



基于多元线性回归的动车组蓄电池性能预警方法研究

栗木功¹, 燕春光¹, 孙海荣², 蒋洁¹, 吴韵逸³, 郭展玮³
(1. 中车唐山机车车辆有限公司 运修技术中心, 河北 唐山 063000;
2. 成都地铁运营有限公司, 四川 成都 610058;
3. 台州市轨道交通建设开发有限公司, 浙江 台州 317599)

摘要: 为提高高速动车组蓄电池系统的运行安全性并降低故障风险, 提出一种基于远程监控数据的蓄电池性能预警方法。研究综合考虑电池组温度、充放电时长、电池组电压、电池组电流及同车其他电池组电量等因素, 融合多元线性回归与孤立森林算法, 构建动态预警模型。其中, 多元线性回归用于定量分析各因素与电池组电量间的线性关系, 建立动态预测模型; 孤立森林算法则可识别数据中的异常点并捕捉电量突变。同时, 借助残差分析曲线动态监测预测误差变化, 当残差异常时触发性能预警。试验结果表明, 该方法能有效挖掘电池组电量变化规律, 预警准确率达95%, 突破了单一方法在复杂工况下的局限, 为动车组蓄电池的智能化运维提供了可靠的技术支持, 对保障铁路运输安全具有重要的实际意义。

关键词: 高速动车组; 蓄电池; 性能预警模型; 多元线性回归; 孤立森林

中图分类号: U292.91⁴

文献标志码: A

文章编号: 1000-128X(2025)04-0130-07

0 引言

动车组蓄电池是在列车没有外接电源或受电弓未受流的情况下, 为动车组的控制电路、车内照明等低压直流设备供电的一种储能设备。当动车组有外接电源或受电弓正常受流时, 由充电机为动车组蓄电池、控制电路、车内照明等设备供电, 蓄电池进入充电状态。

动车组的蓄电池亏电时, 将导致动车组电源系统故障, 无法为负载提供电源, 影响车辆的正常运行。同时, 蓄电池作为辅助供电系统的核心元件, 其工作稳定性直接关系到动车组的安全运行。对蓄电池性能进行预先评估, 避免发生故障造成不必要的损失是铁路领域关注的重点问题之一。文献[1]以新一代动车组为对象, 分析了蓄电池低电压检测的工作原理, 提出了蓄电池低电压检测的最佳调试方法; 文献[2]利用神

经网络算法对高铁列车车载蓄电池剩余使用时长进行预测研究; 文献[3]采用粒子滤波算法对锂离子电池的剩余循环使用寿命进行预测研究。上述研究都是基于实验室数据对蓄电池的寿命或性能进行预测评估, 但其依赖的部分数据在动车组远程监控系统中不易获取。文献[4]针对CRH380B动车组的远程数据提出了动车组蓄电池性能预测方法。然而, 受到远程数据的限制, CRH380B蓄电池的性能也只能通过电池组电压来衡量, 而电压受多个因素影响, 这会对蓄电池性能产生较大影响。随着动车组不断升级改造, 现在CR400BF动车组的远程数据中已可提取电池组剩余电量, 为更准确的性能预警提供了条件。

通过动车组运用部门的调研发现, 当前对蓄电池故障或性能的监控仅能将采集的原始数据呈现给工作人员, 并对电池组电压、充电电流、电池组温度等数据的峰值进行标注, 其缺点是在运行过程中蓄电池出现突发故障时, 工作人员难以快速响应。因此, 如果

作者简介: 栗木功, 男, 高级工程师, 主要研究方向为轨道车辆制造、运维及检修技术; E-mail: limugong.ts@crrecg.cc

引用格式: 栗木功, 燕春光, 孙海荣, 等. 基于多元线性回归的动车组蓄电池性能预警方法研究[J]. 机车电传动, 2025(4): 130-136.DOI:10.13890/j.issn.1000-128X.2025.04.016.

Citation: LI Mugong, YAN Chunguang, SUN Hairong, et al. Research on performance warning for EMU batteries based on multiple linear regression [J]. Electric drive for locomotives, 2025(4): 130-136.DOI:10.13890/j.issn.1000-128X.2025.04.016.

