

分类号: TP183

单位代码: 10363

密 级: 公开

学 号: 2220320176

题 目 锂离子电池的健康状态预测方法研究

英文并列

题 目 Research on state of health prediction methods for
lithium-ion batteries

学生姓名: 何洪伟

指导教师: 郭黎 教授

专 业: 电子信息

研究方向: 人工智能

论文答辩日期: 2025/05/06

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：

何洪伟

日期：2025 年 5 月 13 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：

何洪伟

指导教师签名：

邓静

日期： 2025 年 5 月 13 日

日期： 2025 年 5 月 13 日

锂离子电池的健康状态预测方法研究

摘 要

锂离子电池因其储能密度高和循环寿命较长等主要优点，广泛应用于电动汽车以及储能设备等领域。电池管理系统（Battery Management System, BMS）是保障电池高效安全运行的关键所在。健康状态（State of Health, SOH）是 BMS 重要参数，是衡量锂离子电池当前可用容量的重要标准。准确有效的 SOH 预测有助于电池的日常维护管理和老化电池的及时更换，对提高电池使用寿命和保障系统安全意义非凡。

本论文以锂离子电池为研究对象，选用美国国家航空航天局的 NASA 电池数据集和马里兰大学高级生命周期工程中心的 CALCE 电池数据集共两个公开锂离子电池老化实验数据集，对锂电池的 SOH 预测方法展开研究。主要工作包括以下三个方面：

(1) 针对传统锂离子电池健康因子（Health Indicator, HI）存在单一和与容量相关性较低的问题，本文通过遍历法从不同角度选取多组强相关 HI，全面表征锂离子电池退化程度。从电池循环老化数据中挖掘并构造六组 HI，分别为恒流充电时间、等上升电压充电时间、等充电时间上升电压、电压曲线覆盖面积、等电压下降放电时间、IC 曲线覆盖面积，并通过遍历法确定最佳特征区间。采用 Person 相关性方法和 Spearman 相关性方法共同分析各 HI 与容量之间的相关性。

(2) 针对单一数据驱动模型存在锂离子电池 SOH 点预测精度较低且有效预测周期较短的问题，提出 1DCNN-Informer 框架实现准确且长期有效的 SOH 点预测。通过 Informer 模型捕捉锂离子电池长循环使用寿命时间序列中 HI 的长期依赖关系。为避免其少训练样本预测能力的不足和提高预测精度，引入一维卷积神经网络（One-Dimensional Convolutional Neural Network, 1DCNN）预测当前时间步的 SOH。为避免传统基于验证集下 RMSE 自适应权重，偏向短期预测效果更好的单一模型的问题，引入超参数 λ 调整权重分配，可以根据实际任务需要调整单一模型权重。将 1DCNN 当前时间步和 Informer 模型的下一时间步的预测结果融合，可有效提高所提框架下一时间步的 SOH 点预测精度。所提框架均进行了全面对比验证，证明所

提框架 SOH 点预测的准确性。

(3) 容量再生现象对锂离子电池 SOH 准确预测干扰严重, 针对传统容量解耦预测方法不能有效地应用于 SOH 周期预测的问题, 提出 CEEMDAN-LSTM 框架并关注再生容量的快速下降过程, 实现准确的 SOH 周期预测。通过完全自适应噪声集合经验模态分解 (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMDAN) 解耦电池原始容量退化曲线, 采用 Person 相关性方法和 Spearman 相关性方法共同分析各本征模态函数 (Intrinsic Mode Functions, IMF) 与容量之间的相关性, 剔除不相关的高频子序列。通过构造双向长短期记忆网络 (Bidirectional Long Short-Term Memory, Bi-LSTM) 和长短期记忆网络 (Long Short-Term Memory, LSTM) 分别预测残差分量和相关高频子序列, 最后将预测结果相加集成获得最终 SOH 周期预测结果。所提框架均进行了全面对比验证, 证明所提框架 SOH 周期预测的准确性。

关键字: 锂离子电池, 健康状态, 健康因子, SOH 预测, LSTM, Informer

RESEARCH ON STATE OF HEALTH PREDICTION METHODS FOR LITHIUM-ION BATTERIES

ABSTRACT

Lithium-ion batteries are widely used in electric vehicles and energy storage devices due to their high energy storage density and long cycle life. Battery management system (BMS) plays an important role in the efficient, safe, and reliable operation of batteries. State of health (SOH) is an important parameter of BMS, an important criterion to measure the current usable capacity of Li-ion batteries. Accurate and effective SOH prediction helps in daily maintenance management and timely replacement of aging batteries, which is of great significance in improving battery life and ensuring system safety.

The thesis takes lithium-ion batteries as the research object and selects the NASA battery dataset from the national aeronautics and space administration (NASA) and the CALCE battery dataset from the center for advanced life cycle engineering (CALCE) of the university of Maryland, which are two publicly available experimental datasets on aging of lithium-ion batteries, to study the SOH prediction method of the batteries. The main work includes the following three aspects:

(1) Aiming at the problems of single health indicator (HI) and low correlation with capacity of traditional lithium-ion batteries, this paper selects multiple sets of strong correlation HIs from different perspectives by traversal method to comprehensively characterize the degradation degree of lithium-ion batteries. Six groups of HI are mined and constructed from the battery cycling aging data, which are constant-current charging time, equal-rise-voltage charging time, equal-rise-voltage charging time, voltage curve

coverage area, equal-voltage-drop-discharging time, and IC curve coverage area, and the optimal characteristic intervals are determined by the traversal method. The correlation between each HI and capacity was analyzed using the Person correlation analysis and the Spearman correlation method.

(2) Aiming at the problems of low SOH point prediction accuracy and short effective prediction period of lithium-ion batteries in a single data-driven model, the 1DCNN-Informer framework is proposed to realize accurate and long-term effective SOH point prediction. Capturing the long-term dependence of HI in long cycle lifetime series of lithium-ion batteries by Informer modeling. To avoid the lack of prediction ability with few training samples and to improve the prediction accuracy, a One-dimensional convolutional neural network (1DCNN) is introduced to predict the SOH at the current time step. To avoid the problem of traditional RMSE adaptive weights based on the validation set, which favors a single model with a better short-term prediction effect, hyperparameter λ is introduced to adjust the weight allocation, which can adjust the single model weights according to the actual task needs. Fusing the prediction results of the current time step of 1DCNN and the next time step of the Informer model can effectively improve the SOH point prediction accuracy of the next time step of the proposed framework. The proposed frameworks are comprehensively compared and validated to prove the accuracy of the SOH point prediction accuracy of the proposed frameworks.

(3) The capacity regeneration phenomenon interferes seriously with the accurate SOH prediction of lithium-ion batteries, and in response to the problem that the traditional capacity decoupling prediction method cannot be effectively applied to the SOH cycle prediction, the CEEMDAN-LSTM framework is proposed and focuses on the rapid decline process of regenerated capacity to realize the accurate SOH cycle prediction. The original capacity degradation curve of the battery is decoupled by complete

ensemble empirical mode decomposition with adaptive noise (CEEMDAN), and analyzed by Person correlation analysis and Spearman correlation method. The correlation between each intrinsic mode function (IMF) and capacity was analyzed by Person's correlation analysis and Spearman's correlation method, and the irrelevant high-frequency subsequence was eliminated. Bidirectional long short-term memory (Bi-LSTM) and long short-term memory (LSTM) were constructed to predict the residual components and the related high-frequency subsequences, and then the predicted results were summed up to obtain the final SOH cycle prediction results. The final SOH cycle prediction result is obtained by summing up the prediction results and integrating them. The proposed frameworks are comprehensively compared and validated to demonstrate the accuracy of SOH cycle prediction.

Hongwei He (Electronic Information)

Supervised by Prof. Li Guo

KEY WORDS: Lithium-ion battery; State of health; Health indicator; SOH prediction; LSTM; Informer

目录

摘 要	I
ABSTRACT	III
第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 健康状态定义标准	2
1.3 锂离子电池 SOH 预测国内外研究现状	2
1.3.1 基于模型的方法	3
1.3.2 基于数据驱动的方法	4
1.3.3 基于融合的方法	6
1.3.4 基于容量再生问题的方法	7
1.4 研究现状分析	8
1.5 论文的主要研究内容和章节安排	9
第 2 章 锂离子电池老化特性分析	12
2.1 锂离子电池基本结构和工作原理	12
2.2 锂离子电池容量衰退的影响因素	13
2.2.1 内部因素	14
2.2.2 外部因素	15
2.3 锂离子电池公开数据集	15
2.3.1 NASA 电池数据集	16
2.3.2 CALCE 电池数据集	16
2.4 锂离子电池容量再生现象原因	18
2.4.1 容量再生	18
2.4.2 再生容量的快速退化	19
2.5 本章小结	19
第 3 章 多特征健康因子挖掘与容量再生特性分析	20
3.1 基于充电阶段的 HI 提取	20
3.1.1 基于横流充电电流曲线的 HI 提取	20
3.1.2 基于横流充电电压曲线的 HI 提取	21
3.2 基于放电电压曲线的 HI 提取	23
3.3 基于 IC 曲线的 HI 提取	25
3.4 相关性分析	27
3.4.1 PCC	27

3.4.2 SCC.....	28
3.5 容量再生现象特性分析	29
3.6 本章小结	31
第 4 章 基于 1DCNN-Informer 框架的锂离子电池 SOH 点预测.....	32
4.1 1DCNN 原理.....	32
4.2 Transformer 和 Informer 原理	34
4.2.1 Transformer	34
4.2.2 Informer.....	37
4.3 锂离子电池 SOH 点预测框架	39
4.4 锂离子电池 SOH 点预测实验	41
4.4.1 SOH 点预测结果	42
4.4.2 不同序列预测模式比较	46
4.4.3 集合框架不同权重的比较	47
4.4.4 所提框架与其它框架 SOH 点预测比较	48
4.5 本章小结	49
第 5 章 基于 CEEMDAN-LSTM 框架的锂离子电池 SOH 周期预测	50
5.1 CEEMDAN 原理	50
5.2 LSTM 原理	52
5.2.1 LSTM	52
5.2.2 Bi-LSTM	53
5.3 锂离子电池 SOH 周期预测框架	54
5.4 锂离子电池 SOH 周期预测实验	56
5.4.1 容量解耦与相关性分析	56
5.4.2 SOH 周期预测结果	59
5.4.3 与全 IMFs 周期预测实验结果比较	64
5.4.4 所提框架与其它框架 SOH 预测比较	65
5.5 本章小结	66
第 6 章 总结与展望	67
6.1 总结	67
6.2 后续工作展望	68
参考文献.....	69
攻读学位期间发表的学术论文目录.....	78
致谢	79

第 1 章 绪论

1.1 研究背景及意义

“绿水青山就是金山银山”。面对日益加剧的能源危机和碳排放量过高引发的环境污染问题，我国致力于开发清洁能源及储能设备，以加速经济转型和社会的可持续发展，实现碳中和目标^[1]。锂离子电池拥有诸多优点，例如：储能密度高、充放电效率高、自放电率低、循环寿命长等，在便携式电子设备、新能源电动汽车、储能设备、航天和航空等领域应用广泛^[2]。传统燃油车汽车产业为排碳大户，在产业结构转型过程中扮演重要角色。因锂离子电池成本持续下降，电动汽车引来发展浪潮。我国在《新能源汽车产业发展规划（2021-2035 年）》中指出，到 2035 年末新能源汽车将成为主流车辆，公共领域用车实现全面电动化。目前，锂离子电池在新能源汽车领域中占比最大，常见的有磷酸铁锂离子电池和三元锂离子电池。它们提供更大的能量密度同时充电时间较短，满足客户的日常使用需求^[3]。

然而，锂电池在日常充放电和存储过程中，由于内部副反应的持续发生，不可避免的出现了老化现象，通常表现为电池容量缓慢衰减和内阻缓慢上升^[4]。电池老化不仅会缩短电动汽车的续航里程，还会引起电池工作的安全边界发生变化，发生电解质泄漏和微短路，引发热失控等严重安全隐患问题^[5]。因此，为了保证锂电池高效安全的可靠运行，需要一套精准有效的电池管理系统（Battery Management System, BMS），保障电池的预测和健康管理^[6]（Prognostics and Health Management, PHM）。

BMS 主要关注的状态参数包括荷电状态（State of Charge, SOC）和健康状态（State of Health, SOH）等。SOH 作为表征电池寿命的重要指标反映了电池的老化程度。准确有效的 SOH 预测技术有助于电池的日常维护管理和老化电池的及时更换，对提高电池使用寿命和保障系统安全意义非凡。同时，也有助于及时退役的电动汽车锂离子电池等实现梯次利用^[7]。

SOH 呈现高度非线性和强耦合的趋势，对于 SOH 的长期有效的预测十分具有挑战性。利用可测量电池数据对 SOH 进行准确预测成为当下研究热点。同时，在电池退化过程中，容量下降到阈值时常发生容量再生现象，电池在下一次循环周期的

容量高于当前循环周期，并在之后快速退化^[8]。对于 SOH 准确预测而言，容量再生现象的干扰显著。

1.2 健康状态定义标准

锂离子电池在充放电循环过程中，电池内部会发生不可逆的物化学变化，具体表现为电池容量的持续下降和内阻升高。对于电动汽车而言，锂离子动力电池出厂额定容量为 100%，当锂电池容量低于 80% (即 SOH 小于 80%)，即可认为电池不能正常工作，应及时更换^[9]。

常用于表征电池 SOH 的重要指标是容量和内阻^[10]。从容量的角度 SOH 定义为电池当前最大可用容量占电池出厂额定容量的百分比，根据容量定义的 SOH 计算公式为：

$$SOH = \frac{Q_{now}}{Q_{rated}} \times 100\% \quad (1-1)$$

式中， Q_{now} 为电池目前最大可用容量； Q_{rated} 为电池出厂额定容量。

电池内阻增大是电池老化的直接原因，根据内阻定义的 SOH 计算公式为：

$$SOH = \frac{R_{EOL} - R_{now}}{R_{EOL} - R_{rated}} \times 100\% \quad (1-2)$$

式中， R_{EOL} 为电池循环寿命结束时的内阻， R_{rated} 为电池出厂额定内阻， R_{now} 为电池目前状态下的内阻。

通过容量定义的 SOH 方法更加直接且计算简单。通过内阻定义的计算方法不仅不能提前得知电池在循环寿命结束时的内阻，而且在实际应用中难以准确计算出当前锂离子电池的内阻大小，计算误差较大。通过容量计算 SOH，当前最大可用容量可根据电池易获取的充放电数据测量获得。因此，本文将以容量定义方法计算锂离子电池 SOH，进行后续的 SOH 预测实验。

1.3 锂离子电池 SOH 预测国内外研究现状

近年来，国内外研究人员针对锂电池 SOH 预测展开大量实验。预测方法可以分为三个方向：一为基于模型的方法，二为基于数据驱动方法，三为基于融合的方法^[11]。图 1-1 所示为锂电池 SOH 预测方法总结。

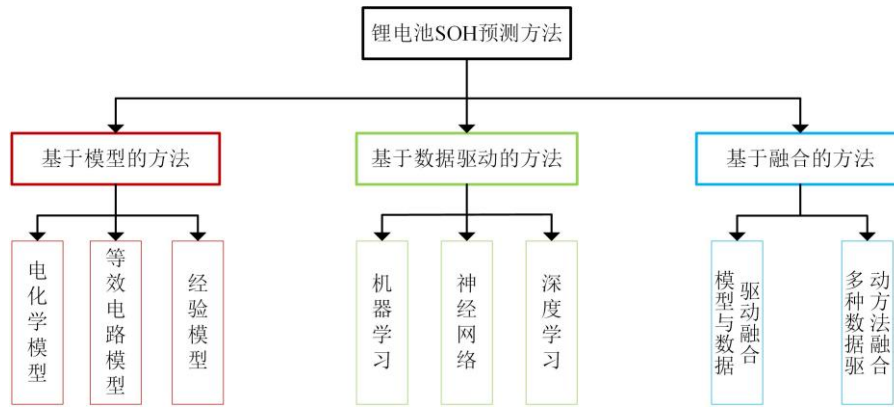


图 1-1 锂电池 SOH 预测方法总结

1.3.1 基于模型的方法

物理模型主要是用于描述电池内部复杂的扩散、迁移和物理化学反应。基于物理模型的方法通过搭建状态空间方程表征电池内部电化机理和外部电行为特征，往往结合滤波技术更新模型参数。常见的物理模型有电化学模型（Electrochemical Model, EM）和等效电路模型（Equivalent Circuit Model, ECM）。经验模型主要是基于大量实验数据和经验知识进行构建。

(1) 电化学模型

电化学模型通过建立动力学偏微分方程，表征锂离子电池内部电化学反应速率、离子浓度等，揭示电池内在反应机理（锂离子的嵌入和拖嵌、固体电解质中间相界面变化、活性材料损耗等）。德国亚琛工业大学林长伟等人提出了一种基于分数阶降阶的电化学模型，把锂离子电池内阻增加和容量衰减作为 SOH 指标。避免了参数量较少而导致的高计算成本，在考虑锂电池不同容量老化水平、动态负载取消和初始化误差下保持了高精度 SOH 估计^[12]。上海交通大学高一钊等人提出一种降阶电化学模型用于锂离子电池 SOH 预测。采用极简电化学模型在线估计锂离子损失、活性材料损失和电阻增加，通过最小二乘识别模型参数以评估锂电池 SOH，平均误差在 2% 左右^[13]。电池的阻抗包含有关电池内部结构和健康状态的丰富信息。加拿大麦克马斯特大学张文林等人提出了基于电化学阻抗谱（Electrochemical Impedance Spectroscopy, EIS）测量结合无迹卡尔曼滤波的锂离子电池 SOH 估计，减少数据要求并提高稳定性。状态预测模型会根据电池已经历的循环次数来估算 SOH，预测模型则会依据所估算出的 SOH 来预测预期的健康因子（Health Indicator, HI），滤波器会基于预测 HI 和实测 HI 之间的差值对估算值进行修正，SOH 估计误

差低于 2%^[14]。

(2) 等效电路模型

等效电路模型使用电阻、电容等电路元件，通过捕获电池的电流、电压和瞬态行为描述电池。常见的有一阶 RC 等效电路模型和二阶 RC 等效电路模型等^[15]。西南大学陈璐等人一种基于温度相关二阶 RC 等效电路模型的改进灰狼优化和双自适应扩展卡尔曼滤波算法，以减少噪声干扰和测量缺陷的影响。在 $-5^{\circ}\text{C} \sim 35^{\circ}\text{C}$ 的复杂工况下验证了算法的可行性，SOH 估计绝对误差低于 1%^[16]。天津大学韩乔妮等人提出基于一阶 RC 等效电路模型和改进高斯过程回归的锂离子电池 SOH 预测框架。该框架通过提取容量增量曲线中的 HI，通过主成分分析，然后建立锂电池老化的改进高斯过程回归模型进行 SOH 预测。并建立锂电池一阶状态空间模型，结合改进粒子滤波算法实现 SOH 的长期预测^[17]。上海交通大学李超凡等人提出基于改进等效电路模型的电化学阻抗谱锂电池 SOH 预测方法。构建附加电容器等效电路模型，根据电化学阻抗谱数据确定了模型参数并作为高斯过程回归 (Gaussian Process Regression, GPR) 的输入并估计 SOH，均方根误差 (Root Mean Square Error, RMSE) 为 1.77%^[18]。

(3) 经验模型

基于经验模型的锂电池 SOH 预测关注电池重要的寿命衰退信息，研究各种性能参数与老化周期的变化趋势来预测电池寿命。西安交通大学吕海潮等人将锂离子电池的老化过程分为固体电解质界面膜增长和锂沉积机制主导的两个阶段，快速识别老化过程两阶段的转折点，采用两步非线性拟合方法建立了基于锂电池电化学过程模型的两阶段老化过程经验模型。实验表明，可以有效预测不同参数条件下电池的两阶段老化过程^[19]。英文考文垂华威大学刘凯龙等人考虑锂电池退化的电化学和经验知识用于改进 GPR，能够在一步和多步预测锂电池容量方面取得令人满意的结果^[20]。马里兰大学何伟等人提出一种基于 Dempster-Shafer 和贝叶斯蒙特卡洛的方法估计锂离子电池 SOH。开发了一个基于锂离子电池物理降解行为的经验模型，预测很好地捕捉了电池容量的非线性衰减趋势^[21]。

1.3.2 基于数据驱动的方法

随着人工智能的快速发展，数据驱动技术突破了复杂非线性系统传统建模方法难以应用的困境，已经成为锂离子电池 SOH 预测的主流研究方向^[22]。常见的数据

驱动法可分为三大类：机器学习、神经网络和深度学习。

(1) 机器学习

机器学习是让计算机系统通过对海量数据的观察分析，挖掘数据规律，并利用这些规律来改进自身性能或进行预测和决策的过程^[23]。传统的机器学习模型包括支持向量机 (Support Vector Machine, SVM)、极限学习机 (Extreme Learning Machine, ELM) 和 GPR 等。英国克兰菲尔德大学罗伟等人提出一个基于 SVM 和电化学阻抗谱的离线框架用于锂离子电池老化情况分析。不仅校准了锂电池 SOH 值还分析锂离子电池的老化机制^[24]。吉林大学段文贤等人提出一种基于变遗忘因子在线序贯 ELM 的锂离子电池 SOH 预测方法。采用改进的鲸鱼优化算法优化模型参数，极端随机树算法获取与可用容量高度相关的特征，以降低模型复杂度并提高估计精度，SOH 估计 RMSE 精度低于 0.15%^[25]。南京林业大学李芳等人提出一种基于堆叠集成稀疏 GPR 的锂离子电池快速充电健康状态估计方法，解决单个 GPR 可能会在不同的快速充电电池之间产生性能差异，并且会消耗大量时间的问题。在局部放电判断中提取 HI，并通过遗传算法优化堆叠集成稀疏 GPR 核函数参数。多种工况实验下 SOH 估计的 RMSE 在 1.21% 以内^[26]。

(2) 神经网络

神经网络由众多的神经元可调的连接权值连接而成，具有大规模并行处理、良好的自学习能力等优势。传统的神经网络模型包括多层感知机 (Multi-Layer Perceptron, MLP)、反向传播算法 (Back-Propagation, BP) 和各自人工神经网络 (Artificial Neural Network, ANN)。德国柏林大学弗洛里安·热普卡等人引入了一种考虑滞后序列或时间序列数据的新方法，以加快学习过程，通过 MLP 模型准确预测电池 SOH，均方误差为 2.95%^[27]。太原科技大学吴青峰等人提出一种基于 BP 神经网络的锂电池健康状态估算方法。以电压、输出电流、容量和等压降放电时间四个 HI 作为 BP 神经网络的输入值，引入模拟退火算法优化 BP 神经网络收敛效率低的问题，提高了 SOH 预测精度^[28]。清华大学王欢等人提出一种用于锂离子电池健康状态估计的仿生尖峰时空注意力框架。为解决传统 ANN 存在数据依赖性、高能耗和认知能力不足等问题，提出了一种生物启发的尖峰时空注意力神经网络框架，赋予其高效的时空特征处理能力和低功耗。利用全生命周期 EIS 进行锂电池 SOH 监测，框架有助于精确建模 EIS 数据和 SOH 之间的非线性关系，并实现有竞争力

的性能^[29]。

(3) 深度学习

相对于传统的机器学习方法,深度学习能从原始数据的低级属性中学习并抽取更高层次的有意义性特征^[30]。在锂离子电池 SOH 预测领域,热门方向有卷积神经网络 (Convolutional Neural Network, CNN)、长短期记忆网络 (Long Short-Term Memory, LSTM)、门控循环单元 (Gated Recurrent Unit, GRU) 和 Transformer。西安交通大学柯雷思等人提出一种基于特征优化和 CNN 的锂电池 SOH 预测方法。使用主成分分析和模拟退化算法构建了 HI 优化器,并提供 CNN 估计 SOH, RMSE 关键指标增加了 30%^[31]。台湾科技大学王福冠等人提出了一种改进的时间融合转换器方法,该方法基于双向长短期记忆编码器-解码器层,引入基于树结构 Parzen 估计器的新型可解释超参数调整方法,提高了 SOH 预测精度^[32]。印度悉达多理工学院 K. Dhananjay Rao 等人提出一种基于最优调优的 GRU 锂离子电池健康状态预测方案。使用高级傅里叶重采样方法对数据进行重采样,并通过最佳超参数调整的 GRU 预测 SOH, RMSE 为 0.3%^[33]。印度尼西亚穆罕默德·里夫奇·福兹等人提出了一种使用具有季节性和增长嵌入的指数平滑的 Transformer 预测锂离子电池的 SOH 方法,与多种 Transformer 变体相比较,突出了增长和季节性嵌入在提高 SOH 预测准确性方面的价值^[34]。

1.3.3 基于融合的方法

融合方法出发点是解决单一模型的缺点,并整合各模型的优势,提高锂离子电池 SOH 预测精度。基于融合的方法可分为基于模型与数据驱动融合和多种数据驱动融合两大类。

(1) 基于模型与数据驱动融合的方法

模型法预测精度不高,但可解释电池内在机理;而数据驱动法准确性和泛化性较高,但没有机理性描述。因此,基于模型与数据驱动融合的方法,同时兼顾精度、泛化性及模型的可解释性,是主流研究方向之一。郑州大学秦东辰等人提出了一种在电化学模型和双层前馈神经网络之间交替的复合渐进方法,以起到消除锂离子电池容量跳水现象对 SOH 预测精度差的影响。与单一复合和数据驱动方法相比,复合渐进方法在容量跳水间隔内将 RMSE 提高 50% 和 36%,在整个周期内提高了 38% 和 41%^[35]。上海工程技术大学孙国庆等人将等效电路模型与 LSTM 相结合,利用

锂离子电池的电压反弹特性确定二阶 RC 等效电路的蓄电池的参数，从模型和数据集中提取的特征被用作 LSTM 神经网络的输入来预测 SOH。结果表明，该方法的 RMSE 为 0.79%^[36]。哈尔滨理工大学梁金莲等人提出一种集成单粒子模型和神经网络的混合模型方法预测锂离子电池的 SOH。从改进的单粒子模型取了七个电化学特征，构建了一个结合时间卷积网络和双向长短期记忆网络的混合模型预测 SOH，预测误差 RMSE 低于 1.085%^[37]。

(2) 基于多种数据驱动融合的方法

多种数据驱动方法组合，使单一数据模型存在的不足得以补充，可提升 SOH 预测精度和泛化性，是未来的主流研究方向之一。北京航空航天大学程倩等人提出一种用于动态工作条件下锂离子电池多态估计的 CNN-SAM-LSTM 混合神经网络，结合了卷积神经网络、自注意力机制和长短期记忆神经网络，共同预测 SOH。此外，还开发了一个考虑同方差不确定性的联合损失函数，以优化与三个状态参数相关的训练损失的权重调整。应用于随机分割的数据时，所提出的模型表现出稳健性，有效地处理了与随机排放段的偏差^[38]。同济大学岳翔等人提出了一种基于动态放电曲线的半监督学习进行锂离子电池 SOH 预测方法。使用少量定期校准数据从电池动态放电曲线中提取特征进行容量估计，基于双向门循环单元和结构化核插值 GPR 的半监督深度学习方法，RMSE 误差低于 1.91%^[39]。中北大学牟桂香等人提出一种基于特征融合的堆叠模型的锂电池容量预测方法。使用 PCA 融合健康因素以降低模型的计算复杂性，鉴于单个模型在不同数据集上的性能差异，提出由 GPR、SVR 和 LSTM 三种不同结构的机器学习模型组成堆叠模型，以增强堆叠模型的鲁棒性和泛化能力，实验证明了集成方法的准确性和泛化能力^[40]。桂林电子科技大学李紫阳等人提出一种基于 Bi-LSTM-Transformer 的锂电池 SOH 预测方法。从恒流充电阶段的电压曲线中提取出与 SOH 高度相关的四个特征，利用提出框架映射 SOH，预测误差都在 0.6% 以内^[41]。

1.3.4 基于容量再生问题的方法

锂离子电池中的存储容量随着充放电循环次数的增加而缓慢降低，但在充放电过程中处于静态状态时，反应产物会消散，导致锂离子电池在降解过程中的电化学性能与上一个周期相比相对恢复，从而增加下一个周期的可用容量^[42]。容量再生直接影响短期内锂电池的 SOH，使容量在短周期内出现较大回弹，导致短期 SOH 预

测偏离实际目标^[43]。此外，容量再生问题也会严重影响到 RUL 的长期精确预测，引起较大预测误差^[44]。

对于锂电池 SOH 准确预测而言，需特别注意容量再生现象造成的干扰，这也是近年来热点方向之一。合肥工业大学王晓华等人提出一种基于混合模型 MFE-GRU-TCA 的锂离子电池 SOH 预测方法。提出时间卷积注意力 TCA 模块，它可以自适应地处理历史电池特性对当前容量信息的影响。该模块还考虑了时间方面，减少了容量再生对 SOH 预测的影响。实验结果表明，RMSE 降低了 5.53% - 33.44%^[45]。重庆理工大学袁子凡等人提出一种基于变分模态分解去噪的混合神经网络预测锂电池 SOH 的方法。原始容量数据经过变分模态分解平滑过程，以消除容量再生和部分噪声信号。CNN 用于特征提取，采用修改后的 Transformer 模型来捕获时间序列中的内在相关性，并将特征映射到未来的 SOH 值^[46]。合肥工业大学刘征宇等人提出一种将经验模态分解 (Empirical Mode Decomposition, EMD) 与 GRU 和差分自回归移动平均模型 (Autoregressive Integrated Moving Average Model, ARIMA) 相融合的锂电池 SOH 预测方法。首先，利用 EMD 将锂电池原始 SOH 序列进行多尺度分解，并通过计算分解子序列的连续均方误差定量出高低频分界点；然后，GRU 预测具有强烈数据波动的高频子序列，ARIMA 预测剩余的低频子序列和残差；最后，将每个子序列的预测结果进行叠加以获得最终预测结果。试验结果表明融合模型具有高预测精度，可捕捉局部容量再生特性^[47]。武汉科技大学丁国荣等人提出一种基于布谷鸟搜索优化变分模态分解和 GRU 的锂电池容量预测方法^[48]；浙江理工大学潘海鹏等人提出了一种新的锂离子电池衰减模型和改进的 SOH 估计粒子滤波算法^[49]；长安大学吴春玲等人提出一种基于变分模态分解和蜣螂优化的支持向量回归的锂离子电池 SOH 预测模型^[50]；北京航空航天大学崔玉轩等人提出使用自适应噪声的完整集成经验模态分解，SVM 和 LSTM 预测锂离子电池 SOH^[51]，这些研究方法都与文献^[47]相类似，本质是将原始容量退化过程解耦，分别预测残差信号和本征模态函数。

