



多健康指标的锂离子电池健康状态无分布区间估计

郭子瑶¹, 张晓桐¹, 庞晓琼¹, 王竹晴²

(¹中北大学计算机科学与技术学院, 山西 太原 030051; ²太原理工大学计算机科学与技术学院, 山西 太原 030024)

摘要: 针对现有锂离子电池健康状态(state of health, SOH)点估计方法难以量化估计不确定性的问题, 本工作聚焦于更具实际应用价值的SOH区间估计方法。现有大多数区间估计方法是基于分布假设的, 当现实数据偏离这些假设时, 会引入估计偏差, 进而影响估计的可靠性。为此, 无分布的上下界估计(lower upper bound estimate, LUBE)方法逐渐受到关注, 但其仍面临若干挑战: 其一, 损失函数不可微, 导致模型优化困难; 其二, 部分研究采用Sigmoid函数将不可微的损失函数转变为可微损失函数, 但Sigmoid函数的引入往往需人工调整斜率参数; 其三, 现有的研究多建立在容量这一理想健康指标上, 然而容量的精确测量成本高昂, 限制了该方法在实际场景中的适用性。为此, 本工作提出一种面向多健康指标的锂离子电池SOH无分布区间估计方法。首先, 采用核主成分分析(kernel principal component analysis, KPCA)方法对提取的多健康指标进行非线性降维。在此基础上, 构建一个双输出神经网络模型, 通过引入一种无需手动调整斜率参数的损失函数, 使模型能够基于降维后的数据, 稳定地输出高质量的预测区间。在公开可用的CALCE数据集上的实验结果表明, 本工作所提方法不仅满足预定义置信水平的要求, 还进一步提升了预测区间的质量。

关键词: 锂离子电池; 健康状态; 无分布区间估计; 上下界估计; 核主成分分析

doi: 10.19799/j.cnki.2095-4239.2025.1031

中图分类号: TM 912

文献标志码: A

文章编号: 2095-4239 (2026) 04-1375-12

Automated distribution-free interval estimation for lithium-ion battery state of health using multiple health indicators

GUO Ziyao¹, ZHANG Xiaotong¹, PANG Xiaoqiong¹, WANG Zhuqing²

(¹College of Computer Science and Technology, North University of China, Taiyuan 030051, Shanxi, China;

²College of Computer Science and Technology, Taiyuan University of Technology, Taiyuan 030024, Shanxi, China)

Abstract: To address the limitations of conventional state of health (SOH) point estimation methods for lithium-ion batteries, this study develops a more practical SOH interval estimation method. Most previous interval estimation methods rely on distribution assumptions. However, when real-world battery data deviate from these assumptions, estimation biases may be introduced, consequently reducing estimation reliability. Consequently, the distribution-free lower upper bound estimation (LUBE) method has gradually attracted attention, although it still faces critical challenges. First, the loss function is non-differentiable, complicating model optimization. Second, several studies employed Sigmoid functions to transform non-

收稿日期: 2025-11-18; 修改稿日期: 2025-12-22。

基金项目: 山西省研究生教育创新计划项目(2024AL20)。

第一作者: 郭子瑶(1999—), 女, 硕士研究生, 研究方向为复杂系统的故障预测与健康管, E-mail: 784465478@qq.com; 通信作者: 庞晓琼, 副教授, 研究方向为复杂系统的故障预测与健康管, E-mail: xqpang@nuc.edu.cn。

引用本文: 郭子瑶, 张晓桐, 庞晓琼, 等. 多健康指标的锂离子电池健康状态无分布区间估计[J]. 储能科学与技术, 2026, 15(4): 1375-1386.

Citation: GUO Ziyao, ZHANG Xiaotong, PANG Xiaoqiong, et al. Automated distribution-free interval estimation for lithium-ion battery state of health using multiple health indicators[J]. Energy Storage Science and Technology, 2026, 15(4): 1375-1386.

differentiable loss functions into differentiable loss functions; however, this approach often requires manual slope-parameter tuning. Third, previous studies mostly rely on capacity as an ideal health indicator; however, accurately measuring capacity is costly, and this limits the real-world applicability of this method. To address these shortcomings, this study proposes a distribution-free SOH interval estimation method for quantifying LIB SOH using multiple health indicators. First, the kernel principal component analysis (KPCA) method is applied to reduce the dimensionality of the extracted health indicators. Based on this, a dual-output neural network model is constructed; this model introduces a loss function that eliminates the need for manual slope-parameter tuning, enabling it to stably output high-quality prediction intervals based on the reduced-dimensional data. Experimental results using the publicly available CALCE dataset demonstrate that the proposed method consistently meets nominal-confidence-level requirements while significantly improving prediction-interval quality.

Keywords: lithium-ion batteries; state of health; distribution-free interval estimation; lower upper bound estimate; kernel principal component analysis

锂离子电池因具有高能量密度、高输出电压、低自放电率以及无记忆效应等^[1-2]优点已被广泛应用于电动汽车^[3]、航空航天^[4]等领域。准确的估计锂离子电池 SOH 对于保障电池使用安全、延长电池寿命均具有重要意义^[5-7]。目前,学术界尚未就 SOH 的定义达成共识。为明确本工作的研究对象,根据 Okoshi 等人^[8]提出的定义,如式(1)所示,将 SOH 表示为当前电池容量 C_{aged} 与电池标称容量 C_0 的比值。

$$SOH = \frac{C_{aged}}{C_0} \times 100\% \quad (1)$$

在锂离子电池 SOH 估计的研究中,现有方法大多采用点估计^[9-11],这种方法虽然能给出一个具体的 SOH 数值,但无法反映估计的不确定性。相比之下,区间估计量化了这种不确定性,通过提供 SOH 可能的波动范围,为电池的安全管理、寿命预测和风险控制等实际应用提供了更有价值的信息。为了实现可靠的 SOH 区间估计,研究人员提出了各种方法,如: Bootstrap 法^[12]、核密度估计法^[13]和高斯过程^[14-15]等。然而,这些方法通常依赖于对数据分布的特定假设。当现实世界的的数据偏离这些假设时,会引入估计偏差。

为了克服这一局限性, Khosravi 等人^[16]在 2010 年提出了无分布的 LUBE 方法。该方法使用具有两个输出的神经网络来直接生成预测区间的上下限。LUBE 方法尽管避免了对数据分布的假设,但也带来了新的挑战。若原始损失函数不可微,则无法直

接应用梯度下降算法,需依赖元启发式优化算法进行参数训练。这带来了高昂的计算成本,尤其是对于大型模型。同时,这类算法也存在收敛速度慢、优化效率低等问题。针对此问题, Pearce 等人^[17]在 2018 年重新设计了损失函数,引入 Sigmoid 函数替代原有的阶跃函数,从而实现了与梯度下降算法兼容。Pang 等人^[18]将该思想应用到锂离子电池 SOH 估计中,并在此基础上考虑了预测区间中心的重要性。然而, Sigmoid 函数本身会引入需要手动调整的斜率参数,不当的选择会影响预测区间的质量。为了解决这个问题,本研究引入了一种新的损失函数消除手动调整斜率参数的需要,同时保持了与梯度下降算法的兼容性。

