



## 基于SVD-SAE-GPR的锂离子电池RUL预测

董渊昌<sup>1</sup>, 庞晓琼<sup>1</sup>, 贾建芳<sup>2</sup>, 史元浩<sup>2</sup>, 温杰<sup>2</sup>, 李笑<sup>1</sup>, 张鑫<sup>1</sup>

(<sup>1</sup>中北大学计算机科学与技术学院; <sup>2</sup>中北大学电气与控制工程学院, 山西太原 030051)

**摘要:** 锂离子电池是重要的储能手段之一, 对其剩余使用寿命(RUL)进行预测具有非常重要的实际意义。本工作首先针对传统特征提取方法依赖参数设置且对于不同锂离子电池数据集适应性差的缺陷, 将电池数据视作矩阵, 并引入奇异值分解(SVD)从测量数据和包含更多退化信息的特征提取对象中提取潜在健康因子(HIs)。其次, 潜在HIs的冗余和不足会影响RUL的预测, 同时考虑到主成分分析(PCA)的缺点, 使用Spearman相关分析和堆叠自编码器(SAE)处理HIs得到一个融合HI。在此基础上, 利用高斯过程回归(GPR)算法构建了融合HI与容量之间的模型, 得到了带有不确定性表达的最终预测结果。最后, 通过NASA提供的四个老化电池数据验证了所提预测模型的可行性和有效性。并额外选取MIT电池数据集验证特征提取方法的适应性。实验结果表明, 所提出的RUL预测框架具有较好的预测性能, SVD特征提取方法在避免参数设置的前提下具有较好的适应性。本工作提取的HI与经过PCA融合的HI、其他HI相比, 预测精度有显著提高。

**关键词:** 锂离子电池; 剩余使用寿命(RUL); 奇异值分解; 堆叠自编码器; 高斯过程回归

doi: 10.19799/j.cnki.2095-4239.2022.0767

中图分类号: TP 206

文献标志码: A

文章编号: 2095-4239 (2023) 04-1257-11

## Remaining useful life prediction of lithium-ion batteries based on SVD-SAE-GPR

DONG Yuanchang<sup>1</sup>, PANG Xiaoqiong<sup>1</sup>, JIA Jianfang<sup>2</sup>, SHI Yuanhao<sup>2</sup>, WEN Jie<sup>2</sup>,  
LI Xiao<sup>1</sup>, ZHANG Xin<sup>1</sup>

(<sup>1</sup>School of Computer Science and Technology; <sup>2</sup>School of Electrical and Control Engineering, North University of China, Taiyuan 030051, Shanxi, China)

**Abstract:** Lithium-ion batteries are an important energy storage sources, and it is of great practical importance to predict their remaining useful life (RUL). First, the battery data are treated as matrices, and singular value decomposition (SVD) is introduced to extract the potential health indicator (HI) from the measured data and the feature extraction objects containing more degradation information. This will address the drawbacks of traditional feature extraction methods that rely on parameter settings and poor adaptability to different lithium-ion battery datasets. Second, the redundancy and deficiency of potential HIs affect the prediction of RUL, and thus, a fused HI is obtained by processing HIs using Spearman correlation analysis and stacked autoencoder, considering the shortcomings of principal component

收稿日期: 2022-12-27; 修改稿日期: 2023-02-21。

基金项目: 国家自然科学基金 (7207011096), 山西省高性能电池材料与器件重点实验室开放基金 (2022HPBMD01002), 中国山西省留学基金委资助课题 (2020-114)。

第一作者: 董渊昌 (1999—), 男, 硕士研究生, 研究方向为复杂系统的故障预测与健康管理, E-mail: dyc36309123@163.com; 通讯作者: 庞晓琼, 副教授, 研究方向为复杂系统的故障预测与健康管理、信息安全与密码学, E-mail: xqpang@nuc.edu.cn。

引用本文: 董渊昌, 庞晓琼, 贾建芳, 等. 基于SVD-SAE-GPR的锂离子电池RUL预测[J]. 储能科学与技术, 2023, 12(4): 1257-1267.

**Citation:** DONG Yuanchang, PANG Xiaoqiong, JIA Jianfang, et al. Remaining useful life prediction of lithium-ion batteries based on SVD-SAE-GPR[J]. Energy Storage Science and Technology, 2023, 12(4): 1257-1267.

analysis (PCA). Accordingly, a model between fused HI and capacity is constructed using the Gaussian process regression algorithm, and the final prediction results with uncertainty expression are obtained. Finally, the feasibility and validity of the proposed prediction model are verified by four aging batteries provided by NASA. The MIT battery dataset is used to verify the adaptability of the feature extraction method. The experimental results show that the proposed RUL prediction framework has good prediction performance and that the SVD feature extraction method has good adaptability while avoiding parameter settings. The HI extracted in this paper has significantly improved the prediction accuracy compared with the HI after PCA fusion and other HIs.

**Keywords:** lithium-ion batteries; remaining useful life; singular value decomposition; stacked autoencoder; gaussian process regression

锂离子电池因其工作电压高、功率密度大、寿命长等优点,已经成为电动汽车、智能电网和储能系统等众多应用领域的主流能源存储器件<sup>[1-2]</sup>。然而,随着充放电次数的增加,电池会通过不同的机制退化甚至失效。如果在其达到失效阈值前不及时采取有效措施,可能会导致设备性能退化甚至灾难性事件<sup>[3]</sup>的发生。因此,研究电池的退化并建立可靠的退化模型以准确预测其剩余使用寿命(remaining useful life, RUL)是非常重要的。

近年来,数据驱动的方法因可以利用历史监测数据直接预测电池的退化趋势而不依赖于复杂的物理模型,在锂离子电池 RUL 预测领域受到了广泛关注<sup>[4]</sup>。目前大多研究定义容量<sup>[5]</sup>或内阻<sup>[6]</sup>作为健康因子(health indicator, HI)。然而,容量或内阻的检测对实验条件和仪器设备要求较高,难以实现在线采集<sup>[7]</sup>。因此,提取易于计算且能量化电池退化状态的 HI 具有重要的实用价值。Khaleghi 等<sup>[8]</sup>设置电压上下边界并指定固定的采样率从部分充电电压曲线中提取 HIs 并用来估计电池健康状态。此外,短放电时间间隔内的电流变化<sup>[9]</sup>、固定电压间隔的时间差<sup>[10]</sup>等也可以作为反映锂离子电池健康状态的特征。虽然这些方法<sup>[8-10]</sup>在提取 HI 方面展现了优异的性能,但需要设定固定的时间或电压等参数,这不仅依赖于人为经验,还有可能使特征提取变得复杂。Liu 等<sup>[11]</sup>分别在 NASA 和 CALCE 的电池数据集上设置了不同的参数提取相同 HI 并进行 RUL 预测,结果表明不同锂电池特征提取参数需求可能不同。Pang 等<sup>[12]</sup>提取了更敏感的 IC 曲线的峰值和峰下的区域面积作为 HIs,并对 NASA 提供的锂离子电池数据集进行 RUL 预测;而 Pan 等<sup>[13]</sup>在 CALCE 提供

