

文章编号:1005-0523(2026)01-0101-13



基于WPD-GA-BP的电动汽车动力电池健康 状态预测方法

黄亮亮^{1,2}, 郭拯诗¹, 何峰¹, 胡鑫^{1,3}, 李毓书¹, 段兴兵^{1,3}, 曾建邦^{1,3}

(1. 华东交通大学机电与车辆工程学院, 江西南昌 330013; 2. 南昌市消防救援支队, 江西南昌 330038;
3. 江西江铃集团新能源汽车有限公司, 江西南昌 330114)

摘要:动力电池健康状态(SOH)的精准预测对于延长电动汽车使用寿命和保障行车安全至关重要。针对BP神经网络存在的特征提取能力有限、对初始参数敏感以及易陷入局部最优等问题,基于某车企监控平台数据,提出了一种基于WPD-GA-BP的电动汽车动力电池SOH预测方法。首先,基于容量增量分析法提取平台数据特征参数,通过皮尔逊相关系数筛选出与SOH显著相关的特征作为模型输入;其次,为丰富特征参数维度,采用小波包分解对标签值进行多尺度重构;最后,采用遗传算法优化BP神经网络的初始权重和阈值,利用更广泛的搜索空间进行全局优化,有效避免局部最优,从而实现对动力电池SOH的精准预测。结果表明:WPD-GA-BP模型与WPD-BP和BP模型相比,最大估计误差低于1.5%,预测性能显著提升。相较于SVR和LSTM模型,WPD-GA-BP模型拟合优度(R^2)最高,且MAE和RMSE均为最小,表现出更强的预测精度与稳定性,进一步验证了该方法在动力电池SOH预测中的有效性。

关键词:动力电池;健康状态;小波包分解;遗传算法;BP神经网络

中图分类号:U466

文献标志码:A

本文引用格式:黄亮亮,郭拯诗,何峰,等.基于WPD-GA-BP的电动汽车动力电池健康状态预测方法[J].华东交通大学学报,2026,43(1):101-113.

Battery Health State Prediction for Electric Vehicles Based on WPD-GA-BP

Huang Liangliang^{1,2}, Guo Zhengshi¹, He Feng¹, Hu Xin^{1,3}, Li Yushu¹, Duan Xingbing^{1,3}, Zeng Jianbang^{1,3}

(1. East China Jiaotong University Key Laboratory of Conveyance and Equipment, Ministry of Education, Nanchang 330013, China; 2. Nanchang Fire Rescue Brigade, Nanchang 330038, China; 3. New Energy Vehicle Corporation, Jiangxi Jiangling Motors Group, Nanchang 330114, China)

Abstract: Accurate prediction of the state of health (SOH) of power batteries was essential for extending the service life of new energy vehicles and ensuring driving safety. To address the limitations of BP neural networks in feature extraction, sensitivity to initial parameters, and local optima issues, a WPD-GA-BP-based prediction method was proposed. First, feature parameters were extracted using capacity increment analysis, and key features related to SOH were selected using Pearson correlation. Next, wavelet packet decomposition was applied to

收稿日期:2025-04-19

基金项目:国家自然科学基金(51206171);江西省重点研发项目(20223BBE51016,20243BBG71009);江西省自然科学基金项目(20242BAB25273,20244BAB28006);江西省研究生创新专项资金项目(YC2023-S505);载运工具与装备教育部重点实验室基金(KLCE2022-09)

multi-scale reconstruct the label values, enriching the feature set. Finally, a genetic algorithm optimizes the BP neural network's initial weights and thresholds, avoiding local optima and improving prediction accuracy. The results show that, compared to the pre-improvement WPD-BP and BP models, WPD-GA-BP reduces the maximum error to less than 1.5%, significantly improving prediction performance. It outperforms SVR and LSTM models, achieving the highest R^2 and the smallest MAE and RMSE, demonstrating stronger accuracy and stability in predicting power battery SOH.

Key words: power battery; SOH; wavelet packet decomposition; genetic algorithm; BP neural network

Citation format: HUANG L L, GUO Z S, H F, et al. Battery health state prediction for electric vehicles based on WPD-GA-BP[J]. Journal of East China Jiaotong University, 2026, 43(1): 101-113.

准确预测动力电池 SOH, 对提高电动汽车的驾驶体验和驾驶安全具有重要意义^[1-2]。目前, 动力电池 SOH 的预测方法主要分为三类。

第一类是基于模型的方法, 这种方法是通过建立电化学模型或等效电路模型, 利用卡尔曼滤波(Kalman filter, KF)或高斯滤波实现动力电池 SOH 预测。Yang 等^[3]构建了考虑热效应的锂离子电池参数不一致性电化学仿真模型。该模型通过集成热模块进行温度场计算, 并采用双扩展卡尔曼滤波算法实现动力电池 SOH 的准确预测。Xu 等^[4]利用多尺度 DEKF 预测动力电池的 SOH。然而, 基于模型的方法高度依赖于模型本身的准确性, 而电池模型参数易受老化、工况和环境等因素影响, 进而导致动力电池 SOH 预测精度不足。

第二类是基于数据驱动的方法。近年来, 在大数据与人工智能技术迅猛发展的推动下, 数据驱动方法已成为当前学术研究的重要方向之一。该方法通过对电压、电流及温度等关键参数进行实时监测与特征分析, 结合机器学习与深度学习算法构建动力电池 SOH 与多维输入特征之间的非线性映射模型, 从而实现电池 SOH 的快速精准预测^[10-12]。长短期记忆网络和卷积神经网络等深度学习模型已被广泛用于动力电池 SOH 预测, 均表现出良好的性能^[7]。Yang 等^[8]提出了基于卷积神经网络和随机森林的锂离子动力电池 SOH 预测方法。然而, 基于数据驱动的方法存在小样本学习能力不足、泛化性能差以及缺乏物理可解释性等局限性, 从而使其预测精度受到限制^[9]。

第三类是基于混合的方法, 这种方法是指系统性地融合多种数据驱动法预测动力电池的 SOH。He 等^[10]设计了一个将 LSTM 和 GRU 相结合

的复杂系统, 利用 LSTM 在处理高维数据方面的出色能力及 GRU 在处理大量数据方面的优势来预测动力电池 SOH。Zheng 等^[11]提出了一种基于自回归滑动平均(ARMA)模型和 Elman 神经网络(ARMA-ENN)的融合模型来预测动力电池 SOH, 该模型考虑了局部小容量的恢复和充电过程中复杂信息的变化。

与它们相比, BP 神经网络结构简单, 计算方便, 很多学者对其进行了改进。此外, BP 神经网络具有误差反向传播的特性, 可以通过误差来调整参数, 使得预测结果更加准确。BP 神经网络具有较强的泛化能力, 适用于动力电池 SOH 预测, 但需要注意的是, 该方法的预测性能易受参数初始化策略的影响, 且容易陷入局部最优解; 此外, 其在特征提取能力方面也存在一定的局限性^[12]。

为解决上述问题, 本文提出一种基于 WPD-GA-BP 的电动汽车动力电池 SOH 预测方法, 主要分为三个部分: 首先, 以实际车辆运行数据为研究对象, 采用容量增量分析法(incremental capacity analysis, ICA)提取特征参数; 其次, 为丰富特征参数的多样性, 采用小波包分解(wavelet packet decomposition, WPD)对标签数据进行特征扩维; 最后, 利用遗传算法(genetic algorithm, GA)优化 BP 神经网络的初始权重和阈值, 通过更广泛的搜索空间进行全局优化, 有效避免陷入局部最优, 从而实现动力电池 SOH 的准确预测。