1.4 研究现状分析

综上所述的近年来锂离子电池 SOH 预测国内外现状，对 SOH 预测存在的问题进行分析总结。基于物理模型的方法，在电池建模准确的前提下，可以准确的预测

电池 SOH。但是，需要对电池的电化学特性有着全面和深入的了解，所搭建模型不仅十分复杂而且计算成本高，耗时长。模型的参数辨识不仅仅依赖于滤波算法的收敛性，同时受环境温度、电池放电倍率等条件影响大，参数的动态辨识准确度不能得到有效保障。经验模型预测精度受衰减因子影响，其无法包括电池的所有退化路径，所以基于经验模型的方法鲁棒性较低。基于数据驱动的方法只需通过测量电池工作电流、电压和温度等数据，提取与 SOH 高度相关的 HI，并将 HI 作为历史训练数据构建预测模型，从而实现 SOH 的准确预测。相比于基于物理模型的方法，数据驱动法不需要对锂电池内部复杂的电化学机理进行分析与建模，具有较高的可迁移性。但单一数据驱动方法存在无法避免的局限性，往往对 SOH 特征选择具有较高的要求而且适用场景有限。因此本文采用基于多数据驱动模型集成的方法对锂电池 SOH 预测展开研究。

锂离子电池 SOH 预测已取得卓有成效的研究成果，但仍有一些问题值得思考：

(1) 基于数据驱动的方法需从大量老化数据中提取能反映电池退化的 HI，从不同数据维度挖掘多个健康样子可以充分利用数据，避免传统数据驱动方法需要对特征进行二次优化过程。

(2) 锂离子电池循环寿命较长，不少研究只在单一且循环寿命较短的数据集进行验证，无法肯定模型的 SOH 长期预测的准确性。寻求锂离子电池 SOH 短期和长期的有效且准确的预测意义重大。

(3) 容量再生问题是影响锂电池 SOH 预测准确性的最主要干扰因素。研究人员往往关注准确的预测容量再生趋势，而忽视了再生容量的快速退化特性。同时，传统的基于容量再生问题解耦预测是基于实验室理想数据下的，无法验证实际应用的可行性。

(4) 锂离子电池 SOH 的预测往往只关注点预测。考虑到容量退化速率的不一致性，开展必要的 SOH 周期预测更有应用价值。

1.5 论文的主要研究内容和章节安排

为探讨解决在上述章节中锂离子电池 SOH 预测存在的问题，本文考虑采用于多数据驱动模型集成的方法对锂电池 SOH 预测展开研究。一方面基于锂离子电池多维特征，挖掘健康因子（Health Indicator, HI）并结合一维卷积神经网络（One-

Dimensional Convolutional Neural Network, 1DCNN) 和 Informer 模型展开长期且准确的 SOH 点预测; 另一方面考虑锂离子电池退化过程中容量再生问题和退化速率的不一致性, 采用完全自适应噪声集合经验模态分解 (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMDAN) 解耦原始容量退化曲线并计算皮尔斯相关系数 (Pearson Correlation Coefficient, PCC) 和斯皮尔曼相关系数 (Spearman Correlation Coefficient, SCC) 以剔除不相关的高频子序列。分别构造双向长短期记忆网络 (Bidirectional Long Short-Term Memory, Bi-LSTM) 和长短期记忆网络 (Long Short-Term Memory, LSTM) 分组进行预测, 最后将预测结果相加获得最终 SOH 周期预测结果。各大章节详细工作内容如下:

第一章 绪论。本章首先阐述了锂离子电池 SOH 预测的研究背景及意义, 然后介绍了两种常用的 SOH 定义, 接着对近年来 SOH 预测的国内外研究现状进行全面具体的阐述, 最后对 SOH 预测研究现状进行总结并提出一些潜在面临探讨的问题。

第二章 锂电池老化分析。本章首先介绍了锂离子电池基本结构和工作原理, 然后从内部因素和外部因素两个方面介绍了影响电池容量衰退的原因, 接着介绍了两个锂离子电池公开数据集, 分别是 NASA 数据集和 CALCE 数据集, 最后分析了锂离子电池退化过程中出现容量再生现象和再生容量快速退化的原因。

第三章 多特征健康因子挖掘与容量再生特性分析。通过分析锂离子电池充放电循环老化曲线特性, 从充电电流曲线、充电电压曲线、放电电压曲线以及容量增量曲线中提取六组 HIs, 并通过 PCC 和 SCC 定量分析所提 HIs 与原始容量的相关性。还对容量再生特性进行分析, 指出传统实验室环境下针对容量再生问题提出的分解-集成预测方法并不能有效的适用实际 SOH 周期预测应用所需。

第四章 基于 1DCNN-Informer 框架的锂离子电池 SOH 预测。本章主要介绍了 1DCNN、Transformer、Informer 原理。基于锂电池长循环使用寿命特点, 以长序列时序预测性能优异的 Informer 模型为出发点。为解决 Informer 模型在少样本下预测能力的不足和模型存在预测波动性较大的问题, 引入 1DCNN 预测并修正 Informer 预测结果。为避免传统基于验证集下 RMSE 自适应权重偏向短期预测效果更好的单一模型的问题, 引入超参数 λ 调节权重分配。将 1DCNN 当前时间步和 Informer 模型的下一时间步的预测结果融合, 可有效提高所提框架下一时间步 SOH 点预测精度。在不同数据集实验和与其他框架对比, 验证所提框架的 SOH 点预测的准确性。

第五章 基于 CEEMDAN-LSTM 框架的锂离子电池 SOH 预测。本章主要介绍了 LSTM、Bi-LSTM、CEEMDAN 原理。基于传统的容量再生解耦-集成不能有效应用于 SOH 长周期预测，更加关注再生容量的快速下降过程。通过 CEEMDAN 解耦容量，经相关性分析舍弃不相关高频子序列。分别构造 Bi-LSTM 和 LSTM 预测残差分量和相关高频子序列，最后将预测结果相加集成获得最终 SOH 周期预测结果。在不同数据集实验和与其他框架对比，验证所提框架的 SOH 周期预测的准确性。

第六章 总结与展望。本章总结了本文完成的研究内容以及创新点并对本文存在的不足之处进行探讨，最后对锂离子电池 SOH 预测未来工作方向进行展望。

第2章 锂离子电池老化特性分析

锂离子电池拥有高储能密度、高充放电效率、低自放电率、长循环周期等优点，在便携式电子设备、储能设备、航天和航空等领域分布广泛。在新能源电动汽车领域，锂电池同样占据着主导地位^[52]。随着锂离子电池的日常充放电使用，其内部会发生不可逆的物理化学过程，电池性能下降，表现为容量和 SOH 的衰减。锂离子电池 SOH 是评估其性能老化的关键指标，因此对锂离子电池内部工作原理和容量衰退因素进行研究是保证 SOH 预测方法合理性的前提和关键之处。本章节对锂离子电池工作原理和容量衰退影响因素进行分析，并介绍了两种可靠的电池老化实验数据集，同时分析容量再生问题的原因，为后续锂离子电池 SOH 预测方法研究提供理论基础。

2.1 锂离子电池基本结构和工作原理

锂电池属于复合材料，基本结构主要由正负极材料、电解质、隔膜和外壳共五部分组成^[53]。正极材料通常为锂金属氧化物，常见的有锂铁磷酸盐 (LiFePO_4)、锂钴氧化物 (LiCoO_2) 和锂镍锰氧化物 (LiNiMnCoO_2)。正极材料是影响锂离子电池容量的关键部分，其在充放电循环过程中吸收和释放锂离子的能力决定了锂离子电池的基本性能。负极材料通常由碳材料石墨构成，锂离子在充放电过程中可以在其结构间隙中进行自由脱嵌而不破坏材料结构。电解质是电池中正负极之间的介质，起到传输锂离子的作用，通常由溶解在有机溶剂中的锂盐组成。选择合理的电解质是使锂离子电池具有高能量密度和稳定安全性能的关键。隔膜是一层绝缘材料，可防止电池正负极相互接触引起短路故障，同时应该足够薄，有利于锂离子的正常通过。外壳起到封装保护的作用，主要材质为铝，重量轻、硬度高、足够稳定。

锂电池的工作原理是在电池正负极间因锂离子浓度差驱动的迁移、脱嵌与嵌入过程。当锂电池连接外界电源充电时，正极材料中的锂离子在电场力的作用下，从正极晶格中脱嵌，进入电解液，形成可自由移动的锂离子。通过扩散和迁移，锂离子穿过隔膜微孔，向负极移动并嵌入负极材料的晶格结构中。随着充电的进行，足够多的锂离子嵌入负极，使得负极处于富锂状态，电池的充电容量也随之增加。当锂电池与外部电路连接形成闭合回路时，由于电池内部存在电位差，电子会从负极

流出，通过外部电路流向正极，从而形成电流，为外部设备提供电能。负极中的嵌入态锂离子会从负极材料的晶格中脱嵌，负极的锂离子浓度增加，并通过电解液和隔膜向正极方向迁移，重新嵌入到正极材料的晶格中。与正极附件的电子结合，形成锂化合物。随着放电的持续进行，负极中的锂离子不断脱嵌并向正极迁移，直到电池的电压下降到一定程度，或者锂离子几乎全部从负极脱出，此时电池放电结束。以钴酸锂电池为例，其工作原理如图 2-1 所示。

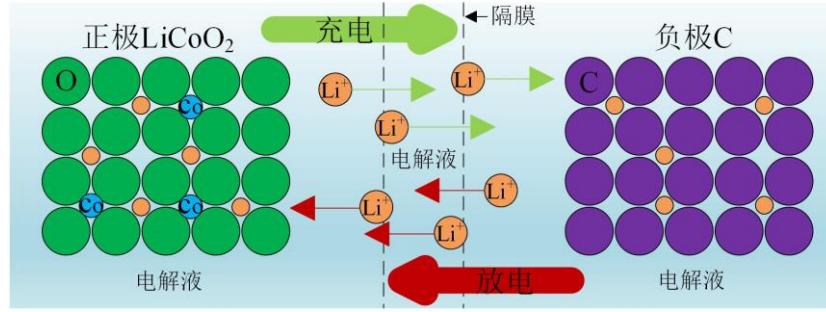
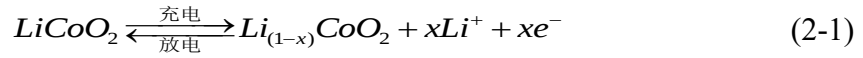


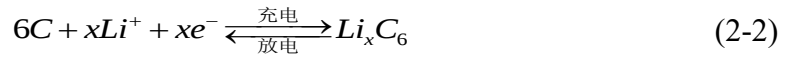
图 2-1 锂离子电池工作原理

电池内部的化学反应过程如下：

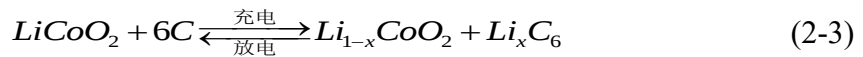
正极反应式为：



负极反应式为：



电池总反应式为：



2.2 锂离子电池容量衰退的影响因素

在日常的使用和存储过程中，由于电池内各个部位均会发生副反应，锂离子电池不可避免的出现老化现象，表现为电池容量的衰减和 SOH 的缓慢降低。影响锂离子电池 SOH 衰退的因素可以分为两个方面。一是内部因素，如析锂、固体电解质界面（Solid Electrolyte Interface, SEI）膜生长、电解液分解和电极活性材料的消耗等。二是外部因素，如环境温度、充放电倍率、过充与过放等。充分了解这些因素的影响过程，有利于对锂离子电池展开 SOH 预测和日常维护。

2.2.1 内部因素

(1) 析锂

析锂是指锂离子从电解液中以金属锂单质沉积到负极表面的现象。锂离子的交互保证了电池的充放电循环的正常运行，析锂会导致不可逆的锂离子总量减少，从而使电池可用容量降低。析锂是锂离子电池老化的主要因素之一，同时锂枝晶在生长过程中可能会刺穿隔膜，导致正负极直接接触引发电池内部短路，引发热失控等严重危害。文献^[54]表明在低温环境下，电池负极电位降低，锂离子在电池负极嵌入速率下降，此时在正常的充电电流下锂离子也可能来不及嵌入而发生析锂。

(2) SEI 膜生长

SEI 膜是指锂电池在首次充电时，负极材料表面与电解液发生系列复杂化学反应而产生钝化膜，将电解液与负极隔开，避免电解液墩、锂离子与负极表面材料持续消耗。文献^[55]表明在循环过程中，负极材料表面本身存在膨胀与破裂现象，导致其表面的 SEI 膜承受应力发生形变、变薄与破裂，导致 SEI 膜持续的形成与增长，增加了锂离子与电解液的消耗。

(3) 电解液分解

电解液起到传递锂离子的作用，随着循环次数增加，电解液会发生不可逆的氧化或分解反应而消耗。电解液损失导致电解液中的有机溶剂和锂盐的浓度降低，锂离子在电池内交互传递的传输路径减少，导致锂离子从正负极材料脱嵌和嵌入速率变慢，进而导致锂电池最大可用容量的降低。文献^[56]表明当电池最大工作电压大于电解液的电化学稳定窗口后，电解液会与正极材料间发生氧化分解反应。

(4) 电极活性材料损失

充放电循环过程中，锂离子不断从正负极中嵌入和脱嵌，造成电极材料发生体积膨胀和收缩变化。在放电过程中，负极材料由于脱锂体积收缩，正极材料由于嵌锂体积膨胀。文献^[57]指出，硅材料负极体积膨胀可达 300% 以上，显著的体积变化造成了硅颗粒破碎和粉化。反复的体积膨胀和收缩会导致电极材料内部产生应力，使材料结构逐渐破碎、粉化，进而造成活性材料从电极上脱落，导致活性材料损失。电极活性材料是锂离子存储和释放的场所，活性材料损失直接导致能够参与电化学反应的锂离子数量减少，从而使电池的容量下降。

2.2.2 外部因素

(1) 环境温度

环境温度对于锂离子电池在性能和安全方面影响显著。锂离子在电解液中扩散迁移，而温度对离子扩散速度产生影响，进而影响到电池容量。文献^[58]研究指出，锂离子电池适合在 15-35℃ 的温度周期内工作。在低温环境下（低于 0℃），电解液的黏度增加，离子的迁移速率变慢，锂离子在正负极之间的传输受阻，同时电极材料的电化学反应活性降低，锂离子的嵌入和脱嵌变得困难。在高温环境下（高于 40℃），虽然锂离子迁移速度加快，但会加速电池内部的副反应，如电解液分解和电极活性材料结构的破坏等。

(2) 充放电倍率

充放电倍率是指电池在单位时间内充进或放掉其额定容量相对应电流的倍数。由于单体电池的不一致性，充放电倍率会加速电池模组中容量较小的单体电池的老化，在高充放电倍率下，会频繁的发生过充和过放电现象，加速电池退化。同时，高倍率充电时，锂离子从正极快速脱出并向负极移动，但负极材料可能来不及接收这么多锂离子，导致部分锂离子无法嵌入负极而形成“死锂”，从而造成电池容量损失。文献^[59]指出，大倍率充放电会加速电池性能衰退，主要是由于负极表面 SEI 膜的生长和电极活性材料的损失。

(3) 过充与过放

过充与过放是指锂离子电池在充放电过程中，电池电压分别超过了电池最高截止充电电压和低于电池最低截止放电电压。过充与过放均会导致正极材料结构的变化、负极出现析锂现象和加速电解液的氧化分解反应，从而引起锂电池容量下降和热失控等安全问题。文献^[60]研究指出过程与过放会使得电荷转移阻抗急剧增加、电极分解、开裂和变形、负极材料石墨厚度下降和正极材料表面的钝化层变厚等，电池会严重退化。

2.3 锂离子电池公开数据集

不同类型电池在循环寿命、退化速率等方面存在一定差异，对后续 HI 提取以及 SOH 预测会产生影响。为了充分验证所提 SOH 预测框架的准确性，本文选用两个公开电池老化实验数据集，分别是美国国家航空航天局的 NASA 电池数据集^[61]和

马里兰大学的高级生命周期工程中心 CALCE 电池数据集^[62]。

2.3.1 NASA 电池数据集

NASA 研究中心以额定容量为 2Ah 的 18650 型锂离子电池为实验对象，在不同温度和放电倍率下进行电池循环充放电寿命和电化学阻抗测试。NASA 电池数据集记录每一次充放电循环下的测试数据，包括电池的内阻、端电压、电流和温度信息，以及一次完整循环后的容量结果，数据结构清晰直观。

本文主要选取在室温 24°C 的环境下进行充放电循环老化实验的 B0005、B0006、B0007 共三节电池，其充放电标准如表 2-1 所示。三节电池以 1.5A 的恒定电流在恒流模式下充电，直到电池电压达到充电截止电压 4.2V，然后以恒压模式继续充电，直到充电电流下降到 20mA 停止充电。然后以 2A 的恒流放电，直至电池电压降至放电截止电压。由于反复的充电和放电循环会加速锂离子电池老化，当电池达到寿命终止标准时，实验将停止，此时额定容量下降 30%，从 2Ah 降至 1.4Ah。三种锂离子电池的循环寿命均为 168 次。

表 2-1 NASA 电池循环老化实验标准

电池	容量 (Ah)	循环 周期	充电电流 (A)	充电截止 电压(v)	充电截止 电流(mA)	放电电流 (A)	放电截止 电压(V)
B0005	2	168	1.5	4.2	20	2	2.7
B0006	2	168	1.5	4.2	20	2	2.5
B0007	2	168	1.5	4.2	20	2	2.2

图 2-2 为 NASA 数据集三节电池的容量衰减曲线。从图中可以直观的看出容量衰减曲线具有非线性，其总体呈现下降趋势，但在衰减过程中会出现短暂性的局部容量增加，但在之后的循环周期会快速退化。

2.3.2 CALCE 电池数据集

CALCE 电池数据集以额定容量为 1.1Ah 的 INR18650-20R 型锂离子电池为研究对象，进行不同温度下多种电池循环寿命测试。相比于 NASA 电池数据集，CALCE 电池数据集数据量庞大，数据的提取与处理难度较高，但是实验测试条件更贴合实际，具有更高的实验价值。

本文主要选用电池编号为 CS235、CS236、CS237 共 3 块电池数据，其充放电标准如表 2-2 所示。三节单体电池均以 0.5C 的恒定电流充电直到电池电压达到充电截止电压 4.2V，之后保持 4.2V 的电压充电直到电流下降到 50mA。充满电后，电

池以 1C 的恒定电流放电直到电压低于放电截止电压 2.7V。考虑到实际应用需求，取电池前 65% 区间的 SOH 的循环数据进行实验，CS235、CS236、CS237 的三块电池的循环寿命分别为 717、694、758 次循环，明显高于 NASA 数据集。

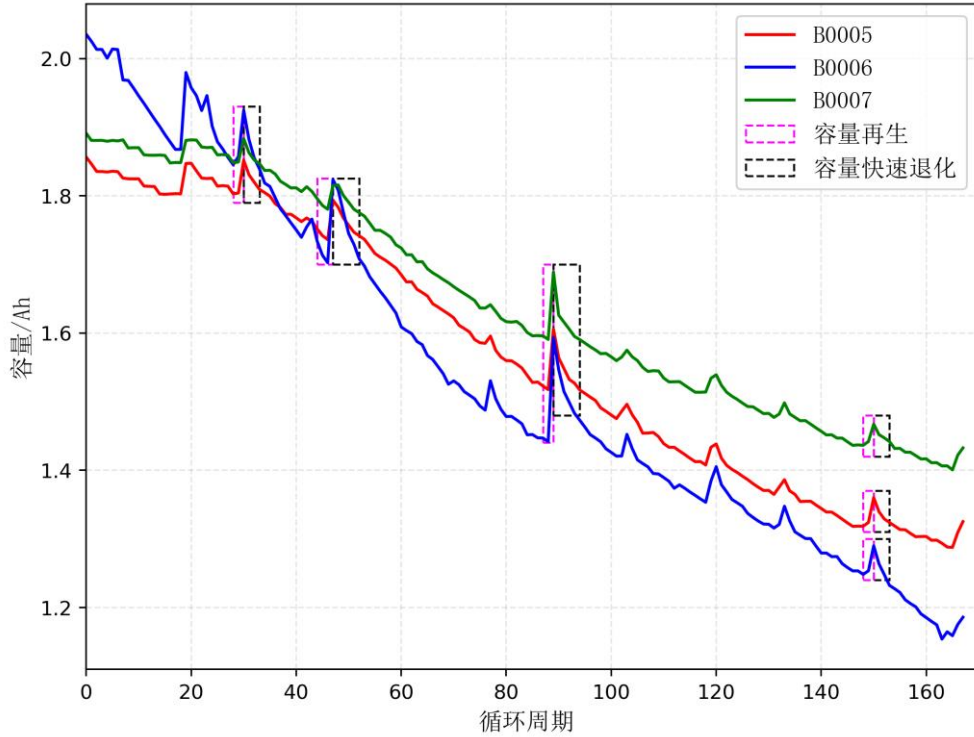


图 2-2 NASA 电池的容量衰减曲线与容量再生现象

表 2-2 CALCE 电池循环老化实验标准

电池	容量 (Ah)	循环 周期	充电电流 (A)	充电截止 电压(v)	充电截止 电流(mA)	放电电流 (A)	放电截止 电压(V)
CS235	1.1	717	1	4.2	50	2	2.7
CS236	1.1	694	1	4.2	50	2	2.7
CS237	1.1	758	1	4.2	50	2	2.7

图 2-3 为 CALCE 数据集三节电池的容量衰减曲线。与图 2-2 相类似，容量衰减曲线具有非线性，总体呈现下降趋势，衰减过程中会出现短暂性的局部容量增加，但在之后的循环周期会快速退化。文献^[63,64]研究指出，在经过一定次数的充放电循环后，锂离子电池通常会以更快的速率退化，快速退化的开始可称之为拐点。以 CS235 电池为例，去掉化学状态不稳定初始阶段，可以大概的把其循环周期分为三段。第 100-400 循环周期容量下降缓慢；第 400-620 周期容量加速下降；最后第 620-717 周期，容量出现断崖式下降趋势。综上所述，锂离子退化过程中的容量再生问题和退化速率的不一致性给锂电池准确的 SOH 预测带来挑战性。

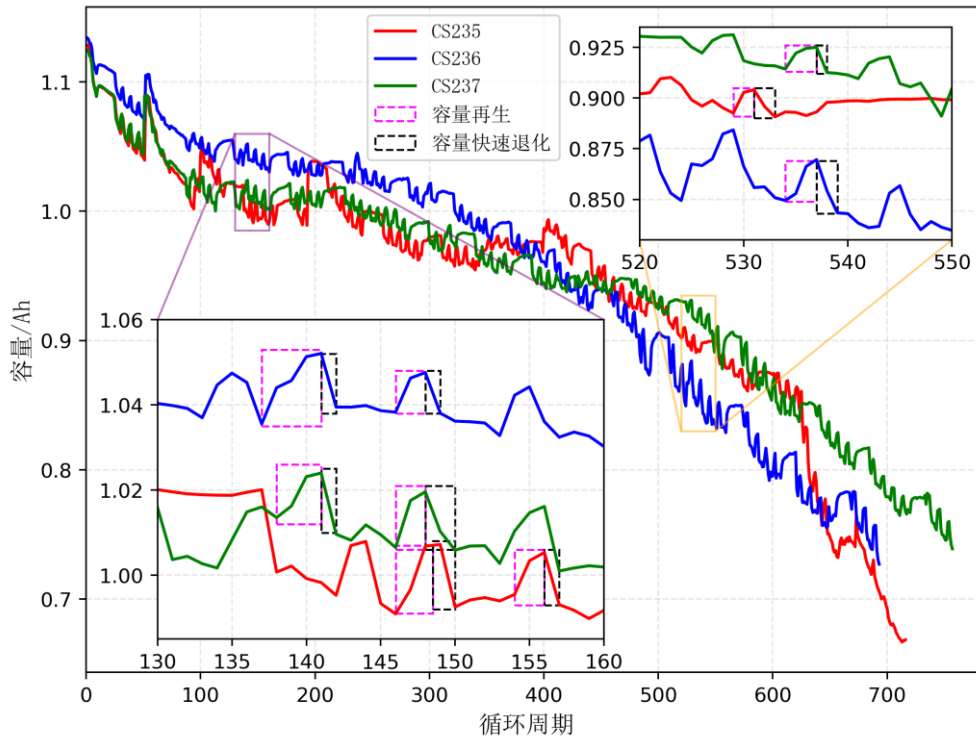


图 2-3 CALCE 电池的容量衰减曲线与容量再生现象

2.4 锂离子电池容量再生现象原因

大多数研究论文包括本文将电池容量用作 SOH 的评判标准，那么 SOH 预测的问题就转换为容量预测的问题。松弛效应是锂离子电池退化过程中的典型特征，如果电池的休息时间比退化过程中的正常中断时间长得足够长，则会导致电池的恢复过程，这可能会增加下一个循环的可用容量，这种自充电现象称之为容量再生^[65]。松弛效应引起的恢复过程可分为两个阶段，在图 2-2 和图 2-3 可以明显看出：第一阶段是休息时间的容量再生过程；第二阶段是再生容量的降级过程^[66]。在实际应用中，由于长时间静止引起的松弛效应对锂离子电池的退化规律影响很大，对锂电池 SOH 的准确预测为主要干扰因素，目前这个问题尚未得到很好的解决。

2.4.1 容量再生

在电池处于休息状态下，由于活性物质和梯度松弛的再平衡现象，从而使得电池的容量短暂上升。容量再生过程部分原因如下：当电池处于休息状态时，电极材料中的锂离子会在浓度梯度的驱动下，通过扩散作用重新分布，使得部分原本难以脱嵌的锂离子能够回到电极表面，增加了可参与反应的锂离子数量；放电时，电解质中的锂离子会向电极表面迁移并参与反应，导致电解质中锂离子浓度分布不均匀。

在休息期间，锂离子会在电解质中重新扩散，以达到更均匀的分布状态，使得电极表面能够接触到更多的锂离子；休息阶段，SEI 膜可能会发生一定程度的自我修复和重构，使得其更加稳定和有利于锂离子的传输。这些现象会导致短时间内自由锂离子数量的回升，从而有助于容量的恢复。

2.4.2 再生容量的快速退化

容量再生只是一个短暂的过程，其不会改善电池的老化状态。再生容量会在之后的循环中快速消退，使得锂电池容量退化呈现短暂不规则的波动现象。再生容量的快速退化的部分原因如下：活性物质的不可逆损失，电极材料中的活性物质会逐渐发生不可逆的化学反应，活性物质的总量减少，从而使电池能够存储的锂离子数量减少；SEI 膜的持续生长和增厚，会阻碍锂离子的传输，降低电池的充放电效率；电极材料的结构退化，颗粒团聚、孔隙率降低，减少锂离子的嵌入和脱嵌点位。尽管容量再生增加了自由锂离子数量，但难以阻挡电池内部的不可逆的副反应趋势，这些自由锂离子会被快速的消耗殆尽。

2.5 本章小结

本章节首先探讨了锂离子电池的基本结构和工作原理，接着讨论了锂电池容量退化的内部和外部影响因素。为了充分验证所提 SOH 预测框架的准确性，分别介绍了 NASA 电池数据集和 CALCE 电池数据集并分别选取其中具有代表性的三块电池合计共六块电池。最后从容量再生和再生容量的快速退化两个角度解释了容量再生问题，为后续锂离子电池 SOH 预测奠定工作基础。

第 3 章 多特征健康因子挖掘与容量再生特性分析

锂离子电池的容量衰退受复杂的电化学反应引起的内部不可逆副反应和环境温度、充放电倍率、过充与过放等外部因素共同影响。然而，这些因素几乎无法直接测量而得，只能根据电池工作的充放电电压、电流、温度等可测量参数构造衰老特征与 SOH 之间的映射关系^[67]。基于数据驱动的方法通过挖掘这些可测量参数，结合大量的老化数据进行分析，构造与电池容量退化高度相关的健康因子（Health Indicator, HI）从而实现 SOH 准确预测。美国阿拉巴马大学等人^[68]研究发现 HI 多样化会显著提高驱动模型的预测能力。因此，本文将构造多个锂离子电池退化的 HI，有利于全面表征锂离子电池退化程度。此外，对锂电池容量再生现象的特性进行讨论分析，引出本文更加关注其快速退化的特性，从而舍弃对容量再生部分的预测。

3.1 基于充电阶段的 HI 提取

对于实际生活中锂离子电池应用而言，电池充电过程较为稳定，电池充电模式一般为恒流-恒压充电，因此在充电过程中选择多个 HIs 往往更具有实用价值。为具体说明电池的特征变化，本文以 NASA 电池数据集的 B0005 电池和 CALCE 数据集 CS235 电池为例，分析 HI 的挖掘过程和结果展示。

3.1.1 基于横流充电电流曲线的 HI 提取

图 3-1 为两节电池在不同循环周期下的充电电流曲线。从图 3-1 可直观的观察，横流充电电流曲线随着充放电循环周期的不断增大，曲线在慢慢的向左移动，恒流充电时间在不断减少。

