DOI：10.13234/j.issn.2095-2805.2026.4.205 **中图分类号**：TM912 **文献标志码**：A

基于改进高斯过程回归的锂离子电池健康状态估计

邓鸿杨[1,2]，吴松荣[1,2]，刘 齐[1,2]，张翰文[1,2]，柳 博[1,2]，付 聪[1,2]
(1.磁浮技术与磁浮列车教育部重点实验室，成都 610031；2.西南交通大学电气工程学院，成都 611756)

摘要：准确估计锂离子电池的健康状态 SOH(state-of-health)是保障电池系统安全、稳定和高效运行的基础。基于此，提出 1 种基于健康因子提取和改进高斯过程回归 GPR(Gaussian progress regression)的锂离子电池 SOH 估计方法。首先，从充放电电流电压曲线中提取健康因子，使用 Pearson 相关分析法筛选出与容量相关性高的健康因子并通过自编码器深度学习算法对健康因子降维；其次，利用 GPR 建立 SOH 估计模型并通过组合核函数方式来构造 GPR 的核函数；然后，使用非洲秃鹫优化算法对组合核函数中的超参数进行寻优；最后，采用 NASA 锂离子电池数据集进行模型验证，实验结果表明本文所提方法具有较高的准确性。

关键词：锂离子电池；健康状态；高斯过程回归；非洲秃鹫优化算法

State-of-health Estimation of Lithium-ion Batteries Based on Improved Gaussian Process Regression

DENG Hongli[1,2], WU Songrong[1,2], LIU Qi[1,2], ZHANG Hanwen[1,2], LIU Bo[1,2], FU Cong[1,2]
(1. Key Laboratory of Magnetic Suspension Technique and Maglev Train of Ministry of Education, Chengdu 610031, China;
2. School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: The accurate estimation of the state-of-health (SOH) of lithium-ion batteries is a basis for ensuring the safe, stable and efficient operation of battery systems. On this basis, a method for estimating the SOH of lithium-ion batteries was proposed based on health indicator extraction and improved Gaussian process regression (GPR). First, the health indicators were extracted from the charge/discharge current and voltage curves, and those with a high correlation with capacity were screened out by Pearson correlation analysis, which were further reduced by an autoencoder deep learning algorithm. Second, GPR was used to establish an SOH estimation model, and the kernel function of GPR was constructed by combinatorial kernel functions. Third, the African vultures optimization algorithm was used to optimize the hyperparameters in the combinatorial kernel function. Finally, the NASA lithium-ion battery dataset was used for model verification, and experimental results show that the proposed method has high accuracy.

Keywords: Lithium-ion battery; state-of-health (SOH); Gaussian process regression (GPR); African vultures optimization algorithm

随着世界能源危机的不断升级，清洁能源的开发和利用成为了全球关注的焦点[1]。锂离子电池具有高能量密度、长循环寿命、低自放电率等优点，在清洁能源的储存和使用中发挥着重要的作用[2]。然而在锂离子电池的持续使用过程中性能会逐渐恶化，如果在失效前无法及时更换电池，将会导致设备运行异常甚至引起灾难性事故的发生[3-5]。因此，准确估计锂离子电池的健康状态 SOH(state-of-health)对于确保设备运行的安全性和效率具有重要意义。

SOH 反映了当前电池与出厂时的状态相比提供指定性能的能力，本文使用电池当前状态下的满

收稿日期：2023-11-28；**修回日期**：2024-01-27；**录用日期**：2024-02-19；**网络首发日期**：2024-04-26

充容量与额定容量之间的比值作为电池的 SOH。目前，锂离子电池的 SOH 估计方法主要包括基于模型的方法和基于数据驱动的方法。基于模型的方法主要通过构建电化学机理模型和等效电路模型来描述锂离子电池的退化过程。文献[6]提出了 1 种基于分数阶电化学模型的估计方法，实现了电池内阻和容量的同时估计；文献[7]使用 1 个基于 2-RC 的动态等效电路模型和 Arrhenius 方程实现了电池 SOH 的准确估计。然而基于模型的方法需要考虑到电池内部复杂的退化机制，存在建模困难、精度易受模型参数影响等缺点[8]。数据驱动的方法无须了解电池内部复杂的退化机理，仅通过建立从实时采集的电池运行数据中提取的健康因子 HI(health indicator)与电池容量之间的映射关系便可实现 SOH 的准确估计。文献[9]从电池充放电数据中提取 4 个健康因子建立支持向量回归 SVR (support vector regression)模型实现了 SOH 的准确估计，但样本数量的增加会加大 SVR 模型的计算复杂度导致运行时间过长；文献[10]通过分析容量增量曲线上不同峰值区间与 SOH 的关联性，提取最优健康因子构建基于高斯过程回归 GPR(Gaussian process regression)的电池 SOH 估计模型，其 SOH 估计的均方根误差在 2%以内。相比其他数据驱动方法，GPR 训练模型简单，能给出输出结果的置信区间使预测值具有不确定性表达，尤其对于解决非线性问题有突出的优势[11]。现有研究在采用传统 GPR 来建立 SOH 估计模型时，普遍采用单一核函数，并通过共轭梯度法调整核函数中的超参数[12-13]。然而，这种做法面临着一些局限性。首先，单一核函数可能限制了模型的泛化能力，影响其在不同或新的数据集上的适应性。其次，共轭梯度法在超参数优化过程中高度依赖于初始值的选择，可能导致优化过程陷入局部最优，从而使 SOH 的估计不够准确。