能基于蓄电池的多变量历史数据建立综合预测模型, 提前评估蓄电池的性能, 可以防止由于蓄电池性能恶化造成更为严重的故障及经济损失^[5-8]。本文以CR400BF动车组远程数据为基础, 综合考虑电池组温度、电池组电压、充电电流、电池组充电/放电时长、同车其他电池组电量等因素, 利用多元线性回归方法对充/放电模式下电池组剩余电量进行预测, 并采用统计方法实现评估预警功能。另外, 利用孤立森林算法对浮充模式下蓄电池性能进行异常检测。

1 动车组蓄电池概述

CR400BF动车组的8节编组配备4组蓄电池, 16节编组配备8组蓄电池, 每组蓄电池均由多节单体蓄电池通过串并联方式构成。在供电电源切断时, 蓄电池可通过内置控制机制, 将供电时间和最低电压控制在合理范围, 确保在电压极限内为车上负载提供紧急供电, 保障动车组在断电情况下的运行安全。

1.1 单体电池

动车组单体电池由正极、负极、隔膜、电解液、电池壳等构成, 通过将电能转变为化学能实现电能存储(充电), 再通过化学能转变为电能实现电能释放(放电)。单体电池类型为钛酸锂电池, 它具有寿命长、安全性高、充放电快速等优点。

1.2 蓄电池组

CR400BF动车组的蓄电池组为额定110 V、400 Ah的电池包, 由450节单体钛酸锂电池组成。具体结构为: 45节单体钛酸锂电池串联成一个小电池组, 10个小电池组并联组成一组蓄电池组。

1.3 动车组蓄电池工作模式

在系统故障场景下, 例如接触网电压缺失、牵引变压器或变流器故障、蓄电池充电机故障等情况时, 动车组蓄电池需向重要负载持续供电。针对CR400BF动车组蓄电池的实际使用工况及充电特性, 开展其电池组电压与电量变化规律研究发现: 当动车组充电回路断开时, 蓄电池进入放电状态, 电压持续下降至预设固定值; 在充电过程中, 蓄电池先以恒流模式充电至均充电压, 随后保持均充电压进入恒压充电阶段, 此时充电电流逐渐减小; 当充电电流降至设定阈值时, 充电电压切换为浮充电压, 系统进入浮充状态。基于上述特征, 可将蓄电池工作模式划分为放电、充电、浮充3种, 以下将对它们进行具体分析。

1) 放电模式

放电模式是指高压充电断开, 蓄电池持续为关键设备供电维持动车组正常运行, 电压由最大值逐步变

小至最小值, 放电电流表现为充电电流负值(因远程监控仅传输充电电流, 故充电电流为负时对应放电)。所以, 当充电电流小于0 A时为放电模式。

2) 恒流-恒压充电模式

恒流充电是指电池组剩余电量不足时, 对其进行持续高压充电, 此时充电电流保持在96 A, 电池组电压持续升高, 直至达到恒定值114 V; 恒压充电是指恒流充电后, 电池组电压会持续保持在114 V并继续充电, 充电电流持续减少直到8.5 A。所以, 满足充电电流大于8.5 A条件时为恒流-恒压充电模式。

3) 浮充模式

浮充是指在恒流-恒压充电模式后, 蓄电池以浮充电压保持充电, 此时充电电流在8.5 A以下。所以, 当充电电流小于8.5 A且大于0 A时, 为浮充模式。

根据动车组蓄电池工作原理, 分别研究蓄电池在放电、恒流-恒压充电、浮充3种模式下的电量变化规律, 并建立相应的蓄电池性能评估方法。

2 蓄电池性能预警方法

传统方法对蓄电池性能的预警是通过规则进行的, 这些规则来源于模拟试验或者实际经验总结, 在动车组运行过程中, 影响因素众多, 造成规则预警准确性下降。随着大数据和人工智能算法在工业应用的加深, 越来越多的规则预警被算法模型替代。本文针对蓄电池远程历史数据, 利用多元线性回归算法拟合出不同状态下蓄电池的性能评估标准, 以提升预警准确性。