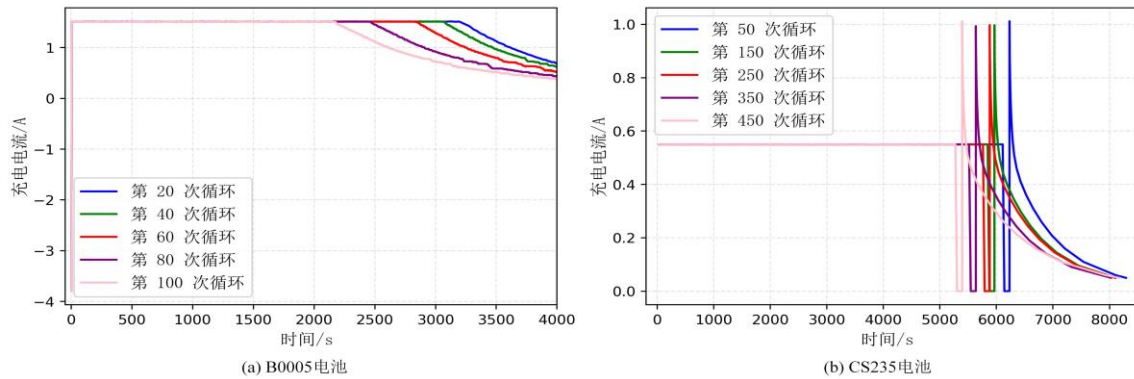


图 3-1 不同循环周期下的充电电流曲线

随着使用时间和充放电次数的增加，锂离子电池内部极化程度加深，横流充电

的持续时间相应减小。令横流充电时间为 Δt_l ， Δt_l 越大，说明横流充电时间越长，电池充电容量越多，反之则相反。将横流充电持续时间构造为健康因子 HI1，公式为：

$$HI1 = \Delta t_l | (I_N = 1.5 \& I_C = 0.55) \quad (3-1)$$

式中， I_N 和 I_C 分别表示 NASA 数据集和 CALCE 数据集电池端横流充电电流的值。

图 3-2 为 HI1 的曲线图。从图 3-2 可以看到，HI1 同电池的容量曲线衰减趋势一致且十分贴合。

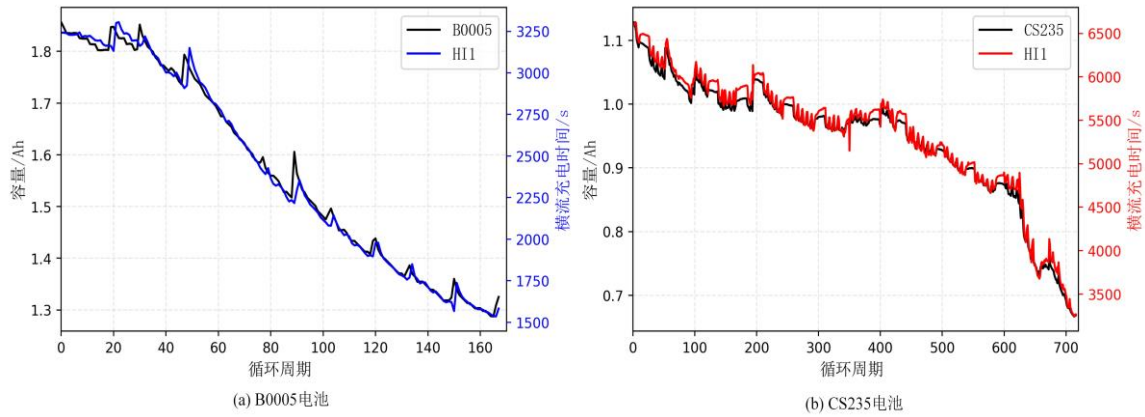


图 3-2 横流充电时间 HI1 曲线

3.1.2 基于横流充电电压曲线的 HI 提取

图 3-3 为两节电池在不同循环周期下的充电电压曲线。从图 3-3 可以看到，电压曲线达到截止电压的时间与充电电流曲线类似，持续向左慢慢移动。同时在充电过程中，电压曲线的斜率也在变化，不仅在电压初始位置的斜率越来越大，同时在此后较稳定的缓慢上升电压平台中，斜率并不相同。这意味着单位时间内电压的变化率的不一致，反应了电池的老化特性。

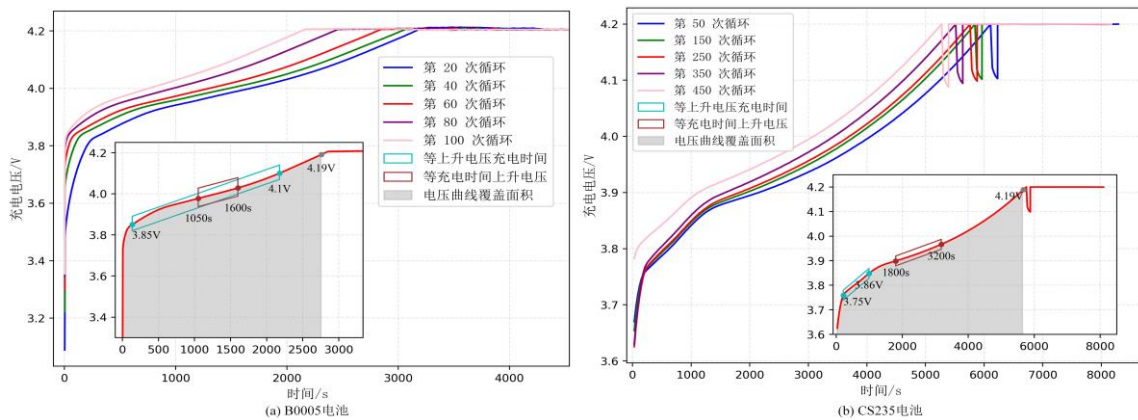


图 3-3 不同循环周期下的充电电压曲线

本文从锂电池充电电压曲线中提取三个不同衡量指标的 HI：等上升电压充电时

间、等充电时间上升电压、电压曲线覆盖面积。为了更好的确定最佳周期，提高 HI 与原始容量衰减曲线的相关性。本文使用遍历法，通过调整电压与时间间隔，遍历充电电压曲线上的相关周期。利用 3.4 节的相关性计算方法，保留相关系数大于 0.9 的作为备选 HI，最后选出相关系数最大的电压周期和时间周期所对应的特征参数构造对应的 HI。

将等上升电压充电时间构造为健康因子 HI2，公式为：

$$HI2 = t_{v_max} - t_{v_min} \quad (3-2)$$

式中， t_{v_max} 表示恒流充电电压周期上限电压对应的时间， t_{v_min} 表示恒流充电电压周期下限电压对应的时间。

将等充电时间上升电压构造为健康因子 HI3，公式为：

$$HI3 = V_{t_max} - V_{t_min} \quad (3-3)$$

式中， V_{t_max} 表示时间周期上限时间对应的恒流充电电压， V_{t_min} 表示时间周期下限时间对应的恒流充电电压。

如图 3-3 所示，以 B0005 电池为例，本文确定的最佳等上升电压周期为 [3.85V, 4.10V]，最佳等充电时间周期为 [1050s, 1600s]。图 3-4 和图 3-5 分别为 HI2 和 HI3 的曲线图。从图 3-4 可以看出，CS235 存在一定的波动和噪声，但大体上 HI2 都与对应电池的容量曲线衰减趋势一致。从图 3-5 可以看出，HI3 与对应电池的容量衰减曲线呈现负相关。随着循环次数的增加，单位时间内上升电压越来越多。侧面反映了电池的老化导致内阻增大，在充电电流不变的情况下，内阻增大，电池内部的电压降增加，从而导致单位时间内上升的电压变多。

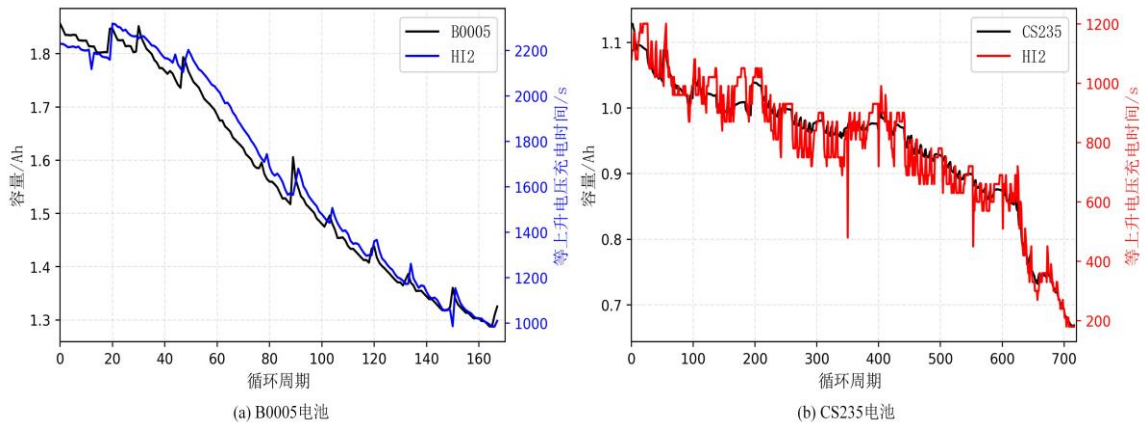


图 3-4 等上升电压充电时间 HI2 曲线

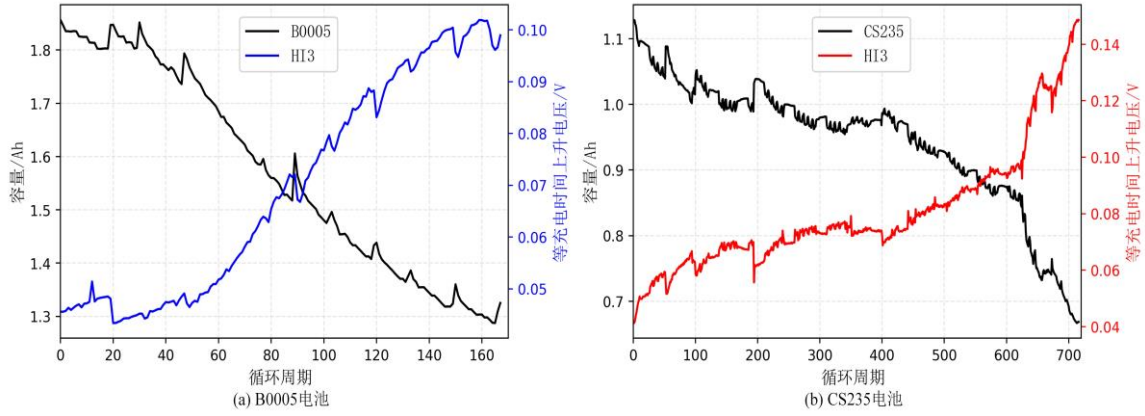


图 3-5 等充电时间上升电压 HI3 曲线

相比于达到恒定电压的充电时间这一点的变化，横流充电过程中的电压曲线覆盖面积考虑了整个横流充电过程中电压的变化率，更能全面体现容量的变化率。如图 3-3 所示，将横流充电下电压从充电开始到 4.19V 所覆盖的面积作为 HI4，公式为：

$$HI4 = \int_{t_0}^{t_v} V(t) dt \quad (3-4)$$

式中， $V(t)$ 表示恒流充电过程中的电压曲线， t_0 和 t_v 表示初始时间和电压 4.19V 所对应的截止时间。

图 3-6 为 HI4 的曲线图。从图 3-6 可以看出，HI4 紧密的贴合了容量衰减曲线，与容量衰减趋势高度一致。

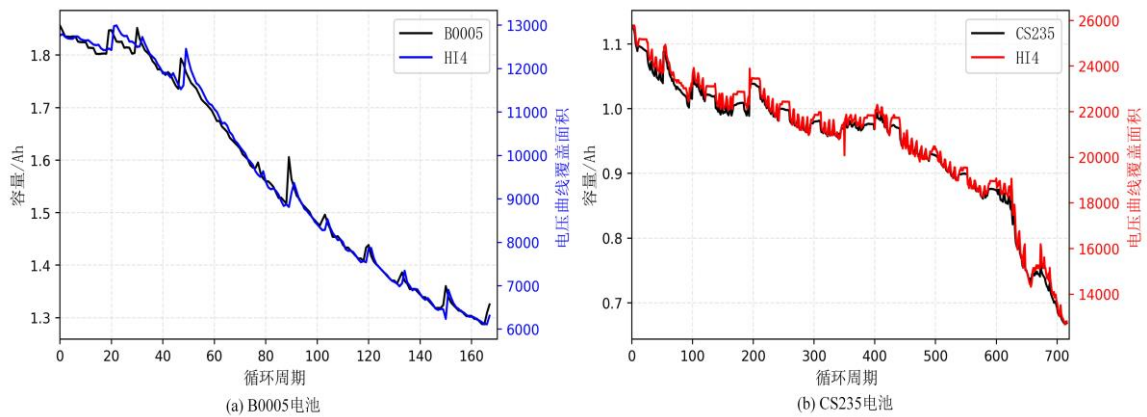


图 3-6 恒流充电电压曲线覆盖面积 HI4 曲线

3.2 基于放电电压曲线的 HI 提取

在锂电池放电过程中，通过保持较稳定的放电电流可以在较长的一段时间内获得持续稳定的输出电压。如电动汽车行驶过程中，需要稳定的电压来驱动电机，保证车辆的动力性能和续航里程。稳定的放电电流可以使锂电池组为电机提供持续稳

定的电力供应，确保车辆在不同路况下的行驶稳定性和安全性。所以放电电压数据在实际应用中也有挖掘 HI 的潜力。图 3-7 为两节电池在不同循环周期下的放电电压曲线。从图 3-7 可以看出，在放电中期，有一段稳定的电压平台。随着充放电循环周期的增加，平台的斜率在慢慢增大，电压到达放电截止电压的时间在不断缩短。

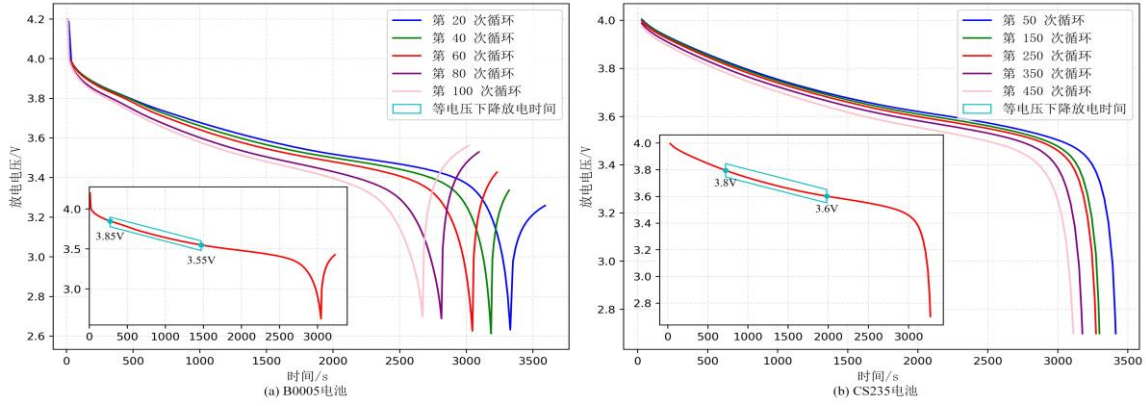


图 3-7 不同循环周期下的放电电压曲线

实际生活中，电池放电过程很少达到完全放电的情况。为保证所提 HI 的合理性，2 节电池的放电截止电压为 2.7V 左右，本论文将遍历周期下限设定为 3.5V。将等电压下降时间构造为 HI5。如图 3-7 所示，通过遍历法，确定 B0005 电池的电压下降周期为 [3.55V, 3.85V]，CS235 电池的电压下降周期为 [3.6V, 3.8V]，电压周期符合实际应用放电工况，保证所提 HI 的有效性。HI5 的提取过程公式如下：

$$HI5 = t_{v_min} - t_{v_max} \quad (3-5)$$

式中， t_{v_min} 表示放电电压周期下限电压对应的时间， t_{v_max} 表示放电电压周期上限电压对应的时间。

图 3-8 为 HI5 的曲线图。从图 3-8 可以看出，CS235 电池虽然相对于 B0005 电池稍有偏差，但也与容量衰减趋势一致，证明了 HI5 的提取准确性。

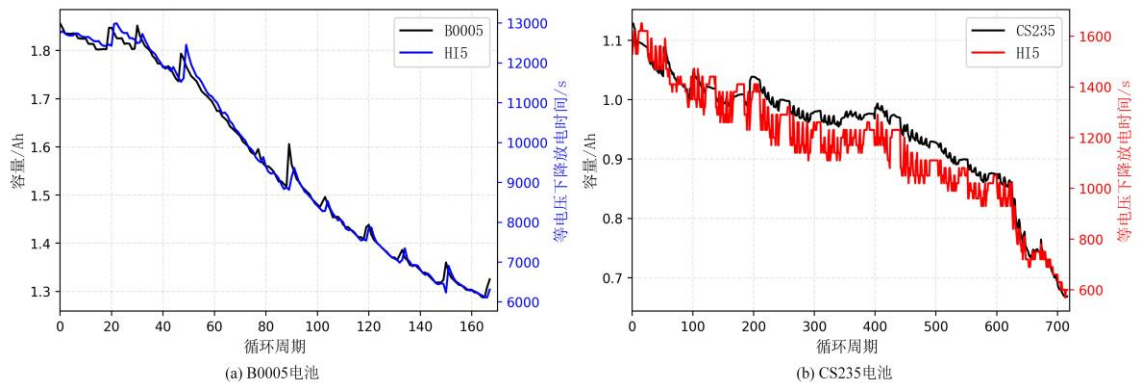


图 3-8 等电压下降放电时间 HI5 曲线

3.3 基于 IC 曲线的 HI 提取

锂离子电池增量容量分析法 (Incremental Capacity Analysis, ICA) 是一种数学模型方法。该方法能够将锂电池充放电循环过程中平缓的电压平台转化为容量增量 (Incremental Capacity, IC) 曲线中的直观的峰值, 这些与电池容量高度相关的特征被广泛应用于锂离子电池 SOH 的预测^[69]。

在横流充电过程中, 如图 3-3 所示, 随着循环周期的增大, 电压平台会发生缓慢的变化。但这些变化往往比较缓慢, 难以直观的从中捕捉 HI 用于锂离子电池 SOH 预测。ICA 通过测量电池在恒流充电过程中的容量变化, 利用数学方法将电池电压平台周期内缓慢上升的电压与容量进行非线性变换, 得到 dQ/dV 的变化曲线, 即 IC 曲线。IC 曲线通过 n 个采样点的有限差分获得, 具体计算公式如下:

$$\frac{dQ}{dV} = \frac{I \cdot dt}{dV} = I \frac{dt}{dV} \quad (3-6)$$

式中, Q 为容量, V 为锂电池端测量电压, I 为恒流充电电流, t 为充电时间。

在实验室数据中和实际应用中, 电流电压数据往往传感器在一定时间间隔下采样的。为了便于计算, 采用增量数值代替微分值计算, 公式如下:

$$\frac{dQ}{dV} \approx \frac{\Delta Q}{\Delta V} = \frac{I_k(t_k - t_{k-1})}{V_K - V_{K-1}} \quad (3-7)$$

式中, I_k , t_k , V_K 分别代表第 K 次采样的的电池充电电流, 时间和充电电压。

由于数据集本身存在一定的误差和错误数值。在 NASA 数据集中, 各循环之间的时间跨度并不一致, 以采样时间间隔计算 IC 曲线不能取得满意的结果。经多次实验, 本文以 10mv 为间隔, 提取相应的时间和电流计算 IC 曲线。CALCE 数据集相对于 NASA 数据集数据更加准确更加复杂, 本论文以 30s 为时间间隔, 采取对应的电流和电压数据计算 IC 曲线, 电压间隔大概为 3mv, IC 曲线更加准确。图 3-9 为 B0005 电池和 CS235 电池对应的 IC 曲线。

从图 3-9 和图 3-3 不同循环周期下的充电电压曲线对比而言, IC 曲线实现了将锂离子电池充电循环过程中平缓的电压平台转化为直观的峰值。从图 3-9 可以看出, IC 曲线的峰值和曲线随循环次数的增加呈下降趋势, 这表明电池活性物质损失和老化。常见的 IC 曲线特征有最大峰值和位置, 峰值下的面积等, 与电池内部的锂化和脱锂过程的相变和相平衡密切相关。然而由于测量和噪声误差, 就 CS235 电池而言,

本论文发现 IC 曲线并不是平滑的，这很影响峰值最值和位置的提取。文献^[70]采用移动平均平滑滤波器加高斯平滑滤波器的双重滤波方法对原始 IC 曲线进行处理，进而提取曲线的峰值特征。文献^[71]采用鲁棒高斯滤波平滑 IC 曲线。为了充分体现多 HIs 和所提框架的优越性，同时考虑到实际应用中数据也存在噪声干扰，本论文不对 IC 曲线进行滤波平滑处理，将 IC 周期曲线覆盖面积作为 HI6，减免需要确定峰值而必须的滤波步骤。HI6 的提取过程公式如下：

$$HI6 = \int_{V_{min}}^{V_{max}} IC(v)dv \quad (3-8)$$

式中， IC 为充电 IC 曲线， V_{min} 和 V_{max} 分别代表充电电压周期的下限周期和上限周期。

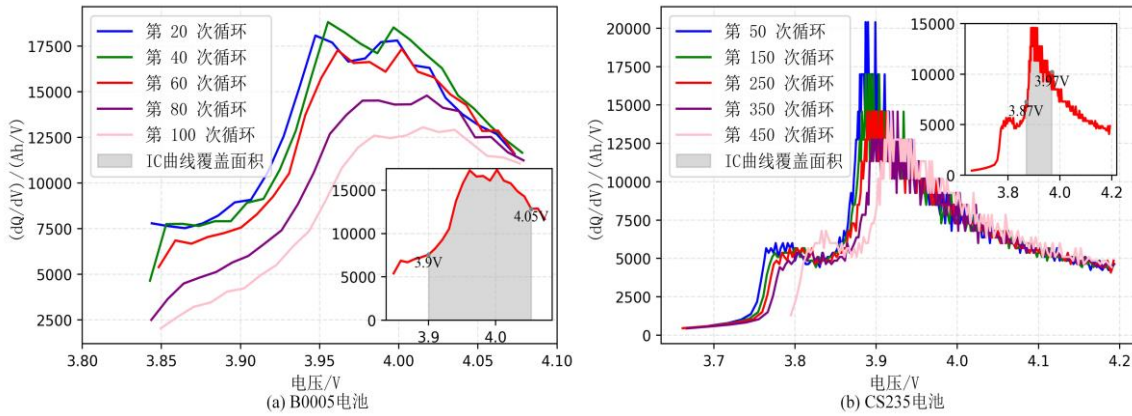


图 3-9 不同循环周期下的充电 IC 曲线

通过遍历法，如图 3-9 所示，B0005 电池 IC 曲线覆盖的最佳周期为 [3.9V, 4.05V]，CS235 电池最佳周期为 [3.87V, 3.97V]。图 3-10 为 HI6 的曲线图。从图 3-10 中可以看出 HI6 都十分贴合各自锂离子电池容量退化趋势，特别是未经滤波平滑优化处理的 CS235 电池，其 HI6 也能非常好的表征电池的衰退特性，与容量退化曲线贴合紧密度很高。

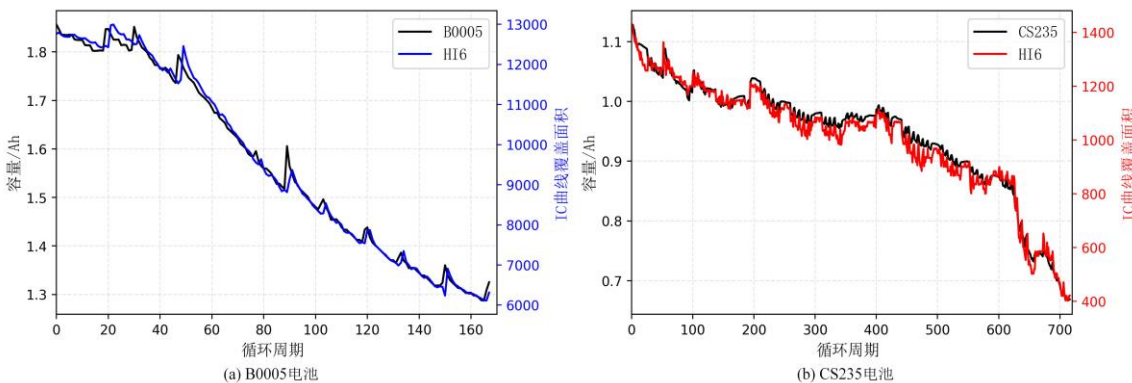


图 3-10 充电过程中 IC 曲线覆盖面积 HI6 曲线

3.4 相关性分析

准确的 HIs 是数据驱动法准确预测 SOH 的前提。由于提取的 HIs 的变化规律和单位尺度都不一样，很难直观判断每个 HI 与锂离子电池容量的相关程度。为了进一步验证所提 HIs 的有效性，本文采用皮尔逊相关系数（Pearson Correlation Coefficient, PCC）和斯皮尔曼相关系数（Spearman Correlation Coefficient, SCC）综合定量分析所提 HIs 与容量之间的相关性程度大小。传统论文单独使用 PCC，只能捕捉 HIs 与容量之间的线性依赖关系，可能遗漏掉有用的非线性信息；引入 SCC，可以补充单调非线性依赖的捕捉，同时提升相关性分析的鲁棒性和全面性，最终筛选出更科学合理的 HIs 用于后续建模。

3.4.1 PCC

PCC 是一种线性相关系数，用于衡量两个连续变量之间的线性相关程度^[72]。它通过计算两个变量的协方差与它们标准差的乘积之比得到，取值范围在 -1 到 1 之间。当 PCC 绝对值为 1 时，表示两个变量完全相关； $|PCC|$ 为 [0.6, 1) 时，表示两个变量高度相关，有较为明显的线性关联； $|PCC|$ 为 [0.2, 0.6) 时，表示两个变量一般相关，线性关系较弱； $|PCC|$ 为 [0, 0.2) 时，表示两个变量几乎不存在明显的线性关联。PCC 的数学计算公式如下：

$$p_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3-9)$$

式中， p_{xy} 代表两个变量之间的 PCC， x_i 和 y_i 分别代表两个变量的第 i 次的样本值， $\bar{x}$ 和 $\bar{y}$ 分别代表两个变量的样本均值。

表 3-1 锂离子电池 HIs 与容量之间的 PCC

电池编号	HI1	HI2	HI3	HI4	HI5	HI6
B0005	0.9965	0.9931	-0.9843	0.9967	0.9973	0.9926
B0006	0.9918	0.9885	0.9952	0.9916	0.9971	0.9865
B0007	0.9958	0.9849	-0.9901	0.9952	0.9982	0.9952
CS235	0.9967	0.9723	-0.9957	0.9969	0.9724	0.9926
CS236	0.9976	0.9796	-0.9976	0.9978	0.9913	0.9934
CS237	0.9968	0.9760	-0.9970	0.9970	0.9891	0.9935

表 3-1 计算了上述六组 HIs 与各自电池容量的 PCC。根据计算结果可知 PCC 的绝对值均高于 0.97，无限接近于 1。这表明所提 HI 与容量的相关性极高，有很强的

线性关系，能有效表征电池的退化性能。同时，除了 HI1 和 HI4 两个完整的横流充电周期，其余 4 个有限周期对应的 HI 的 PCC 也说明了遍历法的优越性，能在最大程度上尽可能提高 HI 的准确性。

3.4.2 SCC

SCC 是一种非参数统计量，用于衡量两个变量之间的单调关系，不依赖于变量的具体分布，而是基于变量的秩次（即排序后的顺序）计算^[73]。PCC 更适用于连续的且服从正态分布具有一定线性关系的变量。当变量之间不满足前述条件时，SCC 能更好的评估 HIs 与锂离子电池容量之间的相关性。SCC 对样本数据的异常值不敏感，受异常值影响微弱，可用于判断两个变量之间的单调相关性强弱。当 SCC 绝对值为 1 时，表示两个变量完全单调相关； $|SCC|$ 为 $[0.6, 1)$ 时，表示两个变量高度相关，变化趋势在单调意义上非常接近； $|SCC|$ 为 $[0.2, 0.6)$ 时，表示两个变量一般相关，存在一定程度的单调关系，但不是特别紧密； $|SCC|$ 为 $[0, 0.2)$ 时，表示两个变量几乎不存在明显的单调关联。PCC 的数学计算公式如下：

$$S_{xy} = \frac{\sum_{i=1}^n (R(x_i) - \overline{R(x)})(R(y_i) - \overline{R(y)})}{\sqrt{\sum_{i=1}^n (R(x_i) - \overline{R(x)})^2 \sum_{i=1}^n (R(y_i) - \overline{R(y)})^2}} \quad (3-10)$$

式中， S_{xy} 代表两个变量之间的 SCC， $R(x_i)$ 和 $R(y_i)$ 代表第 i 个样本数据 x_i 和 y_i 在各自变量中经过排序后所对应的秩次， $\overline{R(x)}$ 和 $\overline{R(y)}$ 代表变量 x 和 y 所有样本数据秩次的平均值。

表 3-2 计算了上述六组 HIs 与各自电池容量的 SCC。根据计算结果可知 PCC 的绝对值均高于 0.97，无限接近于 1。表面每组 HI 都与对应电池容量具有强单调性，再次证明所提 HI 方法和提取 HI 的准确性。通过 PCC 和 SCC 的综合定量分析，充分验证了所提取的 HIs 的能够有效反映锂离子电池的容量的退化特性。

表 3-2 锂离子电池 HIs 与容量之间的 SCC

电池编号	HI1	HI2	HI3	HI4	HI5	HI6
B0005	0.9907	0.9859	-0.9809	0.9909	0.9956	0.9824
B0006	0.9932	0.9925	0.9951	0.9933	0.9990	0.9904
B0007	0.9912	0.9822	-0.9870	0.9902	0.9960	0.9902
CS235	0.9940	0.9612	-0.9959	0.9950	0.9786	0.9919
CS236	0.9971	0.9759	-0.9975	0.9976	0.9895	0.9930
CS237	0.9952	0.9619	-0.9965	0.9959	0.9857	0.9892

3.5 容量再生现象特性分析

前文已对容量再生的现象、原因和研究现状进行分析总结。在电池处于长时间休息状态下，由于活性物质和梯度松弛的再平衡现象，从而使得电池的容量短暂上升。然而容量再生只是一个短暂的过程，其不会改善电池的老化状态。再生容量会在之后的循环中快速消退，使得锂电池容量退化呈现短暂不规则的波动现象。为直观的说明容量再生的特性，计算电池容量的一阶导数。离散一阶导数计算公式如下：

$$\frac{dQ}{dC} \approx \frac{\Delta Q}{\Delta C} = \frac{Q_i - Q_{i-1}}{C_i - C_{i-1}} \quad (3-11)$$

