based on a data-driven algorithm. An air braking force prediction model was developed based on neural network learning focusing on the variation rules of air braking performance across different operating states, to incorporate differences in air braking characteristics across different trains and varied braking system states on same trains. Then, an operation optimization algorithm was designed, establishing a reward function that prioritized speed following features, based on reinforcement learning. This algorithm incorporated constraints including train traction/electric braking characteristics, braking application and release times of train pipes, speed limits, and operational smoothness, to optimize the operation strategy for heavy-haul combined trains in long and steep downhill sections by reinforcement learning. The proposed air braking simulation model and operation optimization algorithm were verified using data collected from operation of real trains, confirming their feasibility and rationality. The results show the effectiveness of the proposed air braking force prediction model in predicting performance of air braking systems on running trains. Compared to manual

基金项目：国家重点研发计划项目（2021YFB2601504）
通信作者：王青元（1984—），男，博士，高级工程师，硕士生导师，主要从事轨道交通电气与自动化方面的研究；E-mail: wangqy@swjtu.edu.cn
引用格式：王建华, 王春毅, 曾周, 等. 基于强化学习的重载组合列车长大下坡操纵优化研究[J]. 机车电传动, 2023(6): 139-146.
Citation：WANG Jianhua, WANG Chunyi, ZENG Zhou, et al. Research on operation optimization of heavy-haul combined trains in long and steep downhill sections based on reinforcement learning[J]. Electric drive for locomotives,2023(6): 139-146.

driving, the optimized operation strategy plays an effective role in reducing longitudinal impulse and maximum coupler force to ensure the safety during train operation.

Keywords: heavy-haul combined train; air braking; long and steep downhill; neural network; reinforcement learning; heavy-haul rallway

0 引言

铁路货运需求量的不断增加，推动着我国铁路货运列车朝着长编组、大载重的方向发展[1]。重载组合列车的运行区间大多经过山区路段，线路条件复杂，往往伴随着连续长大下坡道，需要使用“空气制动–缓解–再制动–再缓解”的循环制动操纵方式[2]。现阶段，我国已开行的重载组合列车运行控制主要依赖于司机的手动驾驶，缺乏灵活可靠的在线操纵指导，长大下坡区段停缓现象时常发生，严重地影响重载铁路运输效益。因此，开展重载组合列车长大下坡区段操纵优化研究，为司机操纵提供动态指导，对提高司机操纵水平、降低重载铁路运输成本，具有重要指导意义。

针对列车运行优化问题，文献[3-4]从极值原理出发，推导了列车节能运行的最优控制工况集。文献[5]利用二次规划算法实现了重载组合列车多目标优化策略的研究，实现了一种满足安全、准点、平稳、节能的货运列车多目标预测控制算法。文献[6]将重载列车的运行操作分为连续的牵引、电制动力控制和离散的空气制动控制，并分别设计不同的机器学习算法实现对不同控制量的智能控制。

针对长大下坡区段的操纵困难问题，文献[7]对重载组合列车在长大下坡区段运行时如何满足周期性制动所需的缓解充风时间这一问题做了论述，并验算了不同情况下能够满足缓解充风时间约束的最大允许运行坡道千分数。文献[8-9]提出在长大下坡区段使用“动力制动+空气制动”联锁制动，能够提高制动调速效果，优化周期性制动次数，降低闸瓦磨耗与减小闸瓦温升。文献[10-11]以大秦线为背景，通过仿真和现场试验结果表明：长大下坡区段空气制动减压量通常为−50 kPa，并且缓解初速度不得低于30 km/h，制动初速度考虑限速，最高不得超过70 km/h。文献[12-13]指出在变坡点进行空气制动时，列车所受的纵向冲动值最大，提高制动波的传递速度可以有效降低列车纵向冲动。

目前，针对重载组合列车长大下坡问题的研究大多关注纵向冲动和空气制动系统的建模，对其操纵优化研究大多停留在非长大下坡区段，鲜有考虑运行过程中重载组合列车空气制动性能的变化。因此，本文拟从工程应用出发，以实车运行数据为基础，针对长大下坡道操纵优化问题，考虑重载组合列车在运行过程中由于前序空气制动操纵引起的空气制动系统性能变化，基于神经网络构建空气制动力预测模型，用以求解列车运行中的空气制动力；在该模型基础上，基于强化学习的方法，设计考虑重载组合列车空气制动性能时变性的长大下坡操纵优化算法，求解重载组合列车在长大下坡区段的操纵控制策略，以解决2万t重载组合列车在长大下坡区段循环空气制动操纵困难问题。

1 重载组合列车运动学建模

重载组合列车编组长，其整车长度可达2.6 km，在线路上运行时往往会覆盖多个坡道或多条曲线。为了能够更加准确地计算列车所受的附加阻力，本文采用刚性多质点模型（如图1所示），该模型将每一节车厢视作1个刚性质点，分别计算其受到的附加阻力。在越过变坡点时，整列车所受附加阻力不会发生大的突变，计算结果精度相对较高；同时，相对于纵向动力学模型，计算简便，能够满足在线实时计算需求。