的电池数据集上提取的 IC 曲线有两个可识别的峰,此特征提取方法显然不适用于不同的锂离子电池数据集。文献[11-13]的特征提取方法虽然可以用于提取 HI,但由于电池参数或实验环境的差异,可能不适用于其他锂离子电池数据集。如上所述,传统的特征提取方法通常需要设定参数且适应性较差,因此,本工作的第一个动机是寻找一种能减少或避免参数设定,且对于不同的锂离子电池数据集适应性强的特征提取方法。

基于以上的分析,本工作对电池的原始测量数据重新进行分析,尝试从中获取隐藏的特征信息。根据电池数据的特点,同时考虑到特征值是矩阵的重要信息,本工作将电池数据看作矩阵,从矩阵中计算特征值,然后判断该特征值是否可以作为重要信息来表征电池的退化状态。但电池数据不是方阵,无法获得其特征值,而奇异值分解<sup>[14]</sup>(singular value decomposition, SVD)可以从非方阵中提取奇异值来表示矩阵的重要信息。因此,本文选择 SVD 从电池数据中计算奇异值作为 HIs。同时,考虑到提取的 HIs 会有一定的冗余性,当作为训练数据时,其维度过高和不足会使模型训练变得复杂且不利于 RUL 预测的准确性。因此,本工作的第二个动机是消除 HIs 不相关信息的负面影响和降低数据维数。在数据去噪和降维方面,主成分分析<sup>[15]</sup>(principal component analysis, PCA)表现出了较好的性能,但它适合于服从高斯分布的数据且只能执行线性变换,缺乏一定的灵活性。堆叠自编码器<sup>[16]</sup>(stacked autoencoder, SAE)可以自动从无标注数据中提取学习特征,给出比原始数据更好的特征描述。因此,选择 SAE 用于 HIs 的在线融合。

此外,模型和数据的不确定性可能导致预测可靠性差,因此,在研究锂离子电池的RUL预测时考虑预测结果的不确定性是非常有必要的。高斯过程回归<sup>[17]</sup>(gaussian process regression, GPR)能够实现任意线性或者非线性系统动态行为特征的建模预测,并能以概率的形式解释预测结果的不确定性。考虑到锂离子电池的退化过程是一个复杂的、动态的、非线性的电化学过程,GPR方法适合于建立锂离子电池RUL预测模型。

综上所述,本文提出了一种不需要参数设定且对于不同锂离子电池数据集有良好适应性的特征提取策略,从电池数据中提取潜在HIs并进行相关性分析,选择了极强相关性的HIs进行SAE数据降维,并将融合的HI结合GPR进行RUL预测。通过两组实验验证了本文所提方法的优越性,一是通过不同的锂离子电池数据集验证了SVD特征提取方法具有良好的适应性;二是将本文所提的HI、经过PCA融合的HI与其他HI分别训练GPR模型并进行对比,验证了本文所提的HI具有更好的预测性能。

## 1 特征提取与分析

### 1.1 实验数据

本工作中选择的电池退化数据集来自NASA PCoE 研究中心<sup>[18]</sup>。该数据集包含在相同实验环境中收集退化状态数据的4个18650型锂离子电池。4个电池的标称容量为2.2 Ah,标称电压为3.7 V。在充电阶段,以1.5 A恒流(CC)模式充电,直到电池电压达到4.2 V,然后继续以恒压(CV)模式充电,直到充电电流降至20 mA。在放电阶段,以2 A的CC进行放电,直到B0005、B0006、B0007和B0018电池的电压分别降至2.7 V、2.5 V、2.2 V和2.5 V。图1为电池实际容量衰减曲线。

### 1.2 基于SVD的特征提取

在本小节中,首先介绍奇异值分解<sup>[19]</sup>方法,该方法在本工作中用于特征提取。SVD是一种矩阵分解方法,它可以将复杂矩阵表示为更小更简单的子矩阵的乘法,如图2所示。

假设矩阵 $\mathbf{A}$ 是一个大小为 $p \times q$ 的矩阵,则定义矩阵 $\mathbf{A}$ 的SVD如式(1)所示:

$$\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (1)$$

矩阵 $\mathbf{U}$ 的秩为 $r$ , $\mathbf{U}$ 是大小为 $p \times p$ 的正交矩阵, $\mathbf{U}$ 的列向量称为左奇异向量。 $\mathbf{\Sigma}$ 是一个大小为

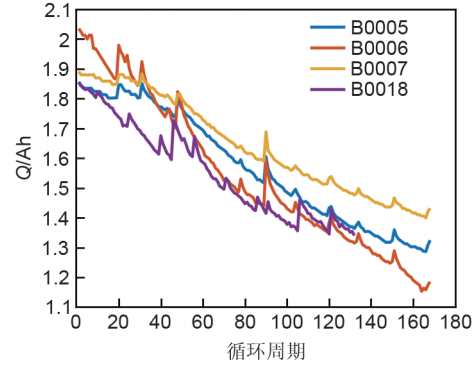


图1 4个电池的容量退化曲线  
Fig. 1 Capacity degradation curves of four batteries

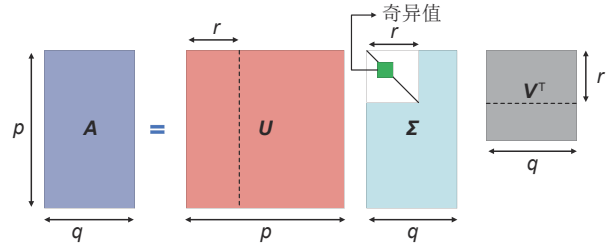


图2 矩阵的SVD  
Fig. 2 SVD of matrix

$p \times q$ 的矩阵,除了主对角线外,所有元素都为0,它是一个对角矩阵。 $\mathbf{\Sigma}$ 主对角线上的每个元素称为 $\mathbf{A}$ 的奇异值,记为 $\sigma_i (i = 1, \dots, r)$ 。 $\mathbf{V}$ 是一个大小为 $q \times q$ 的正交矩阵, $\mathbf{V}$ 的列向量是右奇异向量。将 $\mathbf{A}$ 乘以其转置矩阵 $\mathbf{A}^T$ ,得到式(2),同理可得与 $\mathbf{A}\mathbf{A}^T$ 相关的式(3)。

$$\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{\Sigma} \mathbf{U}^T \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T = \mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^T \quad (2)$$

$$\mathbf{A}\mathbf{A}^T = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{V} \mathbf{\Sigma} \mathbf{U}^T = \mathbf{U} \mathbf{\Sigma}^2 \mathbf{U}^T \quad (3)$$

可以看出, $\mathbf{A}^T \mathbf{A}$ 和 $\mathbf{A}\mathbf{A}^T$ 的特征向量分别组成了SVD中的矩阵 $\mathbf{V}$ 和 $\mathbf{U}$ 。由于 $\mathbf{A}^T \mathbf{A}$ 和 $\mathbf{A}\mathbf{A}^T$ 的特征值均为 $\lambda_1, \dots, \lambda_r$ ,如式(4)所示。因此,将矩阵 $\mathbf{A}$ 的奇异值记为 $\sigma_i = \sqrt{\lambda_i} (i = 1, \dots, r)$ 。