式中， Q 为容量， C 为循环周期， Q_i 和 C_i 分别为第 i 次容量和循环周期。

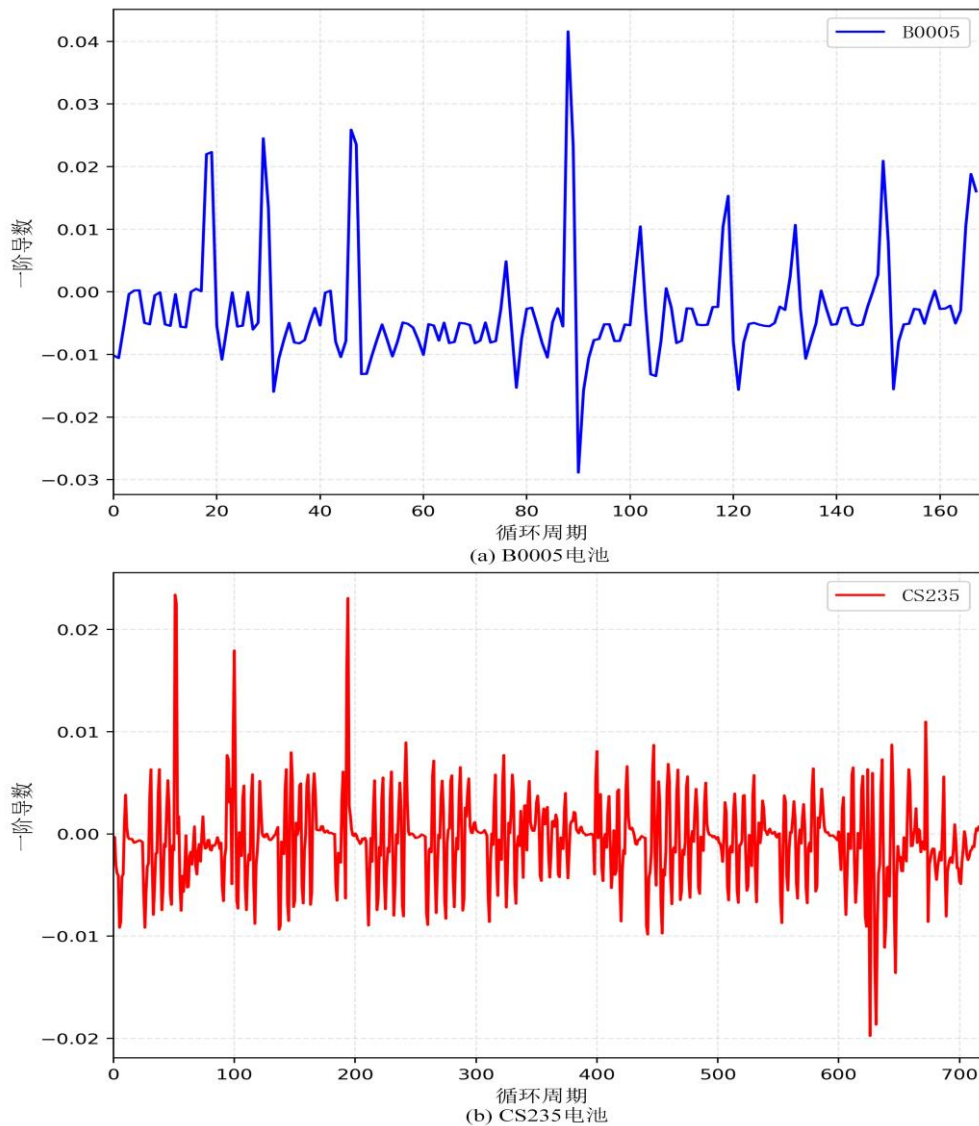


图 3-11 电池容量一阶导数序列图

图 3-11 为 B0005 和 CS235 两节电池的容量一阶导数图。可以看出，如果去除

容量再生波动的影响,两节电池的一阶导数都低于 0, B0005 电池总体大概在 $[-0.01, 0]$ 之间, CS235 电池总体大概在 $[-0.005, 0]$ 之间, 这意味着下一循环周期低于当前循环周期容量, 电池容量整体呈现下降趋势。图中每一个峰值意味着电池容量再生, 而紧随其后的是迅速出现的负峰值, 这意味再生容量的快速退化。就 CS235 电池而言, 由于处于实验环境下, 工况周期较为固定, 容量再生波动呈现一定的规律性。但本论文也发现, 其容量再生的峰值存在差异性。在初始阶段第 0-200 周期内, 有 3 次容量再生峰值较高; 在 600 周期后, 虽然容量再生峰值较为稳定, 但其再生容量的快速退化的负峰值明显开始变大。就 B0005 电池而言, 由于其工况没有 CS235 那么严谨, 休息中断时间没有周期性, 其呈现的一阶导数也不再具有规律性。可以明显看出在第 0-50 周期、50-90 周期、90-168 周期在内三个周期之间的容量再生波动毫无规律性。对于第 50-90 周期段, 在第 50-75 周期内, 由于休息时间较短, 没有容量再生出现, 电池整体呈现稳定下降趋势。而在第 87 周期时, 由于休息时间充足, 其呈现出一个非常显著的峰值, 容量较上一周期提升 0.04 左右, 而在后续 2-3 个周期内, 其容量快速下降 0.03 左右, 且后续 10 个周期持续以 0.08 左右的容量下降。结合图 2-2 其容量衰减曲线图, 可以明显得出这个突出的容量波动并没有改变 B0005 容量下降趋势, 而是快速的消散了。综上所述, 就算在工况周期较为单一稳定的实验室环境下, 锂离子电池容量再生波动发生的周期, 容量再生的数值和再生容量快速退化下降的数值都不能表现出很强的规律性, 相反而是具有较大的随机性。

容量再生波动直接影响短期内锂离子电池的 SOH, 甚至会使容量短期内出现较大回升, 导致短期 SOH 预测偏离实际较大。同时, 容量再生问题也会严重影响到 RUL 的长期精确预测, 引起较大预测误差。目前针对锂离子电池长期 SOH 预测大量研究工作中, 本质是将原始容量退化过程解耦, 分别预测残差信号和本征模态函数, 最后将预测结果相加得到总的 SOH 预测结果。分解算法主要是根据数据自身的频率、幅值等特征将其分解为不同频率段的分量, 它们之间是相互独立, 共同构成了原始数据。然而这具有局限性。锂电池容量再生需要满足充足的休息时间, 这在实际应用中并没有规律可言。以电动汽车为例, 不同用户之间以及个体用户而言, 其充放电时间和闲置时间都具有很大的差异性, 这意味着车载锂电池发生容量再生的周期和波动峰值毫无规律性可言。以前 50% 周期预测后 50% 周期 SOH 而言, 这意味着分解后的 IMFs 分量缺少规律性, 前 50% 周期可能会缺失关键的容量再生阶段的部分信息或者容量退化趋势转折的信息, 无法准确预测后 50% 周期容量再

生波动。除了车辆的充电与停放时间的不同，大量研究表明，不同充放电策略（如倍率变化、休息段设置、SOC 窗口调整）对容量再生特性的影响显著，且表现出明显的不确定性。例如，在低倍率充电或加入休息段时，容量再生频率和幅度均高于高倍率连续充放电条件；浅放电循环比深度放电循环更易触发容量回升。此外，温度变化亦对容量再生行为产生显著影响。这些差异表明，容量再生受到多种外部条件叠加影响，具有显著的随机性和复杂性。总之，传统实验室环境下可行的分解-集成预测方法并不能有效的适用实际应用所需。

3.6 本章小结

本章主要阐述了电池充放电循环周期过程中典型特征的变化趋势进行 HI 的提取和容量再生特性的分析。首先，分析电池老化过程中充电电流、充电电压、放电电压、充电 IC 曲线的变化趋势并提取共六组 HIs，分别为恒流充电时间、等上升电压充电时间、等充电时间上升电压、恒流充电电压曲线覆盖面积、等电压下降放电时间、充电过程 IC 曲线覆盖面积，其中涉及有限周期的特征采取遍历法，提取最佳周期特征。其次，采用 PCC 和 SCC 共同验证所提 HIs 与电池原始容量不仅具有高度线性相关性同时也具有高度单调相关性，能够有效的表征锂离子电池容量退化状态。最后，具体分析了容量再生特性，指出现阶段相关研究工作存在的局限性。综上，为后续章节的 SOH 点预测和周期预测奠定基础。

第 4 章 基于 1DCNN-Informer 框架的锂离子电池 SOH 点预测

对于 SOH 的准确预测，除了所提 HIs 的相关度外，选择合适的数据驱动方法也很重要。锂离子电池较长的循环使用寿命需要模型具备长期有效的预测能力，能够在较长的预测周期保持较低的预测误差。以 Transformer 模型为基础的改进模型 Informer 模型通过引入稀疏自注意力机制可以更高效的处理长序列数据，充分挖掘 HIs 与 SOH 的相互关系，捕捉时间序列中的长期依赖关系，预测精度较高。但单一驱动模型存在局限性，如 Informer 模型在训练样本较少时，预测能力相对不足。因此，引入一维卷积神经网络（One-Dimensional Convolutional Neural Network, 1DCNN），利用其强大的局部特征提取能力实时估计锂电池当前循环周期 SOH，但 1DCNN 也存在长期预测精度下降的问题。为避免传统基于验证集下均方根误差（Root Mean Square Error, RMSE）自适应权重，存在偏向短期预测效果更好的单一模型的问题，引入超参数 λ 调整权重分配，可以根据实际任务需要调整单一模型权值。最后，将 1DCNN 当前时间步和 Informer 模型的下一时间步的预测结果融合，可有效提高所提框架下一时间步的 SOH 点预测精度。所提框架能够实时预测锂电池当前周期和下一周期的 SOH，帮助用户及时识别性能下降的电池并更换，避免因电池故障引发安全事故。

4.1 1DCNN 原理

卷积神经网络（Convolutional Neural Network, CNN）在图像识别、目标检测、自然语言处理等领域取得巨大成功^[74]。由于应用的目的和场所不同，所使用的 CNN 的维度也不相同。目前，二维卷积常用于处理图像类数据。而一维卷积是卷积神经网络的一种变体，与传统的全连接神经网络相比，它能更好地处理数据中的局部关系，提取数据中的局部时序特征，在各种诊断和预测任务中表现出色^[75]。通过嵌套卷积操作，1DCNN 可以实现对时序特征的逐层提取。随着卷积网络深度的增加，提取的高维特征包含越来越丰富的信息，这些信息经过聚合和传递产生最终输出。一维卷积的原理如图 4-1 所示。

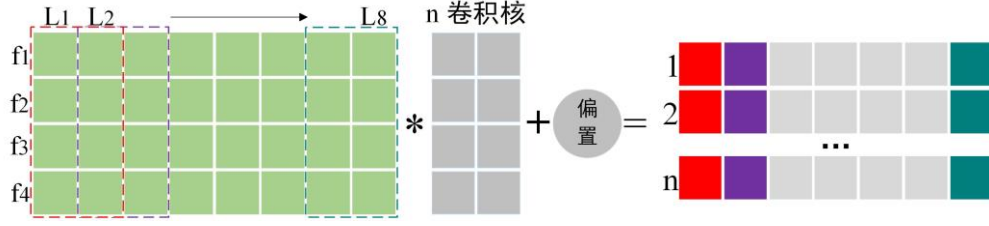


图 4-1 一维卷积计算原理

在图 4-1 中，输入样本维度假定为 (L, f) ，其中 L 是序列长度， f 是特征个数。卷积核大小为 (n, k, f) ，表示大小为 $k \times f$ 的 n 个卷积核。因此，一维卷积运算公式定义如下：

$$y_n(t) = \sum_{i=0}^{k-1} \sum_{j=0}^{f-1} x(t \cdot s + i, j) \cdot W(i, j, n) + bn \quad (4-1)$$

式中， x 为输入矩阵， t 为时间步长， s 为步长大小， W 为卷积核， bn 为每个卷积核的独立偏置值。 $t \cdot s + i$ 为卷积核在当前时间步 t 时在输入矩阵上的具体位置， $x(t \cdot s + i, j)$ 为当前卷积核在输入矩阵中的覆盖区域值， $W(i, j, n)$ 为第 n 个卷积核在位置 (i, j) 上的权重。 $y_n(t)$ 是 n 个卷积核在时间步长 t 时处生成的输出特征图的值。

本文提出的 1DCNN 结构如图 4-4 右下角的一维卷积模块所示，由特征提取模块和预测模块组成。一方面，特征提取模块由两个卷积层组成，目的是在时间维度上有效地提取输入序列中的局部特征和依赖关系。在相邻的卷积层之间，嵌入了激活函数层和最大池化层，以引入非线性映射，同时保证输入序列时间维度的统一性。另一方面，预测模块由两个线性层组成，其目的是将高维特征表示转换为最终预测输出。

卷积结果需要通过激活函数引入非线性因素，增强网络的表达能力。它能够使网络学习到更加复杂的特征和模式，解决线性模型无法处理的非线性问题。ReLU 函数计算简单，收敛速度快，在一定程度上缓解了梯度消失问题，有助于提高训练效率。其表达式如下：

$$f(x) = \max(0, x) \quad (4-2)$$

式中， $f(x)$ 为 Relu 输出，当输入 x 大于 0 时，输出为 x ；当输入 x 小于或等于 0 时，输出为 0。

最大池化层用于对卷积特征的降维，减少数据量和计算量的同时，保证输入序列时间维度的统一性。本文选用池化窗口大小为 2，步长为 1，其表达式如下：

$$f(x) = \sum_{i=1}^{n-1} \max(x_i, x_{i+1}) \quad (4-3)$$

式中, $f(x)$ 为最大池化层输出, 池化窗口以步长为 1 在序列长度方向滑动, 每次取 x_i 和 x_{i+1} 中的较大值为输出。

线性层位于卷积之后, 用于将卷积提取的特征进行整合, 将高维的特征向量映射到低维的输出空间。线性层对特征进行线性组合, 学习特征之间的复杂关系, 从而实现对不同类别或数值的预测。其表达式如下:

$$f(x) = \sum_{i=1}^n Wx_i + b_i \quad (4-4)$$

式中, $f(x)$ 为线性层输出, W 为权重矩阵, 将输入维度 n 映射到设定的输出维度, x_i 为每个神经元值, b_i 为每个神经元对应的随机偏置。

线性层通常为多层, 最后一层输出维度为 1, 即为最终预测结果。为避免神经网络过拟合, 提高模型的鲁棒性, 在相邻的线性层中往往会引入 Dropout 层, 以一定概率随机丢弃神经元的值, 置为 0。本文选择的丢弃率为 0.05。

4.2 Transformer 和 Informer 原理

4.2.1 Transformer

传统的序列转化模型如循环神经网络 (Recurrent Neural Network, RNN) 在处理输入序列时, 其内在的顺序性使得在训练样本内难以并行化计算, 尤其在长序列情况下, 内存限制了跨样本的批处理, 导致训练效率受限。注意力机制虽然已成为许多序列建模和转换模型的重要组成部分, 但以往大多数模型是将其与 RNN 结合使用, 仍未摆脱 RNN 的固有缺陷。长序列时序预测 (Long Sequence Time-sequence Forecasting, LSTF) 要求模型高效捕捉输出与输入之间精确的长距离依赖耦合。2017 年, 谷歌提出了 Transformer 模型, 这是一种具有注意力机制的深度学习模型。该模型可以根据输入数据的重要性为不同部分分配不同的权重, 从而提高预测效果^[76]。Transformer 是一种完全基于注意力机制的模型, 以克服上述问题, 提高模型的并行性和训练效率, 并提升在序列转换任务上的性能。图 4-2 为 Transformer 网络结构。如图 4-2 所示, Transformer 模型是一个由编码器 (Encoder) 和解码器 (Decoder) 组成的序列到序列模型。特征序列先经过词嵌入 (Embedding) 编码和位置编码 (Positional Encoding) 后, 经过编码器编码和解码器解码, 最后通过线性层和软最

大化层 (Softmax) 预测输出结果。

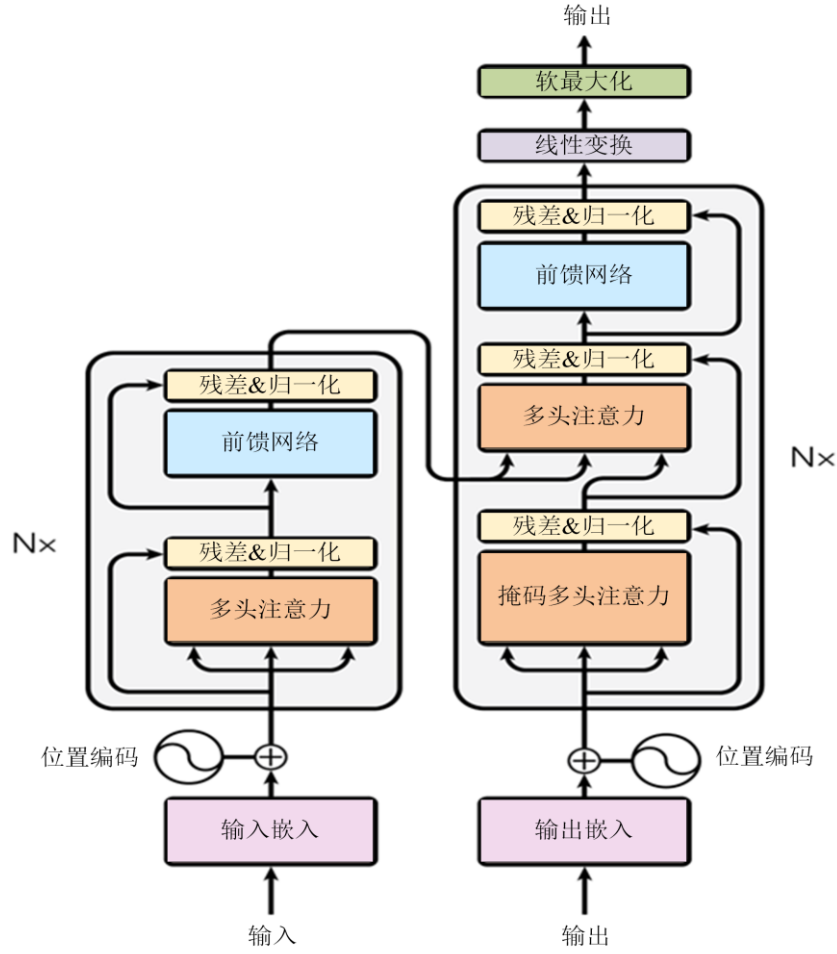


图 4-2 Transformer 模型结构

当输入时间序列时，由于 Transformer 模型基于自注意力机制，无法识别时间序列的位置信息。时间序列往往在时间维度上有重要关系，因此需要通过位置编码来为模型提供关于当前序列在序列中相对或绝对位置的信息。其计算公式如下：

$$PE_{(pos, 2i)} = \sin(pos / 10000^{2i/d_{model}}) \quad (4-5)$$

$$PE_{(pos, 2i+1)} = \cos(pos / 10000^{2i/d_{model}}) \quad (4-6)$$

式中， pos 表示位置， i 表示维度， d_{model} 表示词嵌入的维度。

当输入序列进入编码器中后，其会被三个权重矩阵映射为查询 (Query, Q)、(Key, K) 和值 (Value, V) 三个矩阵，其计算公式如下：

$$Q, K, V = XW_q, XW_k, XW_v \quad (4-7)$$

式中， $Q \in \mathbb{R}^{d_q \times d_k}$ 、 $K \in \mathbb{R}^{d_q \times d_k}$ 和 $V \in \mathbb{R}^{d_q \times d_v}$ 分别代表查询矩阵、键矩阵和值矩阵，

$X \in \mathbb{R}^{d_q \times d_{\text{model}}}$ 为输入向量, $W_q \in \mathbb{R}^{d_{\text{model}} \times d_k}$ 、 $W_k \in \mathbb{R}^{d_{\text{model}} \times d_k}$ 和 $W_v \in \mathbb{R}^{d_{\text{model}} \times d_v}$ 分别代表三个向量的权重矩阵。

与传统注意力机制不同, Transformer 仅依靠点积和多头注意力实现自注意力机制。在时序序列预测任务中, 通过 Q 和 K 的点积操作获得序列不同时间步的关联矩阵, 通过激活函数 Softmax 得到每个时间步的权重矩阵; 最后注意力权重矩阵与 V 相乘, 得到最终输出矩阵。点积计算公式如下:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4-8)$$

式中, softmax 为概率归一化函数; d_k 为 K 向量维度, 缓解点积过大导致的梯度问题, 防止归一化之后概率分布差异较大。

为防止单一注意力过度关注某些局部特征联系, 采用多头注意力机制。通过将 Q 、 K 和 V 进行多次不同的线性投影, 并行地执行注意力函数, 然后将结果拼接并再次投影得到最终输出。这种机制允许模型从不同的表示子空间联合关注信息, 相比单头注意力机制, 能够更好地捕捉输入和输出序列中不同位置之间的复杂关系, 提升模型的表达能力。通过将不同头的 Attention 拼接起来后进行一次线性变换, 多头注意力过程如下:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^0 \quad (4-9)$$

$$\text{head}_i = \text{Attention}(XW_i^q, XW_i^k, XW_i^v) \quad (4-10)$$

式中, $W^0 \in \mathbb{R}^{hdv \times d_{\text{model}}}$, 为线性变换权重矩阵; h 为注意力头数; W_i^q 、 W_i^k 、 W_i^v 分别为每个头输入向量 X 映射为的 Q 、 K 、 V 的权重矩阵。

再计算完成多头注意力之后, 每个编码器的输出结果还会经过残差连接、前馈网络和层归一化。残差链接目的是缓解深度神经网络中的梯度消失或梯度爆炸问题, 使得网络能够更容易地训练深层结构。层归一化是对每个样本的每个特征维度进行归一化, 使其均值为 0, 方差为 1, 有助于稳定模型的训练过程。前馈网络进一步对输入进行特征提取和变换, 增加模型的表达能力。三者计算过程如下:

$$H' = \text{LayerNorm}(\text{MultiHead}(X) + X) \quad (4-11)$$

$$H = \text{LayerNorm}(H' + \text{FFN}(H')) \quad (4-12)$$

式中, X 为输入向量; FFN 为全连接函数; $LayerNorm$ 为层归一化函数; H 为编码器输出结果。

解码器与编码器结构有所不同。解码器自注意力机制时需要进行掩码操作。在生成第 i 个位置的输出时, 要防止其关注到后续位置的信息, 以保证模型的自回归特性。除自注意力之外, 解码器还会与编码器输出进行交互注意力操作。在解码方面, 解码器生成输出序列时是一个动态的、自回归的过程。在每个时间步, 它会根据之前生成的输出和编码器的输出, 以及自身的掩码机制和注意力机制来生成下一个位置的输出, 这个过程是逐步进行的。

原始 Transformer 在 LSTF 任务上表现不佳, 主要包括: 自注意力机制平方级的计算复杂度, 运行速度慢; 堆叠多层编码器与解码器网络, 空间复杂度与序列长度 L 呈二次方增长, 内存占用高; 逐步解码预测, 速度慢。

4.2.2 Informer

针对上述 Transformer 存在问题, 北京航空航天大学周浩义等人^[77]提出 Informer 模型: 概率稀疏自注意力 (ProbSparse Self-attention) 筛选出活跃的 Q , 降低计算复杂度; 自注意力蒸馏 (Self-attention Distilling) 进行下采样操作, 减少维度和网络参数的数量; 生成式解码器 (Generative Style Decoder), 采用全连接层一步得到所有预测输出。Informer 模型有效降低了模型计算的时间和空间复杂度, 提高了基于生成式解码器的预测速度, 并解决了时序数据的长期依赖性问题。图 4-3 为 Informer 模型结构图。

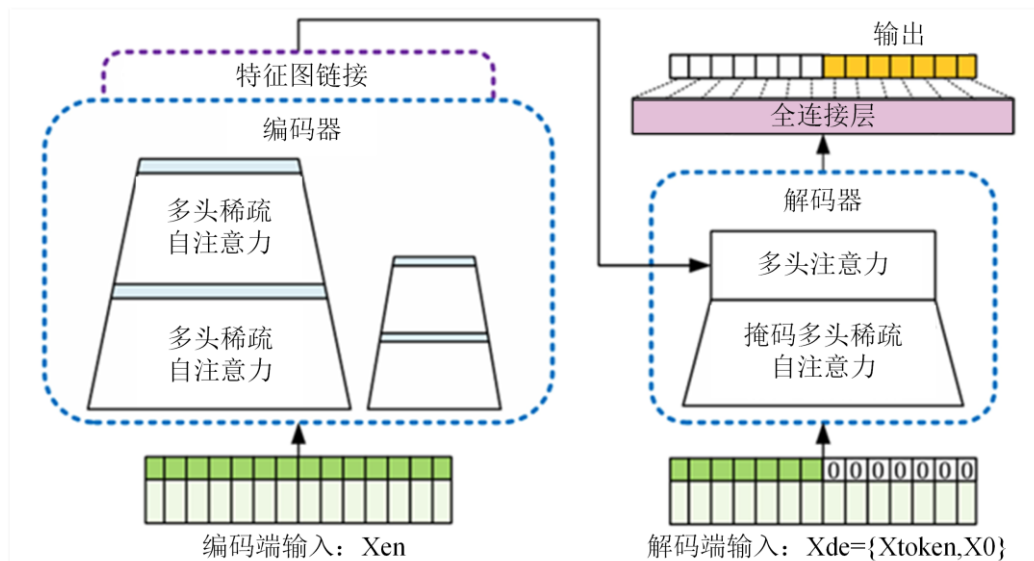


图 4-3 Informer 模型结构

传统的自注意力机制计算复杂度为序列长度的平方级。然而，研究发现，查询向量 Q 呈现长尾分布，即少数活跃的 Q 的点积注意力会有较大贡献，而其他懒散的 Q 的点积注意力呈现平均分布。这些均匀的点积注意力与 V 相乘，无太大区分度。通过计算 Kullback-Leibler 散度可把活跃的和懒散的 Q 区分开。计算公式为：

$$\overline{M}(Q_i, K) = \max\left\{\frac{Q_i K_j^T}{\sqrt{d_k}}\right\} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{Q_i K_j^T}{\sqrt{d_k}} \quad (4-13)$$

式中， L_K 为选取 K 的个数； $\max\left\{\frac{Q_i K_j^T}{\sqrt{d_k}}\right\}$ 为每个 Q_i 与 L_K 个 K 的最大得分；

$\frac{1}{L_K} \sum_{j=1}^{L_K} \frac{Q_i K_j^T}{\sqrt{d_k}}$ 为所有 Q_i 与 L_K 个 K 内积的均匀得分； $\overline{M}(Q_i, K)$ 为每个 Q 最高点积得分。

假设序列长度为 96，其对应 96 个 Q 、 K 、 V 值。Transformer 中需要计算 96 个 Q 和 96 个 K 求内积相关性。Informer 只需随机取出 25 个 K 值，使用 96 个 Q 与 25 个 K 计算点积得分，就能得到一个均匀分布得分；然后，取每个 Q 的最高得分计算与均匀得分的差值。

最后取 $TOP-u$ 个最高差值对应的 Q 即为活跃 Q ，其余 Q 即为懒散 Q 。如果 $TOP-u$ 为 30，那么只需计算 30 个 Q 与 96 个 K 的点积，进而与 V 相乘得到自注意力得分。其余懒散的 Q 对应的 V 值得分只需填充 V 的平均向量即可。这样模型的计算复杂度就从 $O(96^2)$ 下降到 $O(96 \times 30)$ ，大大节约内存。活跃的 Q 计算公式如下：

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{\overline{Q} K^T}{\sqrt{d_k}}\right) V \quad (4-14)$$

式中， $\overline{Q}$ 为 $Top-u$ 个活跃 Q 。

由于懒散的 Q 的得分为 V 值的平均向量填充，那么编码特征是具有冗余的值组合。而编码端通常堆叠多层编码单元，导致这些填充 V 值也需要继续堆叠计算，计算复杂度和内存使用率并没有在下层编码单元中得以降低。因此，Informer 在每一次计算完成自注意力矩阵后，都会进行一次下采样蒸馏操作。蒸馏操作不仅能显著降低输入序列的时间维度，还能使自注意特征图更加突出。第 n 层在第 $(n+1)$ 层的蒸馏过程如下：

$$X_{n+1}^t = \text{MaxPool}(\text{ELU}(\text{Conv1d}([X_n^t]_{AB}))) \quad (4-15)$$

式中, X_{n+1}^t 为第 $(n+1)$ 层多头概率稀疏自注意力输出, Conv1d 为 1DCNN, ELU 为激活函数, MaxPool 为最大池化函数, $[X_n^t]_{AB}$ 为第 n 层多头概率稀疏自注意力特征图。

与传统的 Transformer 分步解码不同, 生成式解码器通过一个全连接层直接输出多步预测。在解码端, 输入首先经过带掩码的多头稀疏概率自注意操作, 然后与编码器的输出特征图进行交叉注意操作, 最后通过全连接层一次性输出所有预测值。解码端的输入为:

$$X_{de} = \text{Concat}(X_{\text{token}}, X_0) \in \mathbb{R}^{(L_{\text{token}}+L_{\text{pre}}) \times d_{\text{model}}} \quad (4-16)$$

式中, $X_{\text{token}} \in \mathbb{R}^{L_{\text{token}} \times d_{\text{model}}}$ 为先验序列; $X_0 \in \mathbb{R}^{L_{\text{pre}} \times d_{\text{model}}}$ 为目标序列并填充零, 防止当前序列聚焦于未来序列, 从而避免自回归; $\text{Concat}()$ 表示将两个序列在时间维度上拼接; X_{de} 为解码器输入。

本文使用的 Informer 结构如图 4-4 的堆叠编码器模块和解码器模块所组成。受实验数据集长度限制, 为了增强 Informer 模型在少样本数据下的特征提取能力, 本文中的编码器由三个堆叠模块组成, 第一个堆栈由三层多头概率稀疏自注意层和两层蒸馏层组成。第二层堆栈由两层多头概率稀疏自注意层和一层蒸馏层组成。第三层是一层多头概率稀疏自注意层。分别将输入时间序列折半两次进行编码, 最后将 3 次编码结果拼接输出, 编码特征分别关注原始序列不同时间尺度的时序信息, 有利于解码器挖掘时序数据的长期依赖关系。

4.3 锂离子电池 SOH 点预测框架

为了验证所提出的 SOH 点预测框架的性能, 本论文在配备英特尔(R)酷睿(TM)i7-8550U CPU @ 1.80 GHz 2.00 GHz、8 GB 内存和 64 位 Windows 10 家庭版操作系统的笔记本电脑上进行实验。软件平台为 PyCharm Community Edition 2022.1.3 和 MATLAB 2020b。

Informer 模型在长序列时序预测方面具有明显优势, 但在训练样本较少的情况下表现不佳。此外, 随着预测周期长度的增加, 模型预测误差逐渐增大。单一框架无法获得令人满意的预测结果。相比之下, 1DCNN 可以在少样本下获得较好的预测精度, 并能非常准确地预测当前时间步的容量。本论文将 1DCNN 当前时间步预