在锂离子电池 SOH 的无分布区间估计研究中,通常将容量作为健康因子进行研究。然而,容量的精确测量往往依赖昂贵且精密的设备,在实际应用场景中并不总是直接获得,需借助电流、电压、放电时长等间接特征进行推算。因此,直接以容量作为健康指标在实际应用中存在一定局限性。基于此,本研究选取更易于测量且与容量高度相关的健康指标,以推进锂离子电池 SOH 的无分布区间估计研究。

基于上述分析,本研究提出了一种面向多健康指标的锂离子电池 SOH 无分布区间估计方法,为电池安全管理、寿命预测和风险控制等实际应用提供了更具可靠性与实用性的区间估计方案。本研究的贡献可以总结为以下几点。①提出多健康指标的

电池 SOH 无分布区间估计框架。针对锂离子电池 SOH 估计, 选取更易于测量且与容量高度相关的多健康指标, 既不依赖于数据分布的假设, 又解决了以容量为单一健康因子时测量繁琐、实施受限的问题。有效克服了基于分布方法的局限性, 提升了所提方法在实际应用场景中的适应性。②使用 KPCA 对多健康指标进行非线性降维。此方法能够从存在复杂相关性的原始指标中, 提取出少数几个互不相关且具有强表征能力的综合主成分。这些主成分作为去相关、低噪的特征输入, 为后续区间估计模型提供了高质量的输入数据。③构建双输出神经网络模型并引入新型损失函数。构建以 KPCA 降维后的主成分为输入, 下一时刻 SOH 真实值为输出目标的双输出神经网络模型, 同时引入了一个兼顾区间覆盖概率与区间宽度的损失函数。该损失函数消除了对传统 LUBE 方法中不可微损失函数的依赖, 也避免了手动调整 Sigmoid 函数斜率参数的局限性, 实现了与梯度下降算法的兼容, 能够直接生成高质量的 SOH 预测区间。

1 特征提取

1.1 实验数据

为验证本工作所提方法对电池退化特征的捕捉能力, 研究选用 CALCE 电池数据集^[19-21]进行验证分析。该数据集选取了 CS2_35、CS2_36、CS2_37 和 CS2_38 共 4 组电池样本开展实验。4 节电池的真实 SOH 退化曲线如图 1 所示, 其额定容量均为 1.1 Ah。充电阶段采用 0.5 C 恒流充电至 4.2 V 后转为恒压模式, 直至电流降至 50 mA 以下; 放电阶段则以 1 C 恒流放电至截止电压 2.7 V。表 1 汇总了 CALCE 数据集在充放电阶段的关键参数。

1.2 健康因子的提取

锂离子电池恒流充电阶段, 电压达到一定值后增长变缓慢, 这被称为电压的平台期。为更直观地

表 1 CALCE 数据集中充电和放电阶段的具体参数

Table 1 Specific parameters of the charging and discharging phase in the CALCE dataset

电池	充电截止 电压/V	放电截止 电压/V	充电恒定 电流/C	放电恒定 电流/C
CS2_35	4.2	2.7	0.5	1
CS2_36	4.2	2.7	0.5	1
CS2_37	4.2	2.7	0.5	1
CS2_38	4.2	2.7	0.5	1

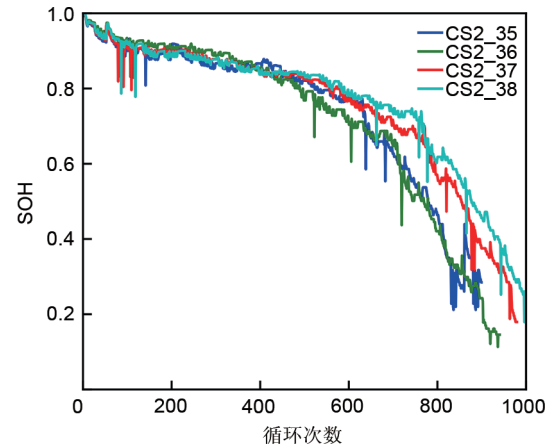


图 1 锂离子电池的 SOH 退化曲线

Fig. 1 SOH degradation curves of lithium-ion batteries

反映老化过程, 容量增量分析法 (incremental capacity analysis, ICA) 方法^[22]被引入, 表达式如下。

$$\frac{dQ}{dV} \approx \frac{\Delta Q}{\Delta V} = \frac{Q_k - Q_{k-1}}{V_k - V_{k-1}} \quad (2)$$

式中, Q_k 、 Q_{k-1} 为第 k 和 $k-1$ 时刻已充入的电池容量, V_k 、 V_{k-1} 为第 k 和 $k-1$ 时刻的电池端电压。电池的老化程度不同, 绘制出的 IC 曲线也会有明显差异, 因此可从 IC 曲线上提取电池老化特征。

本部分选取 CS2_36 电池作为研究对象, 从其对应的 IC 充电曲线中提取健康指标。如图 2(a) 所示, 受噪声及测量误差影响, 直接进行数值微分通常难以获得平滑可靠的结果。为克服这一问题, 本研究采用高斯滤波器对原始 IC 曲线进行平滑处理, 结果如图 2(b) 所示。观察可知, 随着电池老化, IC 充电曲线整体呈现向右下方偏移的趋势。基于该趋势, 本研究提取了 3 个关键健康指标, 分别为: IC 曲线的最高峰值、峰值对应的电压值, 以及峰值所覆盖的区域面积, 分别记作 HI1、HI2 和 HI3, 如图 3 所示。

1.3 KPCA 特征降维

在数据分析中, 各变量之间往往存在信息重叠, 这不仅会增加计算负担, 也会提升问题分析的复杂度。因此, 理想的特征集合应满足以下条件: 既能充分反映电池容量退化的关键信息, 又具备低冗余度和低维度的特点。考虑到数据中存在的非线性关系, KPCA 在此背景下适用于数据的非线性特征降维, 具体的算法步骤如下。