1 特征参数提取与扩维

1.1 数据预处理方法

本研究数据来自某车企新能源服务与管理中心, 该管理中心按照国家标准的规定, 以不低于 20 s

的频率采集车辆数据。本研究采集了8类关键运行参数,包括:数据采集时间、行驶里程、SOC、总电压、总电流、单体电池电压、单体电池温度以及充电状态。所使用的数据来源于某款A级纯电动汽车,其电池包的相关参数详见表1。对数据进行清洗,排除异常值和缺失值,方法如下:

1) 缺失值处理:使用缺失值的上一行或者下一行来填充;

2) 异常值处理:采用箱型图法进行异常值检测,对识别出的异常值直接进行剔除处理,以提高数据质量和模型稳定性^[13],异常值定义方法参考文献[14];

3) 车辆行驶状态划分:本文主要选取停车充电片段且片段时间大于10 min的数据集进行研究。

表1 某款电动汽车动力锂离子电池包详细参数

Tab.1 Detailed parameters of the power lithium-ion battery pack for a certain type of electric vehicle

参数	数值
电池类型	三元锂电池
总额定电压/V	393.12
总额定容量/Ah	160.5
电池包能量/kWh	63.09
单体充电截止电压/V	4.22
单体放电截止电压/V	3.00
电池包结构	3P108S

1.2 锂离子动力电池SOH计算与修正

目前动力电池SOH预测主要基于容量衰减率和内阻增长率两类参数。考虑到容量参数可直接反映电池储能能力的变化,本研究采用容量衰减法定义动力电池SOH,并利用安时积分法来计算动力电池容量,其计算公式如下

$$SOH = \frac{C_p}{C_R} \times 100\% \quad (1)$$

$$C_p = \frac{\int Idt}{\Delta SOC} \quad (2)$$

式中: C_p 为动力电池在当前状态下的可用最大容量; C_R 为电池的额定容量; I 为选取的充电片段电流; ΔSOC 为选取的充电片段SOC差值。

为全面覆盖充电过程中的各阶段,并确保所选片段在电压、电流变化方面具有良好的稳定性,本文选取SOC位于35%~75%区间时对应的总电压[392.5, 417.7] V区间筛选充电片段,计算固定电压

区间充入的部分容量,记为 C_p 。基于某车企一辆刚出厂车辆的完整充电数据片段,提取总电压区间[392.5, 417.7] V的充电数据,并计算电压区间充入容量为61.797 Ah,并将其作为额定容量。由于慢充过程中电流较为稳定,电化学反应也更加充分。因此,本文选取了3辆行驶里程接近10万km,慢充片段多的车辆作为研究对象。3辆车的慢充充电数据如表2所示。

表2 3辆车的慢充充电数据

Tab.2 Slow-charge charging data for 3 electric vehicles

车辆编号	运行时间	数据量/条	行驶里程/km	达标充电片段/个
1	2021.10.29—2023.2.28	1 152 246	100 486	178
2	2021.11.04—2023.2.28	1 177 044	100 775	140
3	2021.11.05—2023.2.28	1 015 200	85 069	107

基于所选取的3辆实车充电数据,依据动力电池SOH公式(3)计算3辆车的动力电池SOH,计算结果如图1所示。由图1可知,计算所得车辆1、车辆2和车辆3在固定电压区间内计算所得的动力电池SOH下降趋势明显且缓慢。3辆车在使用5 A左右小电流充电时,计算所得固定电压区间的动力电池SOH存在较大波动。其中,车辆1计算所得动力电池SOH最大为100.86%,最小为78.52%;车辆2计算所得最大动力电池SOH为93.09%,最小为70.85%;车辆3计算所得最大动力电池SOH为100.21%,最小为78.26%。同时,计算所得车辆1、车辆2和车辆3初始动力电池SOH在95%左右,不符合实际情况。因此,需对计算所得的动力电池SOH进行修正。本研究采用了一种基于实车充电数据的动力电池SOH修正方法^[15]。该方法所依赖的车辆数据(采集频率、参数类型、车型)与本文使用的实车数据完全一致,适用性良好,可满足动力电池SOH修正需求。由于大电流充电过程中电流波动较大且车企监控平台数据采样周期为20 s,基于大电流充电片段数据估算容量易造成较大计算误差。为此,本文筛选慢充数据片段用于动力电池SOH分析。该修正方法以选定电压区间内的慢充片段为基础,通过分析充电电流均值与充电时间对数据进行区间划分。在每个划分区间内,选择充入容量的最大值,记为 C_s ,用于对动力电池SOH进行修正。其计算公式如式(4)所示

$$SOH = \frac{C_{\beta}}{61.797} \times 100\% \quad (3)$$

$$SOH = \frac{C_{\beta}}{C_s} \times 100\% \quad (4)$$

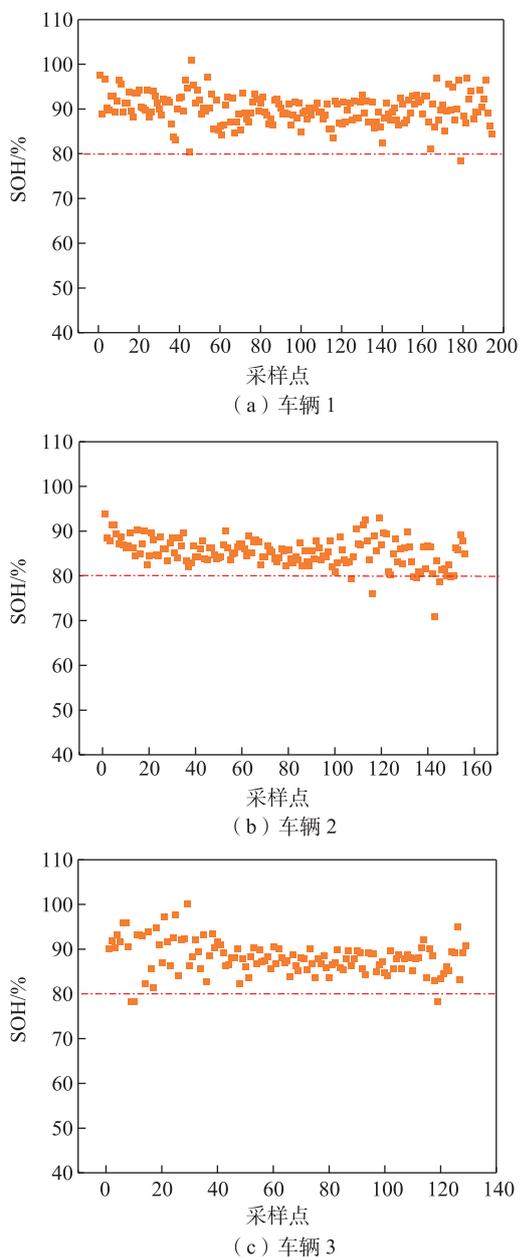


图1 车辆1、车辆2和车辆3修正前的动力电池SOH
Fig. 1 The SOH of No.1, No.2 and No.3 vehicle before the correction

将3辆车的动力电池SOH值代入式(4)进行修正,修正后的动力电池SOH结果如图2所示。从图2可知,修正后的动力电池SOH整体呈平稳衰减趋势,未出现异常波动,3辆车的动力电池SOH状况整体良好。

