

# 基于稀疏高斯过程回归的锂电池 SOH 预测

周润玺 王奕翔 杨涛

陕西省长安区西京学院 陕西西安 710123

**摘要:** 本文基于弛豫电压提出基于融合 HIs 与 SGPR 的锂电池 SOH 预测方法。在该算法中, 采用了单个核函数的 SGPR 模型作为预测锂电池 SOH 模型, 并使用 NCA 数据集对模型进行了验证。在实验过程中, 以融合 HIs 作为 SGPR 模型的输入, 分别选择 SE、Matern52 两种不同类型的核函数预测锂电池的 SOH, 且讨论了不同核函数对预测的影响。实验结果表明, 两种核函数下, 所建立的模型在预测四种锂电池的 SOH 时, 均能将误差控制在 3% 之内。

**关键词:** 弛豫电压; 数据驱动; 稀疏高斯过程回归; 健康状态预测

## 1. 引言

随着环境污染问题日益加剧以及化石能源日渐枯竭, 全球范围内的能源转型迫在眉睫, 尤其是在交通运输领域, 新能源的应用已成为不可忽视的趋势<sup>[1]</sup>。电动汽车 (EV) 作为新能源交通工具的代表, 凭借其零排放、低能耗的特点, 近年来得到了广泛推广。作为电动汽车的核心, 锂离子电池凭借其高能量密度、较长的使用寿命以及相对环保的特性, 成为了主流的能量存储设备。然而, 锂离子电池在长期使用过程中性能不可避免地会出现退化, 表现为可用容量和输出功率的减少<sup>[2]</sup>。电池健康状态 (State of Health, SOH) 是衡量锂电池退化程度的重要指标, 直接关系到电池的工作效率和寿命。SOH 的准确评估不仅对电池管理系统 (Battery Management System, BMS) 的优化设计和参数校准至关重要, 还为电动汽车的安全性和可靠性提供了保障<sup>[3]</sup>。

由于锂离子电池内部的电化学反应复杂, 且外部环境多变, 开发准确、高效的 SOH 估计方法面临着诸多挑战<sup>[4]</sup>。相较而言, 数据驱动方法不依赖于电池内部机制, 更多地依托历史数据进行建模与预测, 因而受到研究人员的青睐。

在 GPR 模型中, 核函数的选择是影响模型性能的关键因素。然而, 对于电池 SOH 估计中最优核函数的选择, 仍然存在争议<sup>[5]</sup>。同时, 随着数据量的增加, 计算复杂度的上升成为 GPR 应用的一个限制因素。因此, 针对不同核函数的选择及其对预测性能的影响, 开展深入研究具有重要的现实意义。

## 2. 豫电压分析与健康指标提取

### 2.1 健康指标的提取

SOH 通常由电池的容量或电阻定义。在本研究中, 电池容量被用来定义 SOH, SOH 表示当前容量与额定容量的比值<sup>[6]</sup>, 即:

$$SOH = \frac{C_{Now}}{C_{Rated}} \times 100\%$$

式中:  $C_{Now}$  表示当前电池容量,  $C_{Rate}$  表示电池的额定容量。如果容量下降到初始额定容量的 71%, 则应更换电池。根据这一定义, 估计 SOH 相当于估计容量。

### 2.2 基于弛豫电压的 HIs 提取

在实际的电动交通应用中, 电池充电是确保车辆持续运作的重要环节。与充电过程密切相关的电压特性成为了研究人员关注的重点。根据最新的研究<sup>[7]</sup>, 基于电压的特征提取技术主要可分为三大类别: 基于恒定电流 (Constant Current, CC) 充电电压的提取, 基于恒定电流 - 恒定电压 (Constant Current-Constant Voltage, CC-CV) 充电电压的提取, 以及基于静止电压的提取。相比于这些基于特定充电阶段的电压特征提取方法, 完全充电后的弛豫电压因其独特优势, 近年来受到了越来越多的关注。图 1 中示出了 NCA 电池使用 0.5C 充电速率的一个完整循环曲线, 其包括五个过程, 即, CC 充电、CV 充电、充电后弛豫、CC 放电、放电后弛豫。