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T \mathbf{A} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r \end{bmatrix} \quad (4)$$

矩阵奇异值的大小与矩阵中隐藏的重要信息密切相关,信息的重要性与奇异值的大小正相关。可以发现,SVD在矩阵分解过程中不涉及参数设置,可以实现简单的计算过程,降低特征提取的复杂性,同时避免了对人为经验的依赖。因此,选择SVD用来提取能表征电池退化状态的HIs。



通常情况下,锂离子电池的电压、电流和温度等参数都是可测的,因此首先从12组测量数据中通过SVD提取各自的奇异值作为12个HIs(详细描述见表1)。同时考虑到特征提取对象包含的退化信息对提取出来的HI量化电池退化状态的能力有重要影响,额外计算了三组特征提取对象,分别为dQ/dV, dV/dQ和dT/dV矩阵。锂离子电池的退化是一个动态非线性过程,降解模式多种多样,如电极材料的机械降解、锂库存损失、活性材料损失等。所有这些退化模式也可以相互影响。dQ/dV、dV/dQ和dT/dV常被用来分析上述相关的退化模式<sup>[20-22]</sup>,因此可以作为反应电池退化状态的特征提取对象,从其中分别提取了HI13、HI14和HI15。dQ/dV、dV/dQ和dT/dV的计算公式如下。

表1 四个电池15个潜在HIs与容量的Spearman相关分析

Table 1 Spearman correlation analysis of 15 potential HIs and capacity of four batteries

特征提取对象	HIs	Spearman相关系数			
		B0005	B0006	B0007	B0018
放电电压	HI1	0.9943	0.9982	0.9953	0.9876
放电电流	HI2	0.9873	0.9986	0.9949	0.9888
放电温度	HI3	0.8118	0.6690	0.8864	0.9688
放电负载电压	HI4	0.9953	0.9986	0.9945	0.9917
放电负载电流	HI5	0.9870	0.9987	0.9945	0.9892
放电数据测量时间	HI6	0.9972	0.9976	0.9951	0.9929
充电电压	HI7	-0.3348	-0.3289	-0.3518	0.9707
充电电流	HI8	0.9908	0.9908	0.9912	0.9734
充电温度	HI9	-0.2286	-0.2668	-0.2123	0.9567
充电负载电压	HI10	-0.3144	-0.3237	-0.5934	0.9634
充电负载电流	HI11	0.9908	0.9907	0.9913	0.9734
充电数据测量时间	HI12	-0.3560	-0.3241	-0.3674	0.9349
dQ/dV	HI13	0.9975	0.9990	0.9959	0.9949
dV/dQ	HI14	-0.5057	-0.7387	-0.3497	0.6355
dT/dV	HI15	0.9833	0.9871	0.9730	0.9929

$$\frac{dQ}{dV} = \frac{I \times dt}{dV} = I \times \frac{dt}{dV} \quad (5)$$

$$\frac{dV}{dQ} = \frac{dV}{I \times dt} = \frac{1}{I} \times \frac{dV}{dt} \quad (6)$$

$$\frac{dT}{dV} = \frac{dT}{dt} / \frac{dV}{dt} \quad (7)$$

其中 $I$ ,  $V$ ,  $t$ 和 $T$ 分别为放电的电流、电压、时间和温度。

以从B0005电池的放电电压中提取HI1为例,使用SVD从电池数据中提取HI的过程为:首先,需要将所有放电周期的电压数据看作矩阵并构建

矩阵集 $\{\mathbf{A}(i)\}_{i=1}^c$ ,其中 $c$ 为总放电周期。其次,将所有循环周期的电压数据即矩阵 $\mathbf{A}(i)$ 依次进行SVD。由于每次SVD的处理对象均为单个周期,所以根据矩阵SVD原理,所有放电周期都可以得到一个代表其特征信息的奇异值 $\sigma_i$ 。最后,从所有放电周期提取出来的奇异值 $\sigma_i$ 则可以构成HI1,表示为 $\{\sigma_i\}_{i=1}^c$ 。其他特征提取对象的SVD与之同理。

### 1.3 特征评估与特征选择

特征融合是HI构建的关键一步,为了得到与容量相关性更强的HI,在融合之前需要对提取出来的潜在HIs进行定量评估和特征选择。考虑到本工作使用的数据不符合正态分布,选择Spearman相关系数来衡量HIs的相关性。15个潜在HIs与容量之间的相关性分析表为表1。根据表1中的相关系数绝对值,选择相关性极强(1~0.9)的HIs进行之后的特征融合。B0005、B0006和B0007选择HI1、HI2、HI4、HI5、HI6、HI8、HI11、HI13和HI15这9个HIs进行融合。B0018选择除了HI14之外的14个HIs进行融合。

### 1.4 基于SAE的特征融合

SAE也叫深度自动编码器,是可以用于降维和特征提取的无监督神经网络。SAE是多个自动编码器(autoencoder, AE)的叠加,它克服了单层编码器不能有效提取复杂特征的缺点。使用SAE对1.3节中选择的相关性极强的HIs进行处理,以降低维数和噪声。SAE的具体结构如图3(a)所示。将两个SAE堆叠成一个栈式自编码器进行二阶参数融合,结构如图3(b)所示。具体的融合过程为:首先让初始参数作为图3(a)SAE结构的输入和输出,进行网络编码与解码,得到输入层到 $h^{(1)}$ 层的网络参数,训练完后留图3(b)所示的输入层至 $h^{(1)}$ 层,隐藏层输出为一阶融合结果。将其结果作为图3(a)所示结构的输入与输出,隐藏层节点数设定为1,使融合结果为一个序列。融合过程与一阶相似,去除解码层后的隐藏层输出即为二阶融合HI。

图中 $\mathbf{x} = [x_1, x_2, \dots, x_n]$ 和 $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n]$ 分别为网络的输入和输出; $w_{ij}$ 和 $w_{jk}$ 为连接权值。隐藏层节点输出 $h_j$ 和输出层节点 $\hat{x}_k$ 分别为