2.1 蓄电池远程监控数据采集与处理

为实现对动车组蓄电池性能预警, 首先需要获取蓄电池远程监控数据。通过动车WTDS系统获取蓄电池的充电电流、电池组电压、电池组温度、电池组剩余电量等监控数据, 数据传输频率为每30秒一次。由于蓄电池监控数据使用不同的传感器进行采集, 传输时间会有差异, 整理的原始数据会产生缺失值和异常值, 需要对原始数据做预处理, 步骤如下:

①剔除异常值。在原始数据电池组电压、充电电流、电池组温度、电池组剩余电量出现明显不符合逻辑的异常值时, 例如电压会出现3 685 V, 这种由于数据传输或者解析转换造成的脏数据, 按照主机厂提供的蓄电池参数说明, 并且根据各变量历史数据统计分析峰值, 制定以下剔除标准: a. 蓄电池温度高于100 °C或蓄电池温度低于-20 °C; b. 电池组电压大于130 V[额定电压为(115±1) V]; c. 充电电流大于150 A或充电电流小于-150 A[电池组最大充电电流为(95±2) A]。

②数据缺失处理。远程监控数据的传输偶尔存在丢包和延迟的情况, 造成获取的数据存在部分缺失值, 考虑到充电电流、电池组电压、电池组温度、电池组

剩余电量等变量不是连续变化,不宜采用插值法进行填充。本文采用邻近填充法,对前后无值的缺失数据进行删除处理。

③计算电池组平均温度。远程监控数据提供电池组最高温度及最低温度,模型研究的是电池组的整体状况,取电池组最高温度和最低温度的平均值作为电池组温度指标。

④计算同车其他电池组平均电量。通过对CR400BF电池组故障数据的研究表明,如果单一电池组出现故障,该电池组的电流、电压和剩余电量会与其他电池组存在显著的差异。因此,在研究单一电池组性能时,应考虑加入同车其他电池组平均电量,作为重要参考因素。

2.2 放电模式性能预警模型

针对放电模式下WTDS蓄电池数据,建立蓄电池数据分析预警模型,采用多元线性回归拟合电池组剩余电量变化的标准曲线,通过电池组剩余电量变化实际曲线与标准曲线的对比,找出蓄电池的异常点,准确识别蓄电池放电性能情况。

在放电模式下,电池组剩余电量受电池组电压、充电电流、电池组温度、放电时长及同车其他电池组平均电量等多因素影响而波动,以这些影响因素为自变量,以电池组剩余电量为目标变量,通过多元线性回归拟合构建剩余电量变化标准曲线,再对比标准曲线与实际曲线以评估蓄电池性能。

1) 设定回归方程

设电池组剩余电量的多元线性回归模型为

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \quad (1)$$

式中: y 为电池组剩余电量; x_1 为电池组电压; x_2 为充电电流; x_3 为电池组平均温度; x_4 为放电时长; x_5 为同车其他电池组平均电量; β_0 为常数项; β_1 、 β_2 、 β_3 、 β_4 、 β_5 分别为 x_1 、 x_2 、 x_3 、 x_4 、 x_5 的回归系数。

2) 回归参数估计

回归方程建立后,需要根据样本对回归参数进行估计,得到回归参数的估计值 $\hat{\beta}_0$ 、 $\hat{\beta}_1$ 、 $\hat{\beta}_2$ 、 $\hat{\beta}_3$ 、 $\hat{\beta}_4$ 、 $\hat{\beta}_5$ 。电池组剩余电量的预测值 $\hat{y}_i$ 与实际值 y_i 的残差平方和 Q 为

$$Q = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i} - \hat{\beta}_4 x_{4i} - \hat{\beta}_5 x_{5i})^2 \quad (2)$$