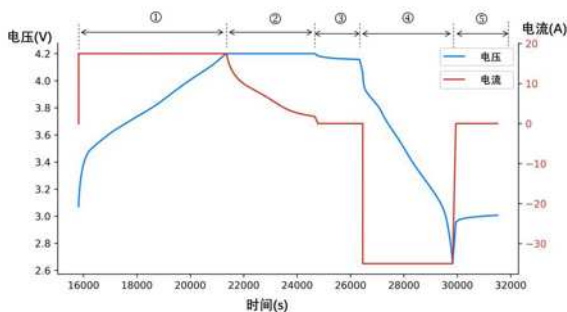


图 1 NCA 电池第一个循环的电压和电流曲线

### 2.3 特征工程构建

为了有效利用弛豫电压时序数据进行电池健康状态 (SOH) 的预测, 本文构建了 6 个特征, 包括最大值、最小值、标准差、均方根、形状因子和方差。这些特征能够捕捉电压信号的不同统计属性, 从而为后续的 SOH 估计提供有力支持。具体如表 1 所示:

表 1 基于时域统计特征

名字	公式	名字	公式
最大值	$MAX = \max(x_1, x_2, \dots, x_n)$	最小值	$MIN = \min(x_1, x_2, \dots, x_n)$
标准差	$STD = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$	均方根	$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$
形状因子	$SF = \frac{PVD}{RMS}$	方差	$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$

## 3. 方法

### 3.1 高斯过程回归

高斯过程回归 (Gaussian Process Regression, GPR) 是一种基于统计学和贝叶斯先验知识的非参数模型<sup>[8]</sup>。本文提出了稀疏高斯过程回归 (Sparse Gaussian Process Regression, SGPR) 方法, 以降低时间复杂度<sup>[9]</sup>。SGPR 通过引入诱导点来减少计算需求。这些诱导点通过等间隔采样的方法进行选择, 本文将诱导点的采样间隔设置为 10。诱导点用于生成目标函数  $f(x)$  的联合分布, 从而降低模型复杂度。SGPR 中的诱导点是基于等间隔采样选取的, 本文将间隔设置为 10, 诱导点用于生成  $f$  的联合分布。

$$p(\mathbf{f}, \mathbf{f}^*) = \int p(\mathbf{f}, \mathbf{f}^* | \mathbf{u}) p(\mathbf{u}) d\mathbf{u}$$

其中  $\mathbf{u}$  是诱导点。

### 3.2 核函数

在高斯过程回归 (GPR) 中, 核函数是其核心组成部分,

决定了模型如何衡量数据点之间的相似性。不同的核函数适用于不同类型的数据分布和模式, 因此选择合适的核函数对于 GPR 的性能至关重要。常用的核函数包括平方指数核 (SE Kernel) 和 Matern 核函数<sup>[10]</sup>。本文在模型中选用了 SE 核函数和 Matern 核函数, 以探讨不同核函数对锂电池健康状态 (SOH) 预测的影响。

#### SE 核函数

平方指数核 (Squared Exponential Kernel, SE Kernel) 是 GPR 中最常用的一种核函数, 因其具有较强的拟合能力, 适用于各种非线性数据。SE 核函数假设输入点之间的相似性随着它们之间的距离增大而呈指数衰减, 因此能够平滑地描述复杂的函数关系。

$$k_{SE}(r) = \sigma_f^2 \exp\left(-\frac{r}{2l^2}\right).$$

其中  $l$  为特征长度尺度, 控制了函数整体的平滑程度,

$\sigma_f^2$  为信号方差。

#### Matern 核函数

Matern 核函数是高斯过程回归 (GPR) 中另一种常用的核函数, 它在衡量数据点相似性时比平方指数 (SE) 核函数更具灵活性和适应性。与 SE 核函数 (或 RBF 核函数) 相比, Matern 核函数可以更好地处理不同平滑度和复杂度的函数关系, 因此在一些更复杂或粗糙的实际应用场景中表现出色。其公式表示为:

$$k_{Matern} = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \text{Bigg}\left(\frac{\sqrt{2\nu r}}{l} \text{Bigg}\right)^\nu K_\nu \text{Bigg}\left(\frac{\sqrt{2\nu r}}{l} \text{Bigg}\right)$$