(1) 原始数据获取与标准化。将提取到的健康

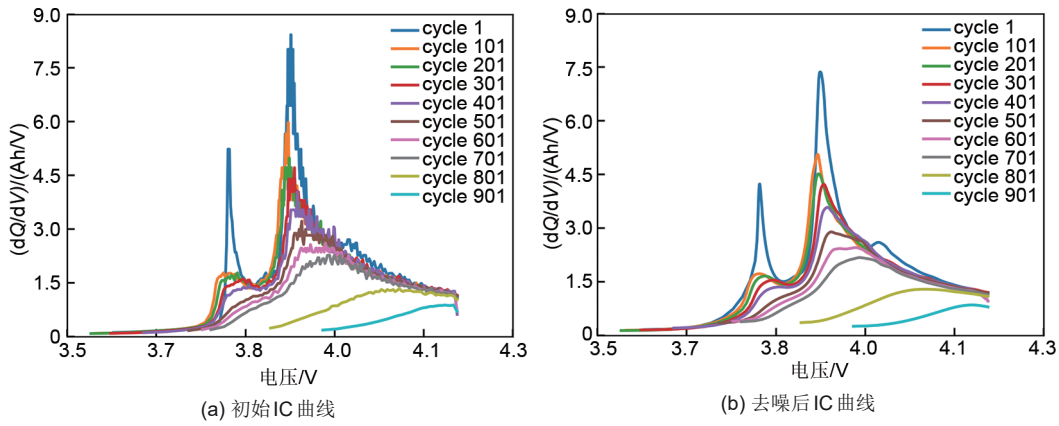


图2 CS2_36电池在不同周期的IC曲线

Fig. 2 IC curves of CS2_36 battery at different cycles

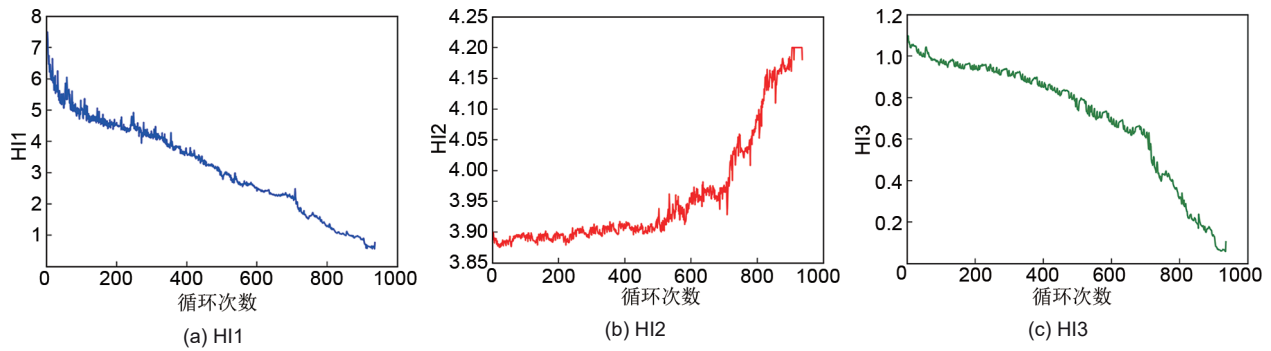


图3 3个关键健康因子

Fig. 3 Three key health indicators

指标组成矩阵并进行标准化, 得到矩阵 \mathbf{X} , 如式(3)所示。其中, m 代表电池健康指标数量, n 代表采集的电池样本数。

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} = [\mathbf{x}'_1, \mathbf{x}'_2, \cdots, \mathbf{x}'_n] \quad (3)$$

(2) 将RBF核函数作为KPCA的非线性核, 根据式(4)计算核矩阵 \mathbf{K} 。其中 \mathbf{x}_k 与 \mathbf{x}_g 是参数矩阵中的样本, $k_{\text{RBF}}(u, v) = \exp(-\|u - v\|^2 / \sigma^2)$ 为RBF核函数。

$$K_{ij} = k_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}_j) \quad (4)$$

(3) 对核矩阵 \mathbf{K} 进行中心化, 得到 \mathbf{K}' , 可以表示为式(5):

$$\mathbf{K}' = \mathbf{K} - \frac{1}{n} \mathbf{e} \mathbf{e}^T \mathbf{K} - \mathbf{K} \frac{1}{n} \mathbf{e} \mathbf{e}^T + \frac{1}{n} \mathbf{e} \mathbf{e}^T \mathbf{K} \frac{1}{n} \mathbf{e} \mathbf{e}^T \quad (5)$$

其中, \mathbf{e} 是所有元素为1的列向量。

(4) 通过引入非线性函数 $\phi(\mathbf{x}_k)$, 将数据空间

映射到高维空间, 并获得协方差矩阵 \mathbf{C} :

$$\mathbf{C} = \frac{1}{n} \sum_{k=1}^n \phi(\mathbf{x}_k) \cdot \phi(\mathbf{x}_k)^T \quad (6)$$

(5) 求解协方差矩阵的特征值 λ 和特征向量 \mathbf{V} , 其中, 特征值和特征向量满足以下条件:

$$\lambda \mathbf{V} = \mathbf{C} \mathbf{V} \quad (7)$$

引入 $\phi(\mathbf{x}_k)$, 式(7)可写为:

$$\lambda \phi(\mathbf{x}_k) \mathbf{V} - \phi(\mathbf{x}_k) \mathbf{C} \mathbf{V} = 0 \quad (8)$$

结合式(6)和(7), 特征向量 \mathbf{V} 可由公式(9)获得:

$$\mathbf{V} = \sum_{k=1}^n \alpha_k \phi(\mathbf{x}_k) \quad (9)$$

(6) 引入中心化核矩阵 \mathbf{K}' , 获得样本在非线性映射 $\phi(\mathbf{x})$ 上的投影, 引入 $\mathbf{K}' = [\phi(\mathbf{x}_k), \phi(\mathbf{x}_g)]$ 并结合式(6)~(9), 将式(7)简化为:

$$n \lambda \alpha = \mathbf{K}' \alpha \quad (10)$$

其中 α 是 \mathbf{K}' 的特征向量, 则样本在非线性映射 $\phi(\mathbf{x})$ 上的投影可表示为:

$$\mathbf{V}^T \cdot \phi(\mathbf{x}) = \sum_{k=1}^n \alpha_k \mathbf{K}(\mathbf{x}_k, \mathbf{x}) \quad (11)$$

(7) 根据式(12)计算主成分的累计贡献率 D_Σ ，并选择满足 $D_\Sigma \geq 85\%$ ^[23]的主成分作为融合HI。

$$D_\Sigma = \sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i, \quad (k = 1, 2, \dots, n) \quad (12)$$

1.4 定量评估

在统计学中，皮尔逊相关系数和斯皮尔曼相关系数常被用来评估变量之间的相关程度。通常，皮尔逊相关分析更适用于呈正态分布的连续变量。而斯皮尔曼相关系数则不受上述约束，适用范围更广^[23]。基于此，本研究选择斯皮尔曼相关分析来衡量健康指标与容量之间的相关性。其计算公式如下：

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (13)$$

$$d_i = x_i - y_i \quad (14)$$

式中， n 是样本量； ρ 值介于-1~1之间， ρ 绝对值越大，代表 x 和 y 之间的相关程度越高。

2 无分布SOH区间估计方法

考虑到基于分布方法依赖数据分布的假设，当假设与实际情况不符时，往往会引入估计偏差，进而影响估计结果的可靠性。而无分布LUBE方法^[16]由Khosravi等人于2010年提出，恰好为解决上述问题提供了有效途径。其核心思想是通过直接优化预测区间的关键评价指标，来构建高质量的预测区间。图4展示的是LUBE方法中用于构建预测区间的神经网络结构，这是一个具有双输出的前馈神经网络，其设计目的是直接估计预测区间的上界和下界。其中，输入层接收来自样本的特征向量，隐藏层通过非线性激活函数进行特征变换，以捕捉输入与输出之间的复杂关系。与传统的单输出神经网络不同，该模型的两个输出神经元分别直接估计预测区间的上界和下界。这种设计的优势在于神经网络能够通过优化覆盖概率和区间宽度的组合指标进行训练，从而直接生成高质量的预测区间。

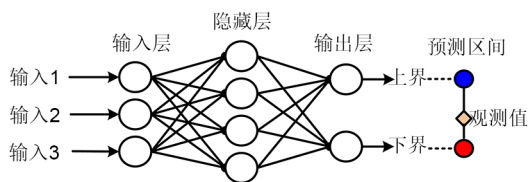


图4 神经网络模型结构

Fig. 4 Neural network model structure

LUBE方法虽然能得到可靠的预测区间，但其损失函数不可微，难以与梯度下降算法兼容，这极大地限制了LUBE在锂离子电池SOH估计中的应用。质量驱动的预测区间(QD_LUBE)方法^[17]通过引入Sigmoid函数，成功解决了损失函数不可微的问题，实现了与梯度下降算法的兼容性。然而，Sigmoid函数本身所存在的斜率问题，又会对预测区间的质量产生影响。