测的容量近似为下一步的容量衰减结果，从而达到修正 Informer 模型预测误差的目的。综上所述，本文提出了一种模型集合框架，命名为 1DCNN-Informer，实现对锂离子电池进行准确的 SOH 预测。锂离子电池的 SOH 点预测框架如图 4-4 所示。

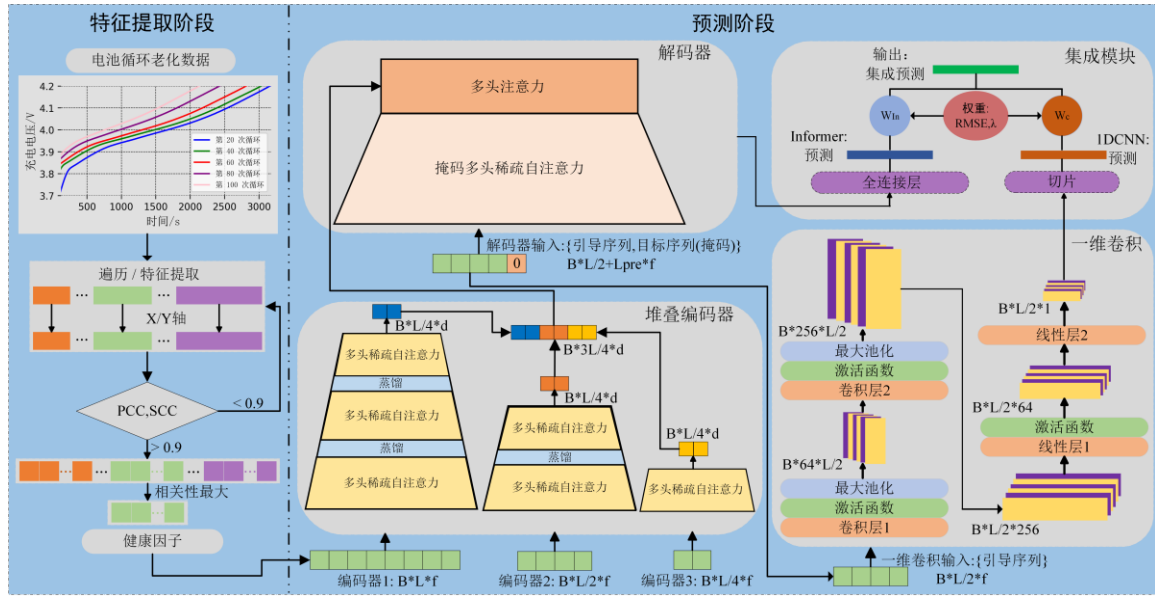


图 4-4 锂离子电池 SOH 点预测框架

如图 4-4 所示，该框架主要包括两个部分，即特征提取阶段和预测阶段。具体步骤如下：

步骤 1，特征提取阶段。

(1) 使用遍历法。通过调整特征间距，在循环老化曲线的特定范围内移动。得到一系列与周期相对应的间接特征。

(2) 利用皮尔逊相关系数（Pearson Correlation Coefficient, PCC）和斯皮尔曼相关系数（Spearman Correlation Coefficient, SCC）对间接特征与容量的相关性进行定量分析，选出 $|PCC \& SCC| > 0.9$ 的候选特征（它间接反映了特征与容量通常具有很强的线性和单调相关关系）。最后，选择 PCC 最大的特征作为 HIs，作为预测模型的输入特征序列。

步骤 2，预测阶段。

(1) Informer 模块。对编码器输入时间序列进行切片，取序列的后半部分得到三个不同尺度的时间序列，分别输入到一个堆栈中进行编码，最后将所有堆栈的特征图连接起来作为编码器的输出。通过解码器解码后，再通过全连接层得到 Informer 模型的 SOH 下一步预测结果。

(2) 1DCNN 模块。去掉解码端输入的掩码，将其作为 1DCNN 的输入。预测每个时间步对应的容量，然后通过切片得到最后一步预测的当前时间步容量。

(3) 集成模块。通过加权层对两个 SOH 预测结果分配权重并求和，得到最终的 SOH 下一步点预测结果。

在加权层中， W_{in} 和 W_c 分别代表 Informer 模型和 1DCNN 的自适应权重，它们的权重之和为 1。模型的预测误差越小，权重越大，反之亦然。 W_{in} 和 W_c 的计算公式如下：

$$W_{in} = 1 - \frac{\lambda Rmse_{in}}{Rmse_{in} + Rmse_c} \quad (4-17)$$

$$W_c = \frac{\lambda Rmse_{in}}{Rmse_{in} + Rmse_c} \quad (4-18)$$

式中， $Rmse_{in}$ 和 $Rmse_c$ 分别代表 Informer 模型和 1DCNN 在验证集上的容量预测结果与目标容量之间的均方根误差（Root Mean Square Error, RMSE）。 λ 是超参数，用于灵活控制模型权重。RMSE 的定义如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2} \quad (4-19)$$

式中， y_i 和 y_i^* 分别代表锂离子电池的实际容量与预测容量， n 为测试总循环周期次数。

所提框架的 SOH 点预测最终结果由两个模型的预测结果相加得出，其计算公式为：

$$SOH = \frac{W_{in} * C_{in} + W_c * C_c}{C_R} \times 100\% \quad (4-20)$$

式中， C_{in} 和 C_c 分别代表 Informer 模型和 1DCNN 预测容量， C_R 为锂离子电池出厂额定容量。

4.4 锂离子电池 SOH 点预测实验

除 SOH 外，剩余使用寿命（Remaining Useful Life, RUL）也是锂离子电池健康状况的关键指标之一。RUL 是指从当前状态到电池寿命结束的剩余充放电循环

次数。虽然表达方式不同，但两个指标的核心要点都是准确预测容量衰减趋势。为了分析预测性能，除公式 (4-19) 的 RMSE 指标外，还将采用预测结果中实际 RUL 与换算后 RUL 的绝对误差 (Absolute Error, AE) 作为 SOH 预测的衡量指标，从预测结果的准确性和实用性两方面来证明所提框架预测结果的准确性。AE 指标的计算公式如下：

$$AE = |RUL_{true} - RUL_{pre}| \quad (4-21)$$

式中， RUL_{true} 为实际剩余使用寿命， RUL_{pre} 为预测容量序列转换的剩余使用寿命。

4.4.1 SOH 点预测结果

本文使用了两组六节锂离子电池数据来验证所提出的框架。每个电池循环寿命周期分为 40%、50% 和 60% 的训练集，5% 的验证集，以及相应的测试集长度。B0005、B0006 和 B0007 的训练集长度分别为 67、84 和 100 个循环周期，CS235 的训练集长度分别为 286、358 和 430 个循环周期，CS236 的训练集长度分别为 277、347 和 416 个循环，CS237 的训练集长度分别为 303、379 和 454 个循环。如前所述，当充电容量下降到额定容量的 80% 时，电池通常会被视为失效。但是，由于电池 B0005、B0006 和 B0007 的循环寿命较短，本论文将 B0005 和 B0006 的额定容量定义为 70%。此外，由于 B0007 的 SOH 大于或等于 70%，本论文将 B0007 的额定容量定义为 71%，以实现 AE 指标的计算。

表 4-1 Informer 模型参数

参数	值	参数	值
轮数	20	编码器层数	[3,2,1]
编码器输入序列长度	20	解码器层数	1
解码器引导/预测序列长度	10/1	批量大小	4
嵌入维度	256	丢弃率	0.05
注意力头数	6	学习率	0.0001

Informer 模型适用于长序列时序预测。考虑到电池数据长度的限制，本文通过堆叠编码器来充分提取输入序列不同时间维度的特征。编码器输入序列的长度为 20，本论文依次堆叠了三个输入序列长度分别为 20、10 和 5 的稀疏概率自注意力编码器。通过对三个堆栈的输入序列进行特征提取，分别得到长度为 5 的三个特征图，然后拼接成长度为 15 的特征图。考虑到锂离子电池 SOH 预测精度的特点，解码器的输入由长度为 10 的先验序列和长度为 1 的带掩码的目标序列组成。注意力

头数为 HIs 个数，通过调整原始堆叠编码器层数和解码器层数，通过实验对比发现的层数的叠加没有明显的精度提升，却大大加大训练时长。最终确定堆叠编码器层数为 [3/2/1]，解码器层数为 1。Informer 模型参数如表 4-1 所示。

由于单一模型对数据的依赖性很强，因此可能会导致估计精度低、泛化性能差等问题。在样本量较小的情况下，Informer 模型表现不佳。随着预测轮数的增加，特征分布的变化也会导致预测误差的增加。1DCNN 可以在训练样本较少的情况下获得良好的预测结果。本文使用 Informer 解码器的先验序列作为 1DCNN 的输入，确保时间尺度的一致性同时滤波器数量与 Informer 嵌入维度保持一致。该模型获得每个时间步的容量预测结果，然后取上一步的容量，近似为下一步容量衰减的结果。在锂电池的长周期衰减期，相邻两步的容量指标非常接近。1DCNN 当前时间步骤的预测结果用于修正 Informer 模型的下一步容量预测结果，以提高预测的准确性。最后，通过权重层分配两个模型的权重，得到锂电池的最终 SOH 预测结果。1DCNN 结构与参数如表 4-2 所示。

表 4-2 1DCNN 结构与参数

网络层	滤波器大小	滤波器/神经元数量	激活函数
卷积层 1	2	64	Relu
池化层 1	2	/	最大池化
卷积层 2	2	256	Relu
池化层 2	2	/	最大池化
线性层 1	/	64	Relu
线性层 2	/	1	/

六组锂电池 SOH 点预测结果如图 4-5 至图 4-10 所示。图 4-5 到图 4-7 为 B0007、B0006、B0007 电池实验结果。可以看出，三组电池的 SOH 预测曲线都较为紧密的贴合真实 SOH 退化曲线。相较于 40% 的训练周期，50% 和 60% 训练周期下的 RUL 指标误差更小。同时，除去 40% 训练周期下因三组电池在第 87 周期左右有较大的容量再生波动带来的预测误差，其余预测的 SOH 误差都在很小的误差范围内波动。图 4-8 到图 4-10 为 CS235、CS236、CS237 电池实验结果。可以直观的看出，随着训练样本的增多，三组电池的 SOH 预测趋势十分紧密的贴合真实 SOH 退化曲线，特别是在电池 SOH 在 80% 之前，三组电池 SOH 预测误差的绝对值小于 0.02。从 RUL 指标上看，预测指标基本与真实指标重合。

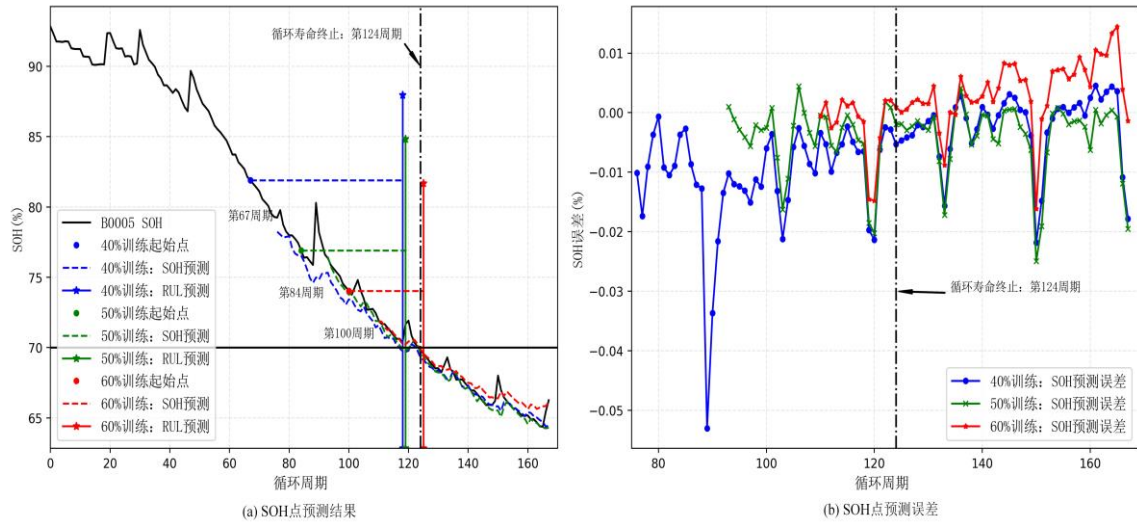


图 4-5 B0005 电池 SOH 点预测结果与误差

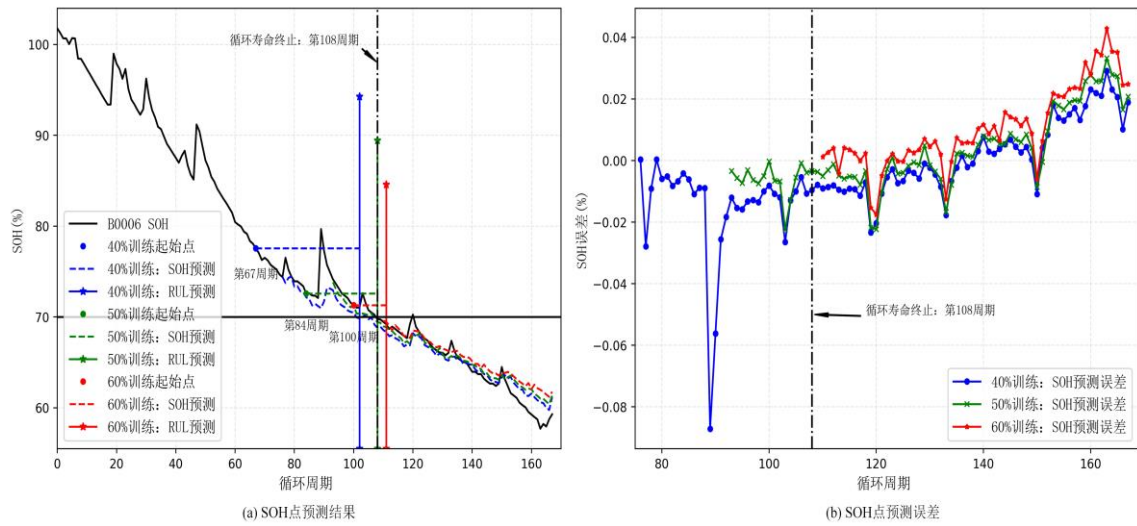


图 4-6 B0006 电池 SOH 点预测结果与误差

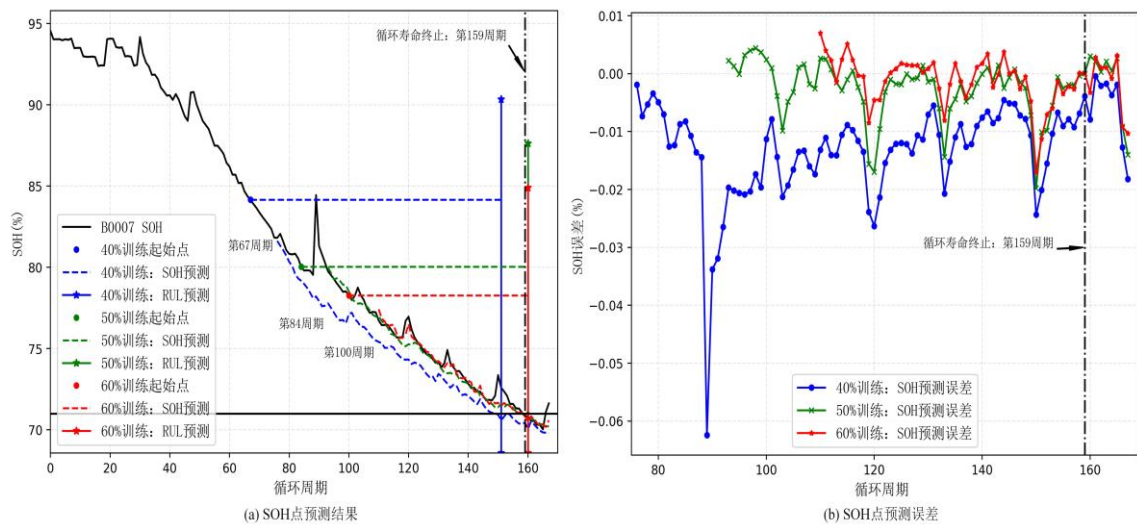


图 4-7 B0007 电池 SOH 点预测结果与误差

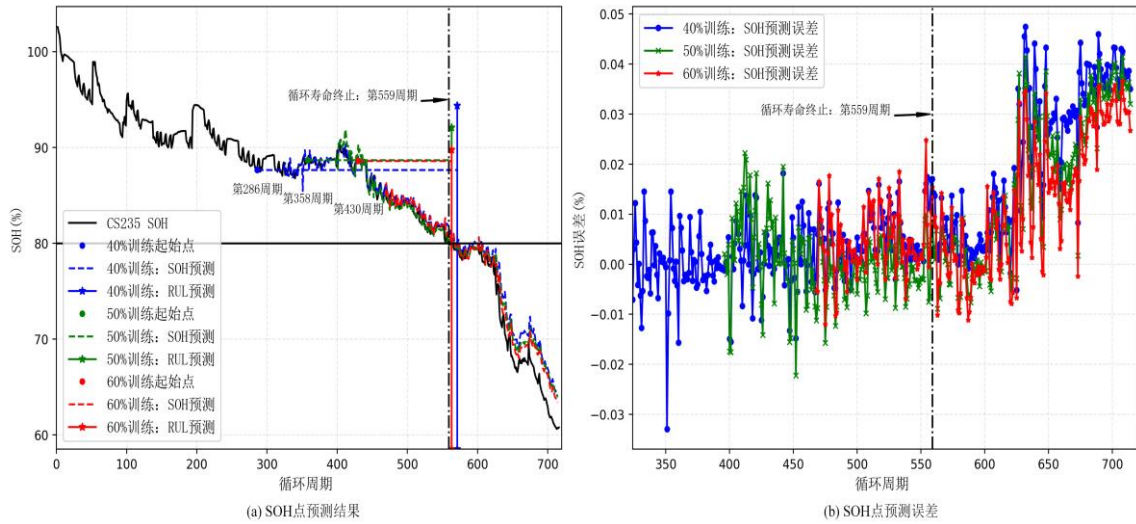


图 4-8 CS235 电池 SOH 点预测结果与误差

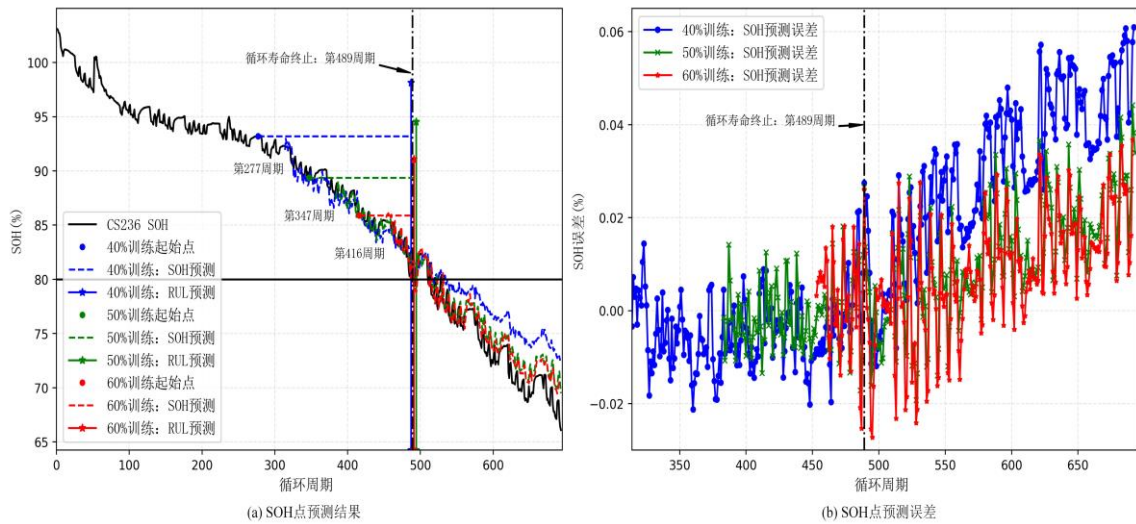


图 4-9 CS236 电池 SOH 点预测结果与误差

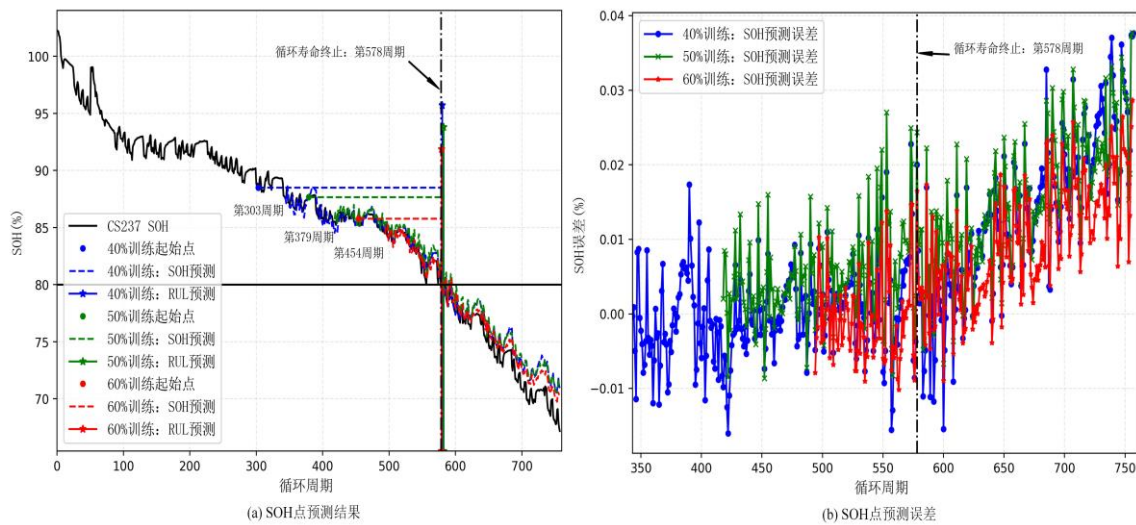


图 4-10 CS237 电池 SOH 点预测结果与误差

4.4.2 不同序列预测模式比较

锂离子电池在经历一定次数的充放电循环后，通常会出现容量衰减速度加快的趋势，其特征分布会发生变化，因此很难根据历史数据准确预测未来的趋势。单一数据驱动模型容易受到短期训练数据的影响，模型的短期训练权重不能有效地适应长序列预测。

为客观评价锂离子电池的 SOH 预测，表 4-3 列出了 RMSE 和 AE 指标。从表 4-3 中可以发现，虽然 1DCNN（预测当前时间步 SOH）和 Informer（预测下一时间步 SOH）模型在 RMSE 指标上都能达到较好的精度，但在 AE 指标上并不稳定，尤其是 Informer 模型，它受到 HIs 特征噪声的影响，在 AE 指标上有较大的跳变。为了更佳直观的说明单一模型和所提框架的预测精度，以六组电池 50% 训练周期示例来说明这一点，如图 4-11 所示。结合表 4-3 和图 4-11 可以看出，在锂离子电池 SOH 预测任务中，1DCNN 的准确率随时间序列的延长而下降。相比于在 NASA 数据集，1DCNN 在 CALCE 数据集上预测精度有所下降。随着迭代预测周期的增加，其精度越发偏离实际值。而 Informer 的预测结果波动较大，这与 CALCE 数据集 HIs 存在波动有关，说明了单一模型存在的局限性。本文提出的框架可以克服单一模型的不足。通过结合两个模型在不同时间步长的预测结果，可以在短期和长期预测任务以及 RMSE 和 AE 指标方面取得更好的结果。

表 4-3 锂离子电池 SOH 点预测的 RMSE 和 AE 结果

电 池	预测起点/ 剩余寿命	预测剩余寿命/ AE / RMSE			电 池	预测起点/ 剩余寿命	预测剩余寿命/ AE / RMSE		
		1DCNN	Informer	1DCNN- Informer			1DCNN	Informer	1DCNN- Informer
B0005	67/57	49/8/0.0152	60/3/0.0240	50/7/0.0109	CS235	286/273	330/57/0.0171	241/32/0.0264	285/12/0.0174
	84/40	34/6/0.0095	37/3/0.0150	34/6/0.0070		358/201	217/16/0.0182	166/35/0.0246	205/4/0.0163
	100/24	27/3/0.0052	19/5/0.0094	25/1/0.0063		430/129	145/16/0.0164	108/21/0.0231	133/4/0.0156
B0006	67/41	35/6/0.0128	23/8/0.0266	34/7/0.0162	CS236	277/212	223/25/0.0300	119/93/0.0363	210/2/0.0261
	84/24	25/1/0.0133	18/6/0.0154	24/0/0.0124		347/142	169/27/0.0127	102/40/0.0287	147/5/0.0147
	100/8	-/-/0.0189	-/-/0.0139	-/-/0.0164		416/73	77/4/0.0131	63/10/0.0232	75/2/0.0146
B0007	67/92	74/8/0.0200	-/-/0.0182	83/9/0.0153	CS237	303/275	278/3/0.0095	247/28/0.0232	277/2/0.0122
	84/75	68/7/0.0061	70/5/0.0077	76/1/0.0055		379/199	208/9/0.0128	200/1/0.0158	203/4/0.0141
	100/59	63/4/0.0040	51/8/0.0063	60/1/0.0044		454/124	127/3/0.0123	92/32/0.0157	125/1/0.0107

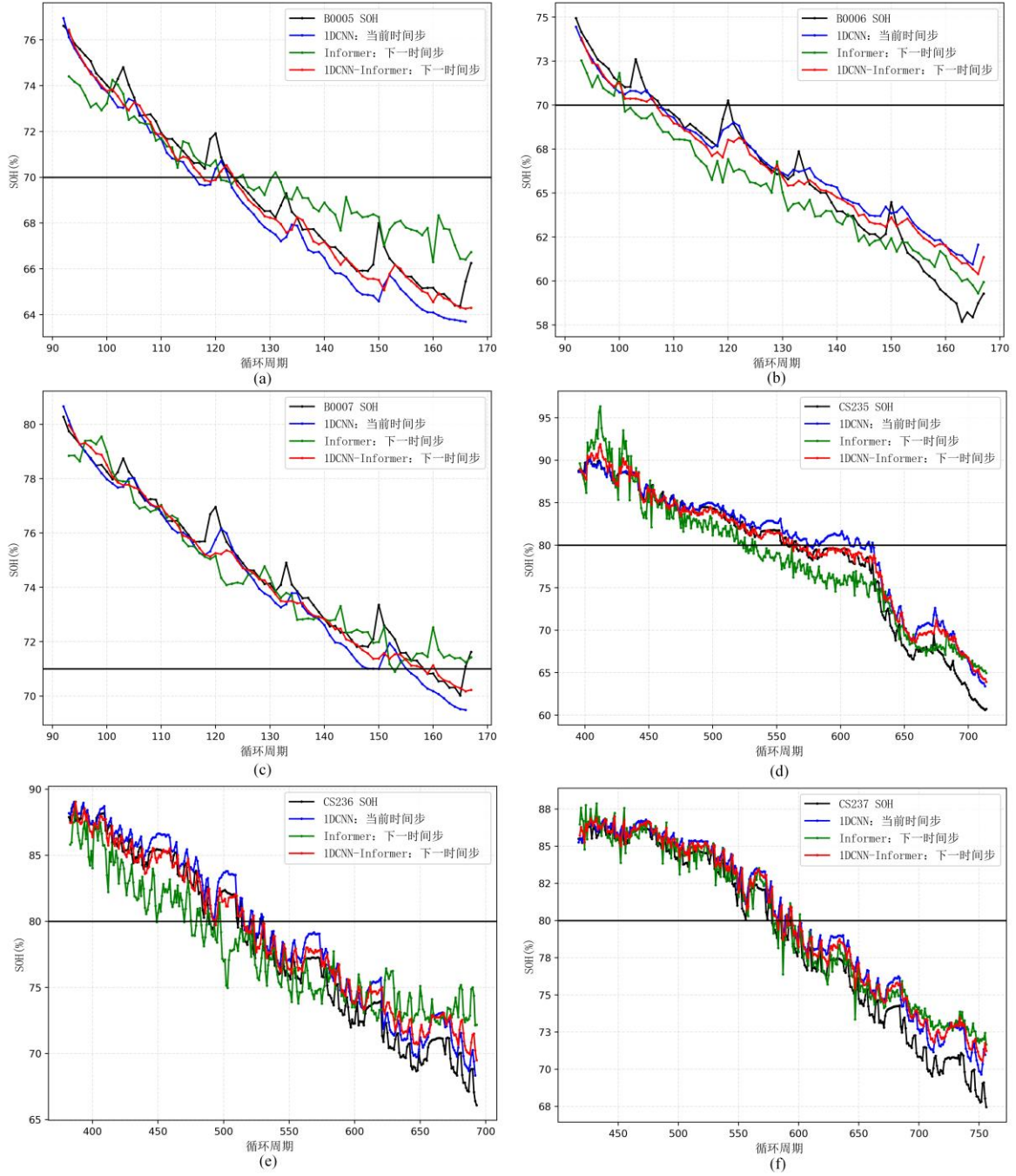


图 4-11 50% 训练周期下单一模型与集合框架 SOH 点预测结果比较

4.4.3 集合框架不同权重的比较

在计算验证集上的均方根误差来分配测试集上的模型权重时，短期性能较好的 1DCNN 往往在权重分配中占据主导地位，忽略了 Informer 模型的长期预测优势。因此，本文引入超参数 λ 来灵活调整集成模型的权重分配，充分考虑了 1DCNN 长期预测能力的偏差和 Informer 模型在样本较少情况下训练能力的不足。

为了说明加权模块的优越性，本文以训练周期为 40% 的 B0006 和 CS236 电

池为例,如表 4-4 所示。从表 4-4 可知,与固定权重和计算均方根误差相比,本文提出的方法灵活且具有更高的预测精度。当训练数据较少、预测周期较短时,1DCNN 的预测精度较高,因此令 λ 稍大于 1,增加 1DCNN 的权重;当训练数据较多、预测周期较长时,Informer 模型的预测精度提高,同时 1DCNN 测试集的精度不能很好地反映其长期预测能力的偏差,因此令 λ 稍小于 1,增加 Informer 模型的权重。通过引入超参数 λ ,可以平衡 1DCNN 短期和长期预测性能的影响,寻找合适的权重分配,适应不同任务所需,提高 SOH 预测精度。