采用最小二乘法对模型参数进行估计,使残差平方和 Q 达到最小。通过微积分方法,得到 Q 最小的必要条件为

$$\frac{\partial Q}{\partial \hat{\beta}_0} = 0, \frac{\partial Q}{\partial \hat{\beta}_1} = 0, \frac{\partial Q}{\partial \hat{\beta}_2} = 0, \frac{\partial Q}{\partial \hat{\beta}_3} = 0, \frac{\partial Q}{\partial \hat{\beta}_4} = 0,$$

$$\frac{\partial Q}{\partial \hat{\beta}_5} = 0$$

由此可以得到回归系数的估计值 $\hat{\beta}_0$ 、 $\hat{\beta}_1$ 、

$$\hat{\beta}_2$$
、 $\hat{\beta}_3$ 、 $\hat{\beta}_4$ 、 $\hat{\beta}_5$ 。

3) 回归模型检验

为了确定模型的精度,需要对模型进行检验。回归模型的检验主要包括对模型的拟合优度检验、回归模型显著性检验及回归系数显著性检验。拟合优度检验采用 R^2 统计量,检验样本数据点聚集在回归线周围的密集程度,从而评价回归方程对样本数据的代表程度。回归模型的显著性检验通过 F 检验,判断被解释变量与所有解释变量间的线性关系是否显著;回归模型的系数显著性检验通过 t 检验,判断被解释变量与单个解释变量间的线性关系是否显著。

4) 放电模式蓄电池性能预警

针对蓄电池性能异常的预警,实际是寻找与标准状态不相符的数据,并对数据做出预警。对放电模式下蓄电池实际电量与拟合标准曲线的残差进行统计分析,经正态性检验,残差符合正态分布特征。基于统计学 3σ 准则,以残差均值为中心、三倍标准差为阈值构建异常值判定标准:残差超出该阈值范围的数据点,判定为异常点。

2.3 恒流-恒压充电模式性能预警模型

在恒流-恒压充电模式下,电池组剩余电量受电池组电压、平均温度、充电电流、充电时长及同车其他电池组平均电量等多因素影响而波动,以这些影响因素为自变量,电池组剩余电量为目标变量,通过多元线性回归构建剩余电量变化标准曲线,再对比标准曲线与实际电量数据,实现对蓄电池性能的评估。

恒流-恒压充电模式下蓄电池性能预警模型的建立,与放电模式下的过程一致,此处不再赘述。

2.4 浮充模式性能预警模型

在浮充模式下,电池组的剩余电量、电压会在一个相对稳定的范围内波动。此时,基于回归的残差分析方法敏感性不足。本文采用孤立森林算法对浮充模式下的蓄电池进行异常检测。

孤立森林算法通过寻找超平面将数据空间逐步切割为2个子空间,然后继续对子空间进行切割,直到每个子空间仅包含一个数据节点,形成孤立树。每个叶子节点都包含一个数据节点。由于离群点的数据密度很低,它们往往能够迅速被孤立到某个子空间中。孤立森林算法计算过程如下:

①训练阶段:搭建孤立森林,目标是造一批孤立树(共 t 棵),每棵树的作用是用来尝试孤立样本。

第一步:选小样本。每棵树只从全量数据里随机抽取 $\mathcal{V}$ 个样本——不用全量,小样本足够反映分布。

第二步:随机切分建树干。从根节点开始,每次随机选一个特征、随机定一个阈值,把样本分成小于阈值(左分支)和大于等于阈值(右分支)两部分;

重复切分,直到分支里只剩1个样本,或树长到设定最大深度,最后形成一棵孤立树。

第三步:复制成森林。重复前面第一步、第二步,共构建*t*棵树,组成孤立森林。

②推理阶段:用森林判定异常(实际检测)。对新样本,用森林算异常分,分高就是异常。

第一步:算“孤立度”。把样本放进每棵树,计算它从根节点到最终叶子节点要走多少步——异常点走的步数少(容易被孤立),正常点走的步数就多。

第二步:求平均步数。把*t*棵树的步数取平均,反映样本整体的孤立难度。

第三步:求标准化异常分。先计算正常样本的平均步数,只和样本量*n*有关,公式是固定的;用公式 $2-(\text{待测样本平均步数}/\text{正常样本平均步数})$ 算分,得分范围是[0,1]。

第四步:判定异常。设置阈值,得分大于阈值是异常,小于等于阈值是正常。

选取电池组电压、电池组剩余电量、充电电流、电池组平均温度、同车其他电池组平均电量作为变量。

3 实例分析

本章采用CR400BF动车组蓄电池远程监控数据搭建模型,分别对浮充、充电、放电3种模式下的预警模型进行实例分析。

3.1 放电模式性能预警模型

基于2021年1月至2025年7月期间的1960895条蓄电池远程监控数据,依据放电模式判定准则(充电电流小于0A)进行数据筛选,提取放电状态下的有效数据。采用随机抽样法,将筛选后的数据以2:8的比例划分为测试集与训练集,为后续建模分析提供基础数据支撑。