鉴于此，本研究引入了一个新的损失函数于锂离子电池SOH估计研究，摒弃了传统上依赖离散的阶跃函数 k_i [式(15)]或通过Sigmoid函数将原先离散的 k_i 转化为连续的 k_{soft} 形式[式(16)]。

$$k_i = \begin{cases} 1, & y_i \leq y_i \leq y_u \\ 0, & \text{其他} \end{cases} \quad (15)$$

$$k_{\text{soft}} = \sigma[s \cdot (y_i - y_l)] \cdot \sigma[s \cdot (y_u - y_i)] \quad (16)$$

式中， σ 是Sigmoid函数， s 为缩放因子，控制斜率。图5对上述离散函数和其通过Sigmoid函数平滑后的连续形式进行了直观对比。然而， s 需要人工进行调控，这会影响模型的估计精度。因此，本研究的损失函数 L 如式(17)，该损失函数由两部分组成。

$$\min_{\theta} L = L_w + L_c \quad (17)$$

式中， θ 代表神经网络的参数集合。

第一部分为宽度损失 L_w ，其定义如下：

$$L_w = \frac{1}{c} \sum_{i=1}^n \left[(y_u - y_i)^2 + (y_i - y_l)^2 \right] \times k_i \quad (18)$$

$$c = \sum_{i=1}^n k_i \quad (19)$$

式中， y_i 为第 i 个样本的观测值， y_u 和 y_l 分别是神经网络估计的预测区间的上界和下界。这种设计不仅间接地驱使区间宽度变窄，更重要的是，支持预测区间的边界紧密地包裹住实际的观测值。一个狭窄但偏离观测值的区间在此损失下会受到较大惩罚，从而促使网络生成高质量的预测区间。

第二部分为覆盖损失 L_c ，该项负责驱动模型满足预先定义的置信水平，定义如下。

$$L_c = UP_{\theta} + LP_{\theta} \quad (20)$$

$$UP_{\theta} = \sum_{i=1}^n \max(0, y_i - y_u) \times \gamma \quad (21)$$

$$LP_{\theta} = \sum_{i=1}^n \max(0, y_l - y_i) \times \gamma \quad (22)$$

式中， UP_{θ} 为上界惩罚项，在 $y_i > y_u$ 时被激

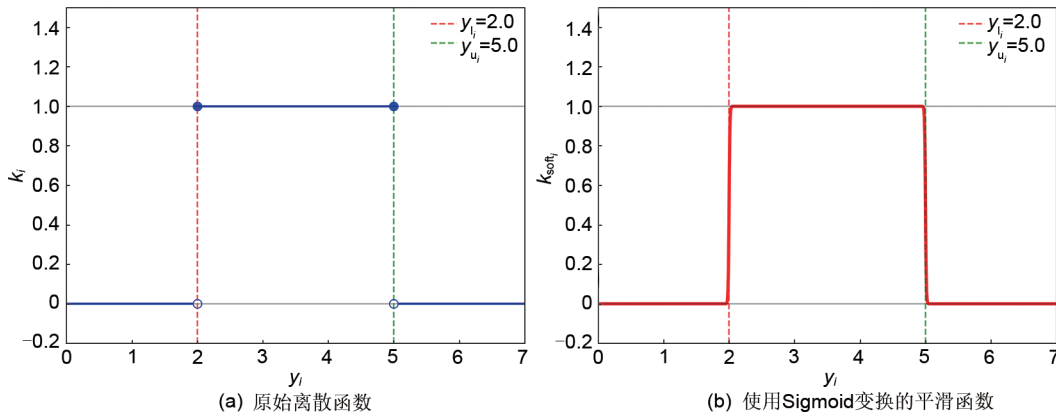


图5 离散函数 (a) 及其转化为连续的Sigmoid平滑形式 (b) 的比较

Fig. 5 Comparison between discrete functions (a) and their transformation into the continuous Sigmoid smooth form (b)

活； LP_{θ} 为下界惩罚项，在 $y_i < y_{li}$ 时被激活。为了确保模型在优化过程中满足预测区间覆盖概率 (prediction interval coverage probability, PICP) 的要求，并在宽度最小化与覆盖最大化之间形成平衡，需要一个自适应权重因子 γ 。 γ 值动态地取决于当前模型的实际PICP与预先定义的置信水平 $1 - \alpha$ 之间的差值。当 $\gamma > 0$ 时，覆盖损失通常远大于宽度损失，这强制优化器优先调整网络参数以扩大覆盖范围，从而快速提升PICP值。一旦实际PICP达到或超过 $1 - \alpha$ ， γ 的值自动变为零，覆盖损失项随之失效。优化器便可专注于最小化宽度损失，努力减小区间宽度， γ 的定义如式(23)。这种损失函数的设计克服了传统LUBE和QD_LUBE方法的固有局限性，从而直接生成高质量的预测区间。

$$\gamma = \lambda \times \max[0, (1 - \alpha) - \text{PICP}_{\theta}] \quad (23)$$

$$\text{PICP}_{\theta} = \frac{\sum_{i=1}^n k_i}{n} \quad (24)$$

3 多健康指标下无分布区间估计框架

根据上述方法原理，本部分设计了一种考虑多健康指标下的无分布区间估计框架，整体流程如图6所示。该策略包括3个主要部分：健康因子的提取与降维、无分布区间估计模型构建，以及模型性能评估，具体步骤如下。

步骤一：健康因子的提取与降维。相较于传统基于充放电曲线的方法，ICA具有更高的灵敏度，可有效提取与电化学性能相关的电池行为信息。基于此优势，本研究从IC曲线中提取了3个健康因

子：IC曲线的峰值、IC曲线峰值对应的电压值，以及峰值下的区域面积。进一步地，对上述多健康因子集合进行了KPCA非线性降维处理，从而获得互不相关且具有更强表征能力的低维特征。

步骤二：无分布区间估计模型构建。构建的双输出神经网络模型以融合后的健康指标序列为基础，采用滑动窗口方法提取时间窗口内的降维特征作为输入，并将下一时间步的实际SOH值作为输出目标。训练过程中，通过最小化所定义的损失函数[式(17)]，动态调整预测区间的覆盖概率与区间宽度之间的关系，在确保覆盖概率满足预先定义置信水平的同时尽可能缩小预测区间宽度。模型训练完成后，可直接输出预测区间的上下限，从而实现不确定性量化估计。

步骤三：模型性能评估。为检验模型性能，使用测试数据对其进行验证。通过以下3个指标对预测区间的质量进行综合评价：PICP、平均预测区间宽度(mean prediction interval width, MPIW)，以及基于覆盖宽度的评价准则(coverage width-based criterion, CWC)。

4 实验过程与结果分析

4.1 特征处理

为严格保证训练集与测试集的独立性并充分利用有限数据，本研究采用留一法交叉验证。具体而言，每次迭代将一个电池的完整数据留作测试集，其余电池数据作为训练集，并循环此过程直至所有电池均被测试一次。