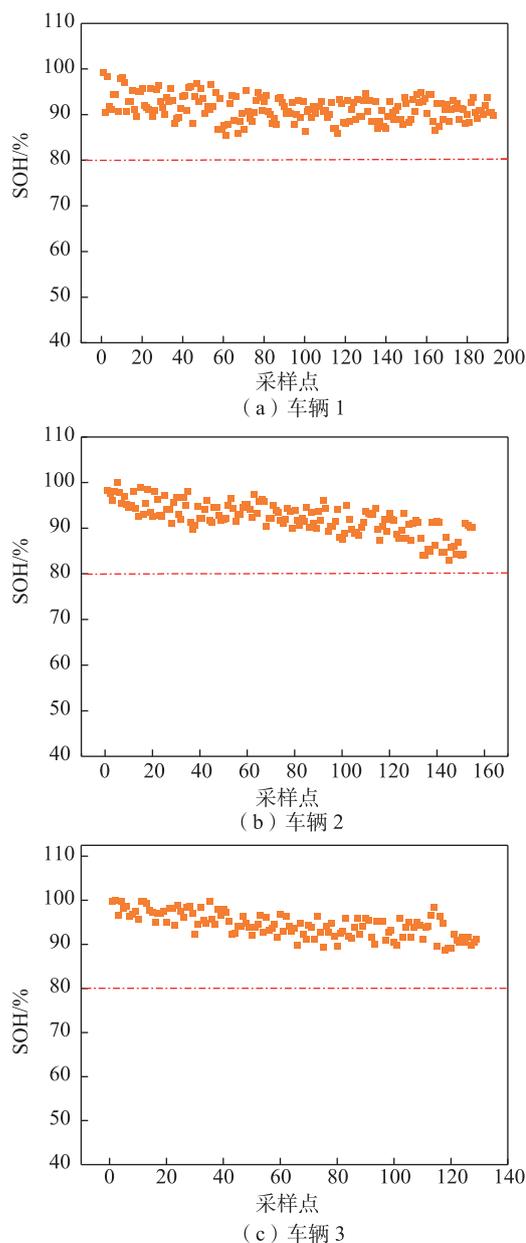


图2 车辆1、车辆2和车辆3修正后的动力电池SOH
Fig. 2 The corrected SOH of No.1, No.2 and No.3 vehicle

1.3 特征参数提取

在锂离子电池长期充放电的循环过程中,电池容量会逐渐衰减,其内部复杂的老化机制难以被直接观测和量化。为了克服这一难题,研究者们开发了容量增量分析法。该方法通过分析电池充放电曲线中容量随电压变化的微小差异,帮助识别并量化不同类型的老化机理^[6]。基于该方法,本文特征参数提取主要分为三个步骤:首先,采用ICA提取特征;其次,对ICA提取的数据曲线进行平滑处理;最后,采用皮尔逊相关系数对ICA提取

的特征进行相关性分析。

1.3.1 IC曲线获取

IC曲线的获取是基于电池充电过程中容量与电压之间的关系,即通过分析容量对电压的变化率来揭示电池内部的老化特征^[7]。目前常用的获取方式主要有两种:一种是对电压-容量曲线进行多项式拟合后再求导,另一种是直接采用微分公式 dQ/dV 。其中,多项式拟合法虽然在一定程度上可以平滑曲线,减小噪声干扰,但其对拟合阶数和数据质量较为敏感,容易引入拟合误差,进而影响IC曲线的精度。因此,本文选择采用微分公式法来提取IC曲线,以更真实地反映原始数据特性。

由于本文所使用的实车数据为每20 s采集1帧,数据本质上是离散的,且电压数据的精度为0.1 V,这可能导致相邻两帧间电压无变化,从而无法计算出有效的IC值。为解决该问题,在处理每一个充电片段时,首先计算相邻两帧的电压差,记为 ΔV ,随后计算该区间内充入的容量,记为 ΔQ ,从而得到相应公式如下

$$\frac{dQ}{dV} = \frac{\Delta Q}{\Delta V} = \frac{I \cdot dt}{dV} \quad (5)$$

式中: ΔQ 为计算所得充入容量,Ah; ΔV 为电压,V; t 为采样时间,s。

由式(5)计算得到的结果如图3所示,I峰及其对应的峰谷在IC曲线上通常表现得较为突出,这主要是因为该区间的信号幅度较大且变化缓慢,相对不易受到噪声的干扰;而II峰则往往处于信号的微小振荡区间,其幅度较小且频谱中含有更多高频成分,因而在实际电池测试过程中极易被叠加的电流扰动、温度波动以及噪声所影响。噪声污染不仅会

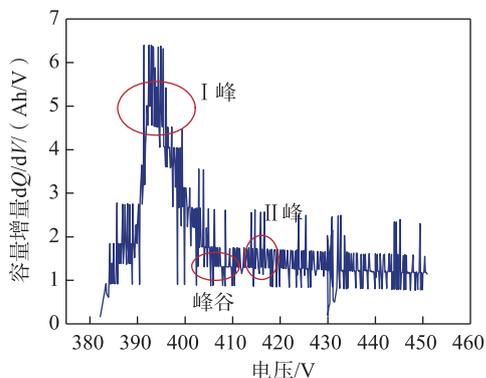


图3 未降噪的容量增量曲线

Fig. 3 Capacity increment curve without noise reduction

削弱II峰的振幅,还会引起峰位偏移,致使该特征难以在原始曲线上被准确辨识和定位。因此,需要对计算结果进行去噪处理,以获得IC曲线中各峰谷值及其对应横坐标电压的精确位置。

1.3.2 IC曲线降噪处理

在数据采集过程中,不可避免地会引入噪声,而这些噪声可能对后续的数据分析产生显著影响。为了获得更加平滑的IC曲线,同时尽可能保留I峰、II峰及波谷的位置和强度等关键特征信息,本文采用S-G滤波方法对原始数据进行平滑降噪处理。该方法通过滑动窗口对时间序列进行多项式拟合,利用卷积操作平滑容量数据,有效去除高频噪声,实现信号的平滑处理。具体计算公式如下^[8]

$$Q(n) = \sum_{i=0}^N a_i n^i \quad (6)$$

式中: $Q(n)$ 为去噪后的容量值; N 为多项式拟合次数; a_i 为多项式的系数。

通过采用不同滑动窗口、选用不同拟合多项式,对容量增量曲线进行平滑处理,平滑后的曲线如图4所示。图4为滑动窗口长度分别取33、43、53以及多项式分别取2次和3次的S-G滤波平滑效果对比图。分别对比图4(a)、图4(b),图4(c)、图4(d),图4(e)、图4(f)可知,滑动窗口长度一致的情况下,2次多项式和3次多项式经过S-G滤波平滑后,结果并没有明显的差别。根据S-G滤波原理可知,多项式次数越多,曲线越不平滑,结合图4的S-G滤波效果图,本文S-G滤波采用2次多项式。

对比图4(a)、图4(c)和图4(e)可知,当滑动窗口为33时,S-G滤波平滑后的曲线整体效果较差,I峰、II峰和波谷等地方存在较多毛刺;当滑动窗口为43时,S-G滤波平滑后的曲线整体平滑效果良好,I峰、II峰和波谷等特征明显,并且平滑后的曲线形状没有发生变形;当滑动窗口为53时,平滑后的曲线整体开始向电压更高处偏移,并且I峰和II峰高度相较于滑动窗口33和43下降较大。综上所述,为了得到更好的平滑曲线,并且保留I峰、II峰和波谷的位置和强度等特征信息,S-G滤波采用滑动窗口为43,多项式次数为2,对容量增量曲线进行平滑降噪处理。