其中  $\nu, l > 0$ ,  $K_\nu$  是一个修正的贝塞尔函数。

Matern 核函数通过引入额外的参数  $\nu$  来控制结果函数的平滑程度, 这使得它更加灵活, 并且能够更好地适应真实的底层函数关联属性。较为常用的  $\nu=5/2$ 。

$$k_{\nu=5/2}(r) = \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{l} + \frac{5r^2}{3l^2}\right) \exp\left(-\frac{\sqrt{5}r}{l}\right).$$

## 4. 结果

### 4.1 评价标准

为了全面评价所构建模型的性能, 本研究采用了三种公认的评价指标: 均方根误差 (RMSE)、平均绝对误差 (MAE) 以及决定系数 ( $R^2$ )。这三者分别从不同角度对模型预测精度进行了量化分析。

## 4.2 锂电池 SOH 预测结果分析

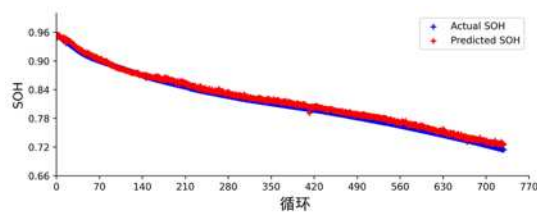
表 2 基于弛豫电压的预测结果对比

电池编号	方法	RMSE	MAE	R <sup>2</sup>
NCA_13	SE	0.0060	0.0045	0.9920
	Matern52	0.0055	0.0040	0.9930
NCA_14	SE	0.0065	0.0050	0.9915
	Matern52	0.0058	0.0048	0.9925
NCA_15	SE	0.0058	0.0043	0.9918
	Matern52	0.0053	0.0038	0.9932
NCA_16	SE	0.0070	0.0055	0.9905
	Matern52	0.0062	0.0050	0.9935

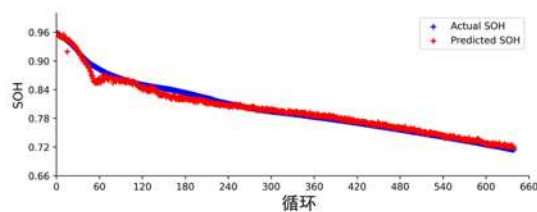
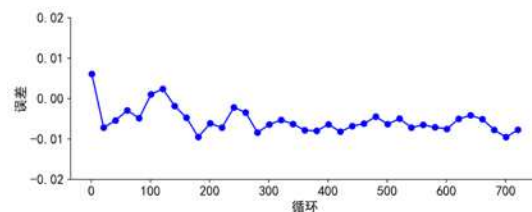
在本研究中，我们对四个锂电池样本（NCA\_13、NCA\_14、NCA\_15、NCA\_16）应用了稀疏高斯过程回归

（SGPR）方法，分别使用平方指数核（SE）和 Matern52 核函数进行锂电池健康状态（SOH）的预测。如表 2 所示。

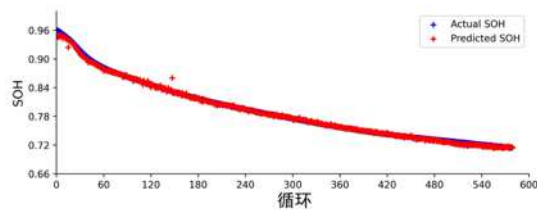
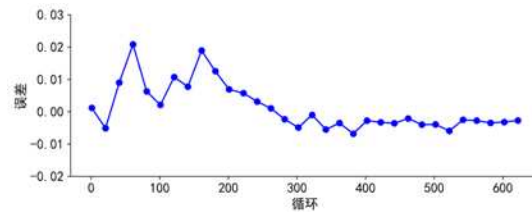
从结果上来看，NCA\_13 电池样本的 RMSE 和 MAE 分别为 0.0060 和 0.0045，使用 Matern52 核函数时，这些指标显著降低至 0.0055 和 0.0040，表明 Matern52 核函数在此样本上提供了更为精准的 SOH 预测。对于 NCA\_14，SE 核函数的 RMSE 和 MAE 分别为 0.0065 和 0.0050，而 Matern52 核函数则降低至 0.0058 和 0.0048，显示出 Matern52 核函数的优势。如图 2 所示。