$$h_j = f\left(\sum_{i=1}^m x_i w_{ij} + b_1\right) \quad (8)$$

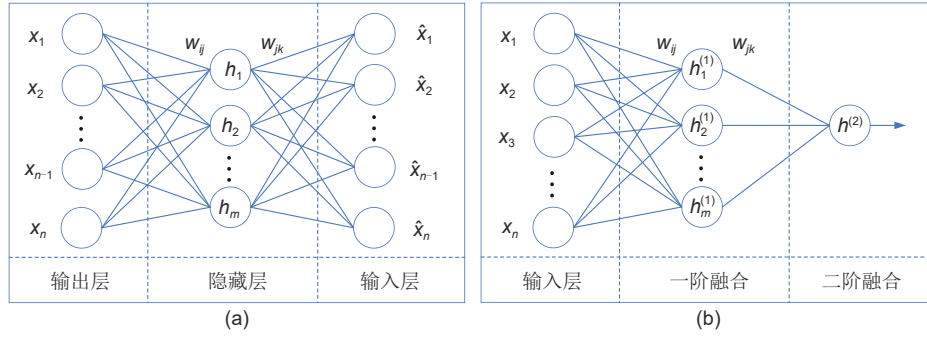


图3 堆栈式SAE: (a)SAE结构; (b)栈式自编码结构

Fig. 3 Stacked SAE: (a)SAE structure; (b) Stack autoencoder structure

$$\hat{x}_k = f\left(\sum_{j=1}^n h_j w_{jk} + b_2\right) \quad (9)$$

其中  $m$  表示融合特征的个数;  $n$  表示输入输出的节点个数;  $f(\cdot)$  表示节点激励函数;  $b_1$  和  $b_2$  均为权值。编码阶段主要是将  $n$  维输入样本  $x_i$  通过编码激励函数映射成  $m$  维的  $h_j$ , 解码阶段是将  $h_j$  重新映射为  $n$  维的  $\hat{x}_k$ 。

为了进一步分析融合 HI 量化电池退化状态的能力, 对融合 HI 进行了定量与定性分析。定量分析结果在表 2 给出, 同时为了检验结果的显著性, 使用  $p$  值法进行了显著性检验。从表中可以看出, 四个电池融合 HI 的 Spearman 相关系数均大于 0.99, 且  $p$  值也远小于 0.01, 表明融合 HI 与容量之间显著相关。定性分析的结果在图 4 中给出。通过分析图 4 可以得知四个电池融合 HI 的曲线与原始容量曲线的退化趋势基本一致。结合定量与定性分析结果, 可以得出结论, 融合 HI 可以很好地表征电池的退化状态。

表2 四个电池融合的HI与容量的Spearman相关分析

Table 2 Spearman correlation analysis of fused HI and capacity of four batteries

Spearman 相关系数	B0005	B0006	B0007	B0018
HI	0.9982	0.9987	0.9984	0.9904
p 值	1.81E-205	6.39E-218	8.07E-209	1.05E-113

## 2 方法论

### 2.1 GPR

GPR 模型是一种非参数模型, 可通过先验知识实现贝叶斯框架下的状态预测, 估计后验分布, 输出预测均值和方差, 并通过 95%(或其他值)的置

信区间来表示预测的不确定性。高斯过程记作如下形式:

$$f(\mathbf{x}) \sim GP[m(\mathbf{x}), k(\mathbf{x}_i, \mathbf{x}_j)] \quad (10)$$

其中  $m(\mathbf{x})$  为均值函数、 $k(\mathbf{x}_i, \mathbf{x}_j)$  为协方差函数。具体如下所示:

$$m(\mathbf{x}) = E[f(\mathbf{x})] \quad (11)$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = E\{[f(\mathbf{x}_i) - m(\mathbf{x}_i)][f(\mathbf{x}_j) - m(\mathbf{x}_j)]\} \quad (12)$$

协方差函数用于捕获不同输入之间的相似性。超参数通常通过边概率的最大似然估计来优化。根据训练数据和测试数据构造先验分布, 利用贝叶斯理论得到预测的后验分布:

$$p(\mathbf{y}^* | \mathbf{x}, \mathbf{y}, \mathbf{x}^*) = N[\mathbf{y}^* | \hat{\mathbf{y}}^*, \sigma^2(\mathbf{y}^*)] \quad (13)$$

其中  $\mathbf{x}$ ,  $\mathbf{y}$  为训练样本的输入和输出,  $\mathbf{x}^*$ ,  $\mathbf{y}^*$  为测试样本的输入和预测输出,  $\hat{\mathbf{y}}^*$  为预测均值,  $\sigma^2(\mathbf{y}^*)$  为预测方差。

通过 95% 的置信区间(CI)用于预测结果不确定度的评估:

$$95\%CI = \hat{\mathbf{y}}^* \pm 1.96 \times \sigma^2(\mathbf{y}^*) \quad (14)$$

### 2.2 基于SVD-SAE-GPR的建模过程

本节将 SVD、SAE 和 GPR 相结合, 提出了一种新的 SVD-SAE-GPR 锂离子电池 RUL 预测方法。具体流程如图 5 所示, 可分为四个步骤。

**步骤 1: SVD 特征提取。** 首先对电池的测量数据进行 SVD, 将各自的奇异值作为 12 个潜在 HIs; 其次根据电流, 时间, 电压和温度等参数, 计算  $dQ/dV$ ,  $dV/dQ$  和  $dT/dV$  矩阵, 并使用 SVD 从中提取 HI13, HI14 和 HI15。

**步骤 2: 特征选择和特征融合。** 通过 Spearman 相关分析从步骤 1 中的 15 潜在 HIs 中选择相关性极强的 HIs 进行特征融合。通过 SAE 处理

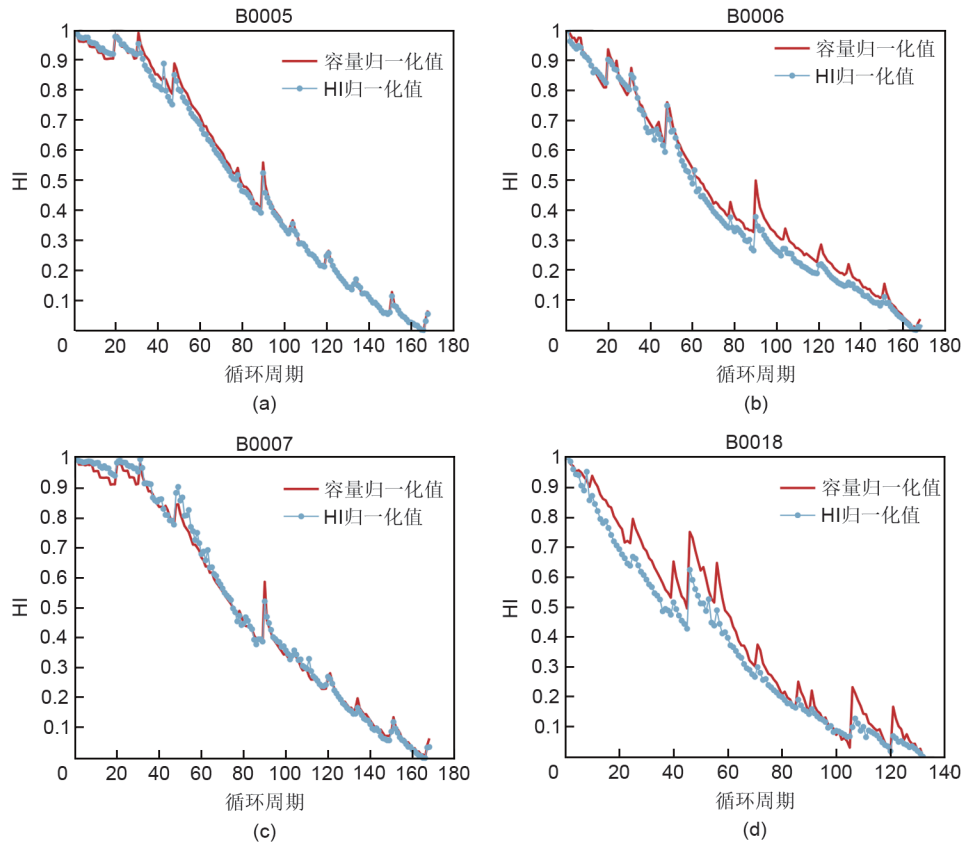
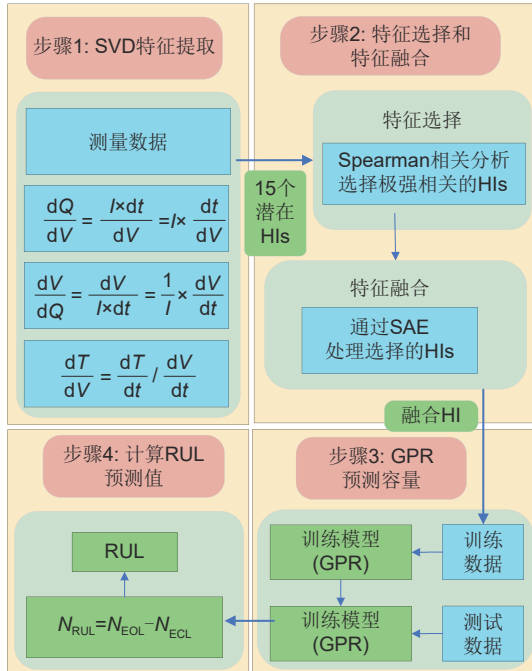