3.1.1 相关性分析

首先对所选取的变量做相关性分析,利用统计软件得到放电状态变量相关系数矩阵,如表1所示。

从表1可以看出,电池组剩余电量与电池组电压、放电时长、同车其他电池组电量均值的相关系数分别为0.9653、-0.8539、0.8794,为高度相关;与充电电

流、电池组平均温度相关系数分别为-0.5361、0.5034,为一般相关。综上分析,可以利用充电电流、电池组电压、放电时长、电池组平均温度、同车其他电池组电量均值对电池组剩余电量进行预测。

表1 放电状态变量相关系数

Table 1 Correlation coefficients of SOD variables			
指标名称	对比指标	相关系数	显著性
电池组剩余电量	充电电流	-0.5361	0.0000
电池组剩余电量	电池组电压	0.9653	0.0000
电池组剩余电量	放电时长	-0.8539	0.0000
电池组剩余电量	电池组平均温度	0.5034	0.0000
电池组剩余电量	同车其他电池组电量均值	0.8794	0.0000

3.1.2 线性回归模型

1) 线性回归方程

通过线性相关分析,以充电电流、电池组电压、放电时长、电池组平均温度、同车其他电池组电量均值为自变量,以电池组剩余电量为因变量,建立线性回归方程如下:

电池组剩余电量 = -184.98 + 0.00012 × 充电电流 + 2.2187 × 电池组电压 - 0.0361 × 放电时长 - 0.0334 × 电池组平均温度 + 1.5721 × 同车其他电池组电量均值

2) 结果分析

为验证放电状态回归模型的可靠性,对模型进行拟合优度检验与方差分析,结果如表2(模型拟合优度指标)和表3(方差分析表)所示。从表2可以看出,放电状态回归模型的拟合优度*R*²为0.9841,修正的*R*²值为0.9822,说明这个模型的拟合优度已达到较高水平;从表3方差分析表可以看出,回归平方和为76092.4821,回归自由度为5,*F*值为5932.8910,显著性*P*值为0.0000,小于0.05,表明充电电流、电池组电压、放电时长、电池组平均温度、同车其他电池组电量均值对电池组剩余电量有显著的线性影响。

表2 放电状态模型拟合优度

Table 2 Goodness-of-fit of SOD model						
参数项目	<i>R</i> 平方	修正的 <i>R</i> 平方	估计的 标准误	对数 似然值	赤池信 息量	施瓦兹 准则
参数值	0.9841	0.9822	1.1489	-3167.2098	3.1098	3.7182

表3 放电状态模型方差分析表

Table 3 Analysis of variance for SOD model					
类型	平方和	自由度	均方	<i>F</i> 值	显著性
回归	76092.4821	5	15218.4964	106796.466	0.0000
残差	5932.8910	41636	0.1425		
总和	82025.3731	41641			

3.1.3 性能预警规则

以回归模型得出的剩余电量的绝对误差为分析序列,依据3σ准则(剩余电量阈值=均值±3×标准差)得

出剩余电量残差的阈值为±3.2639,可以实时计算电池剩余电量的实际值与拟合值的差值,并比较差值与阈值的大小来进行预警。

3.1.4 模型验证

为验证模型可靠性，从原始数据中随机抽取 20% 的数据作为测试集。利用训练的模型得出电池组剩余电量预测值，部分统计相对误差如表 4 所示。

表 4 放电状态模型验证结果表

Table 4 Verification results for SOD model

车型	车号	日期	剩余电量实际值/%	剩余电量预测值/%	差值	是否故障
CR400BF	3001	2022-09-23	91.82	97.25	5.43	是
CR400BF	3001	2022-09-23	91.39	95.53	4.14	是
CR400BF	3001	2022-10-07	93.48	90.08	3.40	否
CR400BF	3001	2023-05-15	85.84	89.85	4.01	否
CR400BF	3022	2022-08-12	87.24	83.63	3.61	否
CR400BF	3022	2024-02-09	95.72	92.44	3.28	否
CR400BF	3022	2024-05-11	92.44	96.45	4.01	是
CR400BF	3022	2024-05-12	91.75	95.52	3.77	是
CR400BF	3022	2024-12-06	94.82	98.55	3.73	否