表 4-4 不同加权方法的 RMSE 比较

电池(训练集占比)	RMSE/权重				
	固定权重			无 λ ($\lambda=1$)	λ
B0006(40%)	0.0197/	0.0184/	0.0210/	0.0177/	0.0162/ $\lambda=1.2$ /
	$W_{in}=W_c=0.5$	$W_{in}=0.4/W_c=0.6$	$W_{in}=0.6/W_c=0.4$	$W_{in}=0.35/W_c=0.65$	$W_{in}=0.22/W_c=0.78$
CS236(40%)	0.0265/	0.0262/	0.0274/	0.0275/	0.0261/ $\lambda=0.75$ /
	$W_{in}=W_c=0.5$	$W_{in}=0.4/W_c=0.6$	$W_{in}=0.6/W_c=0.4$	$W_{in}=0.20/W_c=0.80$	$W_{in}=0.41/W_c=0.59$

4.4.4 所提框架与其它框架 SOH 点预测比较

为了更好地验证拟议框架在锂离子电池 SOH 点预测方面的卓越性能,表 4-5 列出了包括 RESE 和 AE 与其他近年来发表在 SCI 期刊上的主流方法的比较结果。B0005、B0006 和 B0007 与 MFE-GRU-TCA^[45]、SVR-CEEMDAN-LSTM^[51]、DTV-LSTM^[78]、GBLS Booster^[79]进行了比较。CS235、CS236 和 CS237 与 CEEMDAN-Transformer-DNN^[80]、GBLS Booster、Bi-LSTM-PSO-DGPR^[81]、IFO-PSO-ACO-CatBoost^[82]、IPSO-PF^[83]进行了比较(一部分框架只包含部分电池)。

从表 4-5 中可以看出,所提框架在 AE 和 RMSE 方面都能提供有竞争力的预测精度。同其它框架相比,所提框架可全部适用于六组电池。均方根误差指标优于大多数其它框架,当均方根误差指标相近时,所提框架也能提供更准确的均方根误差指标。对比结果证明,所提框架在 SOH 点预测方面具有更高的准确性。

表 4-5 不同加权方法的误差比较

电池	框架	预测起点	AE	RMSE
B0005	MFE-GRU-TCA [45]	60/90	0/1	0.0083/0.0090
	SVR-CEEMDAN-LSTM [51]	60/80/100	2/2/4	0.0291/0.0136/0.0060
	DTV-LSTM [78]	67	/	0.0215
	GBLS Booster [79]	50/70/90	/	0.0205/0.0096/0.0081
	本文方法	67/84/100	7/6/1	0.0109/0.0070/0.0063

续表 4-5 不同加权方法的误差比较

电池	框架	预测起点	AE	RMSE
B0006	MFE-GRU-TCA [45]	60/90	0/1	0.0130/0.0092
	SVR-CEEMDAN-LSTM [51]	60/80/100	10/4/1	0.0664/0.0231/0.0155
	DTV-LSTM [78]	67	/	0.0109
	GBLS Booster [79]	50/70/90	/	0.0107/0.0094/0.0098
	本文方法	67/84/100	7/0/-	0.0162/0.0124/0.0164
B0007	MFE-GRU-TCA [45]	60/90	1/1	0.0117/0.0082
	SVR-CEEMDAN-LSTM [51]	60/80/100	5/0/6	0.0142/0.0087/0.0060
	DTV-LSTM [78]	67	/	0.0133
	GBLS Booster [79]	50/70/90	/	0.01270.0073/0.0094
	本文方法	67/84/100	9/1/1	0.0153/0.0055/0.0044
CS235	GBLS Booster [79]	400/600	9/1	0.0617/0.0181
	CEEMDAN-Transformer-DNN [80]	300/462	1/1	/
	Bi-LSTM-PSO-DGPR [81]	450	/	0.0090
	IPO-PSO-ACO-CatBoost [82]	300	4	0.0126
	本文方法	286/358/430	12/4/4	0.0174/0.0163/0.0156
CS236	GBLS Booster [79]	400/600	9/4	0.0578/0.0181
	CEEMDAN-Transformer-DNN [80]	300/599	11/1	/
	IPO-PSO-ACO-CatBoost [82]	300	2	0.0126
	IPSO-PF [83]	400	7	0.0120
	本文方法	277/347/416	2/5/2	0.0261/0.0147/0.0146
CS237	CEEMDAN-Transformer-DNN [80]	300/581	22/6	/
	Bi-LSTM-PSO-DGPR [81]	470	/	0.0062
	IPO-PSO-ACO-CatBoost [82]	300	1	0.0099
	IPSO-PF [83]	400	5	0.0220
	本文方法	303/379/454	2/4/1	0.0122/0.0141/0.0107

4.5 本章小结

本章主要介绍了 1DCNN、Transformer、Informer 的算法原理，并提出了基于 1DCNN-Informer 框架的锂离子电池 SOH 点预测框架。所提框架充分结合了 1DCNN 在特征提取、短序列预测和 Informer 在长序列预测方面的优势，避免了单一模型的不足，提高了 SOH 点预测的准确性。依次通过六组 SOH 预测实验，50% 训练周期下不同序列预测模式比较，集合框架不同加权方式的比较以及和其它主流 SOH 预测框架的比较，证明所提 SOH 点预测框架不仅能适应不同周期长度的预测任务而且预测精度高。

第 5 章 基于 CEEMDAN-LSTM 框架的锂离子电池 SOH 周期预测

对锂离子电池展开准确的 SOH 点预测需要实时更新电池当前运行数据，在实际应用偏向短期 SOH 预测。周期预测目的是利用电池历史数据一次性预测未来一段周期内的电池 SOH 变化，可帮助用户提前规划电池的更换或维护，同时周期预测可评估电池退役后的剩余价值，促进电池的循环经济，减少资源浪费。容量再生现象是影响电池 SOH 周期准确预测的主要因素。前文已经详细分析了容量再生波动的原因和特性，并指出传统实验室环境下针对容量再生问题提出的分解-集成预测方法并不能有效的适用实际应用所需。为了降低电池容量退化过程中容量再生波动对锂电池 SOH 周期预测的负面影响，本文更加关注再生容量的快速下降过程。首先通过完全自适应噪声集合经验模态分解（Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMDAN）解耦电池原始容量退化曲线并计算皮尔斯相关系数（Pearson Correlation Coefficient, PCC）和斯皮尔曼相关系数（Spearman Correlation Coefficient, SCC）以剔除不相关的高频子序列，然后分别构造双向长短期记忆网络（Bidirectional Long Short-Term Memory, Bi-LSTM）和长短期记忆网络（Long Short-Term Memory, LSTM）分别预测残差分量和相关高频子序列，最后将预测结果相加集成获得最终 SOH 周期预测结果。

5.1 CEEMDAN 原理

锂电池原始容量退化曲线是一个包含多个局部特征的非平稳信号，容量再生波动发生的循环周期和波动大小都是高度随机的。信号分解是将复杂信号分解为若干个简单信号分量的过程，旨在更深入地分析信号的特性、提取有用信息以及便于进行后续的处理和系统分析。

经验模态分解（Empirical Mode Decomposition, EMD）依据信号自身的时间尺度特征将其分解为一系列本征模态函数（Intrinsic Mode Functions, IMFs）和一个残差信号（Residual, RES）。每个本征模态函数具有不同的频率特性，能自适应地分解复杂的非线性、非平稳信号。然而，EMD 在处理非线性和非平稳信号时存在一些

问题，如在 IMF 分解中可能存在模态混叠，使得某个 IMF 中包含了本应属于其他 IMF 的成分。为了克服模态混叠等问题，吴兆华等人^[84]在 EMD 分解中添加平均值为 0 的高斯白色噪声用于辅助分析，即集成经验模态分解 (Ensemble Empirical Mode Decomposition, EEMD)。然而，分解后的 IMFs 可能仍然会包含残余的白噪声，影响后续的信号分析处理。因此，Torres 等人^[85]提出了 CEEMDAN 来解决将白色噪声从高频转移到低频的问题，在分解的每个阶段都加入特定噪声，并计算唯一的 RES 以获得每个 IMF。分解结果是完整的，误差在数值上可以忽略不计。CEEMDAN 分解流程如图 5-3 左边流程图所示，具体计算公式如下：

(1) 输入待分解的信号为 $x(t)$ ，设定迭代次数 N ，集成次数 K 。

(2) 引入 K 次均值为 0 的高斯白噪声 $\varepsilon_j \delta_i(t)$ ($i=1,2,3,\dots,K$)， ε_j ($j=1,2,3,\dots,N$) 为高斯白噪声权值系数， $\varepsilon_j \delta_i(t)$ 为第 i 次添加时的高斯白噪声。形成共 K 个扰动信号 $x_i(t)$ 。

$$x_i(t) = x(t) + \varepsilon_j \delta_i(t) \quad (5-1)$$

(3) 对扰动信号 $x_i(t)$ 进行 EMD 分解，得到第一组模态分量 IMF_1^i 并取其均值作为 $IMF_1(t)$ 并更新余量信号 $r_1(t)$ 。

$$IMF_1(t) = \frac{1}{K} \sum_{i=1}^K IMF_1^i \quad (5-2)$$

$$r_1(t) = x(t) - IMF_1(t) \quad (5-3)$$

(4) 将分解后得到的第 j ($j=1,2,3,\dots,N$) 阶段的余量信号添加白噪声 $\varepsilon \delta_i(t)$ 后，继续进行 EMD 分解。

$$IMF_j(t) = \frac{1}{K} \sum_{i=1}^K E_j(r_{j-1}(t) + \varepsilon_{j-1} E_{j-1}(\delta_i(t))) \quad (5-4)$$

$$r_j(t) = r_{j-1}(t) - IMF_j(t) \quad (5-5)$$

式中， $IMF_j(t)$ 为第 j 个 IMF， E_{j-1} 表示对信号进行 EMD 分解后的第 $j-1$ 个 IMF 分量， ε_{j-1} 表示对第 $j-1$ 阶段余量信号加入噪声的权值系数， $r_j(t)$ 表示第 j 阶段余量信号。

(5) 重复第 4 步骤，直到 RES 无法继续分解出有意义的 IMF，CEEMDAN 分解结束。此时原始信号 $x(t)$ 可表示为：

$$x(t) = \sum_{j=1}^N IMF_j(t) + r_N(t) \quad (5-6)$$

通过 CEEMDAN 分解法将锂离子电池原始容量退化曲线分解为 n 个 IMF 分量和一个 RES 分量。IMF 分量具有特定的频率和振幅，可以描述容量的局部容量再生波动；而 RES 是分解信号的剩余部分，包含了容量退化的主要趋势，可用于进一步 SOH 周期预测。

5.2 LSTM 原理

5.2.1 LSTM

在传统的循环神经网络（Recurrent Neural Network, RNN）中，存在着长期依赖问题，即当序列较长时，RNN 难以捕捉到距离较远的信息之间的依赖关系，导致在处理长序列数据时性能下降。为解决长期依赖问题，德国 Hochreiter 等人^[86]提出 LSTM。LSTM 网络通过引入记忆单元和门控机制，有效地捕捉和利用输入序列中的长期依赖关系。图 5-1 为 LSTM 网络结构。

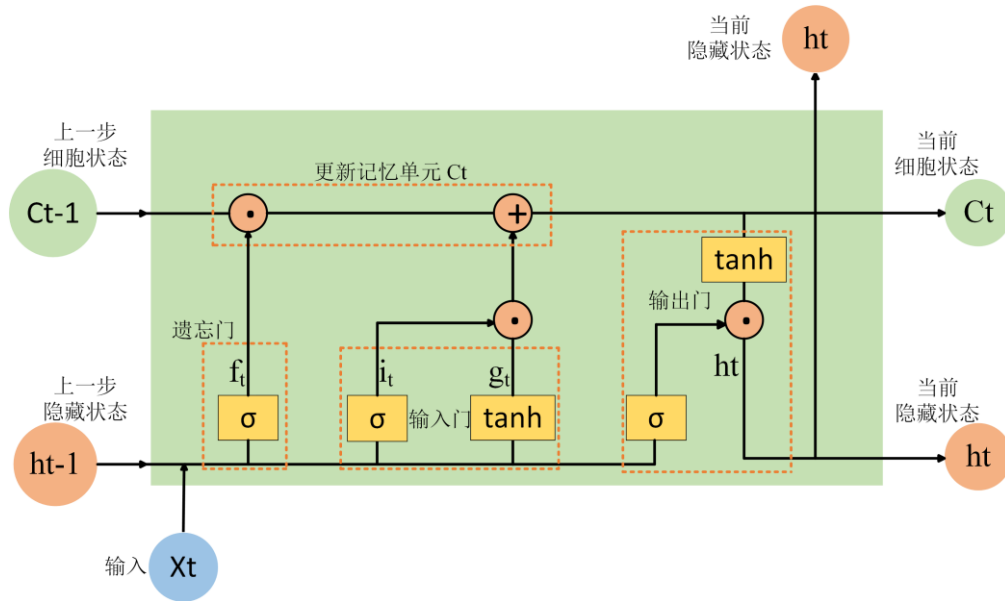


图 5-1 LSTM 网络结构

如图所示，LSTM 输入为当前时间步长 t 输入序列 x_t ，前一个时间步长 $t-1$ 的隐藏状态 h_{t-1} 和细胞状态 c_{t-1} ，输出为当前时间步长 t 的隐藏状态 h_t 和细胞状态 c_t 。隐藏状态包含了序列的信息和特征，细胞状态为记忆信息。 σ 是 sigmoid 激活函数，将输入值压缩到 0 到 1 之间的范围。Tanh 函数也是一个非线性的函数，将输入值压缩到 -1 到 1 之间的范围。

LSTM 网络的核心组成部分是记忆单元（Memory Cell），它能够存储和更新信息，并控制信息的流动。通过遗忘门（Forget Gate）和输入门（Input Gate），LSTM

能够选择性地遗忘和存储记忆单元中的信息。输出门 (Output Gate) 决定了要传递到下一层的输出。公式如下:

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \quad (5-7)$$

式中, W_{if} 和 b_{if} 是输入 x_t 的权重矩阵和偏移项, W_{hf} 和 b_{hf} 是上一时刻隐藏状态的权重矩阵和偏移项。经 sigmoid 激活函数后得到遗忘门输出 f_t , 输出一个值为 0 到 1 之间向量, 表示对前一时刻记忆细胞状态的保留程度。

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \quad (5-8)$$

$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \quad (5-9)$$

式中, W_{ii} , W_{ig} 和 b_{ii} , b_{ig} 是输入 x_t 的权重矩阵和偏移项, W_{hi} , W_{hg} 和 b_{hi} , b_{hg} 是上一时刻隐藏状态的权重矩阵和偏移项。经 sigmoid 激活函数和 tanh 激活函数后得到输入门输出 i_t 和 g_t , 分别输出一个 0 到 1 之间的值和一个 -1 到 1 之间的值。输入门负责更新记忆细胞的状态, 有选择的将信息添加到记忆细胞状态中。

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (5-10)$$

式中, $\odot$ 表示点积, 通过遗忘门, 输入门和上一时刻隐藏状态完成对当前时刻记忆细胞的更新。

$$h_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \odot \tanh(c_t) \quad (5-11)$$

式中, W_{io} 和 b_{io} 是输入 x_t 的权重矩阵和偏移项, W_{ho} 和 b_{ho} 是上一时刻隐藏状态 h_{t-1} 的权重矩阵和偏移项。通过与当前时刻记忆细胞的共同作用得到当前时刻的隐藏状态 h_t 。

5.2.2 Bi-LSTM

Bi-LSTM 是 LSTM 的一种扩展形式, 它同时获取输入序列的前向和后向信息, 能更好地捕捉输入序列数据的特征^[87]。对于 Bi-LSTM 网络来说, 每次的输出都是 LSTM 双向输出的组合。这种结构使 Bi-LSTM 能够更好地捕捉序列中的前后时间关系, 提供更全面的底层特征提取能力。图 5-2 展示了 Bi-LSTM 的简化结构。

前向 LSTM 可表述如下:

$$\vec{h}_t = LSTM(\vec{h}_{t-1}, x_t) \quad (5-12)$$

而后向 LSTM 可以表述如下:

$$\vec{h}_t = LSTM(\vec{h}_{t+1}, x_t) \quad (5-13)$$

因此，Bi-LSTM 模型当前状态输出可表示为：

$$h_t = [\vec{h}_t, \vec{h}_t] \quad (5-14)$$

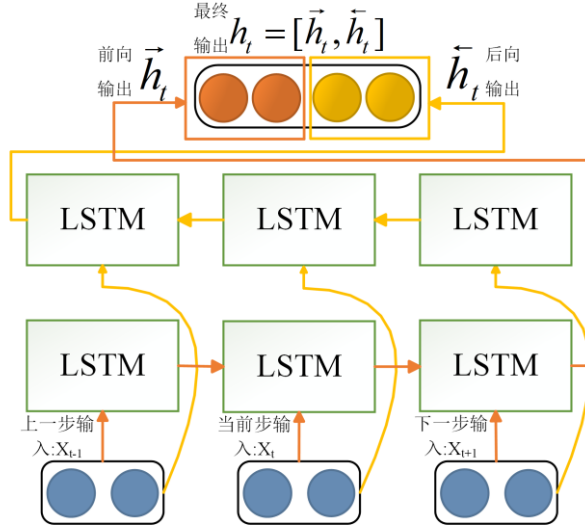


图 5-2 Bi-LSTM 的简化结构

本文选用 Bi-LSTM 网络预测 RES 的衰减趋势，该网络考虑了 RES 的前向和后向信息，能更好地反映电池容量的衰减趋势。PCC 和 SCC 绝对值均大于 0.2 的 IMF 分量可以有效补偿主要衰减趋势与原始电池容量的误差。由于其包含的整体衰减趋势和局部容量再生波动特征较少，因此采用 LSTM 网络预测。

5.3 锂离子电池 SOH 周期预测框架

为了验证所提出的 SOH 周期预测框架的性能，本论文在配备英特尔 (R) 酷睿 (TM) i7-8550U CPU @ 1.80 GHz 2.00 GHz、8 GB 内存和 64 位 Windows 10 家庭版操作系统的笔记本电脑上进行实验。软件平台为 PyCharm Community Edition 2022.1.3 和 MATLAB 2020b。

容量再生只是一个短暂的过程，其不会改善电池的老化状态。再生容量会在之后的循环中快速消退，使得电池容量退化呈现短暂不规则的波动现象。锂离子电池出现容量再生需要满足充足的休息时间，这在实际应用中并没有规律可言。传统实验室环境下可行的分解-集成预测方法并不能有效的适用实际应用所需。基于前文阐述分析（具体见第 2.4 节：锂离子电池容量再生问题原因和第 3.5 节：容量再生现象特性分析），本文提出了一种新的锂离子电池 SOH 周期预测框架，命名为

CEEMDAN-LSTM。与主流预测方法相比，更多地考虑了电池容量快速下降的整体趋势。在本研究中，容量再生波动部分被视为干扰或噪音，应在 SOH 周期预测过程中予以消除。SOH 周期预测实验流程图如图 5-3 所示。

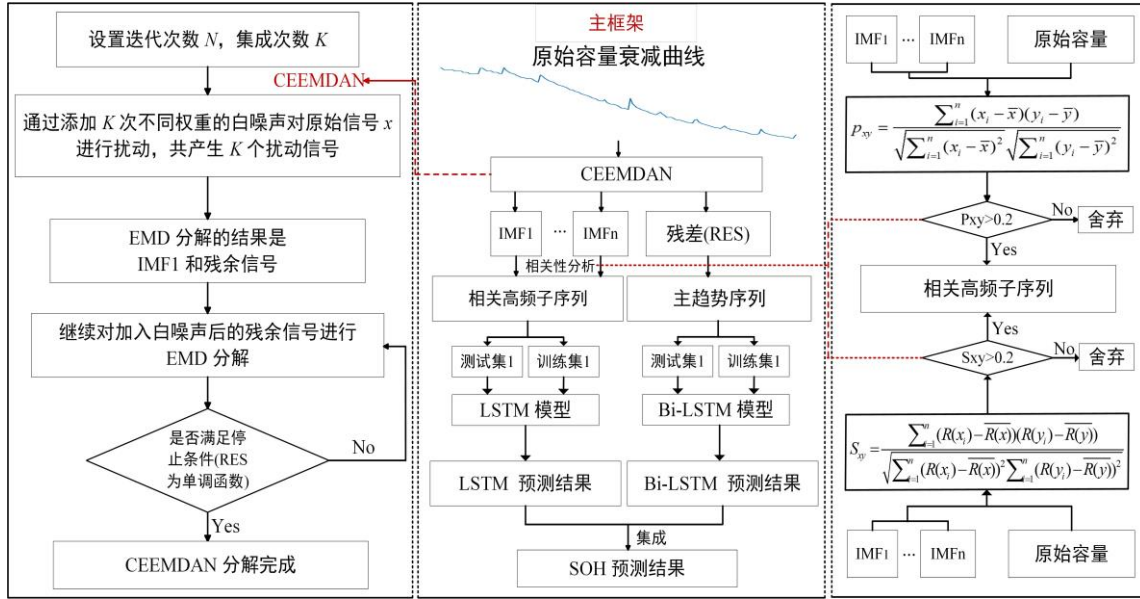


图 5-3 锂离子电池 SOH 周期预测框架

如图 5-3 所示，该框架主要包括三个部分，从左到右依次为 CEEMDAN 分解流程图，SOH 周期预测主框架，PCC 和 SCC 相关性分析。主框架具体步骤如下：

(1) 以电池原始容量衰减曲线为输入，通过 CEEMDAN 对其进行分解，得到 n 个不同的子分量 IMF 和一个残差信号 RES。

(2) 计算 IMFs 与原始容量衰减曲线之间的 PCC 和 SCC。舍弃任一相关性绝对值低于 0.2 的不相关高频子序列，只保留 PCC 和 SCC 绝对值均大于 0.2 的相关高频子序列和主趋势序列 RES。

(3) 将相关高频子序列和残差信号分成不同的训练集，并设定不同的预测起点。分别通过 LSTM 和 Bi-LSTM 模型进行训练，然后分别完成 SOH 周期预测任务。最后，将预测结果进行集成相加，得到最终的电池 SOH 周期预测结果，其计算公式如下：

$$SOH = \frac{\sum_{i=1}^n (BiLSTM(h_{t_0}, r(t)) + LSTM(h_{t_0}, IMF_i(t)))}{C_R} \times 100\% \quad (5-15)$$

式中， h_{t_0} 代表初始隐藏状态， $r(t)$ 代表残差信号， $IMF_i(t)$ 代表保留的相关高频子序列， C_R 代表电池的额定出厂容量。

5.4 锂离子电池 SOH 周期预测实验

为了分析预测性能，除了公式 (4-19) 的 RMSE 指标外，还将以平均绝对误差 (Mean Absolute Error, MAE) 指标横流 SOH 周期预测误差，其计算公式如下：

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_i^*| \quad (5-16)$$

式中， y_i 和 y_i^* 分别代表锂离子电池的实际容量与预测容量， n 为预测循环周期长度。

5.4.1 容量解耦与相关性分析

以 B0005 和 CS235 电池为例，针对容量再生波动现象对 SOH 周期预测精度和稳定性的负面影响，采用 CEEMDAN 将电池的原始容量衰减曲线解耦。B0005 电池解耦结果如图 5-4 所示，由 6 个 IMF 分量和 1 个 RES 分量组成。CS235 电池解耦结果如图 5-5 所示，由 8 个 IMF 分量和 1 个 RES 分量组成。

可以直观的看出，相较于图 2-2 和图 2-3 电池原始容量衰减曲线，RES 曲线光滑，表征了电池容量退化的主要趋势，即锂电池退化过程中容量下降的整体趋势是缓慢下降。IMFs 分量代表了容量退化曲线在不同老化周期尺度上的振荡模式，其频率和振幅会随着老化周期变化，反映了容量再生波动等局部特征。相较于 RES，大部分 IMFs 不仅振幅幅值十分微小，而且前后规律性并不紧密。如图 5-4 中 B0005 电池的 IMF1 分量，其绝大部分波动幅值无限趋近于 0，而其中较为强烈的容量波动发生的周期也无规律可言。这也说明了电池老化过程中容量再生发生的周期和强度并不固定，具有强随机性。而这样的特性使得大多数 IMF 分量并不具有 SOH 周期预测的价值，因为难以取得较为准确的周期预测结果，同时还会浪费大量的预测资源和时间。

为定量分析 IMFs 与原始容量退化趋势之间的相关性，分别计算两者的 PCC 和 SCC。表 5-1 为 NASA 数据集各 CEEMDAN 分量与其原始容量的 PCC 和 SCC，表 5-2 为 CALCE 数据集各 CEEMDAN 分量与其原始容量的 PCC 和 SCC。明显看出 RES 与原始容量 PCC 和 SCC 均大于 0.94，部分相关系数接近 1，这表明 RES 包含了容量主要的整体退化趋势。

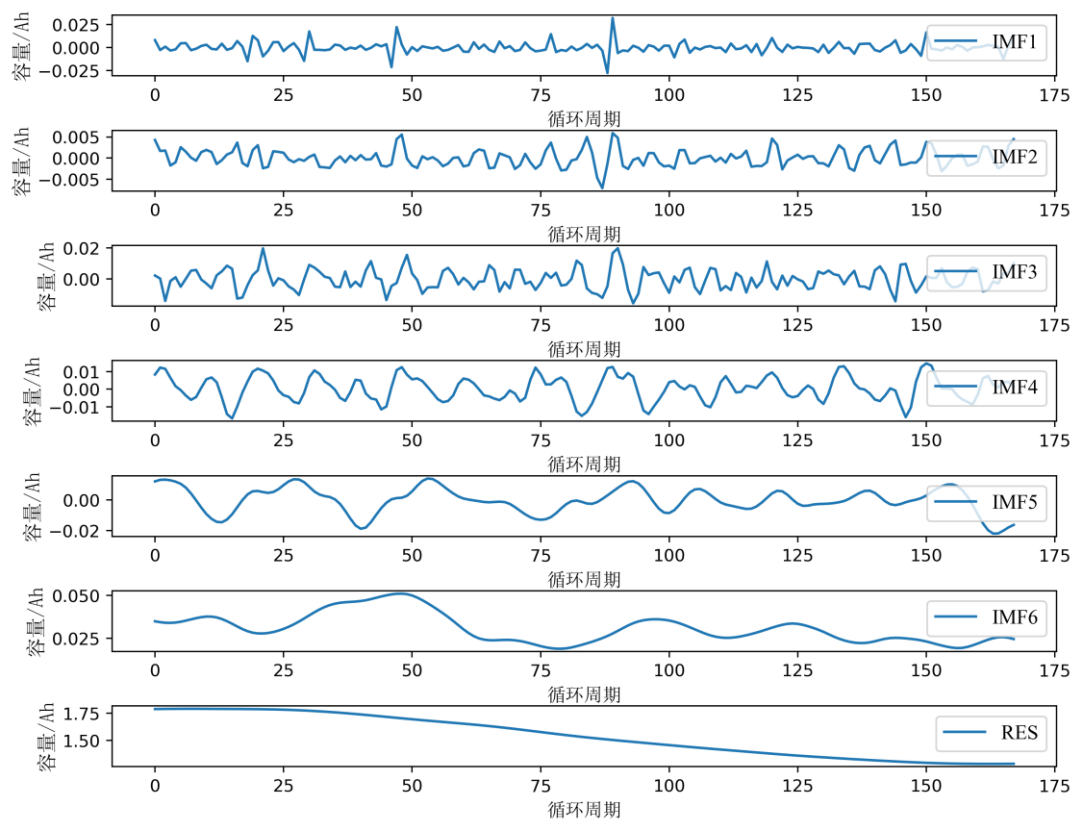


图 5-4 B0005 电池容量 CEEMDAN 分解结果

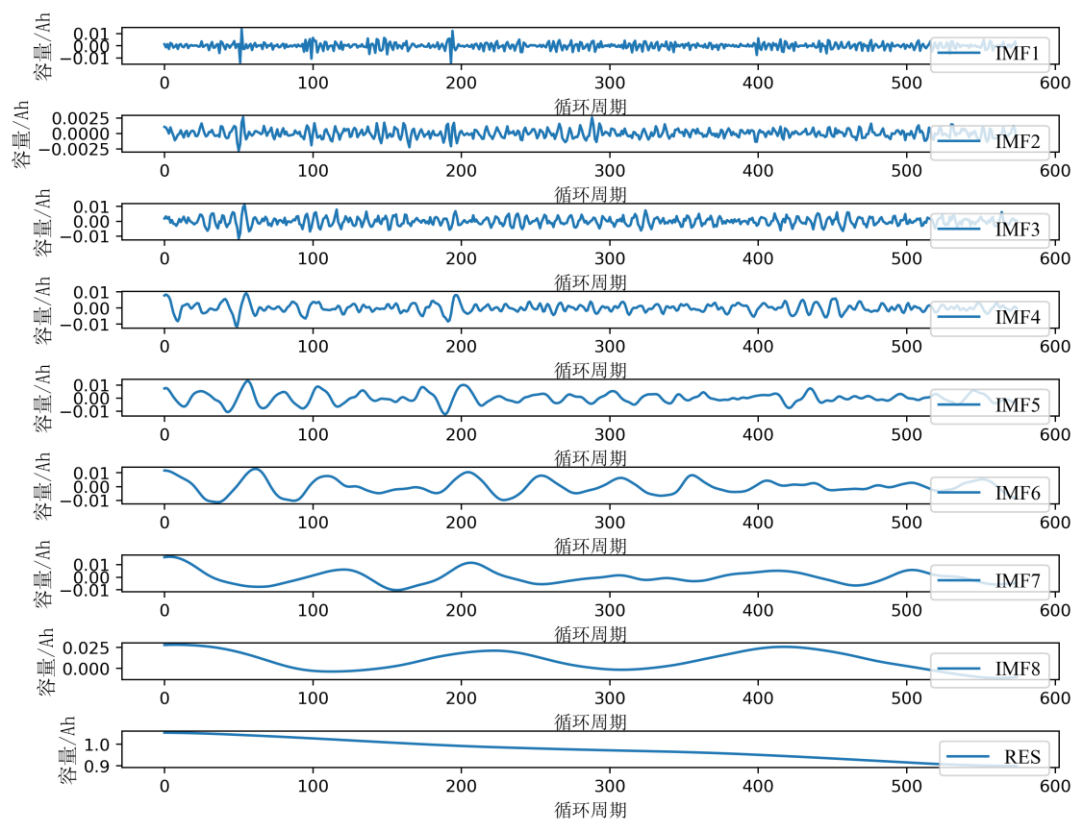


图 5-5 CS235 电池容量 CEEMDAN 分解结果

表 5-1 NASA 数据集各 CEEMDAN 分量与其原始容量的 PCC 和 SCC

电池	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	RES
B0005(PCC)	0.0487	0.0223	0.0330	0.0542	0.1905	0.6172	0.9964
B0005(SCC)	0.0456	0.0455	0.0351	0.0860	0.2238	0.5970	0.9899
B0006(PCC)	0.0568	0.0278	0.0453	0.1580	0.1941	0.2810	0.9909
B0006(SCC)	0.0450	0.0301	0.0273	0.1771	0.1652	0.2238	0.9932
B0007(PCC)	0.0259	0.0073	0.0256	0.0731	0.0467	0.5202	0.9957
B0007(SCC)	0.0223	0.0285	0.0272	0.0677	-0.002	0.5128	0.9920