图4 融合HI定性分析: (a)B0005; (b)B0006; (c)B0007; (d)B0018

Fig. 4 Fusion HI qualitative analysis: (a)B0005; (b)B0006; (c)B0007; (d)B0018

图5 基于SVD-SAE-GPR的锂离子电池RUL预测流程图  
Fig. 5 Flow chart of RUL prediction of Li-ion battery based on SVD-SAE-GPR

选择的HIs融合成一个HI。

步骤3: GPR 预测容量。以这融合特征  $\{HI(i)\}_{i=1}^c$  作为模型的输入, 以电池容量  $\{C(i)\}_{i=1}^c$  作为输出, 构建基于GPR的容量预测模型,  $c$  为总循环周期。训练集记为  $\{\mathbf{x}, \mathbf{y}\}$ , 其中  $\mathbf{x} = [HI(1 : k)^T]$ ,  $\mathbf{y} = [C(1 : k)^T]$ , 预测起始点对应的循环数记为  $k+1$ 。测试集表示为  $\{\mathbf{x}^*, \mathbf{y}^*\}$ , 其中  $\mathbf{x}^* = [HI(k+1 : c)^T]$ ,  $\mathbf{y}^* = [C(k+1 : c)^T]$ 。预测部分以  $\mathbf{x}^*$  为输入, 得到容量预测值  $\hat{\mathbf{y}}^*$ 。

步骤4: 计算RUL预测值。剩余寿命 ( $N_{RUL}$ ) 为电池实际容量下降到阈值时的总充放电循环次数 ( $N_{EOL}$ ) 与电池当前充放电循环次数 ( $N_{ECL}$ ) 之差。计算公式如式(15)。

$$N_{RUL} = N_{EOL} - N_{ECL} \quad (15)$$

### 2.3 预测性能评价指标

为了评价本文所提方法的预测性能, 分别采用均方根误差 (root mean square error, RMSE)、决定系数 ( $R^2$ )、绝对误差 (absolute error, AE) 三种评价指标。

## 3 结果与讨论

### 3.1 锂离子电池RUL预测结果

为了分析该模型的RUL预测性能, 设计了不同预测起始点的实验。实验设置: 选取前59个周期和前79个周期的数据作为训练集, 即预测起点分别为60和80。电池的故障阈值为额定容量的70%~80%。本工作中, 电池B0005、B0006和B0018的容量故障阈值为1.38 Ah。由于B0007电池未达到故障阈值, 因此未给出故障阈值线。

图6分别给出了四个电池在预测起始点60和80处的预测结果及其不确定性表达。从整体上可以看出, 预测曲线与真实的容量退化曲线较为接近, 预测结果受预测起始点的影响较小, 且四个电池在预测起始点80时的预测曲线比在预测起始点60时更接近真实容量退化曲线。

此外, 表3给出了四个电池的不同预测起始点的预测性能评价结果, 从表中可以看出, 预测起始点为80的预测误差均比预测起始点为60的预测误差小, 且四个电池的RMSE均小于0.0301, 说明所提模型的预测性能较好; 四个电池的 $R^2$ 均大于0.93, 说明预测曲线与实际曲线的拟合程度较高; 随着预测起始点的后移, B0005、B0006和B0018三个电池的AE值逐渐减小。在预测起始点为60时, 三个电池的AE值均小于6, 误差值较小; 而且随着预测起始点的后退, 在起始点为80时, AE值逐渐减小, 均不超过2。

### 3.2 SVD对不同电池数据集的适应性

为了验证SVD特征提取方法对不同锂离子电池数据集的适应性, 本小节的实验额外选择了麻省理工学院<sup>[23]</sup>(Massachusetts Institute of Technology, MIT)提供的电池数据集, 其中电池参数和实验环境与NASA提供的电池数据集不同。首先详细描述了用于测试的电池数据集。然后, 仍然使用SVD提取HIs。最后, 利用GPR进行容量预测。

#### 3.2.1 MIT数据集

测试电池是由A123 Systems公司(APR18650 M1A)制造的磷酸锂离子(LFP)/石墨电池<sup>[24-25]</sup>, 放置在30℃的强制对流温度室中, 在充电阶段以不同倍率的恒流快充组合形成72种充电模式, 放电过程统一采用4C倍率。电池标称容量1.1 Ah, 标称电压3.3 V。选取“2017-06-30”第二批数据中的

两组数据Channel 2 (CH2)和Channel 5 (CH5)进行实验。CH2和CH5的原始容量退化轨迹如图7所示。

#### 3.2.2 特征提取、选择与融合

对于CH2和CH5, 使用相同的SVD分别从测量数据、 $dQ/dV$ 、 $dV/dQ$ 和 $dT/dV$ 中提取各自的奇异值作为HIs。 $dQ/dV$ 在数据集中已经给出,  $dV/dQ$ 和 $dT/dV$ 需要计算才能得到。在计算 $dT/dV$ 时, 由于同一周期内某些相邻时间点采集的相同电压值会导致 $dT/dV$ 的值无效, 因此将重复的电压数据和其相应测量时间对应的温度数据去除。对提取出来的潜在HIs进行Spearman相关分析, 分析结果如表4所示。可以看出除了HI8之外, 其余7个潜在HIs与容量之间的相关系数绝对值均大于0.9, 属于极强相关。因此, 根据相关系数选择除HI8之外的7个HIs使用SAE融合成一个HI。为了分析融合HI表征电池退化状态的能力, 计算了它的Spearman相关系数和 $p$ 值, 具体结果在表5中给出。从表5中可以看出, 融合HI与容量之间的Spearman相关系数均大于0.99,  $p$ 值接近于0, 表明融合HI与容量显著相关。