在每次迭代生成的训练集上，为消除特征间的

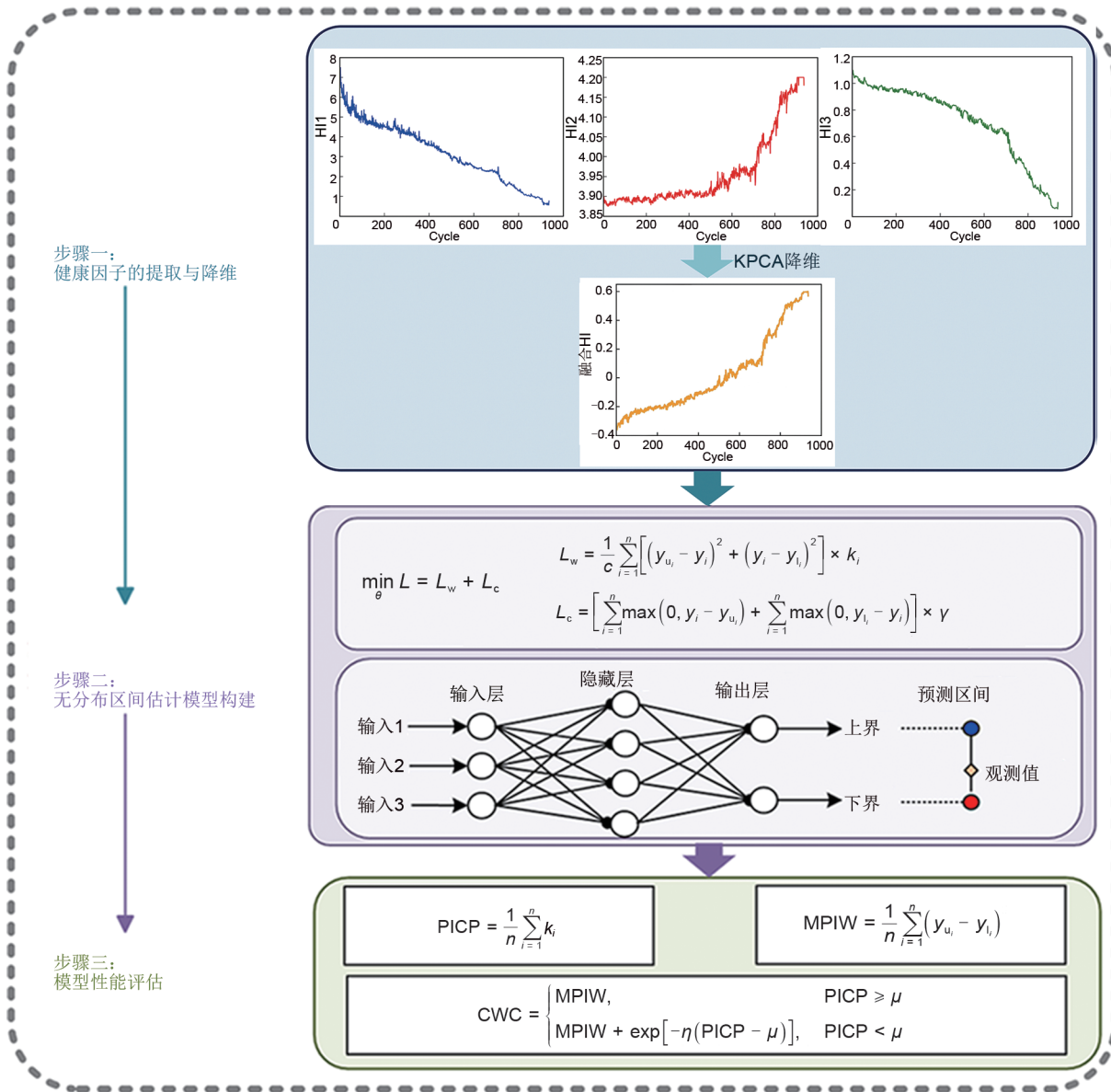


图 6 考虑多健康因子的无分布区间估计框架

Fig. 6 Distribution-free interval estimation of the health indicators with multiple health indicators

信息冗余，本研究采用KPCA算法进行降维处理。如表2所示，融合健康因子与电池容量之间斯皮尔曼相关系数的计算结果，证实了二者存在显著相关性。这进一步验证了第一主成分作为电池健康因子对电池退化状态的表征能力。因此，该融合健康因子可以替代容量，作为表征电池退化的指标。

在此基础上，参考文献[24]的建议，将滑动窗口长度设定为序列长度的5%~10%。据此，本研究对上述健康因子序列采用64的窗口长度进行划分，以构建模型的最终输入。所有模型的性能指标均在90%的置信水平下进行验证。

表 2 相关性分析

Table 2 Correlation analysis

Spearman 相关性	CS2_36, 37, 38	CS2_35, 37, 38	CS2_35, 36, 38	CS2_35, 36, 37
CS2_36,37,38 容量	-0.9892	—	—	—
CS2_35,37,38 容量	—	-0.9919	—	—
CS2_35,36,38 容量	—	—	-0.9907	—
CS2_35,36,37 容量	—	—	—	-0.9908

4.2 评价指标

在锂离子电池SOH区间估计的研究中，预测区间的质量综合评估应基于两个关键维度：覆盖概率和区间宽度。覆盖概率反映预测区间的可靠性，

而区间宽度反映预测区间的清晰度。为了进行定量评估,本研究采用了以下3个核心指标: PICP、MPIW和CWC。

PICP表示落在预测区间上限和下限内的观测值的比例。它是衡量可靠性的关键指标,定义如下:

$$\text{PICP} = \frac{\sum_{i=1}^n k_i}{n} \quad (25)$$

为了捕捉目标观测值的波动范围,需要另一个指标 MPIW。它是衡量预测区间清晰度的关键指标,定义如下:

$$\text{MPIW} = \frac{1}{n} \sum_{i=1}^n (y_{u_i} - y_{l_i}) \quad (26)$$

高质量预测区间应同时具备高 PICP 值和低 MPIW 值。为综合评估所得到的预测区间,引入了 CWC 综合评价指标^[25],定义如下。

$$\text{CWC} = \begin{cases} \text{MPIW}, & \text{PICP} \geq \mu \\ \text{MPIW} + \exp[-\eta(\text{PICP} - \mu)], & \text{PICP} < \mu \end{cases} \quad (27)$$

式中, $\mu = 1 - \alpha$ 为预先定义的置信水平, η 是控制惩罚强度的比例因子。当 PICP 达到置信水平 $1 - \alpha$ 时, $\text{CWC} = \text{MPIW}$, 此时 MPIW 越小, CWC 值越小, 预测区间质量更优; 若 PICP 未达到置信水平 $1 - \alpha$ 时, 则 PICP 越大, CWC 的值越小, 预测区间质量更优。总体而言, CWC 的值越小, 表明预测区间质量越好。

4.3 实验结果与分析

4.3.1 有效性验证

为验证所提出方法的有效性,本研究在 CALCE 数据集上对所提出的 GD_LUBE 方法进行了系统评估。如图7所示,大多数 SOH 退化数据点位于模型输出的预测区间上下界范围之内,表明该方法总体上能够较好地捕捉真实 SOH 变化趋势。然而,仅凭视觉观察尚不足以客观评估预测区间的可靠性,需进一步结合量化指标进行分析。由表3可知,所有 PICP 值均高于 90% 的置信水平,说明所构建的预测区间在统计意义上具备良好的可靠性。