从表 4 可以看出，当模型输出的预测值与实际监测值差值较大时，蓄电池会出现故障。针对此类故障场景，模型的测试结果具有“无漏报”的优势——能够全面捕捉到蓄电池故障的信号，不过在检测过程中仍存在一定的误报情况。为验证模型在实际场景中的表现，研究团队结合现场实际维修反馈数据进行分析，该模型在蓄电池放电阶段的预警准确率达到 95.42%。

3.2 恒流-恒压充电模式性能预警模型

基于同一时期的蓄电池远程监控数据，依据充电状态判定标准（充电电流大于 8.5 A）筛选有效数据后，采用随机抽样法将样本按 2:8 比例划分为测试集与训练集。

3.2.1 相关性分析

对所选取的变量做相关性分析，利用统计软件得到充电状态变量相关系数表，如表 5 所示。

表 5 充电状态变量相关系数表

Table 5 Correlation coefficients of SOC variables

指标名称	对比指标	相关系数	P 值
电池组剩余电量	充电电流	-0.679 3	0.000 0
电池组剩余电量	电池组电压	0.928 5	0.000 0
电池组剩余电量	充电时长	-0.572 3	0.000 0
电池组剩余电量	电池组平均温度	0.517 3	0.000 0
电池组剩余电量	同车其他电池组电量均值	0.931 8	0.000 0

从表 5 可以看出，电池组剩余电量和电池组电压、同车其他电池组电量均值的相关系数分别为 0.928 5、0.931 8，为高度相关；电池组剩余电量和充电电流、电池组电压、充电时长、电池组平均温度的相关系数分别为 -0.679 3、-0.572 3、0.517 3，都为一般相关。综上分析，可以利用充电电流、电池组电压、充电时长、电池组平均温度、同车其他电池组电量均值对电

池组剩余电量进行预测。

3.2.2 多元线性回归模型

1) 回归方程

通过线性相关分析，以充电电流、电池组电压、充电时长、电池组平均温度、同车其他电池组电量均值为自变量，电池组剩余电量为目标变量，建立的多元线性回归方程如下：

电 池 组 剩 余 电 量 = -176.274 - 0.024 × 充 电 电 流 - 2.2671 × 电 池 组 电 压 - 0.0281 × 充 电 时 长 + 0.0309 × 电 池 组 平 均 温 度 + 0.8901 × 同 车 其 他 电 池 组 电 量 均 值

2) 结果分析

充电状态模型拟合优度如表 6 所示。由表 6 可知，充电状态回归模型的拟合优度 R^2 值为 0.937 1，修正的 R^2 值为 0.931 2，表明该模型拟合度较高。充电状态模型方差分析如表 7 所示。由表 7 数据可知，回归平方和为 29 082.387 2，自由度为 5， F 值为 72 796.96，显著性 P 值为 0.000 0（小于 0.05），表明充电电流、电池组电压、充电时长、电池组平均温度、同车其他电池组电量均值对电池组剩余电量有显著的线性影响。

表 6 充电状态模型拟合优度

Table 6 Goodness-of-fit of SOC model

参数项目	R 平方	修正的 R 平方	估计的标准误差	对数似然值	赤池信息量	施瓦兹准则
参数值	0.937 1	0.931 2	1.628 3	-1 319.832	3.823 2	4.631 9

表 7 充电状态模型方差分析表

Table 7 Analysis of variance for SOC model

类型	平方和	自由度	均方	F 值	显著性
回归	29 082.387 2	5	5 816.477 4	72 796.96	0.000 0
残差	3 062.349 1	38 323	0.079 9		
总和	32 144.736 3	38 328			