#### 3.2.3 容量预测

预测部分选用GPR进行建模。以融合HI作为训练数据, 以容量作为输出数据, 构建基于GPR的容量预测模型。根据CH2和CH5的数据大小, 分别选择前500和100个周期的数据进行训练。预测结果如图8所示, 可以发现预测曲线能够相对准确地接近真实容量退化曲线。预测性能评价结果如表6所示。从表中可以看出, CH2和CH5的RMSE均小于0.008,  $R^2$ 均大于0.94, 说明本工作模型具有较好的预测性能, 预测曲线与实际曲线吻合较好。

综上所述, 通过在NASA和MIT提供的电池数据集实验, 可以得出SVD可以用于提取HIs, 并且在避免人为设置参数的情况下, 对不同电池数据集具有良好的适应性。此外, 融合的HI具有较强的相关性和较好的预测性能和准确性。

### 3.3 不同HI的对比试验分析

为了更充分地说明利用SVD方法从测量数据、 $dQ/dV$ 、 $dV/dQ$ 和 $dT/dV$ 中提取潜在HIs, 选择极强相关的HIs使用SAE进行特征融合并结合GPR预测模型(命名为M1)的有效性和优越性, 设计了比较模型M2和M3。其中模型M2通过PCA融合模型



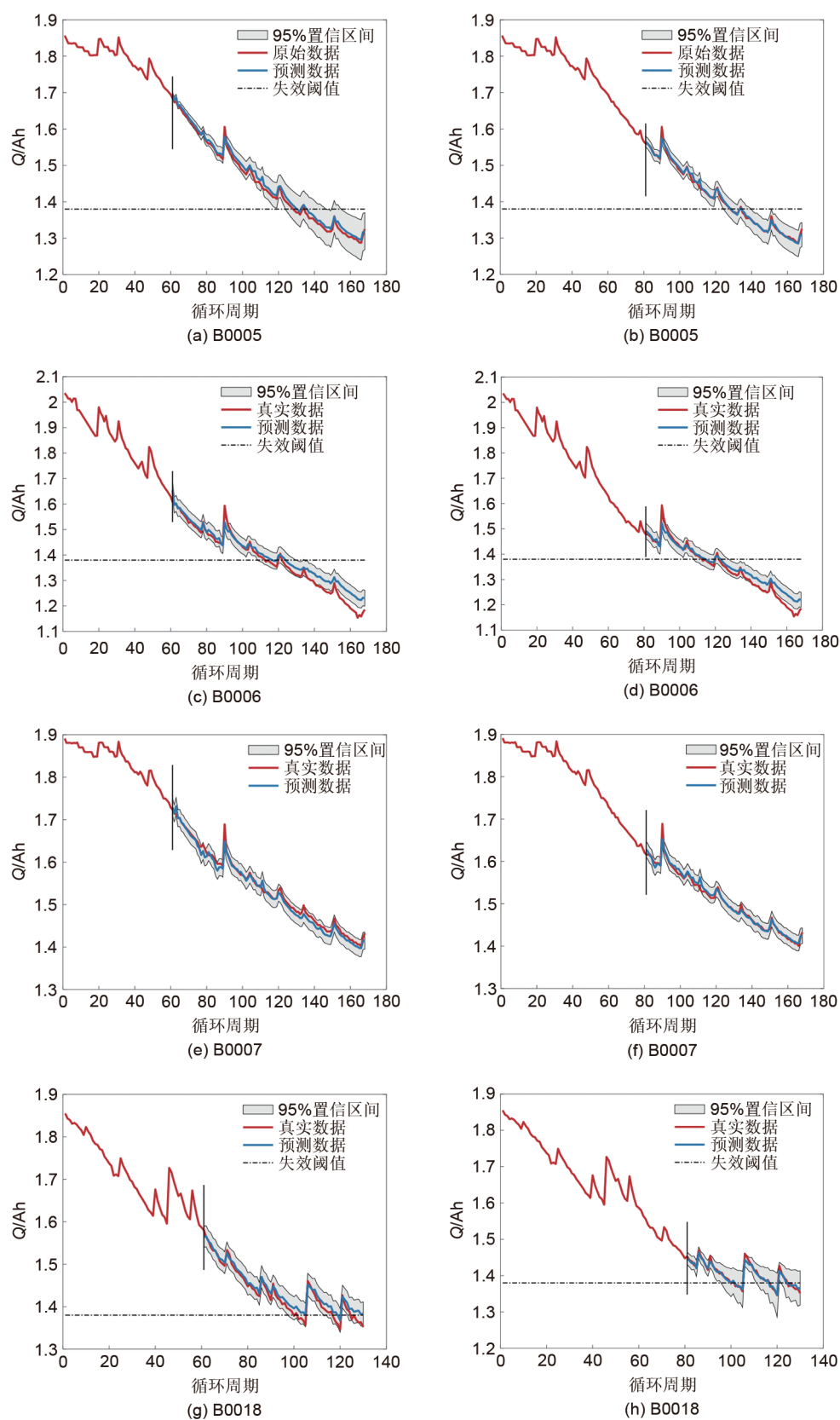


图6 四个电池在不同预测起点的预测结果及不确定性表达

Fig. 6 Prediction results and uncertainty expression of four batteries at different prediction starting points



表 3 四个电池在不同预测起始点的预测性能

Table 3 Prediction performance of four batteries at different prediction starting points

评价指标	预测起始点	B0005	B0006	B0007	B0018
RMSE	60	0.0103	0.0301	0.0198	0.0128
	80	0.0062	0.0267	0.0070	0.0066
$R^2$	60	0.9918	0.9369	0.9501	0.9519
	80	0.9946	0.9400	0.9893	0.9572
AE	60	2	4	—	6
	80	0	2	—	1

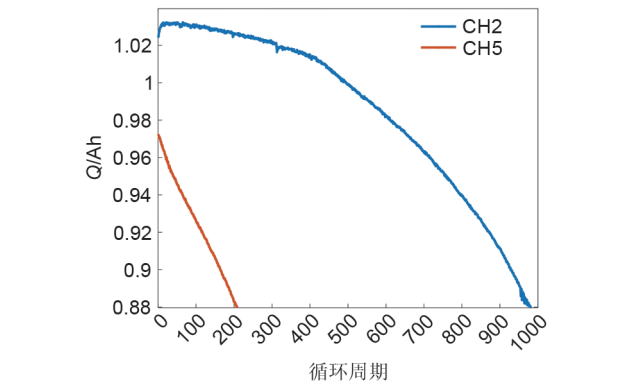


图 7 CH2 和 CH5 的容量退化曲线

Fig. 7 Capacity degradation curves of CH2 and CH5

表 4 CH2 和 CH5 容量与潜在 HIs 的 Spearman 相关分析

Table 4 Spearman correlation analysis between capacity and potential HIs of CH2 and CH5