为进一步验证 KPCA 在捕捉非线性特征方面的优势,本研究对比了 KPCA 降维后所得特征与原始未降维特征所对应的区间估计结果,如表4所示。在 90% 置信水平下,基于 KPCA 降维特征得到的预测区间在综合评价指标 CWC 上表现更优,表明其预测区间质量更高。

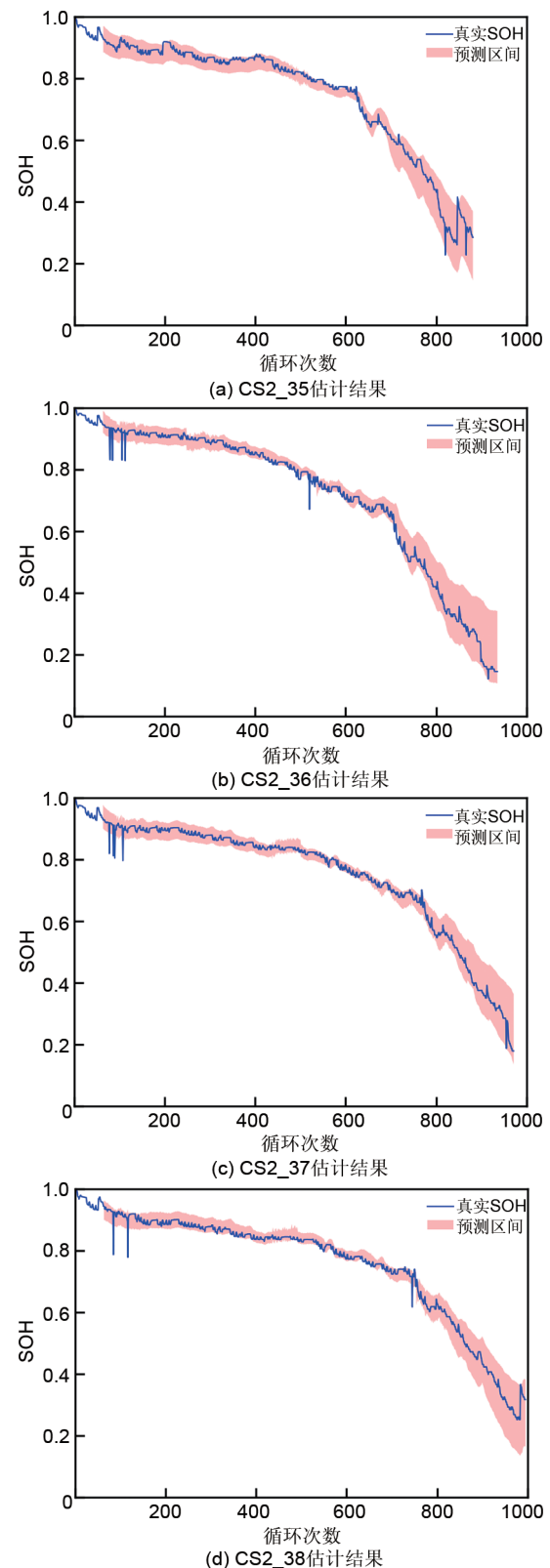


图7 GD_LUBE方法在CALCE数据集上的区间估计结果: (a) CS2_35; (b) CS2_36; (c) CS2_37; (d) CS2_38
Fig. 7 Interval estimation results of the GD_LUBE model on the CALCE dataset: (a) CS2_35; (b) CS2_36; (c) CS2_37; (d) CS2_38

表 3 GD_LUBE 方法在 CALCE 数据集上的性能评估结果
Table 3 Performance evaluation results of the GD_LUBE method on the CALCE dataset

电池	PICP	MPIW	CWC
CS2_35	0.9095	0.0765	0.0765
CS2_36	0.9060	0.0777	0.0777
CS2_37	0.9890	0.0688	0.0688
CS2_38	0.9442	0.0683	0.0683

表 4 KPCA 与无降维区间估计结果对比
Table 4 Comparison of KPCA and non-dimensionality reduction interval estimation results

电池	KPCA 降维预测区间			无降维预测区间		
	PICP	MPIW	CWC	PICP	MPIW	CWC
CS2_35	0.9095	0.0765	0.0765	0.9780	0.0783	0.0783
CS2_36	0.9060	0.0777	0.0777	0.9404	0.0798	0.0798
CS2_37	0.9890	0.0688	0.0688	0.9681	0.0700	0.0700
CS2_38	0.9442	0.0683	0.0683	0.9818	0.0710	0.0710

4.3.2 对比实验与性能分析

为了评估本研究的方法在清晰度和可靠性方面的性能优势，设计并进行了两组对比实验。第一组实验采用基于 Sigmoid 函数的传统 QD_LUBE 方法，作为基准模型；第二组实验则采用本研究提出的 GD_LUBE 方法，该方法整合了一种新型可微损失函数，旨在验证避免人工调节斜率参数的损失函数对预测区间质量的影响。

图 8 展示了两种方法在 CALCE 数据集上对电池 SOH 的区间估计结果。其中，图 8(a) 为基于 Sigmoid 函数的 QD_LUBE 方法所得结果，其区间虽能覆盖真实值，但整体宽度较大；图 8(b) 为使用新型损失函数的 GD_LUBE 方法所得结果，可见其生成的预测区间在覆盖真实 SOH 值的同时，区间宽度有所收窄。

表 5 中的实验数据进一步从量化角度验证了上述观察。在关键评价指标 MPIW 和 CWC 上，GD_LUBE 方法均优于传统 QD_LUBE 方法。具体而言，GD_LUBE 方法得到的区间估计结果在 MPIW 指标上有所下降，反映出区间宽度的有效压缩；而在 CWC 指标上的改善，则说明区间在变窄的同时仍能保持理想的覆盖概率，兼顾了可靠性。这些结果有力证明了本研究所提方法能够更好地平衡区间宽度与覆盖概率之间的关系，从而在保证必

要置信水平的前提下，生成高质量的预测区间。

4.3.3 多步验证

在本部分中，本研究对所提出方法在多步预测任务中的性能进行了系统评估。如表 6 所示，随着预测步长从第 1 步逐步延伸至第 6 步，两种对比模型所得到的 PICP 值均能稳定在 90% 置信水平附近，这从统计层面证明了所有方法都具备了基本的可靠性。在所有预测步长上，所提方法绝大多数在 MPIW 指标上取得了最小值。此外，综合衡量区间宽度与覆盖能力的 CWC 指标提供了更具说服力的证据。从第 1 步到第 6 步，所提出模型的 CWC 值绝大多数都优于对比方法。这一连贯的优势表明，本方法不仅在单步预测中表现优异，在多步预测中同样卓越，能够生成高质量的预测区间。

5 结 论

本研究提出了一种多健康因子的无分布区间估计方法。通过 KPCA 对多健康因子进行降维处理，并结合采用新型损失函数的双输出神经网络，能够有效实现锂离子电池 SOH 的无分布区间估计。该方法创新之处在于考虑了容量作为健康指标在实际应用中的局限性，并解决了传统 LUBE 方法中损失函数不可微、训练效率低下的问题，同时克服了现有可微损失函数需手动调整斜率参数的局限。在 CALCE 数据集上的实验结果表明，所提出的 GD_LUBE 方法在综合评价指标 CWC 上显著优于传统 QD_LUBE 方法，平均下降了 12.07%，验证了该方法的有效性和优越性。此外，在多步预测任务中，该方法同样表现出性能提升，进一步体现出所提方法具有良好预测稳定性。