1.3.3 皮尔逊相关系数分析

锂离子电池内部的化学反应可以通过IC曲线间接反映出来,充电过程中正负极材料在液相与

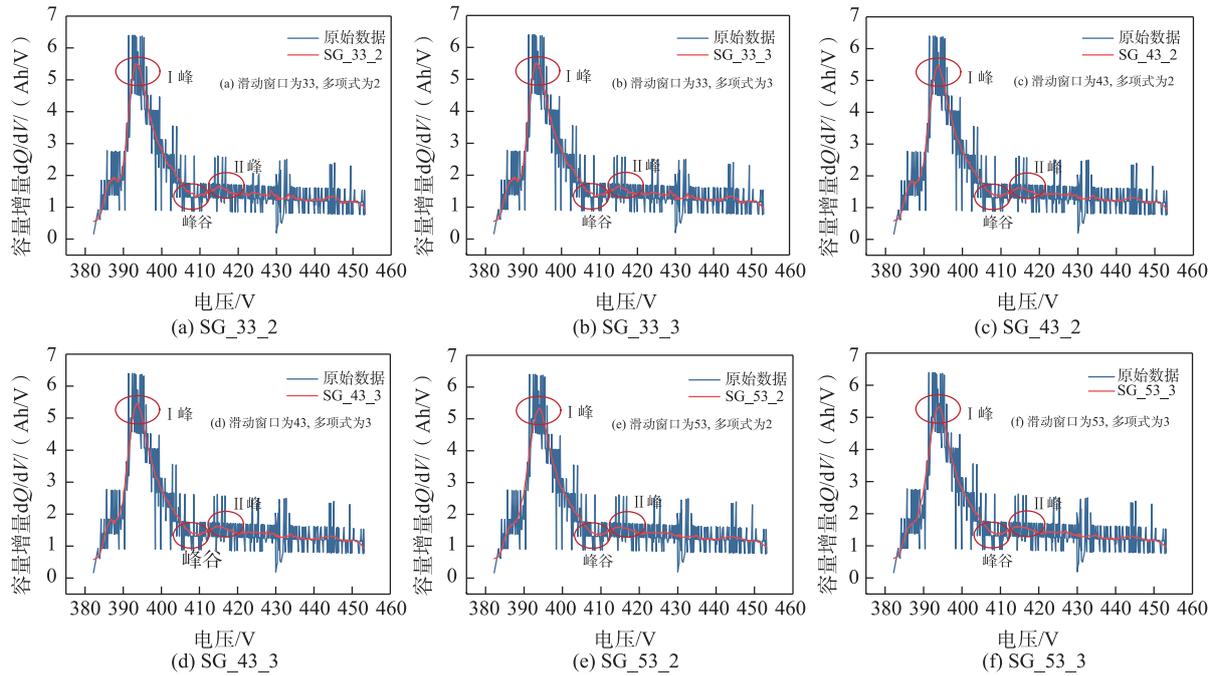


图4 不同滑动窗口长度、不同多项式次数的S-G平滑效果对比

Fig. 4 Comparison of S-G smoothing effect of different moving window length and different polynomial frequency

固相之间的转化行为以定量方式显示。本文通过提取IC曲线中的I峰电压、II峰电压、谷值电压,以及I峰、II峰和谷值的电压幅值共6个特征参数,揭示锂离子动力电池SOH与IC特征之间的潜在关系。为了进一步明确这些特征与动力电池SOH之间的相关性,并为后续预测模型提供有效输入,本文采用皮尔逊相关系数对上述6个参数与动力电池SOH之间的线性相关性进行了分析,为特征选择与建模提供数据支撑。该系数的取值范围为-1至1,其中1表示完全正相关,-1表示完全负相关,0则表示没有线性关系^[9]。皮尔逊相关系数可以帮助理解两个变量之间的关联性,这对于探索数据集中变量之间的关系、进行预测分析、发现潜在趋势以及了解变量如何随时间变化等方面都是非常有用的。皮尔逊相关系数公式如下

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (7)$$

式中: r 为相关系数; x_i 和 y_i 分别为提取的特征以及锂离子动力电池SOH; \bar{x} 和 \bar{y} 分别为 x_i 和 y_i 的平均数。当 $|r| < 0.2$ 时,表示两组变量呈现极弱的相关性; $|r| \in [0.2, 0.4]$ 时,表明两组变量之间存在弱相关性;在 $|r| \in [0.4, 0.6]$ 时,表示两组变量之间存在中等

程度的相关性;当 $|r| \in [0.6, 0.8]$ 时,表明两组变量之间具有强相关性;而当 $|r| \in [0.8, 1.0]$ 时,表示两组变量之间存在极强的相关性,本文将0.7作为判断变量之间是否存在显著关联的参考阈值^[20]。

以车辆1为例,不同特征与电动汽车动力电池SOH之间的皮尔逊相关系数如表3所示。通过表3可知,锂离子动力电池SOH与I峰电压、II峰电压以及谷值电压的相关性系数不超过0.3,而与I峰、II峰和谷值的相关性超过0.7。这说明相对于电压,锂离子动力电池SOH与I峰、II峰和谷值的相关性更高。因此,本文采用IC提取的I峰、II峰和谷值作为后续模型的输入特征。

1.4 基于WPD的特征扩维

实际环境中车载动力电池容易受到环境的影响,采集数据波动性较大,导致使用ICA提取的特征与动力电池SOH的相关性较低,难以准确刻画其变化规律,影响后续模型的预测精度。为提升预测效果,采用WPD方法对动力电池SOH数据进行扩维,将原始动力电池SOH序列分解为8个具有不同频率信息的子序列,以增强特征的时频表达能力。最终将这些子序列与ICA提取的输入特征融合,构建更丰富的输入特征集,以提升模型对动力电池SOH变化趋势的学习能力和预测精度。

表3 不同特征与锂离子动力电池SOH之间的皮尔逊相关系数

Tab.3 Pearson correlation coefficient between different characteristics and the health status of lithium-ion batteries

变量	SOH	I 峰	II 峰	谷值	I 峰电压	II 峰电压	谷值电压
SOH	1.000	0.860	0.790	0.820	-0.200	0.043	-0.040
I 峰	0.760	1.000	0.350	0.260	-0.610	0.230	-0.060
II 峰	0.790	0.350	1.000	0.340	-0.005	-0.001	0.130
谷值	0.720	0.260	0.340	1.000	0.013	0.210	-0.350
I 峰电压	-0.200	-0.610	-0.005	0.013	1.000	0.053	-0.035
II 峰电压	0.043	0.230	-0.001	0.210	0.053	1.000	-0.098
谷值电压	-0.040	-0.060	0.130	-0.350	-0.035	-0.098	1.000

相比于2层WPD分解,3层WPD分解既能更细致地提取动力电池SOH演化过程中的高频突变信息与低频退化趋势,又避免了2层WPD分解时分辨率不足的问题;而与4层WPD分解相比,3层WPD分解在保留足够特征信息的同时,有效控制了特征维度,降低了建模复杂度与过拟合风险^[21]。为此,本文分别对车辆1、车辆2和车辆3的动力电池SOH采用3层WPD处理,在频率分辨率、特征维度与计算成本之间取得了最佳平衡,不同频段下的动力电池SOH分解波形如图5所示。图中展示了3辆车的动力电池SOH信号经3层WPD后的波形图,清晰地将原始信号分解为8个具有不同频率特征的子信号,这些频率分量在不同车辆中的波形差异明显,具备良好的类别区分能力,适用于后续的动力电池SOH预测。

低通滤波器和高通滤波器在频域上相互正交的数学特性,构成了WPD实现多尺度、多层次信号精确分解的理论基础和核心保障,确保了信号分析的完备性,其变换公式如下^[22]

$$\begin{cases} d_{i,j,2m} = \sum_k h(k-2i)d_{k,j+1,m} \\ d_{i,j,2m+1} = \sum_k g(k-2i)d_{k,j+1,m} \end{cases} \quad (8)$$

式中: $d_{i,j,2m}$ 、 $d_{i,j,2m+1}$ 分别为第j层第2m和2m+1个节点在第i个位置的系数, h 和 g 分别为低通和高通滤波器, k 为卷积求和索引。