特征提取对象	HIs	Spearman 相关系数	
		CH2	CH5
时间	HI1	0.9353	0.9868
充电容量	HI2	0.9984	0.9986
电流	HI3	0.9983	0.9978
电压	HI4	0.9405	0.9835
温度	HI5	0.9980	0.9880
dQ/dV	HI6	0.9984	0.9998
dV/dQ	HI7	-0.9973	-0.9922
dT/dV	HI8	-0.2448	0.1870

表 5 CH2 和 CH5 容量与融合 HI 的 Spearman 相关分析

Table 5 Spearman correlation analysis of capacity and fused HI of CH2 and CH5

Spearman 相关系数	CH2	CH5
HI	0.9983	0.9996
p 值	<<0.001	<<0.001

M1 中选择的极强相关的 HIs。模型 M3 使用的 HI 为放电阶段电压从 3.8 V 降到 3.5 V 的等压降放电时

间。在预测阶段，模型 M1、M2 和 M3 均采用 GPR，具体模型描述如表 7 所示。

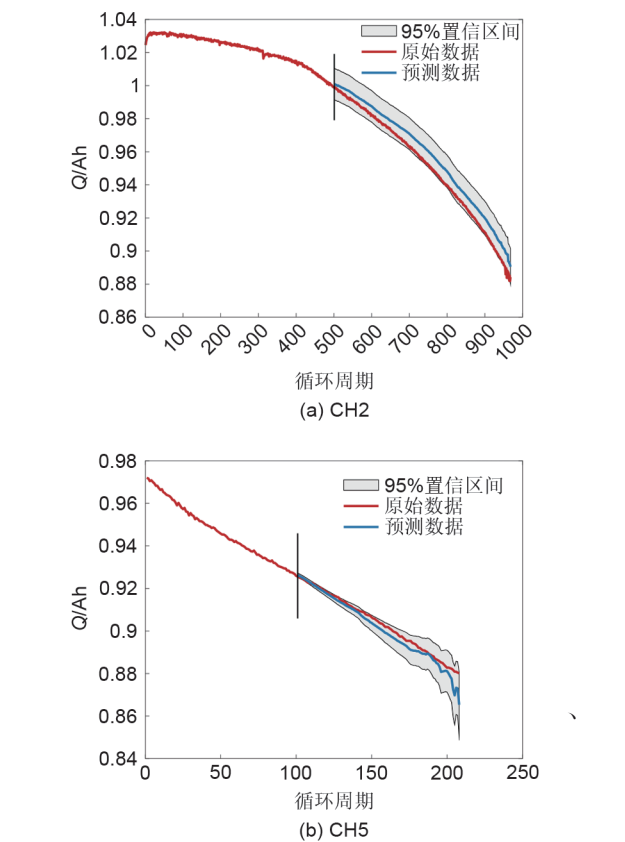


图 8 CH2 和 CH5 的容量预测结果及不确定性表达

Fig. 8 Capacity prediction results and uncertainty expression of CH2 and CH5

表 6 CH2 和 CH5 的预测性能

Table 6 Predicted performance of CH2 and CH5

评价指标	CH2	CH5
RMSE	0.0072	0.0031
$R^2$	0.9505	0.9471

表 7 对比实验中使用的预测模型(M1、M2 和 M3)

Table 7 Prediction models used in comparison experiments (M1, M2 and M3)

模型	模型描述
M1	采用 SVD 方法提取测量数据、dQ/dV、dV/dQ 和 dT/dV 的奇异值作为潜在 HIs，选取相关性极强的 HIs 使用 SAE 融合为一个 HI，结合 GPR 模型进行预测
M2	对于模型 M1 中提取的潜在 HIs，选取相关性极强的 HIs 使用 PCA 融合为一个 HI，结合 GPR 模型进行预测
M3	将放电阶段电压从 3.8 V 降到 3.5 V 的等压降放电时间作为 HI，结合 GPR 模型进行预测

模型 M1、M2 和 M3 对 B0005、B0006、B0007 和 B0018 电池的预测性能如表 8 所示。可以看出,对于模型 M1,电池 B0005、B0006、B0007 和 B0018 的 RMSE 分别为 0.0062、0.0267、0.0070 和 0.0066,四个电池的 RMSE 均小于模型 M2 和 M3,且模型 M2 的 RMSE 也均小于模型 M3。在模型 M1 下,除 B0007 外的三个电池的 AE 值分别为 0、2 和 1,也均小于模型 M2 和模型 M3 的 AE 值。而且,模型 M2 除了 B0018 的 AE 值比模型 M3 大 1 之外,模型 M2 也比模型 M3 具有优势。从 RMSE 和 AE 可以看出,模型 M1 比模型 M2、M3 具有更好的预测精度,且本文提取的 HIs 经过 PCA 融合后的预测性能也优于其他 HI。模型 M1 的四个电池的  $R^2$  分别为 0.9946、0.9400、0.9893 和 0.9572,明显大于模型 M2 和 M3 电池的  $R^2$ ,这充分说明模型 M1 的预测结果比模型 M2、M3 能更好拟合原始退化曲线。此外,除了 B0018 之外,模型 M2 的其余三个电池的  $R^2$  也均明显大于模型 M3,表明经过 PCA 融合后的 HI 的预测结果与原始曲线的拟合程度优于模型 M3 的 HI。对比实验结果表明,同样的 HIs 经过 SAE 融合之后比经过 PCA 融合的 HI 预测性能更好,说明 SAE 用于特征融合的效果优于 PCA。此外,本工作提取的潜在 HIs 经过融合之后的 HI 预测性能优于其他 HI,说明了本工作提取的 HIs 的有效性与优越性。

表 8 不同模型(M1、M2 和 M3)的预测性能  
Table 8 Predicted performance of different models (M1, M2 and M3)

评价指标	模型	B0005	B0006	B0007	B0018
RMSE	M1	0.0062	0.0267	0.0070	0.0066
	M2	0.0176	0.0325	0.0113	0.0195
	M3	0.0270	0.0395	0.0138	0.0075
$R^2$	M1	0.9946	0.9400	0.9893	0.9572
	M2	0.9567	0.8959	0.9719	0.9314
	M3	0.8984	0.8469	0.9583	0.9457
AE	M1	0	2	—	1
	M2	2	3	—	3
	M3	10	4	—	2