虽然本研究为锂离子电池 SOH 提供了一种有效的无分布区间估计策略，但仍有一些需要改进的地方。现有方法假设训练数据与测试数据来自同一分布，然而在实际应用中，电池类型、工作条件和老化模式的差异往往导致数据分布不一致。迁移学习作为一种能够利用已有知识提升模型在新领域性能的范式，为解决此类域偏移问题提供了可能。未来研究可探索无监督域自适应等方法，学习具有域不变性的特征，从而提升模型在未见电池或工况下的泛化能力。

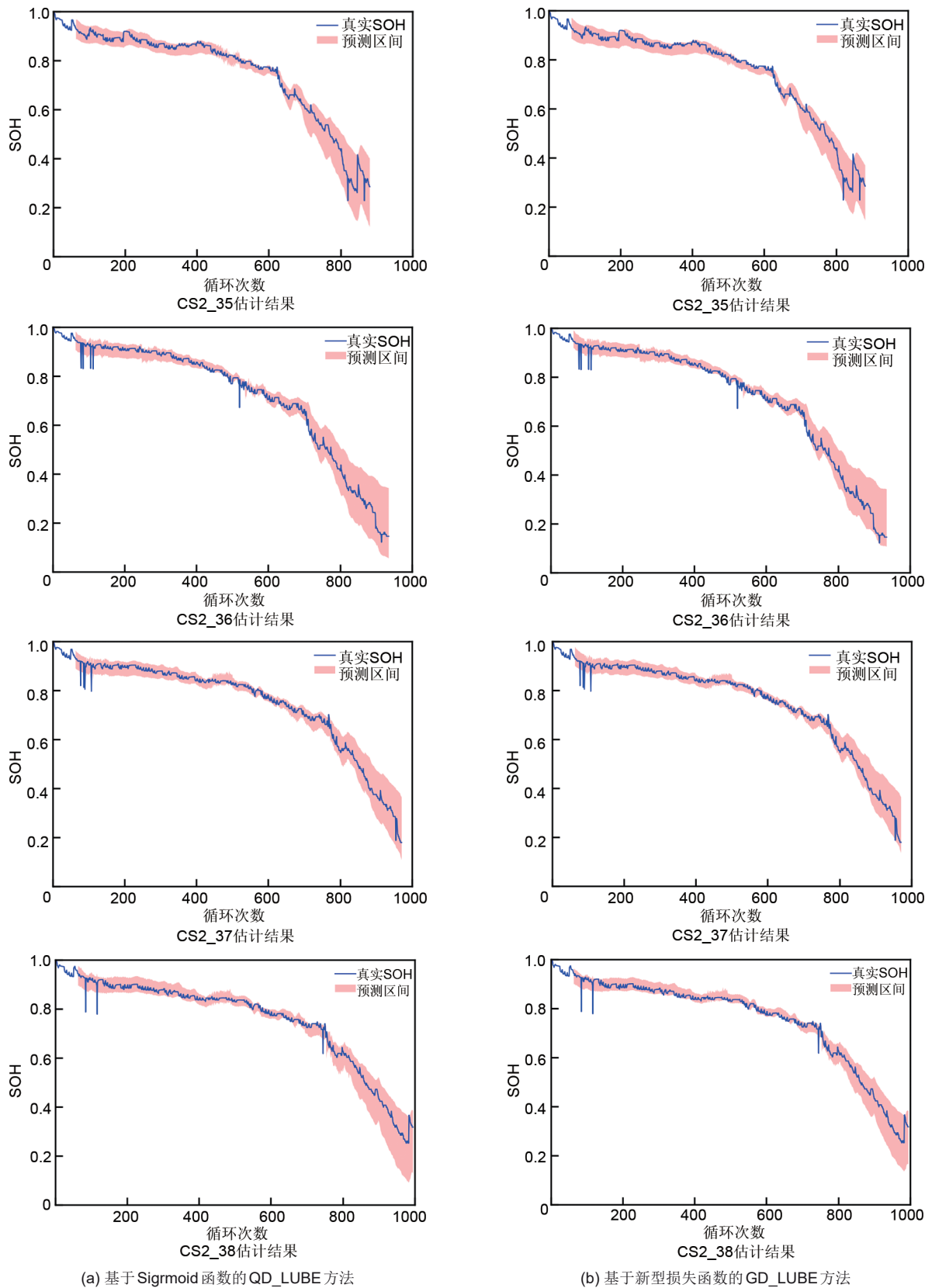


图 8 CALCE 数据集的 SOH 区间估计结果图
Fig. 8 CALCE dataset SOH interval estimation results

表 5 CALCE 数据集的 SOH 区间估计评估结果

Table 5 SOH interval estimation evaluation results for the CALCE dataset

电池	方法	PICP	MPIW	CWC
CS2_35	QD_LUBE	0.9584	0.0850	0.0850
	GD_LUBE	0.9095	0.0765	0.0765
CS2_36	QD_LUBE	0.9358	0.0886	0.0886
	GD_LUBE	0.9060	0.0777	0.0777
CS2_37	QD_LUBE	0.9901	0.0761	0.0761
	GD_LUBE	0.9890	0.0688	0.0688
CS2_38	QD_LUBE	0.9721	0.0817	0.0817
	GD_LUBE	0.9442	0.0683	0.0683

表 6 CALCE 数据集的 SOH 多步区间估计评估结果

Table 6 Evaluation results of SOH multi-step interval estimation for the CALCE dataset

电池	方法	1 步预测			2 步预测			3 步预测		
		PICP	MPIW	CWC	PICP	MPIW	CWC	PICP	MPIW	CWC
CS2_35	QD_LUBE	0.9584	0.0850	0.0850	0.8996	0.0829	1.1014	0.8848	0.0830	2.2208
	GD_LUBE	0.9095	0.0765	0.0765	0.0911	0.0716	0.0716	0.9044	0.0754	0.0754
CS2_36	QD_LUBE	0.9358	0.0886	0.0886	0.9357	0.0955	0.0955	0.9184	0.0953	0.0953
	GD_LUBE	0.9060	0.0777	0.0777	0.8955	0.0885	1.3394	0.9322	0.0872	0.0872
CS2_37	QD_LUBE	0.9901	0.0761	0.0761	0.9879	0.0834	0.0834	0.9923	0.0827	0.0827
	GD_LUBE	0.9890	0.0688	0.0688	0.9901	0.0779	0.0779	0.9857	0.0787	0.0787
CS2_38	QD_LUBE	0.9721	0.0817	0.0817	0.9549	0.0792	0.0792	0.9624	0.0791	0.0791
	GD_LUBE	0.9442	0.0683	0.0683	0.9624	0.0745	0.0745	0.9720	0.0782	0.0782
电池	方法	4 步预测			5 步预测			6 步预测		
		PICP	MPIW	CWC	PICP	MPIW	CWC	PICP	MPIW	CWC
CS2_35	QD_LUBE	0.9055	0.0774	0.0774	0.9263	0.0811	0.0811	0.7909	0.0778	234.0272
	GD_LUBE	0.9043	0.0712	0.0712	0.9080	0.0751	0.0751	0.9079	0.0773	0.0773
CS2_36	QD_LUBE	0.9471	0.0975	0.0975	0.9424	0.0975	0.0975	0.9400	0.0951	0.0951
	GD_LUBE	0.9344	0.0914	0.0914	0.9320	0.0879	0.0879	0.9181	0.0953	0.0953
CS2_37	QD_LUBE	0.9901	0.0784	0.0784	0.9834	0.0830	0.0830	0.9856	0.0832	0.0832
	GD_LUBE	0.9867	0.0742	0.0742	0.9790	0.0750	0.0750	0.9878	0.0766	0.0766
CS2_38	QD_LUBE	0.9709	0.0843	0.0843	0.9591	0.0888	0.0888	0.9364	0.0785	0.0785
	GD_LUBE	0.9580	0.0736	0.0736	0.9515	0.0760	0.0760	0.9622	0.0771	0.0771

参考文献

[1] CHEN Y Q, KANG Y Q, ZHAO Y, et al. A review of lithium-ion battery safety concerns: The issues, strategies, and testing standards[J]. Journal of Energy Chemistry, 2021, 59: 83-99. DOI: 10.1016/j.jechem.2020.10.017.