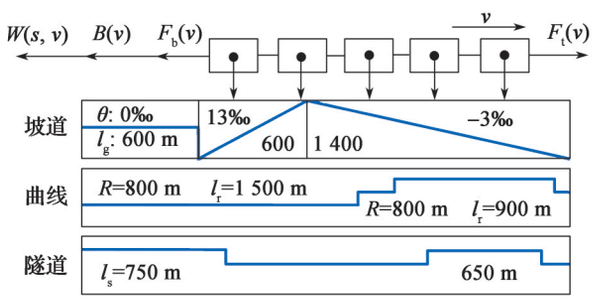


图1 重载列车刚性多质点模型

Fig. 1 Rigid multi-particle model of heavy-haul train

考虑列车运行中所受到的力，包括机车牵引力、机车电制动力、空气制动力、基本运行阻力、附加阻力，建立列车运动学方程，用以解算列车运行状态，运动方程如下：

$$\left.\begin{aligned}\frac{\mathrm{d}v}{\mathrm{d}x}&=\frac{g\left[F_{\mathrm{t}}(v)-W_0(v)-B_{\mathrm{d}}(v)-B_{\mathrm{c}}(v)-W_{\mathrm{j}}(x)\right]}{1000(1+\gamma)v}\\\frac{\mathrm{d}t}{\mathrm{d}x}&=\frac{1}{v}\end{aligned}\right\}\quad(1)$$

式中：x为列车运行位置，m；v为列车运行速度，m/s；γ为列车回转质量系数；$F_{\mathrm{t}}(v)$为列车单位牵引力，N/kN；$B_{\mathrm{d}}(v)$为列车单位电制动力，N/kN；$B_{\mathrm{c}}(v)$为列车单位空气制动力，N/kN；$W_0(v)$为列车单位基本运行阻力，N/kN；$W_{\mathrm{j}}(x)$为列车单位加算坡道阻力，N/kN。

考虑列车长度，加算线路附加阻力W_{j}，其计算公式为

$$W_j=\left(\sum\left(w_{gk}l_{gk}\right)+\sum\left(w_{rk}l_{rk}\right)+\sum\left(w_{sk}l_{sk}\right)\right)\Big/L_t \quad (2)$$

式中：L_t为列车长度，m；w_{gk}, w_{rk}, w_{sk}为列车所覆盖的第k个坡道、曲线和隧道的单位附加阻力，N/kN；l_{gk}, l_{rk}, l_{sk}为列车所覆盖的第k个坡道、曲线和隧道的长度，m。

2　基于神经网络的空气制动力预测模型

空气制动通过机械摩擦的方式将列车的动能转换为热能，然后消散于空气中。空气制动力的大小由列车管的减压量控制，当列车管减压时，空气波通过列车管逐渐传递至全列车，进而使得副风缸为制动缸充风，通过120型空气控制阀作用，并通过旋压制动缸和组合制动梁放大作用驱使闸瓦压紧车轮，形成空气制动力。

在空气制动过程中，由于列车管中空气波的传播不是立即完成的，制动信号到达列车各车辆的时间存在差异，越是靠近机车的车辆接收到空气制动信息指令越早，各车辆制动力随着空气波由前及后依次传递。因此，整列车制动力大小存在一个从小逐渐增强直至稳定发挥的变化过程，如图2中的虚线“*acef*”所示。牵规中为简化计算，将其等效替换为“开展过程+有效制动过程”，如图2中的实线“*ade′f*”所示。

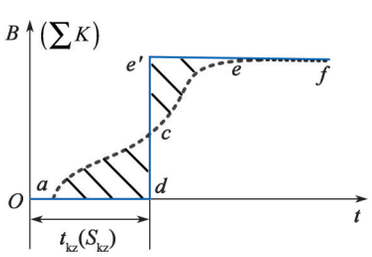


图2　列车空气制动力变化示意图

Fig. 2　Schematic diagram of variations in air braking force

“牵规”同时给出了在该空气制动力模型下的制动力计算方程式[6]，具体如下所示：

$$B_c(t)=\begin{cases}0, & t<t_{kz}\\ \beta_c\left(\sum_{i=1}^{N}\left(\varphi_h K_h\right)\right), & t\geqslant t_{kz}\end{cases} \quad (3)$$

式中：β_c为“牵规”常用制动系数；φ_h为“牵规”闸瓦的换算摩擦因数；K_h为“牵规”闸瓦的换算闸瓦压力，kN；t_{kz}为“牵规”车辆空走时间，s。

在不同减压量下，常用制动系数β_c如表1所示。

表1　常用制动系数β_c

Table 1　Service braking coefficient (β_c)

管压/kPa	减压量/kPa					
	50	60	70	80	90	100
500	0.19	0.32	0.42	0.52	0.60	0.68
600	0.17	0.28	0.37	0.46	0.53	0.60

2.1　经验操纵对空气制动性能的影响

在实际运行中，考虑到过大的减压量带来的纵向冲动问题，重载组合列车司机往往仅采用50 kPa的常用制动减压量。由于列车实际减压量存在波动，以及制动系统状态因前序空气制动操纵而变化等因素的影响，同样是在50 kPa减压作用下，不同列车之间、同一列车不同制动系统状态之间的空气制动系统所发挥的制动能力均存在明显差异[14-15]。