2 基于WPD-GA-BP的动力电池SOH预测方法

2.1 BP神经网络原理

BP神经网络,是由非线性变换单元组成的前馈网络。该网络不需要确定输入输出之间的数学关系,只需要为训练学习提供数据,预测结果的准确性与训练数据量正相关。该模型通过梯度下降法进行参数优化,使目标函数值最小化,并通过一个自适应学习速率将网络收敛控制在一个合适的范围内。BP神经网络模型结构由输入层、隐藏层和输出层三部分组成^[11],其结构如图6所示。该模型内部信号传播模式有两种,分别是信号的前向传

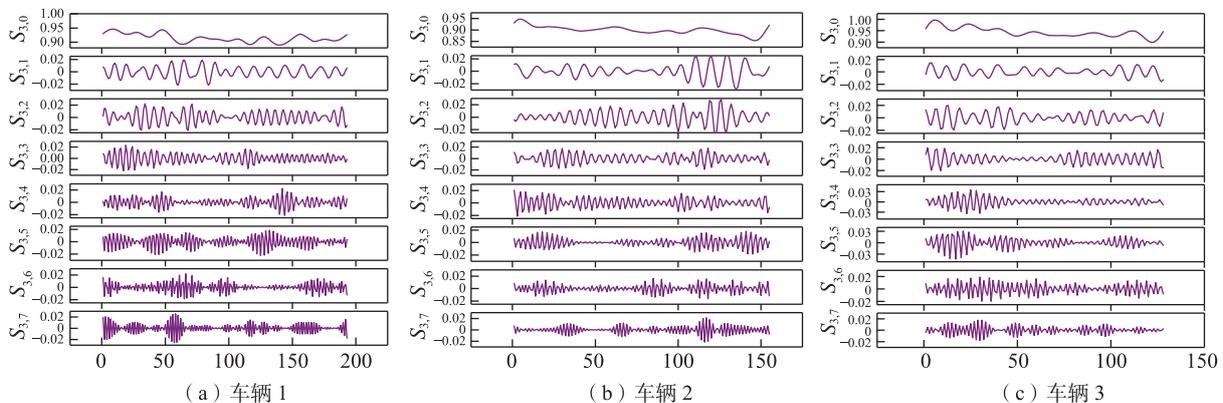


图5 WPD分解后的动力电池SOH波形图

Fig. 5 SOH waveform diagrams after WPD decomposition

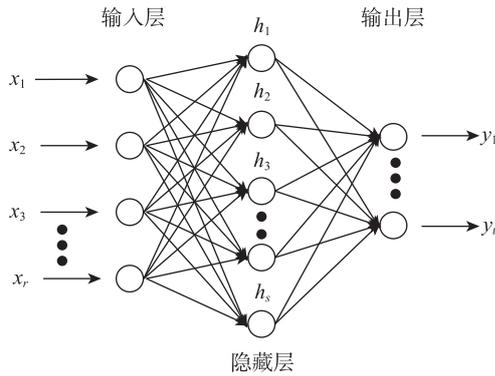


图6 BP神经网络模型结构图

Fig. 6 BP Neural network model structure diagram

播和误差的后向传播。

对于 r 维输入层向量 $X = (x_1, x_2, \dots, x_r)^T$, s 维隐藏层向量 $H = (h_1, h_2, \dots, h_s)^T$, t 维输出层向量 $Y = (y_1, y_2, \dots, y_t)^T$ 。BP神经网络的主要步骤如下:

1) 计算期望输出

$$h_j = f\left(\sum_{i=1}^r w_{ij}^1 x_i + b_j^1\right) \quad (9)$$

$$d_j = f\left(\sum_{i=1}^s w_{ij}^2 h_i + b_j^2\right) \quad (10)$$

式中: w_{ij}^l , $l=1,2$ 为第 l 层的第 i 个神经元和第 $l+1$ 层的第 j 个神经元之间的权重。 f 为激活函数。

2) 计算全局误差

$$E = \frac{1}{2N} \sum_{i=1}^N \sum_{j=1}^t (d_{ij} - y_{ij})^2 \quad (11)$$

式中: N 为训练样本数。

3) 更新权重和偏差

$$w_{ij}^l = w_{ij}^l - \alpha \frac{\partial E}{\partial w_{ij}^l}, \quad l=1,2 \quad (12)$$

$$b_j^l = b_j^l - \alpha \frac{\partial E}{\partial b_j^l}, \quad l=1,2 \quad (13)$$

式中: $\alpha \in (0,1)$ 为学习率。

2.2 GA原理

GA是一种模拟生物进化的全局优化方法,通过自然选择、交叉与变异等机制,在复杂搜索空间中以概率方式寻找最优或近似最优解,具有良好的全局寻优能力和适应性。GA的基本原理包括种群初始化、适应度评估、选择、交叉、变异和替代等步骤^[23]。GA求解的基本步骤如下:

1) 染色体编码:将求解空间的数据转换成GA可处理的形式。常见的编码方式包括二进制编码

和实数编码。二进制编码因其简便易行的特点,以及容易实现遗传操作如交叉和变异,因而被广泛采用。

2) 初始种群的生成:在对解空间中的数据完成编码后,算法随机产生 Y 个串状结构的数据作为染色体,用以表示潜在解。这 Y 个染色体个体共同构成了初始群体,种群记为 $P_j = (x_i | 0 \leq i \leq N)$, j 为目前的进化代数, N 为种群规模。GA进化迭代是以初始种群为起点来进行的。

3) 适应度函数:GA的优化函数需要一个评价指标,于是引入适应度的概念,适应度在GA中用于评估个体接近或达到最优解的程度,高适应度的个体更有可能被遗传到下一代。适应度的选取方式灵活多变,一般根据实际情况来改变,通常来说,有取极大值和极小值两种选取方式。

4) GA通过选择、交叉和变异3种核心操作实现种群的迭代进化。首先,选择操作基于适应度评估结果,优先复制当前群体中适应能力最强的个体作为父代;其次,交叉运算作为GA的关键特征,通过单点、两点或均匀算子对父代染色体进行基因重组,产生具有新特性的子代个体;最后,变异操作以较小概率随机改变染色体特定基因位,既保证了算法的全局搜索能力,又增强了局部优化潜力。这3种遗传操作的协同作用,使得算法能够在解空间中高效地探索最优解。

5) 终止条件判定:通过选择、交叉和变异等操作生成子代群体,而后重复操作3)和4)直到满足结束条件设定最大的迭代次数和适应度值的指标。

2.3 WPD-GA-BP模型构建

基于WPD-GA-BP的动力电池SOH预测方法通过ICA提取特征参数,并利用WPD对其扩维;同时结合GA和BP神经网络构建预测模型。具体预测步骤如下。

1) 对车联网平台数据进行数据预处理,采集预处理后的动力电池电流、电压和SOC等数据,采用安时积分法对采集到的数据进行动力电池SOH计算及修正。

2) 采用ICA获取电池容量增量曲线,并使用S-G滤波进行平滑降噪处理。提取峰值、谷值及对应电压等特征,利用皮尔逊相关系数分析各特征与动力电池SOH的相关性,最终选取与动力电池SOH相关性较强的I峰、II峰和谷值作为模型输入特征。

3) 针对ICA提取的特征维度不足的问题,使用WPD对修正后的锂离子动力电池SOH进行数据扩维,将其分解成8个子序列,与步骤2提取的3个特征一起作为模型输入。

4) 采用GA对BP神经网络的权值和阈值进行优化,借助更广泛的搜索空间实现全局优化,有效避免陷入局部最优,并将步骤3提取的特征参数输入该模型,以实现动力电池SOH的预测。

3 模型参数选择及训练

3.1 评价指标

为了评估所构建模型的预测性能,本文采用了拟合优度(R^2)、均方根误差(Root Mean Squared Error, RMSE)和平均绝对误差(Mean Absolute Error, MAE)3个指标对WPD-GA-BP模型进行评估。上述3个指标计算公式如下, n 为样本数量, y_i 为实际观测值。 \hat{y}_i 为模型的预测值, \bar{y} 为观测值的均值