3.2.3 性能预警规则

以多元线性回归模型得出的剩余电量的绝对误差为分析序列, 利用 3σ 准则得出剩余电量残差的阈值为3.739 1, 可以实时计算电池剩余电量的实际值与拟合值的差值, 并比较差值与阈值的大小来进行预警。

3.2.4 模型验证

为了验证模型的可靠性, 研究团队采用随机抽样的方式, 从原始数据中抽取20%的数据用于模型测试。利用训练的模型得出电池组剩余电量预测值, 部分统计相对误差如表8所示。

表8 充电状态模型验证相对误差统计表
Table 8 Relative error statistics for SOC model verification

车型	车号	日期	剩余电量实际值/%	剩余电量预测值/%	差值	是否故障
CR400BF	3001	2022-09-23	92.11	87.56	4.55	是
CR400BF	3001	2022-09-23	90.92	85.86	5.06	是
CR400BF	3001	2022-09-30	93.89	97.97	4.08	否
CR400BF	3001	2022-09-23	96.97	92.93	4.04	否
CR400BF	3015	2023-11-12	96.97	91.96	5.01	是
CR400BF	3015	2022-11-13	94.93	98.92	3.99	是
CR400BF	3020	2022-10-20	93.88	89.98	3.90	否
CR400BF	3021	2024-05-08	90.87	85.97	4.90	否
CR400BF	3022	2025-02-06	90.85	94.96	4.11	是
CR400BF	3022	2025-02-06	92.91	96.95	4.04	是

从表8可以看出, 当模型输出的预测值与实际监测值差值较大时, 蓄电池会出现故障。针对此类故障场景, 模型的测试结果具有“无漏报”的优势——能够全面捕捉到蓄电池故障的信号, 不过在检测过程中也存在一定的误报情况。为验证模型在实际场景中的表现, 结合现场实际维修反馈数据进行分析, 该模型在蓄电池充电阶段的预警准确率达到94.52%。

3.3 浮充模式性能预警模型

从总数据中, 依据判定标准(充电电流大于0 A而小于8.5 A)提取浮充模式下的数据。利用孤立森林算法进行异常点提取, 孤立树每次采样256个数据点, 构造含有100棵孤立树的孤立森林, 对浮充模式下的数据进行检测, 异常点如表9所示。

表9 浮充模式异常点检测结果
Table 9 Outlier detection results during float charging stage

速度/(km·h ⁻¹)	电池组电压/V	蓄电池充电电流/A	蓄电池组剩余电量/%	结果
230	114.0	4.5	97.99	-1
155	114.1	4.3	97.99	-1
164	114.2	4.2	97.99	-1
189	114.5	4.3	96.99	-1

从表9可以看出, 动车组高速运行时, 蓄电池应处于满电量, 而以上结果显示电池组剩余电量显著偏离100%, 经实际检验为异常点, 测试的准确率达95.68%。

4 结束语

针对动车组蓄电池浮充、恒流-恒压充电、放电3

种典型工作模式, 分别建立了蓄电池性能预警模型。浮充模式下利用孤立森林算法对蓄电池剩余电量进行异常点检测, 恒流-恒压充电和放电模式下采用远程监控数据, 综合考虑电池组电压、电流、温度等变量, 并参考同车其他电池组电量均值, 利用多元线性回归拟合电池组剩余电量的标准曲线, 通过运行数据与标准曲线比对实现蓄电池性能监测预警。通过实证分析, 电池组剩余电量与充电电流、电池组电压、电池组平均温度及同车其他电池组电量均值具有较高相关性, 利用多元线性回归标准曲线进行预警测试的准确率在95%左右。利用此预警方法, 可以提前发现蓄电池性能问题并及时进行处置, 避免动车组因蓄电池性能故障造成不必要的运营损失。

参考文献:

[1] 崔瑛, 栗木功, 崔虎山, 等. 新一代高速动车组蓄电池低电压检测原理分析及调试方法[J]. 大连交通大学学报, 2012, 33(6): 20-22.
CUI Ying, LI Mugong, CUI Hushan, et al. Principle analysis and commissioning solutions of low battery voltage detection in a new high-speed EMU[J]. Journal of Dalian jiaotong university, 2012, 33(6): 20-22.