本文将空气制动力的发挥过程采用数学方程式描述，其数学方程式为

$$f=1\Big/\left(a+\mathrm{e}^{-bt}\right) \quad (4)$$

式中：a, b为描述列车空气制动发挥过程相关的参数，参数a为本次空气制动制动力最大值的倒数，表征列车空气制动实际性能发挥，参数b为影响从开始制动到空气制动有效发挥时间的大小。

对本务机车运行监控记录装置LKJ-2000记录数据进行分析，以列车管压力变化为基准，拆分空气常用制动区段，按式(4)寻找贴合度最高的参数a和b，获得不同常用制动区段的制动力稳态值，统计其中一趟车的制动力数据，具体如表2所示。

表2　某实车空气制动力发挥表

Table 2　Table for air braking force application of a real train

常用制动区段序号	常用制动力/kN	常用制动区段序号	常用制动力/kN
1	2 028	7	1 941
2	1 718	8	1 146
3	1 342	9	1 316
4	993	10	1 479
5	1 055	11	1 166
6	968	12	1 289

分析该本务机车数据可知，在列车运行过程中，不同制动区段之间空气制动力最大值存在明显的差异性，并且制动力的变化并不存在单调性，这种复杂的变化性增加了司机在长大下坡区段的驾驶操作难度。通过对多趟实车数据进行统计，其不同空气制动段制动力的分布如图3所示。

由图3可知，不同列车的初始空气制动性能即存在较大差异，而在后续的运行途中，其空气制动性能的差异性会进一步加大。由于无法掌握空气制动系统的具体作用状况，司机需要依据原有的制动数据与驾驶操作经验，来预测下一次空气制动性能的强弱，但其复杂的变化规律仅凭人的经验很难识别，即使是指导司机也难以对其进行准确预测和预判。若能根据制动操作历史数据来预测下一次的空气制动的制动力数值，将极大地降低重载组合列车在长大下坡区段的操纵难度。

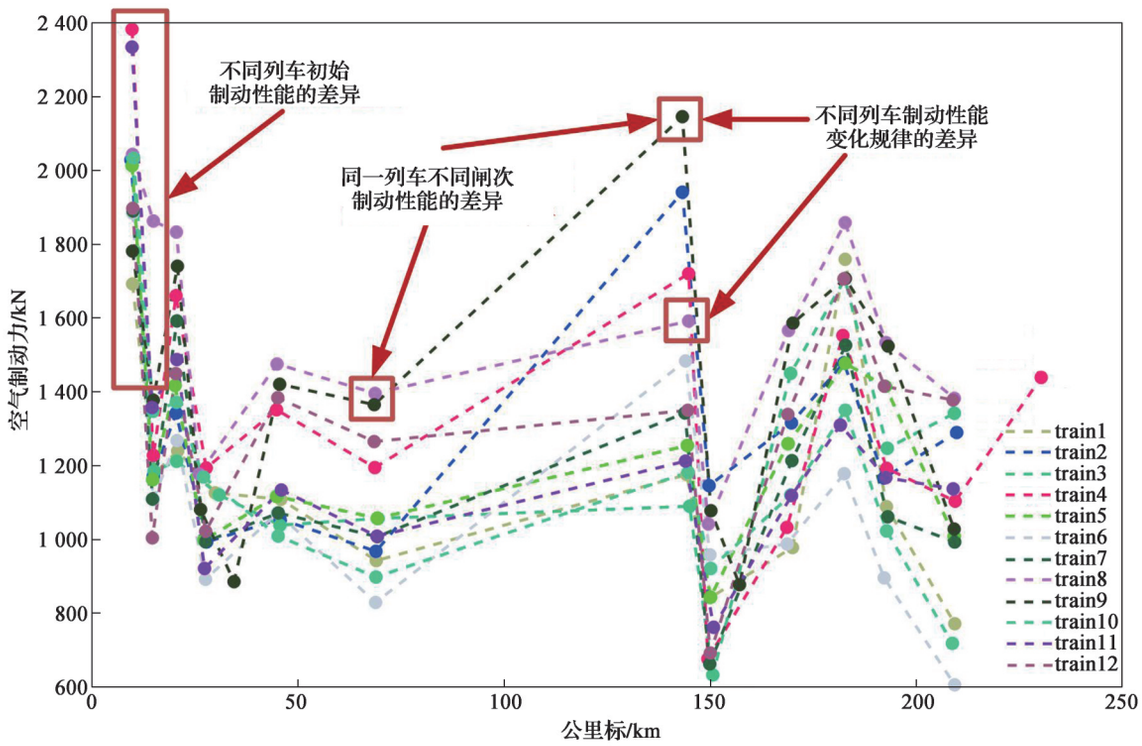


图3 不同列车空气制动力随制动段分布

Fig. 3 Distribution of air braking force in braking sections for different trains