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (15)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (16)$$

3.2 BP神经网络参数的确定

BP神经网络的参数包括隐藏层层数、隐藏层节点、激活函数、损失函数及优化器。

1) 隐藏层层数与隐藏层节点设定:隐藏层与拟合函数能力密切相关,隐藏层过少,函数的拟合性能受影响,层数过多,模型的泛化性能降低。因此,在选择隐藏层时需要谨慎,确保找到适当的层数。本文为了得到良好的预测结果,隐藏层数设置为2层。

隐藏层节点数的选择对BP神经网络性能影响较大。合适的隐藏层节点数,可以充分发挥BP神经网络性能。隐藏层节点数可由经验公式来确定,如下

$$v_h = \sqrt{v_x + v_y + a} \quad (17)$$

式中: v_h 为隐藏层节点数; v_x 为输入层节点数; v_y 为输出层节点数; a 为调节常数,根据经验通常取1~10。经过调试,本文将两层隐藏层节点数都设为7,输出层节点数设置为1。

2) 激活函数设定:相对于Sigmoid、Tanh等激活函数,ReLU函数具有诸多优势。在反向传播更新梯度时,ReLU函数计算更为简便,更适用于处理非线性映射。ReLU函数下,负值直接被置为零,对网络的稀疏性起到增加的作用。因此,本文构建的BP神经网络模型选择采用ReLU作为隐藏层激活函数。

3) BP神经网络利用损失函数评估真实值和预测值之间的误差,并通过梯度下降法来实现参数更新,以实现损失的最小化。本文选择损失函数为MSE,优化器选择Adam。

3.3 GA参数的确定

1) 染色体编码:本研究采用GA对BP神经网络的权值和阈值进行联合优化。在算法实现层面,将网络的权值矩阵和阈值向量直接编码为染色体个体,采用实数编码方案进行基因表示,其中每个基因位对应一个待优化的网络参数。

2) 种群初始化:生成多个代表个体的初始字符串数据,建立GA的初始种群。随着种群个体数量的增加,多样性逐渐增强,但当个体过多时,种群结构复杂,计算负担增加,导致优化时间增加,减慢算法收敛速度。反之,个体数量较少,则GA可能过早收敛。因此,一般建议将初始种群数量控制在20到100之间,本文设定种群数量为100。

3) 适应度函数设置:本文建立的WPD-GA-BP模型完成动力电池SOH的预测,重点在于如何使得预测值与真实值的误差最小。因此,本文适应度函数选取均方根误差,其数学表达式如下

$$F_i = \frac{1}{2} \sum_{i=1}^N (X_i - Y_i)^2 \quad (18)$$

式中: N 为输出层神经元数量; X_i 为神经网络目标输出向量; Y_i 为神经网络前向传播结果。

4) 选择复制:根据一定比例,对种群中排名较高的个体进行复制进化,同时淘汰排名较低的个体。本文采用精英保存法,选择优秀个体的标准为个体适应度的归一化值,如下

$$\begin{cases} f_i = \frac{1}{F_i} \\ P_i = \frac{f_i}{\sum_{i=1}^N f_i} \end{cases} \quad (19)$$

式中: P_i 为被选择概率; N 为种群数,由于适应度函数是均方根误差,适应度值越小越好,在个体选

择前应对适应函数求导数。

5) 交叉与变异操作是遗传算法优化性能的关键因素。交叉概率过低会抑制新个体的产生,降低搜索效率,而过高则可能削弱种群多样性,影响收敛稳定性;同样,变异概率过小会限制搜索空间的探索,过大则易导致算法陷入随机搜索,影响全局寻优能力。为平衡搜索能力与收敛效率,发现当交叉概率取值在0.6~0.8、变异概率位于0.03~0.05区间时,算法在保持种群多样性、提高收敛速度与预测稳定性方面具有良好表现。在此基础上,进一步采用局部微调策略,通过细化步长(如0.02)优化参数,最终将交叉概率确定为0.7、变异概率设为0.04,这与文献[24]结果(交叉概率设置在0.1~1.0之间,变异概率设置在0.01~0.1范围内)完全吻合,有望确保模型性能与收敛效率的有效权衡。

假设有两个染色体分别为 a_k 和 a_l ,两者在 j 位置交叉得到新的个体,则算法实现的公式如下

$$\begin{cases} a_{kj} = a_{kj}(1-\beta) + a_{lj}\beta \\ a_{lj} = a_{lj}(1-\beta) + a_{kj}\beta \end{cases} \quad (20)$$

式中: a_{kj} 和 a_{lj} 分别为交叉后得到的新个体; β 为[0,1]内的随机数。

在变异操作中,假设染色体上 a_k 第 j 个基因变异,则GA的数学表达式如下

$$a'_{kj} = \begin{cases} a_{kj} + (a_{kj} - a_{\max}) \cdot Q(f), \lambda > 0.5 \\ a_{kj} + (a_{kj} - a_{\min}) \cdot Q(f), \lambda \leq 0.5 \end{cases} \quad (21)$$

式中: a'_{kj} 为染色体 a_k 变异后第 j 个基因; a_{kj} 为 a_k 染色体上第 j 个基因;染色体基因的最大、最小值分别用 a_{\max} 和 a_{\min} 表示; f 为迭代次数; λ 为[0,1]内的随机数。

6) 个体评价:将进行过交叉变异等操作得到的新个体进行适应度计算,评估其是否优于上代个体,如果满足设定的评价指标,则进行下一步操作,如果不满足,则再一次从选择复制开始执行,直至满足设定的适应度要求。

7) 将满足适应度函数的最佳权值阈值对BP神经网络的初始权值阈值进行更新替换,再对模型进行训练。

3.4 WPD-GA-BP模型训练

在机器学习领域,对数据集的合理划分至关重要,这会影响到模型的预测精度,多数研究在数据划分阶段,常采用训练集与测试集按照7:3和8:2

的比例划分。本文为了探究不同比例训练集测试集中模型预测的性能表现,分别将训练集测试集按照5:5、6:4、7:3和8:2这4种比例划分,在上述4种比例划分中,车辆1的测试集样本点数分别为97、77、58和40,车辆2的测试集样本点数分别为78、62、47、和31,车辆3的测试集样本点数分别为64、51、38和26,3辆车不同比例测试集预测结果如表4所示。从表4可以看出当测试集的比例为20%, R^2 、MAE及RMSE的结果优于前3种比例。因此,本文采用8:2的比例将3辆车的完整数据集划分为训练集和测试集。

表4 不同比例测试集模型预测性能对比

Tab.4 Comparison of model prediction performance for different proportions of test sets

车辆编号	测试集划分比例	R^2	MAE	RMSE
1	50%	0.871	0.005	0.008
	40%	0.861	0.005	0.009
	30%	0.905	0.004	0.007
	20%	0.927	0.003	0.006
2	50%	0.922	0.006	0.009
	40%	0.920	0.007	0.009
	30%	0.914	0.008	0.010
	20%	0.945	0.006	0.007
3	50%	0.909	0.004	0.007
	40%	0.911	0.005	0.007
	30%	0.893	0.005	0.008
	20%	0.927	0.003	0.006

4 模型应用

为验证本文所构建的WPD-GA-BP模型的预测性能,首先将其与两种基础模型进行对比分析:一是未引入GA优化的WPD-BP模型,二是未进行特征数据扩维的BP模型。对比分析可有效评估GA与WPD在模型性能提升中的作用。基于某车企平台,随机选取车辆4、车辆5和车辆6的测试数据作为模型评估对象,所选车辆均为慢充片段多的车型,且数据规格与训练集所用的车辆1、车辆2、车辆3保持一致,分别对3种模型在上述数据上的预测表现进行对比分析。3种模型的预测结果如图7所示;测试集中预测误差如图8所示;对应的预测性能评价指标(包括RMSE、MAE、 R^2)汇总于表4中。