## 4 结 论

本文提出了一种基于 SVD-SAE-GPR 的锂离子电池 RUL 预测方法。主要贡献如下。首先,为了

解决传统特征提取方法依赖于参数设定且对于不同锂离子电池数据集适应性差的缺点,将 SVD 用于特征提取解决上述问题。然后,从原始测量数据和计算出的包含更多退化信息的特征提取对象中提取了 15 个潜在 HIs。考虑到 HIs 的冗余和不足会影响 RUL 预测,将 Spearman 相关分析和 SAE 用于处理潜在 HIs 得到一个融合 HI。最后,将融合 HI 与 GPR 相结合进行 RUL 预测,并给出结果的不确定性表达。在验证阶段,在 NASA 和 MIT 提供的电池数据集上的实验结果说明了该特征提取策略的有效性和适应性。通过与 PCA 融合 HI 和其他 HI 进行比较,验证了所提 HI 的优越性。综上所述,SVD 特征提取方法具有良好的适应性,提取出的潜在 HIs 使用 Spearman 相关分析和 SAE 得到的融合 HI 具有较好的预测性能,所提出的 RUL 预测框架具有较高的精度且输出概率等优点。本文的方法也具有一定的局限性。比如,锂离子电池的实验室数据根据实验条件和测量仪器的不同,测量出来的数据可能会存在差异,而这些差异是否影响 SVD 特征提取,是否需要数据做进一步的处理,是需要通过具体的测量数据进一步研究的。此外,在电池出厂之后,无法获得实验室数据的情况和在使用阶段比较复杂的工况下,测量数据可能有一定的难度,因此该方法适用性需进一步探究。

## 参 考 文 献

[1] PARGOLETTI E, ARNABOLDI S, CAPPELLETTI G, et al. Smart interfaces in Li-ion batteries: Near-future key challenges[J]. *Electrochimica Acta*, 2022, 415: doi: 10.1016/j.electacta.2022.140258.

[2] 梁新成, 张勉, 黄国钧. 基于 BMS 的锂离子电池建模方法综述[J]. *储能科学与技术*, 2020, 9(6): 1933-1939.

LIANG X C, ZHANG M, HUANG G J. Review on lithium-ion battery modeling methods based on BMS[J]. *Energy Storage Science and Technology*, 2020, 9(6): 1933-1939.

[3] HU X S, ZHANG K, LIU K L, et al. Advanced fault diagnosis for lithium-ion battery systems: A review of fault mechanisms, fault features, and diagnosis procedures[J]. *IEEE Industrial Electronics Magazine*, 2020, 14(3): 65-91.

[4] 陈翌, 白云飞, 何瑛. 数据驱动的锂电池健康状态估算方法比较[J]. *储能科学与技术*, 2019, 8(6): 1204-1210.

CHEN Y, BAI Y F, HE Y. Comparison of data-driven lithium battery state of health estimation methods[J]. *Energy Storage Science and Technology*, 2019, 8(6): 1204-1210.

[5] PANG X Q, ZHAO Z, WEN J, et al. An interval prediction approach based on fuzzy information granulation and linguistic description for remaining useful life of lithium-ion batteries[J]. *Journal of Power*

- Sources, 2022, 542: doi: 10.1016/j.jpowsour.2022.231750.
- [6] CHU A, ALLAM A, CORDOBA ARENAS A, et al. Stochastic capacity loss and remaining useful life models for lithium-ion batteries in plug-in hybrid electric vehicles[J]. Journal of Power Sources, 2020, 478: doi: 10.1016/j.jpowsour.2020.228991.
- [7] ZHOU W L, LU Q, ZHENG Y P. Review on the selection of health indicator for lithium ion batteries[J]. Machines, 2022, 10(7): 512.
- [8] KHALEGHI S, HOSEN M S, KARIMI D, et al. Developing an online data-driven approach for prognostics and health management of lithium-ion batteries[J]. Applied Energy, 2022, 308: doi: 10.1016/j.apenergy.2021.118348.
- [9] JIA J F, YUAN S F, SHI Y H, et al. Improved sparrow search algorithm optimization deep extreme learning machine for lithium-ion battery state-of-health prediction[J]. iScience, 2022, 25(4): doi: 10.1016/j.isci.2022.103988.
- [10] 李旭东, 张向文. 基于主成分分析与WOA-Elman的锂离子电池SOH估计[J]. 储能科学与技术, 2022, 11(12): 4010-4021.
- LI X D, ZHANG X W. State of health estimation method for lithium-ion batteries based on principal component analysis and whale optimization algorithm-Elman model[J]. Energy Storage Science and Technology, 2022, 11(12): 4010-4021.
- [11] LIU J, CHEN Z Q. Remaining useful life prediction of lithium-ion batteries based on health indicator and Gaussian process regression model[J]. IEEE Access, 2019, 7: 39474-39484.
- [12] 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: doi: 10.1016/j.microrel.2021.114405.
- [13] PAN W J, LUO X S, ZHU M T, et al. A health indicator extraction and optimization for capacity estimation of Li-ion battery using incremental capacity curves[J]. Journal of Energy Storage, 2021, 42: doi: 10.1016/j.est.2021.103072.
- [14] WALL M E, RECHTSTEINER A, ROCHA L M. Singular value decomposition and principal component analysis[M]//A Practical Approach to Microarray Data Analysis. Boston: Kluwer Academic Publishers, 2005: 91-109.
- [15] GU X Y, SEE K W, LI P H, et al. A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model[J]. Energy, 2023, 262: doi: 10.1016/j.energy.2022.125501.
- [16] WEI M, YE M, WANG Q, et al. Remaining useful life prediction of lithium-ion batteries based on stacked autoencoder and Gaussian mixture regression[J]. Journal of Energy Storage, 2022, 47: doi: 10.1016/j.est.2021.103558.
- [17] 王瑞洁, 惠周利, 杨明. 基于间接健康指标的高斯过程回归对锂电池SOH预测[J]. 储能科学与技术, 2023, 12(2): 560-569.
- WANG R J, HUI Z L, YANG M. Gaussian process regression based on indirect health indicators for SOH estimation of lithium battery[J]. Energy Storage Science and Technology, 2023, 12(2): 560-569.
- [18] SAHA B, GOEBEL K. Battery data set[R/OL]. [2020-10-20]. <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>.
- [19] TANG T, YUAN H M. The capacity prediction of Li-ion batteries based on a new feature extraction technique and an improved extreme learning machine algorithm[J]. Journal of Power Sources, 2021, 514: doi: 10.1016/j.jpowsour.2021.230572.
- [20] FLY A, CHEN R. Rate dependency of incremental capacity analysis (dQ/dV) as a diagnostic tool for lithium-ion batteries[J]. Journal of Energy Storage, 2020, 29:doi:10.1016/j.est.2020.101329.
- [21] BLOOM I, CHRISTOPHERSEN J P, ABRAHAM D P, et al. Differential voltage analyses of high-power lithium-ion cells[J]. Journal of Power Sources, 2006, 157(1): 537-542.
- [22] SHIBAGAKI T, MERLA Y, OFFER G J. Tracking degradation in lithium iron phosphate batteries using differential thermal voltammetry[J]. Journal of Power Sources, 2018, 374: 188-195.
- [23] SEVERSON K A, ATTIA P M, JIN N, et al. Data-driven prediction of battery cycle life before capacity degradation[J]. Nature Energy, 2019, 4(5): 383-391.
- [24] CLERICI D, MOCERA F, SOMÀ A. Electrochemical-mechanical multi-scale model and validation with thickness change measurements in prismatic lithium-ion batteries[J]. Journal of Power Sources, 2022, 542: doi: 10.1016/j.jpowsour.2022.231735.
- [25] ATTIA P M, GROVER A, JIN N, et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning[J]. Nature, 2020, 578(7795): 397-402.