本文从重载组合列车实际操纵需求出发，利用影响空气制动系统性能发挥的制动操作经验信息，实现基于BP神经网络的空气制动力预测模型。通过制动操作历史数据预测空气制动力，其本质上是通过列车空气制动系统的状态来判断其输出空气制动力的大小。由于这些参数在实际系统中难以采样，本文中选择多趟实际列车运行中LKJ-2000机车运行监控记录设备所记录的相关数据作为特征参数。

2.2 样本数据准备与模型训练

列车空气制动力的大小主要由减压指令传输至后面列车管施加的实际减压量决定，实际减压量的波动主要受2次空气制动间的缓解充风时间影响；同时，制动力的变化还受到列车空气制动系统当前疲劳状态的影响，如前序制动带来的闸瓦温度变化等。

本文对神经网络的输入和输出设计如下：

1）神经网络输入参数

输入参数见式(5)。

$$\boldsymbol{x}=\left[T_{\mathrm{B}}\ D_{\mathrm{B}}\ T_{\mathrm{R}}\ F_{\mathrm{b}}\ F_{\mathrm{m}}\right]^{\mathrm{T}} \tag{5}$$

2）神经网络输出参数

输出参数见式(6)。

$$y=F_{\mathrm{b}}' \tag{6}$$

式中：T_{B}为上一次空气制动过程持续的时间；D_{B}为上一次空气制动过程列车运行距离；T_{R}为上一次空气制动结束后列车管的充风时间；F_{b}为上一次空气制动的制动力稳态值；F_{m}为列车首次空气制动的制动力稳态值；F_{b}'为列车下一次空气制动的制动力稳态值。

3）神经网络训练条件

①网络结构：BP神经网络，其中输入层节点数5、隐藏层节点数32、输出层节点数1；

②优化算法：Levenberg-Marquardt；

③误差指标：均方差；

④数据划分：70%训练集，15%试验验证集，15%测试集；

⑤终止条件：连续6次试验验证网络性能未变化或达到最大训练次数。

对输入参数进行归一化处理，搭建BP神经网络，训练终止后，采用测试集测试该模型，统计预测结果并与实际值对比，计算误差分布，计算结果如表3、图4、表4、图5所示。

表3 绝对误差分布

Table 3 Absolute error distribution

绝对误差/kN	< 100	< 200	< 300
样本比例/%	75	90	96

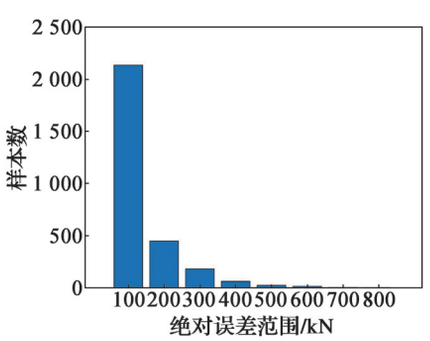


图4 绝对误差分布图

Fig. 4 Chart of absolute error distribution

表4　相对误差分布

Table 4　Relative error distribution

相对误差/%	< 10	< 20	< 30
样本比例/%	50	80	93

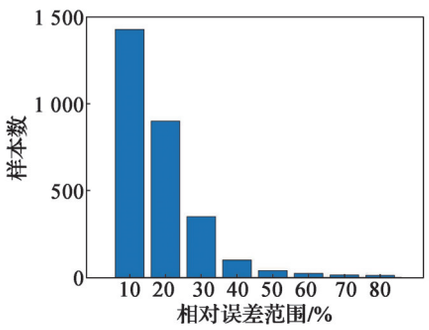


图5　相对误差分布图

Fig. 5　Chart of relative error distribution

如训练结果所示，75% 数据的绝对误差值在 100 kN 内，90% 数据的绝对误差值在 200 kN 内。这表明该空气制动力预测模型对列车运行中空气制动系统制动性能的变化具有良好的预测能力。

3　基于强化学习的重载组合列车操纵优化算法

本节将重载组合列车的操纵问题抽象为标准的强化学习问题，基于Q学习方法设计重载组合列车操纵优化算法，无须驾驶操作经验，能够在满足列车牵引/电制动特性、空气制动缓解时间等多种约束条件下，依据目标奖励函数求解列车操纵控制策略。

3.1　强化学习关键元素设计

①智能体：智能体是强化学习系统中的学习者与决策者，在此问题中，它可以接受环境奖励信号，并根据列车运行状态（如列车位置、运行速度、曲线、平直道及坡道等线路断面等）对列车下一步采取的操纵级位做出决策。从过去的经验中学习出最优价值函数和最优策略，建立完整的操纵策略。

②环境：在此问题中，环境可分为2部分——列车动力学模型和列车运行线路。环境接收智能体给出的动作，根据列车当前的运行状态，计算列车当前所能发挥的牵引力/制动力，以及受到的附加阻力、基本阻力、空气制动力等，进而更新列车下一步的运行状态，并根据奖励函数反馈该步的奖励值。

③动作空间：在此问题中，动作设定为列车手柄级位的加减值，如下式所示：

$$A_k=a,\quad a\in\{a_{\min},a_{\min}-1,\cdots,0,\cdots,a_{\max}-1,a_{\max}\} \tag{7}$$