由图7(a)可知,在车辆4的测试集上,相对于WPD-BP和BP模型,WPD-GA-BP模型更贴近原始

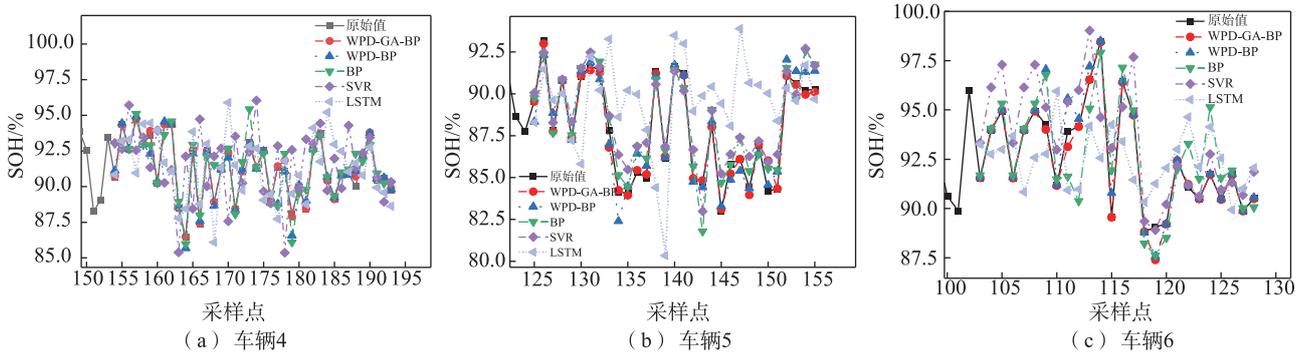


图7 WPD-GA-BP模型与其他4种算法预测结果对比

Fig. 7 Comparison of prediction results of WPD-GA-BP model with those of other 4 algorithms

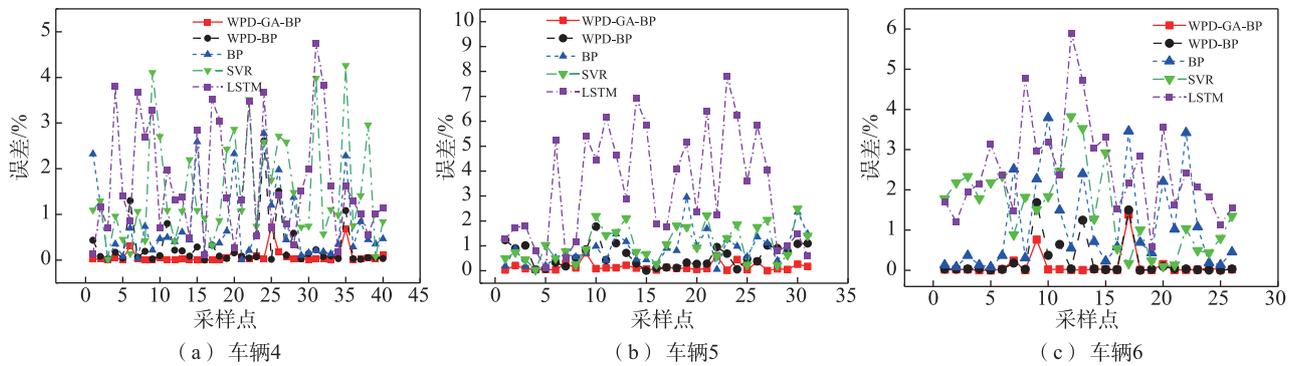


图8 WPD-GA-BP模型与其他4种算法预测误差对比

Fig. 8 Comparison of prediction errors of WPD-GA-BP model with those of other 4 algorithms

值,在WPD-BP和BP模型预测效果误差较大的采样点上,WPD-GA-BP模型的预测值同样能够贴近原始值;由图7(b)可知,在车辆5测试集的中间部分个别点及末尾几个点上,WPD-BP和BP模型的预测结果偏差较大;相比之下,WPD-GA-BP模型在预测精度方面表现出明显改善,预测值与实际值更加接近。由图7(c)可知,在车辆6的测试集上,WPD-GA-BP模型预测值基本与原始值贴合,而WPD-BP和BP模型的预测值存在着与原始值相差较大的情况,这表明WPD-GA-BP模型的预测结果准确度提高了很多。从图8可以看到WPD-GA-BP模型在车辆4、车辆5和车辆6的测试集上的预测误差整体最小,最大预测误差分别为0.730%、0.720%和1.385%。

由表5可知,本文建立的WPD-GA-BP模型在3辆车的性能预测测试中均强于WPD-BP和BP模型。相较于WPD-BP模型,WPD-GA-BP模型的 R^2 指标至少提升了0.036;同时RMSE和MAE指标分别至少降低了0.003和0.002。以上数据说明,相比BP模型,WPD-BP模型能更有效地提取动力电池

表5 WPD-GA-BP模型与其他4种算法预测性能对比

Tab.5 Comparison of prediction performance of WPD-GA-BP model with other 4 algorithms

车辆编号	模型	R^2	MAE	RMSE
4	WPD-GA-BP	0.987	0.001	0.002
	WPD-BP	0.927	0.003	0.006
	BP	0.741	0.007	0.011
	SVR	0.386	0.015	0.019
	LSTM	0.122	0.017	0.021
5	WPD-GA-BP	0.981	0.002	0.004
	WPD-BP	0.945	0.006	0.007
	BP	0.868	0.008	0.010
	SVR	0.735	0.011	0.013
6	WPD-GA-BP	0.979	0.001	0.004
	WPD-BP	0.927	0.003	0.008
	BP	0.599	0.011	0.016
	SVR	0.562	0.015	0.018
	LSTM	-2.847	0.025	0.028

SOH信号的多尺度特征,增强了对时频局部特征的表达能力,从而提升了动力电池SOH演化趋势的预测精度。GA通过全局优化BP神经网络的初始权重和阈值,有效避免了局部最优,进一步提升了WPD-BP模型在复杂非线性关系建模中的准确性和泛化能力。

在验证WPD-GA-BP模型相较于改进前模型的性能优势后,进一步与SVR和LSTM模型进行对比分析。通过在车辆4、车辆5和车辆6的测试集上对比各模型的预测结果,综合评估WPD-GA-BP、SVR和LSTM 3种模型在动力电池SOH预测任务中的表现。WPD-GA-BP、SVR和LSTM 3种模型的预测结果如图7所示;对应测试集上的预测误差如图8所示;各模型的预测性能评价指标汇总于表5中。

由图7可知,WPD-GA-BP模型在3辆车的测试集上表现出较高的预测精度,准确跟踪动力电池SOH变化,展现出良好的泛化性与鲁棒性。相比之下,SVR和LSTM模型预测结果与实际动力电池SOH偏差较大,未能有效捕捉其随时间变化的非线性特征,表现出滞后性和不稳定性。由图8可知,WPD-GA-BP模型在车辆4、车辆5和车辆6上的预测误差整体最小,表现出更高的预测稳定性;相比之下,LSTM模型的预测误差最大,且波动较为显著,反映出其在处理该类数据时的鲁棒性不足。