表 5-2 CALCE 数据集各 CEEMDAN 分量与其原始容量的 PCC 和 SCC

电池	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	RES
CS235(PCC)	0.0701	0.0942	0.0740	0.0981	0.1373	0.1735	0.3091	0.5326	0.9445
CS235(SCC)	0.0740	0.0964	0.0684	0.0831	0.1143	0.1252	0.1496	0.4128	0.9475
CS236(PCC)	0.0524	0.0199	0.0736	0.1035	0.0596	0.1555	0.3784	/	0.9870
CS236(SCC)	0.0302	0.0415	0.0560	0.0523	0.0779	0.0902	0.2174	/	0.9895
CS237(PCC)	0.0607	0.0413	0.0701	0.0106	0.1340	0.0199	0.3040	0.4767	0.9563
CS237(SCC)	0.0290	0.0189	0.0588	0.0795	0.0749	0.0952	0.1661	0.2404	0.9813

以表 5-1 B0005 号电池而言，原始容量与 IMF6 的 PCC 和 SCC 分别为 0.6172 和 0.5970，表示两者之间不仅仅具有较强的线性关系同时而且也具有较强的单调关系。而原始容量与 IMF1-IMF4 之间的 PCC 和 SCC 均低于 0.2，IMF5 虽然 SCC 为 0.2238 但是 PCC 低于 0.2，综合说明原始容量与 IMF1-IMF5 不管是线性关系还是单调关系很弱很小。如图 5-4 所示，IMF1-IMF5 曲线出现了明显的突起，这说明它们捕捉到了 B0005 容量衰减过程中的局部容量再生。但是，由于容量再生具有很强的随机性，这些高频子序列缺乏时间相关性，很难给出准确的预测结果。

由于容量再生波动的随机性较强，大部分 IMFs 分量难以给出准确容量预测的有用信息，因此本文重点关注每次容量再生后的快速衰减趋势。为了提高预测的性能和效率，本文选择 PCC 和 SCC 绝对值均大于 0.2 的相关高频子序列进行 SOH 预测，表明该 IMF 与原始容量之间有一定的线性和单调关系，包含了一定的容量退化趋势。这种选择可以减少容量再生引起的局部突变，降低容量再生问题对 SOH 周期预测的负面影响。以 B0005 和 CS235 电池为例，本论文分别选择只保留 B0005 电池的 IMF6 和 CS235 电池的 IMF8 分量。图 5-6 显示了两个电池舍弃了容量再生波动后的容量退化曲线。

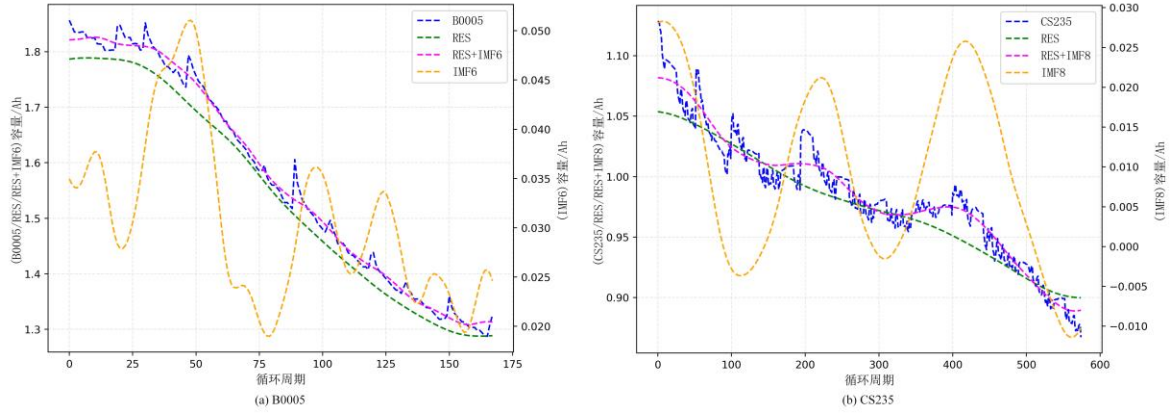


图 5-6 舍弃容量再生波动后锂离子电池容量衰减曲线

在图 5-6 中, Res 曲线有效地反映了电池的整体退化趋势, 但与原始容量退化趋势存在一定的偏差。而 Res+IMF6 曲线和 Res+IMF8 则非常接近于 B0005 和 CS235 电池的原始容量曲线。综上所述, 通过 CEEMDAN 解耦和相关性分析过程可以去除电池老化过程中容量再生特征中不相关的高频子序列, 降低容量再生波动的负面影响, 有利于 SOH 周期预测。

5.4.2 SOH 周期预测结果

大量研究指出, 锂离子电池的 SOH 预测精度会随着循环周期的增加而降低。当一次性预测周期长度较长时, 精度下降更加明显。文献^[63-64]研究指出, 在经过一定次数的充放电循环后, 锂离子电池通常会以更快的速率退化, 快速退化的开始可称之为拐点。以图 2-3 CS235 电池为例, 去掉化学状态不稳定初始阶段, 可以大概的把其循环周期分为三段。第 100-400 循环周期容量下降缓慢; 第 400-620 周期容量加速下降; 最后第 620-717 周期, 容量出现断崖式下降趋势。锂离子退化过程中退化速率的不一致性给锂电池准确的 SOH 周期预测带来挑战性。在本研究中, 结合实际应用需求和精度需要, 本论文利用有限的输入序列进行短周期 SOH 预测, 并利用足够的训练序列进行全局长周期 SOH 预测。同时基于锂电池 SOH 低于 80% 需要回收的特点, 仅取 CALCE 电池数据集 SOH 前 79% 的数据进行 SOH 周期预测实验。因此, 相较于 SOH 点预测, CS235、CS236、CS237 的三块电池的循环寿命分别为 575、516、587 次循环。

针对锂离子电池的 SOH 周期预测, 所提出的框架为基于 Bi-LSTM 网络和 LSTM 网络集成预测的方法。Bi-LSTM 网络预测 RES, 为电池主要的容量衰减趋势。LSTM 网络预测相关高频子序列, 其包含了部分退化趋势, 可以有效弥补 RES

和原始容量之间的偏差。采用尺寸为 20 或 50 的滑动窗口预测 SOH，以测试集上预测的容量值作为滑动窗口输入，一次性预测多步周期 SOH。在 NASA 数据集中，循环周期皆为 168 次。当历史数据小于 50%（84 次循环周期）时，短周期预测长度为 20；当历史数据充足时，进行全局长周期预测。在 CALCE 电池数据集，由于循环周期长且退化速率变化明显。当历史数据小于 400 周期，进行短周期预测，长度为 50 或 100；当历史数据大于 400 周期，进行全局长周期预测。最后，两个模型的 SOH 预测进行集成相加，以获得最终 SOH 预测结果。

在一定范围内通过改变滑动窗口尺寸、LSTM 层数、LSTM 隐藏层维度等，比较 SOH 周期预测实验结果。实验发现，层数的增加没有明显的效果提升，这与特征较为简单有关且训练时间大大加长。最终通过对比寻优，确定合适的网络参数值。表 5-3 为 LSTM 和 Bi-LSTM 网络参数。

表 5-3 LSTM 和 Bi-LSTM 网络参数

模型	层数	轮数	学习率	隐藏层维度	迭代预测长度
LSTM	1	250	0.001	64	1
Bi-LSTM	1	10~50	0.001	32	1

六组锂离子电池 SOH 周期预测结果如图 5-7 至图 5-12 所示。从图中可以看出，短周期预测结果和长周期预测结果都非常符合主体容量下降趋势，证明所提框架的可行性和准确性。在图 5-7 至 5-9 中，对于从第 80 个周期开始的短周期预测结果，由于原始容量序列在第 87 个周期前后出现了显著容量再生波动，这部分波动被分解后舍弃，并没有参与预测，因此与其他结果相比存在相对较大的偏差。不过，如果考虑到这一时期前后的容量整体退化趋势，预测结果还是比较理想的，与主体趋势一致。由于 CALCE 数据集出现容量再生的频率和波动都较高，加上退化速率的不一致性以及循环周期较长，SOH 周期预测具有挑战性。在图 5-10 至 5-12 中，所提出方法的框架在短周期预测和长周期预测都能克服这些问题，呈现有效的预测精度。通过分段 SOH 预测，所提框架可以在历史数据较少的情况下获得可靠的短周期 SOH 预测精度；当历史周期足够长时，所提框架也可获得较为精确的电池 SOH 全周期预测精度，更符合实际情况所需。综上，所提框架在不同数据集和不同预测周期长度都可以取得有效且周期的 SOH 预测结果。

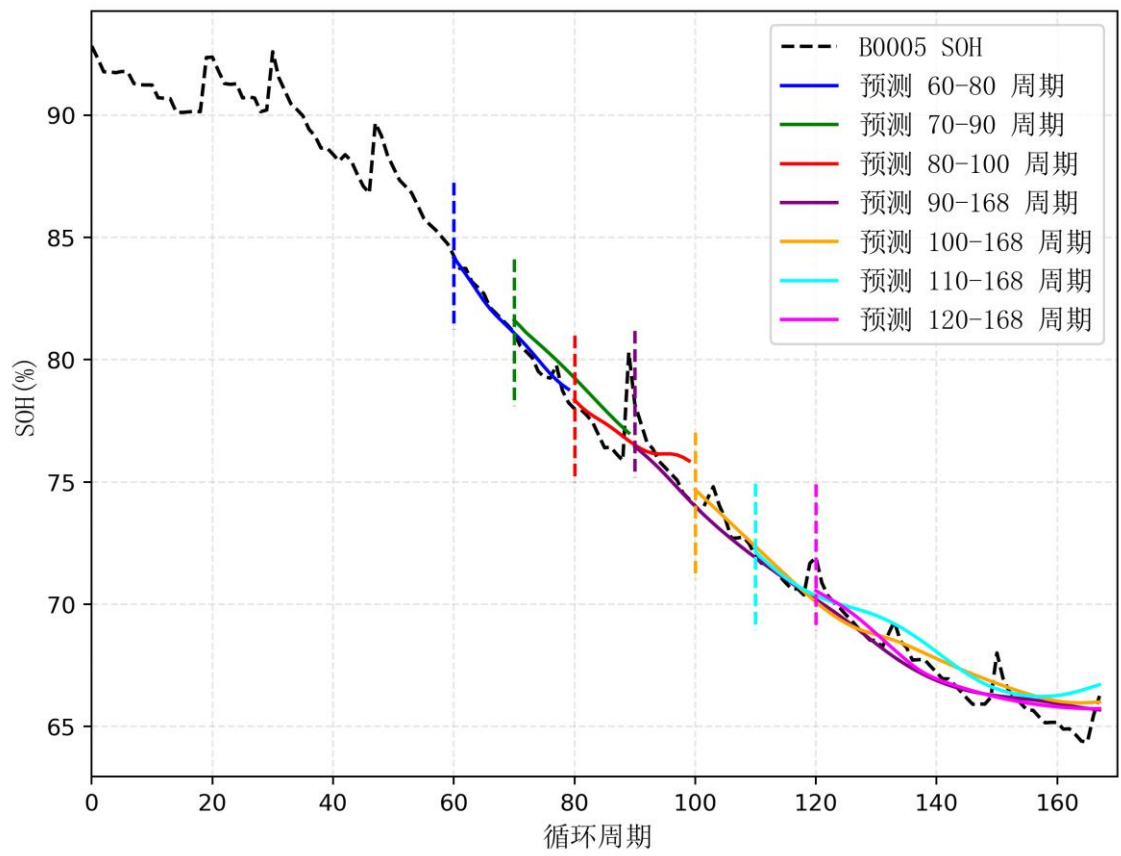


图 5-7 B0005 电池 SOH 周期预测结果

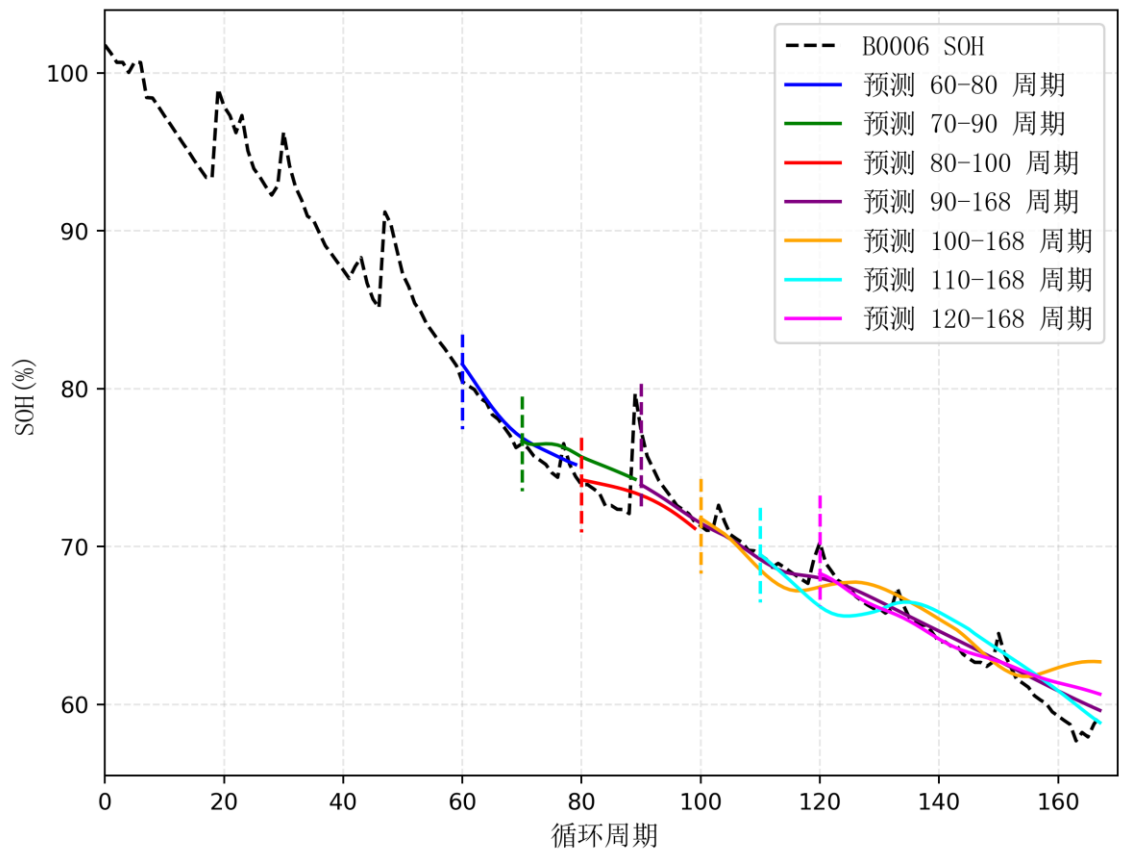


图 5-8 B0006 电池 SOH 周期预测结果

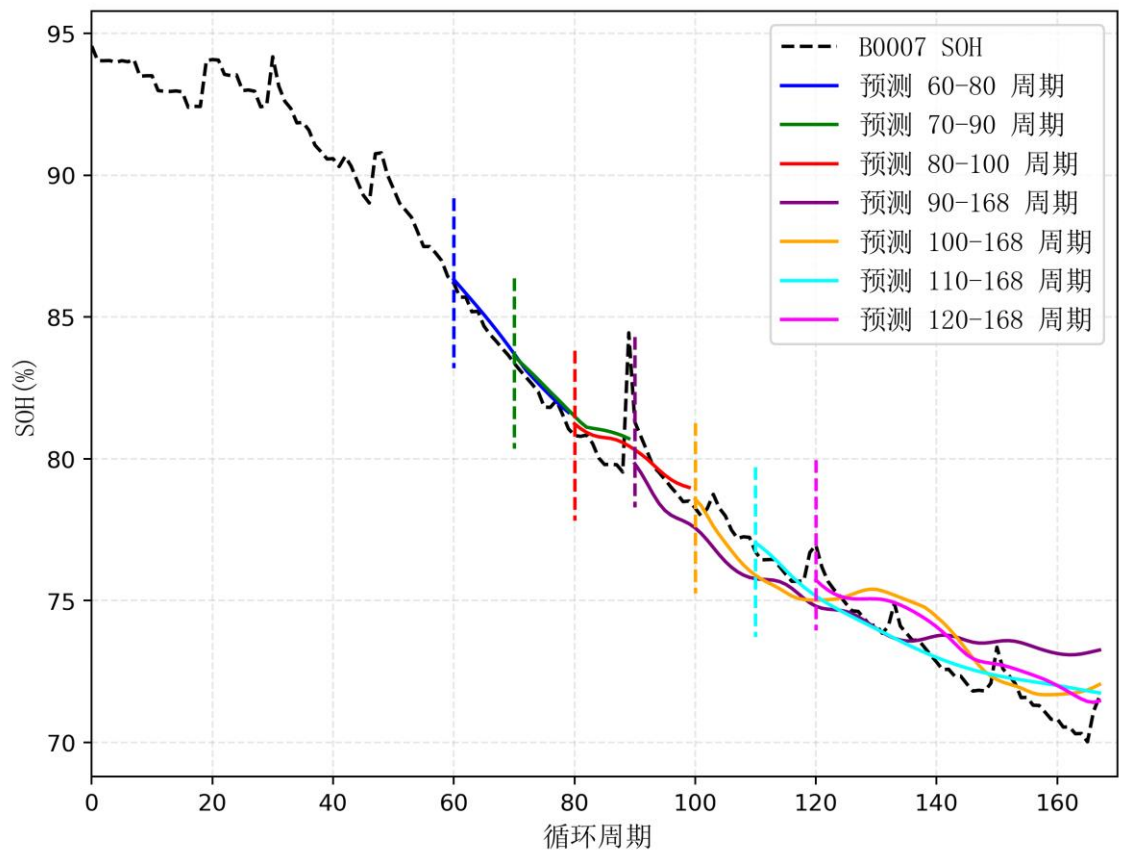


图 5-9 B0007 电池 SOH 周期预测结果

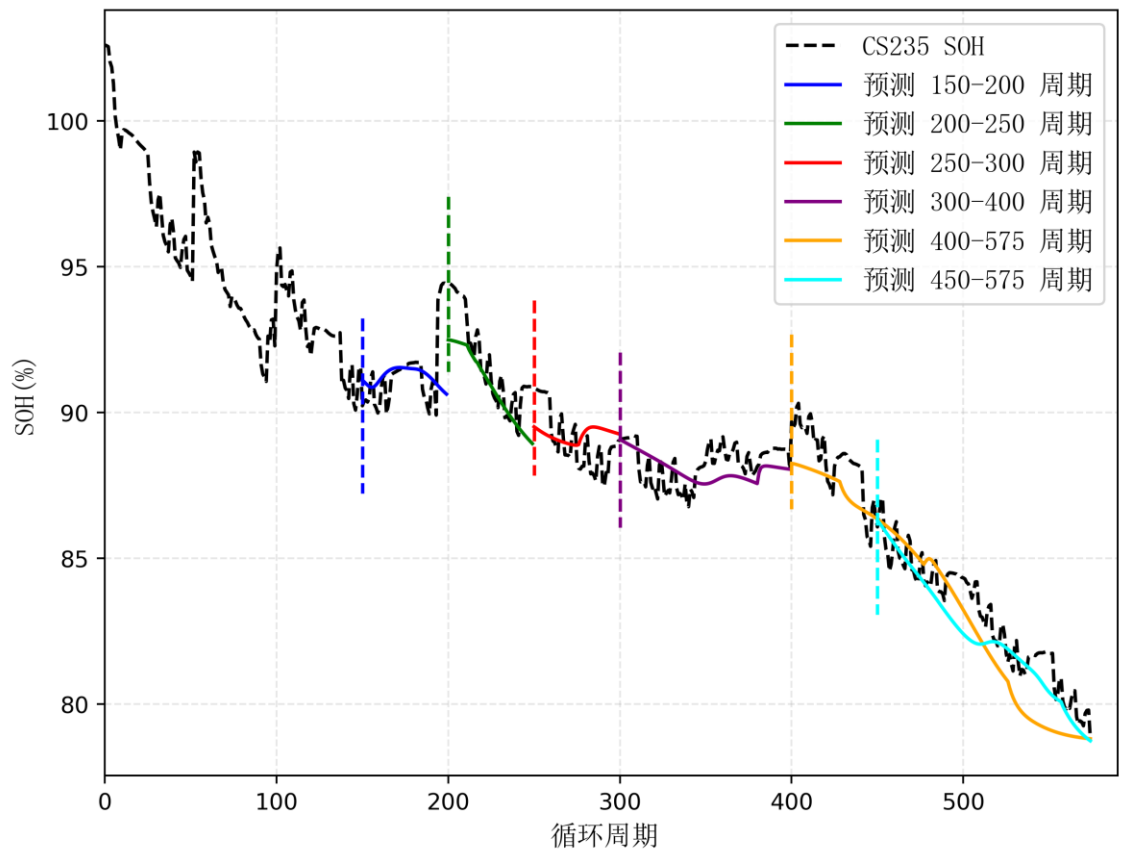


图 5-10 CS235 电池 SOH 周期预测结果

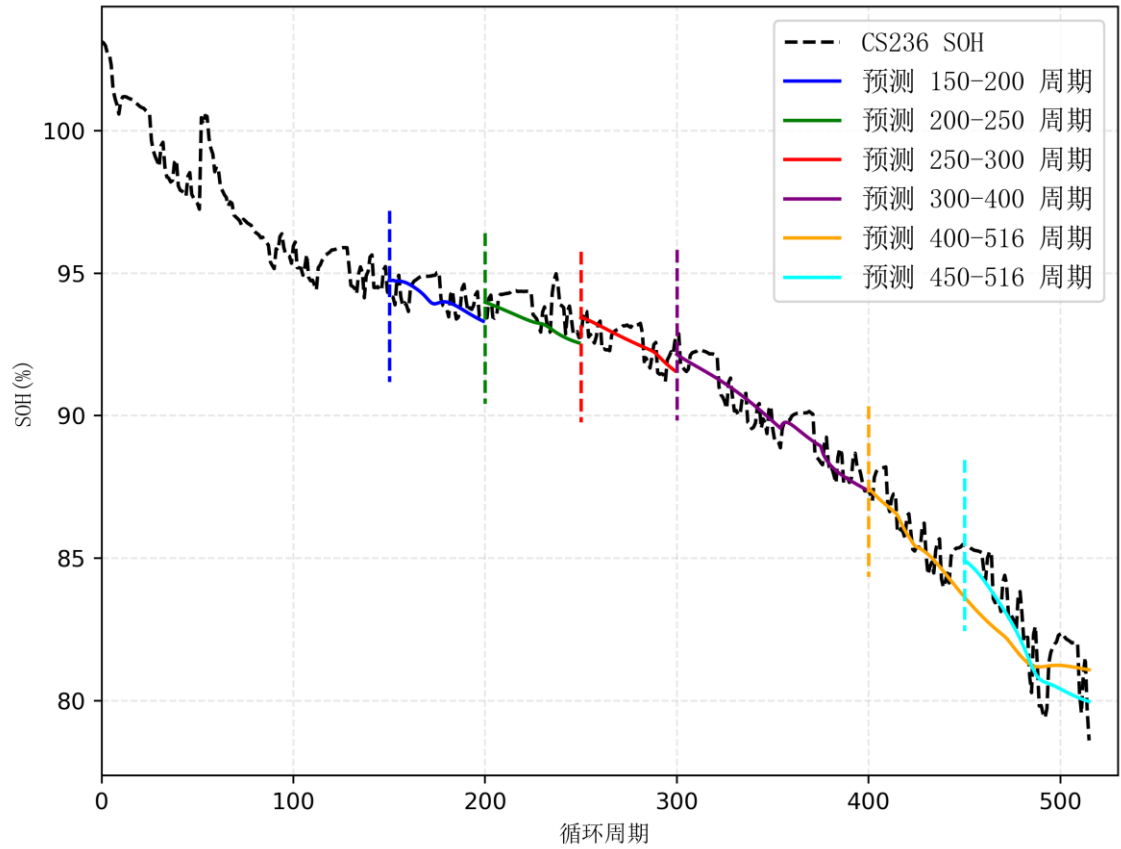


图 5-11 CS236 电池 SOH 周期预测结果

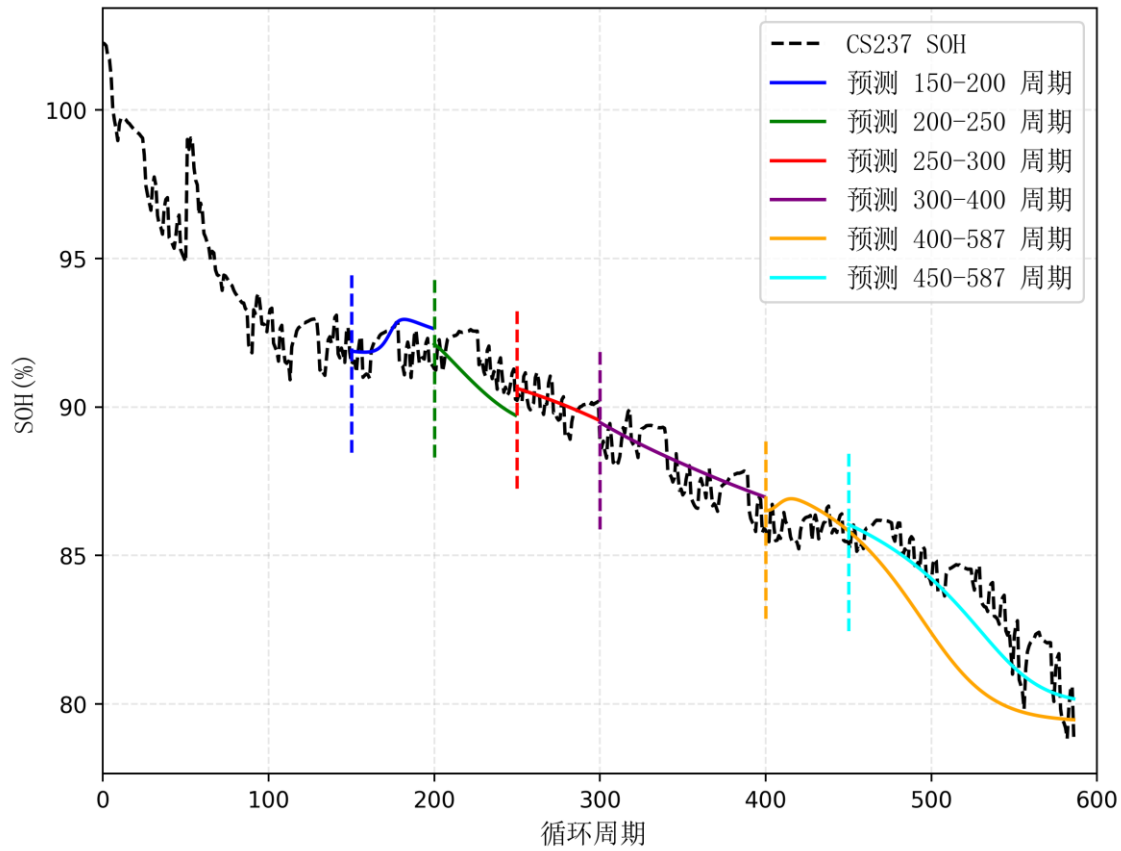


图 5-12 CS237 电池 SOH 周期预测结果

为客观评估锂离子电池的 SOH 周期预测性能，表 5-4 和表 5-5 分别列出两个数据集 MAE 和 RMSE 指标。从表 5-4 和表 5-5 可以看出，MAE 最大值为 1.66%，RMSE 最大值为 2.02%，表明预测性能的优越性。预测框架虽然舍弃一些不相关的高频子序列，但仍能获得可靠的预测结果。本文提出的框架很大程度不受容量再生现象和预测起始点为容量再生点的负面影响，可以提供可靠的 SOH 周期预测结果。

表 5-4 NASA 数据集 SOH 周期预测结果的 MAE 和 RMSE 指标

电池	MAE/RMSE						
	60	70	80	90	100	110	120
B0005	0.0021/0.0027	0.0115/0.0129	0.0089/0.0118	0.0049/0.0067	0.0058/0.0072	0.0074/0.0089	0.0050/0.0063
B0006	0.0065/0.0073	0.0170/0.0203	0.0144/0.0202	0.0071/0.0098	0.0134/0.0175	0.0122/0.0146	0.0086/0.0124
B0007	0.0032/0.0035	0.0077/0.0108	0.0067/0.0106	0.0115/0.0138	0.0089/0.0100	0.0057/0.0075	0.0084/0.0093

表 5-5 CALCE 数据集 SOH 周期预测结果的 MAE 和 RMSE 指标

电池	MAE/RMSE					
	150	200	250	300	400	450
CS235	0.0082/0.0134	0.0091/0.0114	0.0090/0.0105	0.0062/0.0071	0.0108/0.0125	0.0073/0.0092
CS236	0.0049/0.0059	0.0066/0.0080	0.0047/0.0052	0.0048/0.0055	0.0100/0.0119	0.0094/0.0112
CS237	0.0067/0.0084	0.0099/0.0110	0.0044/0.0050	0.0061/0.0070	0.0166/0.0200	0.0088/0.0104

5.4.3 与全 IMFs 周期预测实验结果比较

该实验基于两个关键理念：一是发生容量再生后，再生容量在随后的几个周期内迅速衰减，电池容量回归到总体衰减趋势；二是容量再生需要满足充足休息时间，然而在实际生活中，如电动汽车等，在不同用户之间以及个体用户使用习惯，并无明显规律可言，容量再生发生呈现高度随机性。通过去除不相关的高频子序列(容量再生引起的容量回弹和再生容量快速退化)，退化趋势变得更加平滑，更容易预测。因此，与传统的全模态分解-集成预测相比，本框架采用了更少的 IMFs，不仅提高了 SOH 周期预测框架的效率而且避免了容量再生分解后 IMF 分量高度随机周期预测精度低的问题。