式中：$a_{\min},a_{\max}$表示级位变化的最小值与最大值。

空气制动的施加由动作与当前手柄级位状态共同决定。若当前电制动级位加上动作值小于最低手柄级位限制，并且此时缓解时间满足最小充风时间约束，则施加空气制动；若空气制动已施加，而当前手柄级位加上动作值将大于0级位时，退出空气制动状态，手柄级位恢复至电制动级位。其数学表达为

$$A_{\mathrm{b}_k}=\begin{cases}1 & J_{\mathrm{w}_k-1}+a_{k-1}<J_{\mathrm{w_min}}\&R_{\mathrm{t}_k-1}>180\\ 0 & J_{\mathrm{w}_k-1}+a_{k-1}>0\end{cases} \tag{8}$$

式中：A_{b_k}为第k步是否施加空气制动；J_{w_k-1}为第k−1步时列车的手柄级位，$J_{\mathrm{w}}\in[-12,12]$；$R_{\mathrm{t}_k-1}$为第$k$−1步缓解时间。

④状态空间：在此问题中，状态指所希望观测到的部分列车运行信息，其表征了列车当前的状况，包括列车运行速度、所处位置、当前手柄级位、运行加速度、是否正在施加空气制动等基本信息。同时，为了更加准确地计算空气制动力，还需引入上文提及的空气制动力预测模型的输入信息，如下式所示：

$$\boldsymbol{S}_k=[X_k\ V_k\ J_{\mathrm{w}_k}\ A_{\mathrm{cc}_k}\ A_{\mathrm{b}_k}\ T_{\mathrm{B}_k}\ T_{\mathrm{R}_k}\ D_{\mathrm{B}_k}\ F_{\mathrm{b}_k}\ F_{\mathrm{m}_k}] \tag{9}$$

式中：X_k为第k步列车的位置，km；V_k为第k步列车的速度，km/h；A_{cc_k}为第k步列车的加速度，m/s^2；T_{B_k}为第k步列车空气制动时长，s；T_{R_k}为第k步列车的缓解时长，s；D_{B_k}为第k步列车在空气制动状态下行进的距离，m；F_{b_k}为第k步列车空气制动力的稳态值，kN；F_{m_k}为本车空气制动的标准制动力稳态值，kN。

列车在切换到空气制动状态时，使用上一状态中的制动信息$\left[T_{\mathrm{B}_k-1}\ T_{\mathrm{R}_k-1}\ D_{\mathrm{B}_k-1}\ F_{\mathrm{b}_k-1}\ F_{\mathrm{m}_k-1}\right]$作为空气制动力预测模型的输入参数，得到本次空气制动力预测值F_{b_k}，同时更新空气制动时长T_{B_k}和空气制动距离D_{B_k}为0，并重新计数。列车从空气制动状态结束而开始缓解时，更新缓解时长T_{R_k}为0，并重新计数。

⑤奖励：奖励是强化学习系统的学习目标。在重载组合列车操纵问题中，学习目标主要为准点与安全平稳。重载组合列车的安全平稳性通常使用车钩力来衡量，但由于车钩力的计算复杂，若将其加入到单步奖励中去会极大地增加训练时长。因此，在本文中重载组合列车的平稳性以手柄切换幅度来限制，安全性则由列车的最高限速与最低限速来限制，数学表达式如下所示：

$$\left.\begin{aligned}&R_k=r+r_{\mathrm{v}}\\&r=\begin{cases}r, & v>v_{\min}\&v<v_{\max}\\0, & v>v_{\max}|v<v_{\min}\end{cases}\\&r_{\mathrm{v}}=\begin{cases}r_{\mathrm{v}}\sqrt{\left[1+\mathrm{floor}\left(\mathrm{abs}(v-v^*)/\Delta v\right)\right]}, & v>v_{\min}\&v<v_{\max}\\0, & v>v_{\max}|v<v_{\min}\end{cases}\end{aligned}\right\} \tag{10}$$

式中：R_k为奖励函数；r为单步基础奖励；r_{v}为速度目标奖励；v为实车速度；v^*为目标速度；Δv为速度变化量。

⑥策略：策略决定了智能体在某一状态下选择某

种动作的概率。本文使用变步长的贪心策略，其探索率随着迭代次数的增加而减小，其数学表达为

$$\pi(a'|s)=\begin{cases}1-\varepsilon_0 g^k, & a'=a^* \\ \varepsilon_0 g^k, & a'=\mathrm{rand}\left(A(s)\right)\end{cases} \tag{11}$$

式中：ε_0为随机探索率初值，$\varepsilon_0\in[0,1]$；g为探索率衰减率，$g\in[0,1]$；a'为某一状态下所对应的级位；a^*为最优级位；$A(s)$为与状态相关的动作空间。

在学习初期，智能体应以较大的探索率对环境进行探索以尽快建立粗略的知识体系，其后再以较小的探索率进行优化学习。

3.2 重载组合列车操纵优化算法设计

考虑到本文设计的状态空间维度较大，经验池存储过去的探索经验数据，用来训练神经网络以逼近其q表，能够较好地处理“状态–动作空间”维度较大的情况，并且无须对状态量进行离散处理。算法结构如图6所示。