[2] 郑鹏元, 陈晓龙, 黄绪民, 等. 高铁车载蓄电池剩余寿命预测研究[J]. 科学与技术, 2019(24): 1-8.
ZHENG Pengyuan, CHEN Xiaolong, HUANG Xumin, et al. Study on the prediction of the remaining life of high-speed railway vehicle batteries[J]. Science and technology, 2019(24): 1-8.

[3] 林慧龙, 李赛. 基于粒子滤波的锂离子电池剩余使用寿命预测[J]. 科学技术与工程, 2017, 17(29): 296-301.
LIN Huilong, LI Sai. Remaining useful life prediction of the

- lithium-ion batteries based on particle filter algorithm[J]. Science technology and engineering, 2017, 17(29): 296-301.
- [4] 徐轶垚. 基于远程数据的CRH380B(L)型动车组蓄电池性能预测研究[J]. 上海铁道科技, 2017(3): 106-108.
- XU Yiyao. Study on performance prediction of storage batteries for CRH380B(L) EMUs based on remote data[J]. Shanghai railway science & technology, 2017(3): 106-108.
- [5] 徐广伟, 刘建城, 陈东, 等. 动车组蓄电池供电的应急牵引和应急空调制冷系统[J]. 机车电传动, 2019(5): 116-120.
- XU Guangwei, LIU Jiancheng, CHEN Dong, et al. Emergency traction and emergency air conditioning refrigeration system for battery power supply of EMUs[J]. Electric drive for locomotives, 2019(5): 116-120.
- [6] 杨斌, 罗鑫捷, 曾旺, 等. 城市轨道交通行业电池的使用与维护[J]. 电池, 2023, 53(2): 196-198.
- YANG Bin, LUO Angjie, ZENG Wang, et al. Use and maintenance of batteries in urban rail transit industry[J]. Battery, 2023, 53 (2): 196-198.
- [7] 刘良俊, 高一钊, 朱景哲, 等. 数据驱动的锂离子电池健康状态估计[J]. 电池, 2022, 52 (2): 157-161.
- LIU Liangjun, GAO Yizhao, ZHU Jingzhe, et al. Data-driven SOH estimation of Li-ion battery[J]. Battery, 2022, 52(2): 157-161.
- [8] 李辉, 赵滨, 吴健, 等. 有轨电车的电池/超级电容器能量管理[J]. 电池, 2022, 52(1): 48-52.
- LI Hui, ZHAO Bin, WU Jian, et al. Energy management of battery/supercapacitor for trams[J]. Battery, 2022, 52(1): 48-52.

Research on performance warning for EMU batteries based on multiple linear regression

LI Mugong¹, YAN Chunguang¹, SUN Hairong², JIANG Jie¹, WU Yunyi³, GUO Zhanwei³

(1. Operation and Maintenance Technical Center, CRRC Tangshan Co., Ltd., Tangshan, Hebei 063000, China;

2. Chengdu Metro Operation Co., Ltd., Chengdu, Sichuan 610058, China;

3. Taizhou Rail Transit Construction and Development Co., Ltd., Taizhou, Zhejiang 317599, China)

Abstract: The operational status of the battery system in high-speed EMUs is crucial to train safety. To mitigate failure risks, this study presents a warning method for battery performance based on remote monitoring data. This method employs a dynamic warning model that comprehensively considers factors such as battery pack temperature, charging and discharging duration, voltage, and current, as well as the state-of-charge (SOC) of other battery packs on the same train. It integrates multiple linear regression and the isolation forest algorithm. The former quantitatively analyzes the linear relationships between various factors and the battery pack SOC for dynamic prediction, while the latter identifies outliers to capture sudden SOC changes. Meanwhile, a residual analysis curve is introduced to dynamically monitor fluctuations in prediction errors, and trigger performance warnings whenever anomalies in residuals are detected. Experimental results show that the proposed method effectively captures SOC variation patterns in battery packs, achieving a warning accuracy rate of up to 95%. This method overcomes the limitations of single-method approaches under complex conditions, and provides reliable technical support for the intelligent operation and maintenance of EMU batteries, highlighting its significant value in ensuring railway transportation safety.

Keywords: high-speed EMU; battery; performances warning model; multiple linear regression; isolation forest algorithm