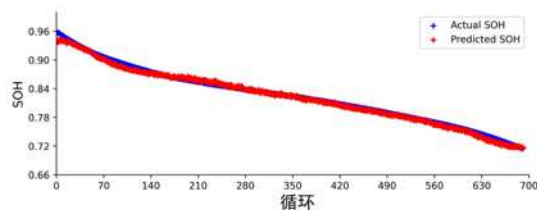
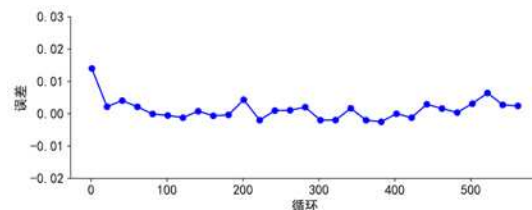
(a)NCA\_13 下健康状态的估计值和真实值之间的误差



(b)NCA\_14 下健康状态的估计值和真实值之间的误差



(c)NCA\_15 下健康状态的估计值和真实值之间的误差



(d)NCA\_16 下健康状态的估计值和真实值之间的误差

图 2 基于 SGPR 的锂离子电池预测

## 5. 结论

本文提出了一项创新性的电池健康状态（SOH）预测方法，即稀疏高斯过程回归（SGPR）模型。该模型以电池的弛豫电压作为输入，实现了对电池 SOH 的高精度估计。为

深入探究不同核函数对模型性能的影响，我们选择了平方指数核（SE）和 Matern52 核函数，并在四个锂电池样本上进行了系统的实验。通过对电池老化过程中收集的弛豫电压数据进行有效处理，我们能够从中提炼出重要的特征信息，

而无需依赖于先前循环的历史数据。我们的实验结果显示, Matern52 核函数在各项评估指标上均优于 SE 核函数, 展现出更好的预测精度和可靠性, 均方根误差 (RMSE) 和平均绝对误差 (MAE) 保持在较低水平, 误差均保持在 3% 之内。实验结果充分表明, 利用 SGPR 模型结合弛豫电压特征, 能够有效提高电池 SOH 的预测能力, 为电池管理系统提供更为精确的决策支持。本研究不仅为锂电池 SOH 的精准估计提供了新的思路, 还为未来电池健康管理领域的研究和应用奠定了坚实的基础。

#### 参考文献:

- [1] 李欣, 刘志强, 魏秀琨. 基于多核高斯过程回归的地铁车辆继电器寿命预测 [J]. 中国铁路, 2024,(10):114-121.
- [2] 周宣昊, 张小龙. 基于多尺度递归定量分析的电池健康状态评估 [J]. 蓄电池, 2024,61(05):201-211.
- [3] 惠周利, 王瑞洁, 冯娜娜, 杨明. 基于集成高斯过程回归的锂离子电池健康状态预测 (英文) [J]. Journal of Measurement Science and Instrumentation, 2024,15(03):397-407.
- [4] 汤烨, 陈庆盈, 周耀华, 李研彪. 基于刚度模型和高斯过程回归模型的重载工业机器人分步标定方法 [J]. 高技术通讯, 2024,34(08):885-894.
- [5] 董业成. 基于神经网络的锂离子电池荷电状态与健康状态预测研究 [D]. 南昌大学, 2024.
- [6] 夏玉杭, 李旭, 王建春. 基于 SFFS 特征选择的锂离子电池健康状态预测研究 [J]. 中阿科技论坛 (中英文), 2024,(05):93-97.
- [7] Jiangong Zhu, Yixiu Wang, Yuan Huang, et al. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. Nature communications, 2022. 13(1): 226 - 234.
- [8] 徐厚宝, 杨承莲, 张永康. 卡尔曼滤波优化的高斯过程回归模型 [J]. 北京理工大学学报, 2024,44(05):538-545.
- [9] 云丰泽, 刘勤明, 汪宇杰. 基于融合神经网络模型的锂离子电池健康状态间接预测 [J]. 信息与控制, 1-13.
- [10] 田煜楷. 基于注意力机制的锂离子电池健康状态预测 [D]. 中北大学, 2023.44-51