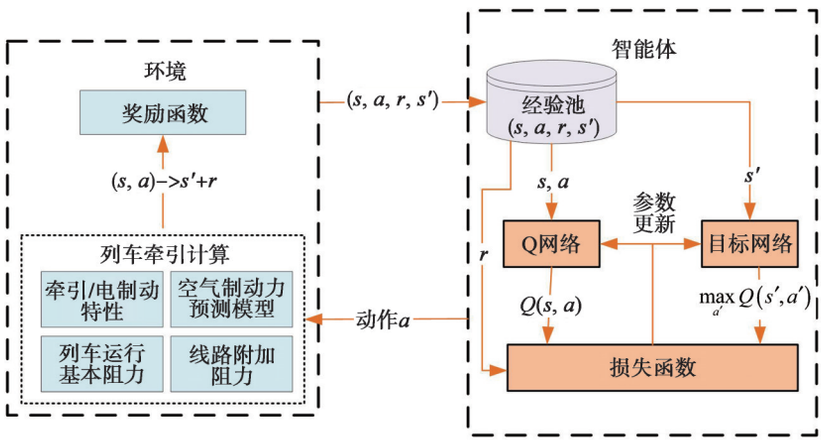


图6 重载列车操纵优化算法结构图

Fig. 6 Structure diagram of operation optimization algorithm for heavy-haul train

算法流程如下：

Step1：超参数设置

环境参数：列车性能参数、运行线路参数、运行目标速度v_t、运行区间$[X_{start},X_{end}]$、运行起始状态S_0、仿真步长Δx、动作步长ΔX等；

智能体参数：Q网络超参数、目标值折扣率γ、探索率初始值ε_0、探索率衰减率g、探索率最小值ε_{min}、经验池参数等；

训练参数：最大迭代次数M、回合最大步数N、终止条件、目标网络更新频率f、Q网络学习输入容量b等。

Step2：初始化Q网络、目标网络、经验池；

Step3：

for (i=1; i <= M; i++)

选择初始化状态 ；

for (k= 1; k <= N; k++)

根据贪心策略，选择动作a；

环境执行动作a；状态从s→s′，并计算奖励r；

环境返回状态转移矩阵[s a r s′]并存入经验池；

从经验池中随机选取b组样本，训练Q网络并更新目标网络；

if Done:

break;

end

end

if rewards > stop Value

break;

end

end

本节给出了基于Q学习的重载列车优化算法，该算法无须任何先验知识，以数据驱动的方法进行学习，给出完整的驾驶策略及对应的速度曲线。

4 仿真试验验证

为验证本文所提出的操纵优化算法，求解长大下坡操纵问题的正确性和有效性，本节在MATLAB环境下，采用真实车辆编组、实际运行线路开展仿真试验，并与实际运行数据进行对比。

以我国某Ⅰ级线路、双线电气化重载铁路为例，其重车方向由西向东，线路长约408 km，坡道条件如图7所示。

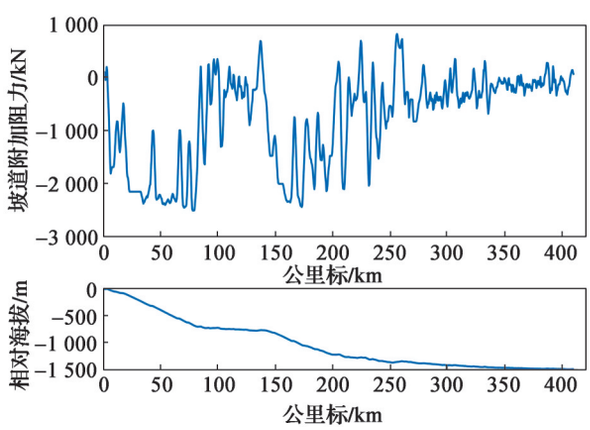


图7 仿真线路条件

Fig. 7 Track conditions for simulation

仿真区段为该线路中的第一个长大下坡区段，公里标为9~90 km。

选取的列车对象为2万t组合列车，列车编组为“2辆HXD1型电力机车+216辆C80货车”。HXD1型电力机车具备牵引和再生制动能力，整列车配备空气制动，具体的列车参数如表5所示，优化算法所使用的仿真参数如表6所示。

表5　列车参数

Table 5　Parameters of trains

参数名称	参数值	参数名称	参数值
HXD1/辆	2	最大牵引力/kN	760
C80/辆	216	最大电制动力/kN	461
总质量/t	21 600	列车长度/m	2 921.6

表6　优化算法参数

Table 6　Parameters of optimization algorithm

参数名称	参数值	参数名称	参数值
机车级位最大值	12	经验池大小	30 000
机车级位最小值	−12	最低速度/(km·h^{-1})	30
探索率初始值	1	探索率衰减率	0.05
探索率最小值	0.05	仿真步长/km	0.05
级位变化最大值	3	动作步长/km	0.25
Q网络学习率	0.01	目标速度/(km·h^{-1})	70
目标网络更新频率	4	最高速度/(km·h^{-1})	80
折扣系数	0.92	最小学习样本数	1 024