[2] NI Y L, SONG K, PEI L, et al. State-of-health estimation and knee point identification of lithium-ion battery based on data-driven and mechanism model[J]. Applied Energy, 2025, 385: 125539. DOI:10.1016/j.apenergy.2025.125539.

[3] TIAN J Q, LIU X H, LI S Q, et al. Lithium-ion battery health estimation with real-world data for electric vehicles[J]. Energy, 2023, 270: 126855. DOI:10.1016/j.energy.2023.126855.

[4] ZHANG D F, LI W C, HAN X D, et al. Evolving Elman neural networks based state-of-health estimation for satellite lithium-ion batteries[J]. Journal of Energy Storage, 2023, 59: 106571. DOI: 10.1016/j.est.2022.106571.

[5] CAI L, LIN J D, LIAO X Y. A data-driven method for state of health prediction of lithium-ion batteries in a unified framework[J]. Journal of Energy Storage, 2022, 51: 104371. DOI: 10.1016/j.est.2022.104371.

[6] 朱冰, 夏天. 多元宇宙优化估算锂离子电池的 SOC 与 SOH[J]. 电池,

- 2024, 54(5): 688-692. DOI: 10.19535/j.1001-1579.2024.05.017.
- ZHU B, XIA T. Estimation of SOC and SOH for Li-ion battery by multi-verse optimization[J]. Battery Bimonthly, 2024, 54(5): 688-692. DOI:10.19535/j.1001-1579.2024.05.017
- [7] LI H J, WANG S K, YANG L, et al. SOH estimation method for lithium-ion battery packs under real-world operating conditions based on a new attenuated model without additional experiments [J]. Energy, 2025, 330: 136802. DOI:10.1016/j.energy.2025.136802.
- [8] OKOSHI T, YAMADA K, HIRASAWA T, et al. Battery condition monitoring (BCM) technologies about lead-acid batteries[J]. Journal of Power Sources, 2006, 158(2): 874-878. DOI:10.1016/j.jpowsour.2005.11.008.
- [9] CHEN L P, XIE S Q, LOPES A M, et al. A new SOH estimation method for lithium-ion batteries based on model-data-fusion[J]. Energy, 2024, 286: 129597. DOI:10.1016/j.energy.2023.129597.
- [10] LIN M Q, YOU Y Q, MENG J H, et al. Lithium-ion batteries SOH estimation with multimodal multilinear feature fusion[J]. IEEE Transactions on Energy Conversion, 2023, 38(4): 2959-2968. DOI:10.1109/TEC.2023.3282017.
- [11] GONG D L, GAO Y, KOU Y L, et al. State of health estimation for lithium-ion battery based on energy features[J]. Energy, 2022, 257: 124812. DOI:10.1016/j.energy.2022.124812.
- [12] GOU B, XU Y, FENG X. State-of-health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method[J]. IEEE Transactions on Vehicular Technology, 2020, 69(10): 10854-10867. DOI:10.1109/TVT.2020.3014932.
- [13] CHEN C, TAO G Y, SHI J T, et al. A lithium-ion battery degradation prediction model with uncertainty quantification for its predictive maintenance[J]. IEEE Transactions on Industrial Electronics, 2024, 71(4): 3650-3659. DOI:10.1109/TIE.2023.3274874.
- [14] HE Y, BAI W Y, WANG L L, et al. SOH estimation for lithium-ion batteries: An improved GPR optimization method based on the developed feature extraction[J]. Journal of Energy Storage, 2024, 83: 110678. DOI:10.1016/j.est.2024.110678.
- [15] WANG Z P, YUAN C G, LI X Y. Lithium battery state-of-health estimation via differential thermal voltammetry with Gaussian process regression[J]. IEEE Transactions on Transportation Electrification, 2021, 7(1): 16-25. DOI:10.1109/TTE.2020.3028784.
- [16] KHOSRAVI A, NAHAVANDI S, CREIGHTON D, et al. Lower upper bound estimation method for construction of neural network-based prediction intervals[J]. IEEE Transactions on Neural Networks, 2011, 22(3): 337-346. DOI:10.1109/TNN.2010.2096824.
- [17] PEARCE T, BRINTRUP A, ZAKI M, et al. High-quality prediction intervals for deep learning: A distribution-free, ensembled approach[C]//Stockholm: Proceedings of the 35th International Conference on Machine Learning, 2018, 80: 4075-4084.
- [18] PANG X Q, GUO Z Y, JIA J F, et al. A distribution-free interval estimation method for lithium-ion battery state of health[J]. Ionics, 2025, 31(8): 7939-7952. DOI:10.1007/s11581-025-06470-3.
- [19] WILLIARD N, HE W, OSTERMAN M, et al. Comparative analysis of features for determining state of health in lithium-ion batteries [J]. International Journal of Prognostics and Health Management, 2013, 4(1): 14-20. DOI:10.36001/ijphm.2013.v4i1.1437.
- [20] 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. DOI:10.1016/j.jpowsour.2011.08.040.
- [21] XING Y J, MA E W M, TSUI K L, et al. An ensemble model for predicting the remaining useful performance of lithium-ion batteries[J]. Microelectronics Reliability, 2013, 53(6): 811-820. DOI:10.1016/j.microrel.2012.12.003.
- [22] WENG C H, CUI Y J, SUN J, et al. On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression[J]. Journal of Power Sources, 2013, 235: 36-44. DOI:10.1016/j.jpowsour.2013.02.012.
- [23] PANG X Q, LIU X Y, JIA J F, et al. A lithium-ion battery remaining useful life prediction method based on the incremental capacity analysis and Gaussian process regression[J]. Microelectronics Reliability, 2021, 127: 114405. DOI:10.1016/j.microrel.2021.114405.
- [24] LV Z H, SONG Y, HE C L, et al. Remaining useful life prediction for lithium-ion batteries incorporating spatio-temporal information [J]. Journal of Energy Storage, 2024, 88: 111626. DOI:10.1016/j.est.2024.111626.
- [25] SHRIVASTAVA N A, KHOSRAVI A, PANIGRAHI B K. Prediction interval estimation of electricity prices using PSO-tuned support vector machines[J]. IEEE Transactions on Industrial Informatics, 2015, 11(2): 322-331. DOI:10.1109/TII.2015.2389625.