B0005 电池和 CS235 电池 IMFs 如图 5-4 和 5-5 所示，分别为 6 组和 8 组。以 B0005 电池预测 60-80 周期和 90-168 周期，CS235 电池预测 150-200 周期和 400-575 周期为例，比较所提框架和传统的全 IMFs 分量预测的预测性能。图 5-13 和图 5-14 分别为 B0005 电池和 CS235 电池单 IMF 分量与全部 IMFs 分量 SOH 周期预测性能对比。同时训练和预测用时和 MAE 和 RMSE 如表 5-6 所示。

表 5-6 所提框架与传统全 IMFs 分量训练与预测耗时以及 MAE/RMSE 对比

电池/ 预测周期	训练与预测耗时/s				MAE/RMSE	
	平均 IMF	RES	All IMFs+RES	本文	All IMFs+RES	本文
B0005/60-80	24.24	14.48	159.92	38.72	0.0065/0.0084	0.0021/0.0027
B0005/90-168	44.53	18.65	285.83	63.18	0.0095/0.0120	0.0049/0.0067
CS235/150-200	67.26	6.60	544.68	73.86	0.0121/0.0163	0.0082/0.0134
CS235/400-575	194.36	14.53	1569.41	208.89	0.0110/0.0130	0.0108/0.0125

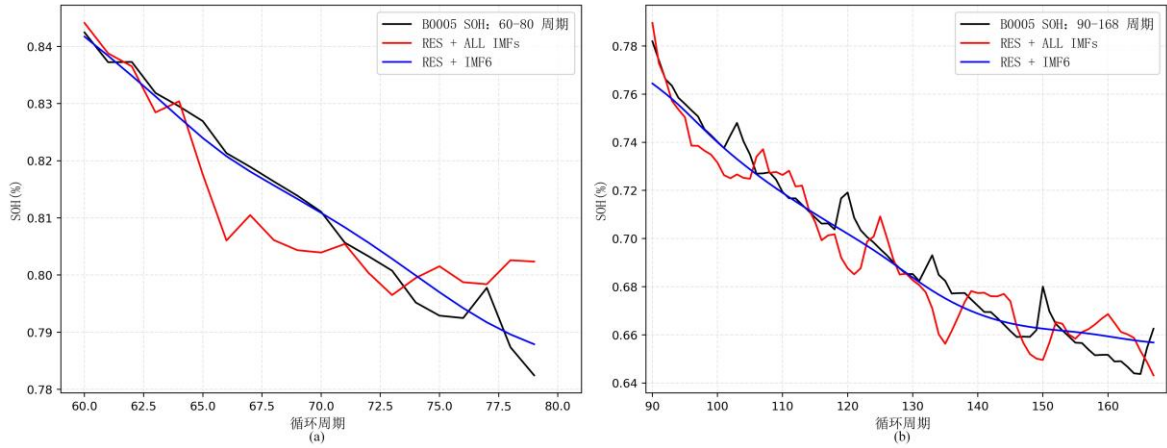


图 5-13 B0005 电池 IMF6 与全 IMFs 分量周期预测性能对比

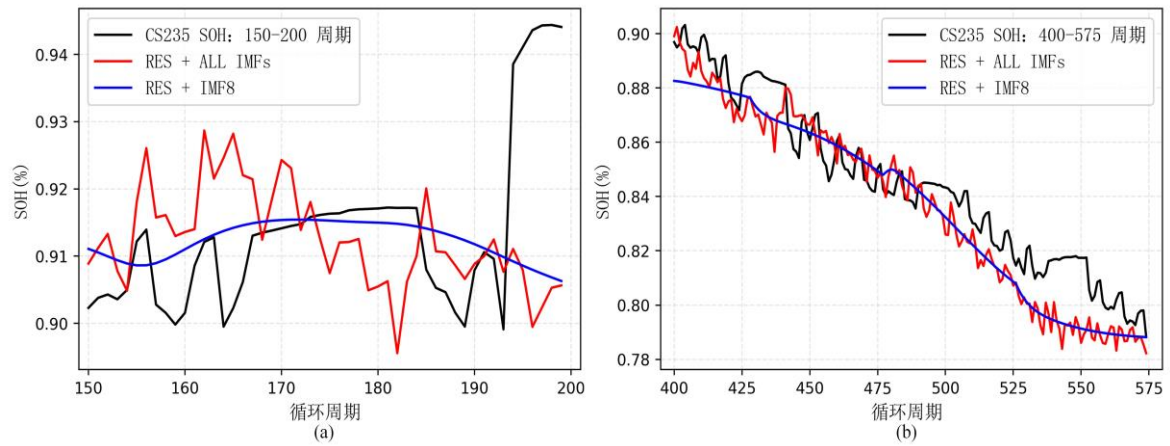


图 5-14 CS235 电池 IMF8 与全 IMFs 分量周期预测性能对比

结合表 5-6，图 5-13 和 5-14 综合来看，本文所提 SOH 预测框架相比于传统的全 IMFs 分量预测相比，不仅仅训练与预测耗时极大幅度缩减，B0005 缩短六分之五左右，CS235 缩短八分之七左右；同时在 MAE 和 RMSE 指标上 SOH 周期预测性能更好。验证了所提框架的合理性和准确性。

5.4.4 所提框架与其它框架 SOH 预测比较

为了更好地验证拟议框架在锂离子电池 SOH 周期预测方面的卓越性能，与其

他近年来发表在 SCI 期刊上的主流方法的进行比较（因为在 CALCE 数据集上进行解耦 SOH 周期预测的论文非常少，且与本文预测起点和预测周期长度不尽相同，故不进行比较展示）。NASA 数据集分别于 VPA^[88], CPR-PF-ARIMA^[89], CEEMDAN-CNN-Bi-GRU^[90], EMD-ARIMA^[91], VMD-GPR-MC^[92]共 5 种方法对比，对比结果如表 5-7 所示。可以直观的看到，与传统 IMFs 分量预测相比，本文所提框架提供有竞争力的预测精度，优于大多数其他模型，验证了所提框架的合理性与准确性。

表 5-7 NASA 数据集 SOH 周期预测不同方法的 RMSE 比较

电池	模型	预测起点	RMSE
B0005	VPA [88]	90/100	0.0104/0.0128
	CPR-PF-ARIMA [89]	90	0.0169
	CEEMDAN-CNN-Bi-GRU [90]	100	0.0104
	EMD-ARIMA [91]	90/100	0.0195/0.0075
	VMD-GPR-MC [92]	100	0.0196
	本文方法	90/100	0.0067/0.0072
B0006	VPA [88]	90/100	0.0169/0.0118
	CPR-PF-ARIMA [89]	90	0.0146
	CEEMDAN-CNN-Bi-GRU [90]	100	0.0115
	EMD-ARIMA [91]	90	0.0285
	VMD-GPR-MC [92]	100	0.0194
	本文方法	90/100	0.0098/0.0175
B0007	VPA [88]	90/100	0.0289/0.0313
	CPR-PF-ARIMA [89]	90	0.0129
	CEEMDAN-CNN-Bi-GRU [90]	100	0.0081
	EMD-ARIMA [91]	90/100	0.0113/0.0094
	VMD-GPR-MC [92]	100	0.0195
	本文方法	90/100	0.0138/0.0100

5.5 本章小结

本章主要介绍 CEEMDAN、LSTM、Bi-LSTM 的算法原理，并提出一种基于 CEEMDAN-LSTM 框架的锂离子电池 SOH 周期预测方法。所提框架基于容量再生波动的高随机性和快速退化性，采用 CEEMDAN 解耦容量曲线，通过相关性分析舍弃多余的包含容量再生波动 IMFs 分量。采用 Bi-LSTM 和 LSTM 分别预测 RES 和相关高频子序列，集成相加获得最终 SOH 周期预测结果。通过与全 IMFs 分量，其他主流模型对比，验证所提框架不仅大大节约时间而且能提供更准确的预测指标。

第 6 章 总结与展望

6.1 总结

针对锂离子电池长循环寿命和容量退化非线性特点,本文以锂离子电池 SOH 准确预测为研究目标,主要研究内容包括两个方面。一方面是从电池循环老化数据中挖掘有效的退化信息并构造 HI,实现准确可靠的 SOH 点预测;另一方面是详细分析了锂电池退化过程中容量再生发生的原因以及特性,实现准确可靠的 SOH 长周期预测。本文具体完成的工作总结如下:

(1) 对锂离子电池的基本结构和工作原理进行介绍,分析讨论了电池容量退化的内外部因素。分别介绍了 NASA 电池数据集和 CALCE 电池数据集,绘制了电池的容量衰减曲线与容量再生波动放大图。从容量再生和再生容量的快速退化两个角度分析了容量再生波动的发生原因。

(2) 绘制锂离子电池循环老化数据中充电电流、电压、IC 曲线以及放电电压曲线,利用遍历法找寻最佳特征区间,挖掘构造六组充分表征电池退化的 HIs,并通过 Pearson 相关系数和 Spearman 相关系数定量分析 HI 的有效性。

(3) 锂离子电池长循环寿命与容量退化的非线性造成 HI 特征分布在不同循环周期具有不一致性,单一数据驱动模型难以长期有效实现 SOH 点预测。提出 1DCNN-Informer 框架集成预测 SOH,采用不同序列预测模式。该框架可以同时预测 1DCNN 的当前时间步和 Informer 模型的下一时间步容量。在加权层中引入超参数 λ 和 RMSE, RMSE 指标分别反映了两个模型的在验证集上预测误差。超参数 λ 平衡了两个模型在长期和短期预测任务中的预测误差。通过加权融合,可有效提高集成框架下一时间步的 SOH 点预测精度和有效预测周期区间长度。

(4) 详细分析了容量再生发生的原因和容量一阶导数序列,指出容量再生是高度随机且不会改变电池容量退化趋势。传统的基于容量再生问题解耦-集成的方法(全 IMFs 预测)不但不符合实际应用且不能实现 SOH 长周期预测。基于再生容量的快速退化特性,利用 CEEMDAN 解耦容量,通过 Pearson 相关系数和 Spearman 相关系数分析 IMFs 与容量相关性,舍弃相关系数绝对值低于 0.2 的 IMFs。分别构造 Bi-LSTM 和 LSTM 模型预测 RES 和 IMF,实现 SOH 长周期准确有效的预测。

6.2 后续工作展望

本论文的出发点旨在解决锂电池 SOH 预测的现实问题出发。就锂电池数据而言，在点预测实验中没有对 HI 进行二次优化，在周期预测实验中消除容量再生波动使框架在实际应用中鲁棒性强，受锂电池容量波动影响较小。两组实验在与近年来多篇 SCI 实验对比中证明了本论文提出的锂离子电池 SOH 预测框架取得了准确的预测效果，但受限于各种因素，也存在不足之处，需要进一步研究：

(1) 使用的数据集是在实验室较为稳定的工况下获得的，而实际工况的环境较为复杂，数据连贯性差，锂电池容量波动较大。所提框架后续需要收集在实际工况运行的锂离子电池循环充放电数据，进一步验证所提框架有效性和准确性。

(2) 本论文各数据驱动模型根据人为经验调参，但 Informer 的模型参数量较大，后续可以加入优化算法自动寻优调参，进一步提高 SOH 预测准确性。

(3) SOH 周期预测只是依靠根据历史数据长度简单的划分短周期和长周期预测起点，没有准确判断电池降解速度加速的拐点。后续可以通过算法寻求这些潜在的拐点，灵活调节 SOH 周期预测的起点和预测长度。

参考文献

- [1] 姚悦, 吕昊东, 彭雪婷, 等. 全球碳管理发展态势与技术前沿进展[J]. 洁净煤技术, 2024, 30(10): 32-40.
- [2] 张涵. 基于多特征健康因子的锂离子电池 SOH 预测研究[D]. 中北大学, 2023.
- [3] 孟君丽. 基于多特征挖掘的锂离子动力电池健康状态估计方法研究[D]. 重庆理工大学, 2024.
- [4] 龚正良, 温芳梅, 汤士军, 等. 锂离子电池日历老化机制的研究进展[J/OL]. 厦门大学学报(自然科学版), 1-16[2025-01-06].
- [5] 郭斌, 刘新华, 何榕, 等. 锂离子电池性能衰减与热失控机制研究进展[J]. 稀有金属, 2024, 48(02): 225-239.
- [6] Lui Y H, Li M, Downey A, et al. Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction[J]. Journal of Power Sources, 2021, 485: 229327.
- [7] 李金东, 古月圆, 王路阳, 等. 退役锂离子电池健康状态评估方法综述[J]. 储能科学与技术, 2019, 8(05): 807-812.
- [8] 熊庆, 邸振国, 汲胜昌. 锂离子电池健康状态估计及寿命预测研究进展综述[J]. 高电压技术, 2024, 50(03): 1182-1195.
- [9] Zubi G, Dufo-López R, Carvalho M, et al. The lithium-ion battery: State of the art and future perspectives[J]. Renewable and sustainable energy reviews, 2018, 89: 292-308.
- [10] 胡晓亚, 郭永芳, 张若可. 锂离子电池健康状态估计方法研究综述[J]. 电源学报, 2022, 20(01): 126-133.
- [11] 赵轩, 李美莹, 余强, 等. 电动汽车动力锂电池状态估计综述[J]. 中国公路学报, 2023, 36(06): 254-283.
- [12] Hosseininasab S, Lin C, Pischinger S, et al. State-of-health estimation of lithium-ion batteries for electrified vehicles using a reduced-order electrochemical model[J]. Journal of Energy Storage, 2022, 52: 104684.
- [13] Gao Y, Liu K, Zhu C, et al. Co-estimation of state-of-charge and state-of-health for

- lithium-ion batteries using an enhanced electrochemical model[J]. IEEE Transactions on Industrial Electronics, 2021, 69(3): 2684-2696.
- [14] Zhang W, Ahmed R, Habibi S. State-of-health estimation for lithium-ion batteries based on electrochemical impedance spectroscopy measurements combined with unscented Kalman filter[J]. Journal of Power Sources, 2025, 625: 235450.
- [15] 郭向伟, 王晨, 陈岗, 等. 基于 RLS 的锂电池全工况自适应等效电路模型[J]. 储能科学与技术, 2023, 12(10): 3230-3241.
- [16] Chen L, Wang S, Chen L, et al. An improved grey wolf optimization–double adaptive extended Kalman filtering algorithm for co-estimation of state of charge and state of health for lithium-ion batteries based on temperature-dependent second-order RC model[J]. Ionics, 2024: 1-16.
- [17] 李谦, 姜帆, 韩乔妮, 等. 基于一阶 ECM-IGPR 的锂离子电池 SOC 及 SOH 联合估计框架[J]. 太阳能学报, 2024, 45(05): 240-250.
- [18] Li C, Yang L, Li Q, et al. SOH estimation method for lithium-ion batteries based on an improved equivalent circuit model via electrochemical impedance spectroscopy[J]. Journal of Energy Storage, 2024, 86: 111167.
- [19] 吕海潮, 康丽霞, 刘永忠. 锂电池的两阶段老化过程经验模型与分析[J]. 高校化学工程学报, 2024, 38(01): 113-120.
- [20] Liu K, Hu X, Wei Z, et al. Modified Gaussian process regression models for cyclic capacity prediction of lithium-ion batteries[J]. IEEE Transactions on Transportation Electrification, 2019, 5(4): 1225-1236.
- [21] He W, Williard N, Osterman M, et al. Prognostics of lithium-ion batteries based on Dempster–Shafer theory and the Bayesian Monte Carlo method[J]. Journal of Power Sources, 2011, 196(23): 10314-10321.
- [22] 金帅, 董静. 基于数据驱动的锂离子电池健康状态估计研究进展综述[J]. 仪器仪表学报, 2024, 45(03): 45-59.
- [23] 何清, 李宁, 罗文娟, 等. 大数据下的机器学习算法综述[J]. 模式识别与人工智能, 2014, 27(04): 327-336.
- [24] Luo W, Syed A U, Nicholls J R, et al. An SVM-based health classifier for offline Li-

- ion batteries by using EIS technology[J]. Journal of The Electrochemical Society, 2023, 170(3): 030532.
- [25] Duan W, Song S, Xiao F, et al. Battery SOH estimation and RUL prediction framework based on variable forgetting factor online sequential extreme learning machine and particle filter[J]. Journal of Energy Storage, 2023, 65: 107322.
- [26] Li F, Min Y, Zhang Y, et al. State-of-health estimation method for fast-charging lithium-ion batteries based on stacking ensemble sparse Gaussian process regression[J]. Reliability Engineering & System Safety, 2024, 242: 109787.
- [27] Rzepka F, Hematty P, Schmitz M, et al. Neural network architecture for determining the aging of stationary storage systems in smart grids[J]. Energies, 2023, 16(17): 6103.
- [28] 吴青峰, 杨艺涛, 刘立群, 等. 基于 GA-SA-BP 神经网络的锂电池健康状态估算方法[J]. 电力系统保护与控制, 2024, 52(19): 74-84.
- [29] Wang H, Li Y F, Zhang Y. Bioinspired spiking spatiotemporal attention framework for lithium-ion batteries state-of-health estimation[J]. Renewable and Sustainable Energy Reviews, 2023, 188: 113728.
- [30] 任利强, 贾舒宜, 王海鹏, 等. 基于深度学习的时间序列分类研究综述[J]. 电子与信息学报, 2024, 46(08): 3094-3116.
- [31] Lin M, Ke L, Wang W, et al. Health prognosis via feature optimization and convolutional neural network for lithium-ion batteries[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108666.
- [32] Gomez W, Wang F K, Chou J H. Li-ion battery capacity prediction using improved temporal fusion transformer model[J]. Energy, 2024, 296: 131114.
- [33] Rao K D, Anand N V, Pandraju T K S, et al. Optimally tuned gated recurrent unit neural network-based state of health estimation scheme for lithium ion batteries[J]. IEEE Access, 2024.
- [34] Fauzi M R, Yudistira N, Mahmudy W F. State-of-Health prediction of lithium-ion batteries using exponential smoothing Transformer with seasonal and growth embedding[J]. IEEE Access, 2024.
- [35] Qin D, Li L, Wu H, et al. A composite progressive approach alternating between

- p>electrochemical model and data-driven model for the state of health assessment of lithium-ion batteries[J]. Journal of Energy Storage, 2024, 97: 112921.
- [36] Sun G, Liu Y, Liu X. A Neural network approach for health state estimation of lithium-ion batteries incorporating physics knowledge[J]. Electronic Materials Letters, 2024: 1-15.
- [37] Zhou D, Liang J, Li F, et al. SOH prediction of lithium-ion batteries using a hybrid model approach integrating single particle model and neural networks[J]. Journal of Energy Storage, 2024, 104: 114579.
- [38] Qian C, Guan H, Xu B, et al. A CNN-SAM-LSTM hybrid neural network for multi-state estimation of lithium-ion batteries under dynamical operating conditions[J]. Energy, 2024, 294: 130764.
- [39] Xiang Y, Fan W, Zhu J, et al. Semi-supervised deep learning for lithium-ion battery state-of-health estimation using dynamic discharge profiles[J]. Cell Reports Physical Science, 2024, 5(1).
- [40] Mu G, Wei Q, Xu Y, et al. Capacity estimation for lithium-ion batteries based on heterogeneous stacking model with feature fusion[J]. Energy, 2024, 313: 133881.
- [41] Li Z, Zhang X, Gao W. State of health estimation of lithium-ion battery during fast charging process based on BiLSTM-Transformer[J]. Energy, 2024, 311: 133418.
- [42] Tang T, Yuan H. A hybrid approach based on decomposition algorithm and neural network for remaining useful life prediction of lithium-ion battery[J]. Reliability Engineering & System Safety, 2022, 217: 108082.
- [43] 赵靖英, 黄麟然, 姚帅亮. 考虑容量再生的锂离子电池健康状态估计方法[J/OL]. 电源学报, 1-14[2025-01-06].
- [44] Zhou Y, Huang M. Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model[J]. Microelectronics Reliability, 2016, 65: 265-273.
- [45] Wang X, Dai K, Hu M, et al. Lithium-ion battery health state and remaining useful life prediction based on hybrid model MFE-GRU-TCA[J]. Journal of Energy Storage, 2024, 95: 112442.

- [46] Yuan Z, Tian T, Hao F, et al. A hybrid neural network based on variational mode decomposition denoising for predicting state-of-health of lithium-ion batteries[J]. *Journal of Power Sources*, 2024, 609: 234697.
- [47] 刘征宇, 张政, 郭乐凯, 等. 基于经验模态分解的锂离子电池健康状态预测[J]. *机械工程学报*, 2024, 60(14): 272-281.
- [48] Ding G, Wang W, Zhu T. Remaining useful life prediction for lithium-ion batteries based on CS-VMD and GRU[J]. *IEEE Access*, 2022, 10: 89402-89413.
- [49] Pan H, Chen C, Gu M. A state of health estimation method for lithium-ion batteries based on improved particle filter considering capacity regeneration[J]. *Energies*, 2021, 14(16): 5000.
- [50] Wu C, Fu J, Huang X, et al. Lithium-ion battery health state prediction based on VMD and DBO-SVR[J]. *Energies*, 2023, 16(10): 3993.
- [51] Cui Y, Chen Y. Prognostics of lithium-ion batteries based on capacity regeneration analysis and long short-term memory network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-13.
- [52] 何文轩, 耿磊, 姚芳. 电动汽车动力锂离子电池可靠性关键技术综述[J]. *电源学报*, 2024, 22(02): 183-196.
- [53] 韩啸, 张成锟, 吴华龙, 等. 锂离子电池的工作原理与关键材料[J]. *金属功能材料*, 2021, 28(02): 37-58.
- [54] Rezvanizani S M, Liu Z, Chen Y, et al. Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility[J]. *Journal of power sources*, 2014, 256: 110-124.
- [55] 闫啸宇, 周思达, 卢宇, 等. 锂离子电池容量衰退机理与影响因素[J]. *北京航空航天大学学报*, 2023, 49(06): 1402-1413.
- [56] 卢威, 张建生, 吴晓东, 等. 高电压下正极表面钝化膜生长过程的原位 AFM 研究[C]. *中国化学会第 29 届学术年会摘要集*, 2014: 2.
- [57] Hu C, Cen Z, Quan Y, et al. SiOC/CNTs composites as anodes for lithium-ion batteries[J]. *Chemical Engineering Journal*, 2024: 152610.
- [58] Yang N, Zhang X, Shang B B, et al. Unbalanced discharging and aging due to

- temperature differences among the cells in a lithium-ion battery pack with parallel combination[J]. *Journal of Power Sources*, 2016, 306: 733-741.
- [59] Dubarry M, Truchot C, Cugnet M, et al. Evaluation of commercial lithium-ion cells based on composite positive electrode for plug-in hybrid electric vehicle applications. Part I: Initial characterizations[J]. *Journal of power sources*, 2011, 196(23): 10328-10335.
- [60] Ouyang D, Weng J, Chen M, et al. Investigation on topographic, electrochemical and thermal features of aging lithium-ion cells induced by overcharge/over-discharge cycling[J]. *Journal of Energy Storage*, 2023, 68: 107799.
- [61] NASA. "Li-ion Battery Aging Datasets." 2024. Available: <https://data.nasa.gov/dataset/Li-ion-Battery-Aging-Datasets/uj5r-zjdb>. Accessed: July 21, 2024.
- [62] Center for Advanced Life Cycle Engineering (CALCE), University of Maryland. "CALCE Battery Data." 2024. Available: <https://calce.umd.edu/battery-data#CS2>. Accessed: July 21, 2024.
- [63] Diao W, Saxena S, Han B, et al. Algorithm to determine the knee point on capacity fade curves of lithium-ion cells[J]. *Energies*, 2019, 12(15): 2910.
- [64] Greenbank S, Howey D. Automated feature extraction and selection for data-driven models of rapid battery capacity fade and end of life[J]. *IEEE Transactions on Industrial Informatics*, 2021, 18(5): 2965-2973.
- [65] Park K, Choi Y, Choi W J, et al. LSTM-based battery remaining useful life prediction with multi-channel charging profiles[J]. *IEEE Access*, 2020, 8: 20786-20798.
- [66] Xu X, Yu C, Tang S, et al. State-of-health estimation for lithium-ion batteries based on Wiener process with modeling the relaxation effect[J]. *IEEE Access*, 2019, 7: 105186-105201.
- [67] Xie Y, Wang S, Zhang G, et al. A review of data-driven whole-life state of health prediction for lithium-ion batteries: data preprocessing, aging characteristics, algorithms, and future challenges[J]. *Journal of Energy Chemistry*, 2024.
- [68] Xia Z, Qahouq J A A. Lithium-ion battery ageing behavior pattern characterization and state-of-health estimation using data-driven method[J]. *IEEE Access*, 2021, 9:

98287-98304.

- [69] He J, Wei Z, Bian X, et al. State-of-health estimation of lithium-ion batteries using incremental capacity analysis based on voltage–capacity model[J]. IEEE Transactions on Transportation Electrification, 2020, 6(2): 417-426.
- [70] Xu Z, Chen Z, 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.
- [71] Chen H, Yue W, Bin G, et al. Filter methods comparison for incremental capacity analysis in lithium-ion batteries health prediction[J]. Journal of Energy Storage, 2024, 101: 113878.
- [72] Pearson K. VII. Note on regression and inheritance in the case of two parents[J]. proceedings of the royal society of London, 1895, 58(347-352): 240-242.
- [73] Spearman C. The proof and measurement of association between two things[J]. 1961.
- [74] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [75] 王逸凡, 井世亮, 夏宇, 等. 基于 1D-CNN 的稻谷石墨烯远红外干燥模型及含水率在线预测[J/OL]. 南京农业大学学报, 1-13[2025-01-06].
- [76] Vaswani A. Attention is all you need[J]. Advances in neural information processing systems, 2017.
- [77] Zhou H, Zhang S, Peng J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C]. Proceedings of the AAAI conference on artificial intelligence. 2021, 35(12): 11106-11115.
- [78] Ma B, Yu H Q, Wang W T, et al. State of health and remaining useful life prediction for lithium-ion batteries based on differential thermal voltammetry and a long and short memory neural network[J]. Rare Metals, 2023, 42(3): 885-901.
- [79] Yang P, Yang H D, Meng X B, et al. Joint evaluation and prediction of SOH and RUL for lithium batteries based on a GBLS booster multi-task model[J]. Journal of Energy Storage, 2024, 75: 109741.
- [80] Cai Y, Li W, Zahid T, et al. Early prediction of remaining useful life for lithium-ion

- p>
batteries based on CEEMDAN-transformer-DNN hybrid model[J]. Heliyon, 2023, 9(7).
- [81] Zheng D, Man S, Guo X, et al. Joint prediction of state of health and remaining useful life for lithium-ion batteries based on health features optimization and multi-model fusion[J]. Ionics, 2024, 30(10): 6239-6252.
- [82] Zhou Y, Wang S, Xie Y, et al. Remaining useful life prediction and state of health diagnosis of lithium-ion batteries with multiscale health features based on optimized CatBoost algorithm[J]. Energy, 2024, 300: 131575.
- [83] Pang H, Chen K, Geng Y, et al. Accurate capacity and remaining useful life prediction of lithium-ion batteries based on improved particle swarm optimization and particle filter[J]. Energy, 2024, 293: 130555.
- [84] Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in adaptive data analysis, 2009, 1(01): 1-41.
- [85] Torres M E, Colominas M A, Schlotthauer G, et al. A complete ensemble empirical mode decomposition with adaptive noise[C]//2011 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2011: 4144-4147.
- [86] Hochreiter S. Long Short-term Memory[J]. Neural Computation MIT-Press, 1997.
- [87] Tian Y, Wen J, Yang Y, et al. State-of-health prediction of lithium-ion batteries based on cnn-bilstm-am[J]. Batteries, 2022, 8(10): 155.
- [88] Lyu G, Zhang H, Zhang Y J, et al. An interpretable remaining useful life prediction scheme of lithium-ion battery considering capacity regeneration[J]. Microelectronics Reliability, 2022, 138: 114625.
- [89] He N, Yang Z, Qian C, et al. Remaining useful life prediction of lithium-ion battery based on fusion model considering capacity regeneration phenomenon[J]. Journal of Energy Storage, 2024, 85: 111068.
- [90] Lv K, Ma Z, Bao C, et al. Indirect prediction of lithium-ion battery RUL based on CEEMDAN and CNN-BiGRU[J]. Energies, 2024, 17(7): 1704.
- [91] Zhou Y, Huang M. Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model[J]. Microelectronics

Reliability, 2016, 65: 265-273.

- [92] Wei M, Ye M, Zhang C, et al. A multi-scale learning approach for remaining useful life prediction of lithium-ion batteries based on variational mode decomposition and Monte Carlo sampling[J]. Energy, 2023, 283: 129086.

攻读学位期间发表的学术论文目录

- [1] Guo L, He H, Ren Y, et al. Prognostics of lithium-ion batteries health state based on adaptive mode decomposition and long short-term memory neural network[J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107317. (SCI 中科院一区)

致谢

本科 4 年，研究生 3 年。自 2018 年来至今，已在安徽工程大学接近 7 年时间。如今学校生活也快到了尾声，即将步入社会，内心还是感慨良多。回望三年研究生学习生活，收获还是良多。不仅仅是学历，更珍贵的在认知，自主学习能力以及心态的变化。悟已往之不谏，知来者之可追。希望自己可以完美结束学生生涯的同时，在往后的人生路上有别样的精彩。

首先，感谢郭黎老师的帮助与指导。不仅仅在学术上，也在生活上给予了不小的帮助。老师在论文的指导上给予大力帮助，帮助我发表第一篇中科院一区 SCI 论文，从 0 到 1，并不容易。小有收获，如稚童第一次吃冰糖葫芦，喜悦溢于言表。祝愿郭老师工作顺利，步步高升。

其次，感谢三年结识的朋友们。好兄弟魏雄飞，以及任思源，一起游历九华山，上海外滩等经历令人难忘。室友杨贤昊，黎锦涛，在学习，日常生活中互帮互助。同门任以然，张浩，王国栋在学习上的探讨与帮助。当然还有很多萍水相逢的朋友们，祝愿他们平安喜顺。

最后，感谢老爸，老妈以及姐姐的无私帮助与支持理解。家人，家庭永远是令人温暖的存在，祝愿他们身体健康，生活自在。