目标奖励函数具体参数如表7所示。

表7　奖励参数

Table 7　Parameters of reward

参数名称	r	r_v	Δv/(km·h^{-1})
参数值	50	20	5

采用本文所提出的空气制动力预测模型，考虑列车空气制动系统制动性能的变化性，建立重载组合列车长大下坡运行仿真环境，设置迭代次数等目标参数，使用本文提出的重载组合列车操纵优化算法进行学习，得到列车速度曲线与操纵工况，并与实车数据进行对比（该车首次空气制动力为2 383 kN）。仿真结果如图8所示，图8包括3个子图，其中子图1包含了线路限速、实车速度曲线、仿真速度曲线、实车管压曲线、仿真管压曲线、仿真级位曲线；子图2包含了实车电制动力曲线、仿真使用电制动力曲线、仿真空气制动力曲线；子图3包含了线路坡道附加阻力曲线。仿真结果中空气制动部分数据和优化算法所使用的空气制动数据如表8所示。

对比图8子图1中的实车管压与优化管压曲线，可以看出优化算法所给出的驾驶策略较司机操纵少了1次空气制动操作；对比图8子图2中的实车电制动力与优化电制动力曲线，可以看出相较于司机操纵而言，优化算法所使用的电制动力变化更为平缓。

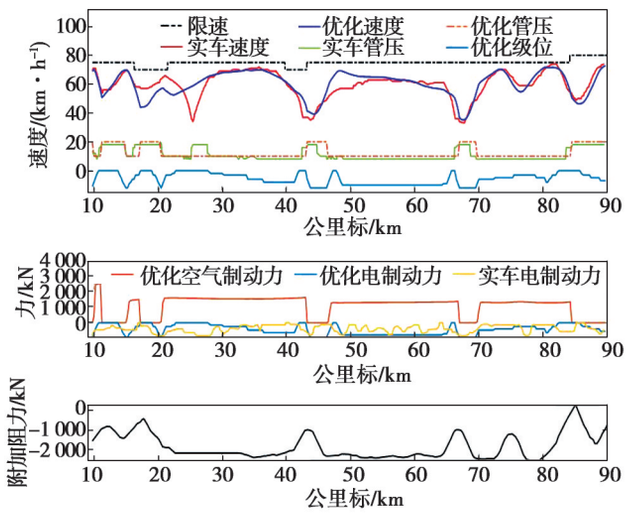


图8　优化算法仿真结果对比

Fig. 8　Comparison of simulation results of optimization algorithm

表8　仿真空气制动数据

Table 8　Air braking data of simulation

参数值	制动段序号				
	1	2	3	4	5
T_B/s	0	69	131	1 282	1 170
T_R/s	0	1 250	2 250	22 750	20 500
D_B/m		227	252	270	248
F_b/kN		2 383	1 382	1 489	1 253
F_b'/kN	2 383	1 382	1 489	1 253	1 260

仿真结果中优化算法所使用的车钩力和实车数据对比结果如图9所示。

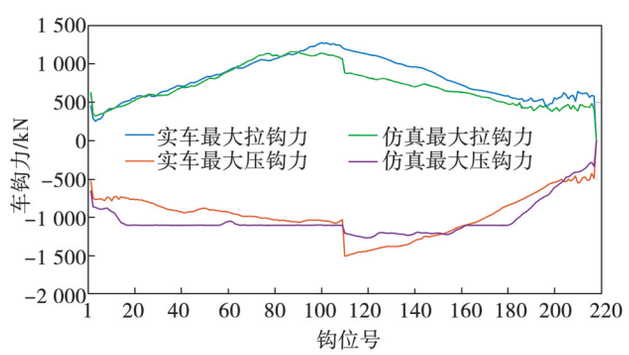


图9　优化算法车钩力仿真结果对比

Fig. 9　Comparison of coupler force simulation results of optimization algorithm

由图9可知，优化算法中车钩力变化趋势与实车数据相同，同时运行过程中最大压钩力和最大拉钩力都小于司机操纵的结果。

表9为仿真结果对比。由表9可知，优化算法所得到的计算结果相较于实车曲线，运行时间减少63 s，最大拉钩力减小8.93%，最大压钩力减小16.00%。

通过对上述仿真结果分析，可以得出以下几点结论：①引入考虑历史制动因素的空气制动力预测模型，使用重载列车操纵优化算法学习重载列车在长大下坡

区段的操纵策略，能够获得与司机操纵相类似的列车运行速度曲线。②引入操纵级位变化率的限制，以及运行目标速度跟随，求解获得的操纵策略相较于司机驾驶电制动力变化更加平稳，能够有效地减小列车运行中的最大车钩力。

表9 仿真结果对比

Table 9 Comparison of simulation results

类型	对比项目		
	运行时间/s	最大拉钩力/kN	最大压钩力/kN
实车曲线	4 962	1 276	1 500
优化算法曲线	4 899	1 162	1 260
优化率/%	1.27	8.93	16.00