从表5的结果也可以看出,WPD-GA-BP模型在各项预测性能指标上均优于SVR和LSTM模型。其中,在车辆5和车辆6的预测中,LSTM模型的 R^2 值出现负值,表明其拟合能力较弱,甚至不如简单的平均值预测。这一异常现象可能由LSTM在处理高噪声或非平稳时间序列数据时泛化能力有限,或未能充分捕捉输入特征与目标变量之间的复杂动态关系导致,暴露出其在复杂工况下的预测局限性。相比之下,WPD-GA-BP模型在 R^2 指标上至少提升了0.246, RMSE和MAE分别至少降低了0.014和0.009,进一步验证了其在动力电池SOH预测任务中的优越性能与更强的泛化能力。

综上所述,相对于WPD-BP、BP、SVR和LSTM 4种模型,本文构建的WPD-GA-BP模型在预测精度和稳定性方面均取得了显著提升,能够更准确地对动力电池SOH进行预测。

5 结论

针对电动汽车动力电池SOH预测的问题,本文

基于车企监控平台运行车辆的电池数据,提出了一种基于WPD-GA-BP的电动汽车动力电池SOH预测方法。主要研究结论如下。

1) 利用WPD扩充数据维度,可以较大丰富BP神经网络预测动力电池SOH的输入特征,从而提高动力电池SOH的预测准确性。

2) 采用GA优化BP神经网络的权重和阈值,借助更广泛的搜索空间实现全局优化,有效避免陷入局部最优,从而提升了该方法在动力电池SOH预测中的稳定性与精度。

本文提出的动力电池SOH预测方法仅对电压、电流和温度等数据进行分析,未来将考虑容量、内阻等多参数对动力电池SOH预测的影响,以进一步提高动力电池SOH预测模型的精确度。

参考文献:

- [1] DINI P, COLICELLI A, SAPONARA S. Review on modeling and SOC/SOH estimation of batteries for automotive applications[J]. Batteries, 2024, 10(1): 34.
- [2] ALSUWIAN T, ANSARI S, AMMIRRUL ATIQUI MOHD ZAINURI M, et al. A review of expert hybrid and co-estimation techniques for SOH and RUL estimation in battery management system with electric vehicle application [J]. Expert Systems with Applications, 2024, 246: 123123.
- [3] YANG F, XU Y X, SU L, et al. State of charge and state of health estimation of lithium-ion battery packs with inconsistent internal parameters using dual extended Kalman filter[J]. Journal of Electrochemical Energy Conversion and Storage, 2024, 21(1): 011004.
- [4] XU P H, HU X Y, LIU B L, et al. Hierarchical estimation model of state-of-charge and state-of-health for power batteries considering current rate[J]. IEEE Transactions on Industrial Informatics, 2022, 18(9): 6150-6159.
- [5] ZHANG M, YANG D F, DU J X, et al. A review of SOH prediction of Li-ion batteries based on data-driven algorithms[J]. Energies, 2023, 16(7): 3167.
- [6] LELLI E, MUSA A, BATISTA E, et al. On-road experimental campaign for machine learning based state of health estimation of high-voltage batteries in electric vehicles[J]. Energies, 2023, 16(12): 4639.
- [7] FENG S Z, SONG M Y, LIN Y J, et al. Convolutional neural network-long short-term memory-based state of health estimation for Li-ion batteries under multiple working conditions[J]. Energy Technology, 2024, 12(2): 2301039.
- [8] YANG N K, SONG Z Y, HOFMANN H, et al. Robust state of health estimation of lithium-ion batteries using convolutional neural network and random forest[J]. Jour-

- nal of Energy Storage, 2022, 48: 103857.
- [9] MERROUCHE W, HARROU F, TAGHEZOUIT B, et al. Improved lithium-ion battery health prediction with data-based approach[J]. e-Prime-advances in electrical engineering, electronics and energy, 2024, 7: 100457.
- [10] HE Z G, GUO X D, HU S, et al. Hybrid estimation strategy for the state of health of lithium-ion batteries with different application needs[J]. International Journal of Electrochemical Science, 2022, 17(8): 220823.
- [11] CHEN Z, XUE Q, XIAO R X, et al. State of health estimation for lithium-ion batteries based on fusion of autoregressive moving average model and Elman neural network[J]. IEEE Access, 2019, 7: 102662-102678.
- [12] MA Y, YAO M H, LIU H C, et al. State of health estimation and remaining useful life prediction for lithium-ion batteries by improved particle Swarm optimization-back propagation neural network[J]. Journal of Energy Storage, 2022, 52: 104750.
- [13] HUANG J, HE T, ZHU W L, et al. A lithium-ion battery SOH estimation method based on temporal pattern attention mechanism and CNN-LSTM model[J]. Computers and Electrical Engineering, 2025, 122: 109930.
- [14] 岳加利, 郝静, 卢海林, 等. 桥梁结构监测数据异常值处理方法[J]. 武汉工程大学学报, 2022, 44(1): 107-111.
YUE J L, HAO J, LU H L, et al. Outliers processing method of bridge structure monitoring data[J]. Journal of Wuhan Institute of Technology, 2022, 44(1): 107-111.
- [15] 张月娅. 基于数据驱动的动力电池故障识别与SOH估计方法研究[D]. 南昌: 华东交通大学, 2023.
ZHANG Y Y. Research on power battery fault identification and SOH estimation based on data drive[D]. Nanchang: East China Jiaotong University, 2023.
- [16] WANG R X, ZHOU X, LIU Y J, et al. Analysis of the aging mechanism of lithium-ion batteries based on the incremental capacity curve peak decomposition method[C]// 2023 6th International Conference on Energy, Electrical and Power Engineering (CEEPE), May 12-14, 2023, Guangzhou, China. New York: IEEE, 2023: 1632-1637.
- [17] XU Z F, CHEN Z W, YANG L, et al. State of health estimation for lithium-ion batteries based on incremental capacity analysis and Transformer modeling[J]. Applied Soft Computing, 2024, 165: 112072.
- [18] LI X, YU H L, WANG J C, et al. Prediction of state-of-health of lithium-ion battery based on CEEMDAN-SG-LSTM combined model[J]. Materials Today Sustainability, 2024, 28: 100999.
- [19] WANG Q, GU Y D, ZHU T, et al. SOH estimation of lithium batteries based on ICA and WOA-RBF algorithm[J]. Energy Engineering, 2024, 121(11): 3221-3239.
- [20] YANG B, QIAN Y C, LI Q, et al. Critical summary and perspectives on state-of-health of lithium-ion battery[J]. Renewable and Sustainable Energy Reviews, 2024, 190: 114077.
- [21] DING P, LIU X J, LI H Q, et al. Useful life prediction based on wavelet packet decomposition and two-dimensional convolutional neural network for lithium-ion batteries[J]. Renewable and Sustainable Energy Reviews, 2021, 148: 111287.
- [22] LIAO L, YANG D, LI X B, et al. Fault diagnosis of lithium-ion batteries based on wavelet packet decomposition and Manhattan average distance[J]. International Journal of Green Energy, 2024, 21(12): 2828-2842.
- [23] HASSANAT A, ALMOHAMMADI K, ALKAFaweEN E, et al. Choosing mutation and crossover ratios for genetic algorithms: a review with a new dynamic approach[J]. Information, 2019, 10(12): 390.
- [24] MA R H, YUAN Y L. Remaining useful life prediction framework of equipment based on improved golden jackal algorithm assisted-LSTM[J]. AIP Advances, 2024, 14: 015102.



第一作者:黄亮亮(1979—),男,高级工程师,研究方向为新能源汽车火灾调查。E-mail:54196869@qq.com。



通信作者:曾建邦(1981—),男,博士,副教授/研究员,研究方向为新能源汽车大数据技术,动力电池主被动安全技术,电化学能源(如锂离子电池、锂空气电池和燃料电池等),多尺度强化传热传质技术。E-mail:jbzeng68@sina.com。

(责任编辑:李根)