5 结论

为解决重载组合列车在长大下坡区段循环制动操纵困难的问题，本文从空气制动性能预测与操纵序列求解这2个方面开展研究，结合真车实线数据进行了理论模型与优化算法的正确性和有效性分析。通过分析，可以得到以下结论：①针对重载组合列车在运行过程中受历史空气制动影响，空气制动性能估算难的问题，采用神经网络学习大量实车空气制动数据，获得基于神经网络的空气制动力预测模型。分析实车数据表明，构建的模型显著提升了空气制动动态过程描述的准确性。② 将重载组合列车的操纵问题转为典型的强化学习问题，设计基于强化学习的重载列车操纵优化算法。结果表明，基于强化学习的重载列车操纵优化算法能够考虑这种空气制动性能的变化性，相较于司机驾驶电制动力变化更加平稳，能够有效地减小列车运行中的最大车钩力。

参考文献：

[1] 蒋立干. 不同操作模式下重载列车长大下坡道曲线地段行车性能分析[D]. 北京: 北京交通大学, 2017.
JIANG Ligan. Analysis on the operation performance of heavy haul trains under different operating modes[D]. Beijing: Beijing Jiaotong University, 2017.

[2] 魏玉晓, 李宵寅. 铁路长大下坡道列车周期制动参数的分析及应用[J]. 铁道运输与经济, 2017, 39(10): 67-72.
WEI Yuxiao, LI Xiaoyin. Analysis and application of train cycle braking parameters on the long steep descending grade [J]. Railway transport and economy, 2017, 39(10): 67-72.

[3] KHMELNITSKY E. On an optimal control problem of train operation[J]. IEEE transactions on automatic control, 2000, 45(7): 1257-1266.

[4] LIU Rongfang, GOLOVITCHER I M. Energy-efficient operation of rail vehicles[J]. Transportation research part A: policy and practice, 2003, 37(10): 917-932.

[5] BAI Baoxue, XIAO Zhuang, WANG Qingyuan, et al. Multi-objective trajectory optimization for freight trains based on quadratic programming[J]. Transportation research record, 2020, 2674(11): 466-477.

[6] 王悉. 基于机器学习的重载列车智能驾驶方法研究[D]. 北京: 北京交通大学, 2017.
WANG Xi. Machine learning based intelligent operation methods for heavy haul train[D]. Beijing: Beijing Jiaotong University, 2017.

[7] 孙中央, 王长明. 长大下坡道区段开行重载列车的安全问题[J]. 铁道机车车辆, 2009, 29(1): 4-7.
SUN Zhongyang, WANG Changming. Security problems of operation heavy haul trains on the long heavy down grade [J]. Railway locomotive & car, 2009, 29(1): 4-7.

[8] 王自力. 列车节能运行优化操纵的研究[J]. 西南交通大学学报, 1994, 29(3): 275-280.
WANG Zili. An optimal operating method of train running for saving energy[J]. Journal of southwest Jiaotong university, 1994, 29(3): 275-280.

[9] 陈清, 王自力. 机车动力制动对操纵方式及限制坡度的影响[J]. 西南交通大学学报, 2005, 40(4): 448-452.
CHEN Qing, WANG Zili. Influences of locomotive dynamic brake on brake operation and limiting grade[J]. Journal of southwest Jiaotong university, 2005, 40(4): 448-452.

[10] 耿志修, 李学峰, 张波. 大秦线重载列车运行仿真计算研究[J]. 中国铁道科学, 2008, 29(2): 88-93.
GENG Zhixiu, LI Xuefeng, ZHANG Bo. Simulation study of heavy haul train operation on Datong—Qinhuangdao railway[J]. China railway science, 2008, 29(2): 88-93.

[11] 张波. 重载组合列车牵引及制动系统的试验与仿真研究[D]. 北京: 中国铁道科学研究院, 2009.
ZHANG Bo. Test and simulation study on traction and brake system of combined heavy haul train[D]. Beijing: China Academy of Railway Sciences, 2009.

[12] 魏伟, 王强. 坡道上重载列车纵向冲动研究[J]. 振动与冲击, 2014, 33(5): 143-148.
WEI Wei, WANG Qiang. Influence of train brake on longitudinal impulse of a heavy haul train passing through a ramp [J]. Journal of vibration and shock, 2014, 33(5): 143-148.

[13] 刘嘉, 杨相健, 罗世辉, 等. 朔黄铁路重载列车线路适应性仿真研究[J]. 振动、测试与诊断, 2020, 40(4): 673-679.
LIU Jia, YANG Xiangjian, LUO Shihui, et al. Heavy-haul-train line adaptability simulation of Shuohuang railway[J]. Journal of vibration, measurement and diagnosis, 2020, 40 (4): 673-679.

[14] 温军刚. 朔黄铁路两万吨重载列车平稳操纵优化方案[J]. 中国高新科技, 2019(6): 58-60.
WEN Jungang. An optimal scheme of smooth operation of the 20, 000-ton heavy haul train of Shuohuang railway[J]. China high-tech, 2019(6): 58-60.

[15] 魏伟, 张益铭. 2万吨重载组合列车操纵优化研究[J]. 铁道机车车辆, 2021, 41(4): 35-40.
WEI Wei, ZHANG Yiming. Operating optimization study of 20,000-ton heavy haul combined train[J]. Railway locomotive & car, 2021, 41(4): 35-40.

收稿日期：2022-09-09
修回日期：2023-10-26