针对上述问题，本文提出 1 种基于改进 GPR 的电池 SOH 估计模型。首先，从电池充放电数据中提取出健康因子，将其作为输入，电池 SOH 作为输出，建立电池老化的 GPR 模型；然后引入非洲秃鹫优化算法 AVOA(African vultures optimization algorithm)并结合构建的组合核函数对模型进行优化；最后进行模型验证，结果表明改进后的 GPR 模型能显著提高 SOH 估计的准确性，实现对电池容量衰减的准确跟踪。

1 电池数据分析及 HI 提取

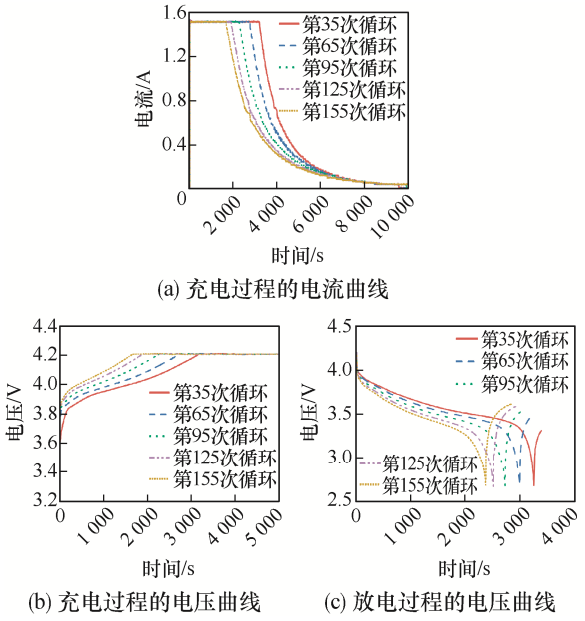
1.1 数据集介绍

本文锂离子电池数据来自美国航空航天局 NASA(National Aeronautics and Space Administration)艾姆斯卓越预测中心，选取其中 4 块 18650 型锂离子电池(编号为 B05、B06、B07、B18)的循环老化数据进行模型验证。4 块电池的额定容量为 2 A·h，均在 24 ℃的环境温度下进行充放电实验，此温度通常被认为是许多电池应用，特别是室内应用中的标准操作温度。因此，本文基于这一数据集的验证结果，可以为这些常见应用场景提供有效的 SOH 估计。在充电过程中，以 1.5 A 的恒定电流对电池充电，当电池电压达到 4.2 V 时转为恒压充电直到电流降至 20 mA，结束充电。在放电过程中，以 2.0 A 的恒定电流对电池放电，当 4 块电池的电压分别降至 2.7 V、2.5 V、2.2 V、2.5 V 时，结束放电。

1.2 HI 选取

使用数据驱动方法估计电池 SOH 不仅取决于算法本身，还依赖表征电池容量退化程度的 HI，因此选择合理的 HI 能显著提升模型的估计效果[14-15]。图 1 为 B05 电池在不同循环次数下的充放电电压、电流曲线。可以看出，随着循环次数的增加，其电压和电流曲线会规律性地向左移动，相同时间内电流电压的增减值、相同电流电压变化所需时间、恒流与恒压充电的时间均呈现出规律性的变化，因此可以基于以上规律从充放电曲线中选取 HI。初步选取 8 个 HI：充电过程中电流从 1.0 A 降至 0.3 A 的时间为 HI1；充电过程中电压从 3.9 V 升到 4.2 V 的时间为 HI2；恒压充电 500 s 后的电流下降值为 HI3；充电过程中第 500 s 的电压减去 3.8 V 的值为 HI4；恒流充电时间为 HI5；放电过程中电

压从 3.8 V 降至 3.5 V 的时间为 HI6；放电过程中第 500 s 到 1 000 s 的电压下降值为 HI7；恒压充电时间为 HI8。



(a) 充电过程的电流曲线

(b) 充电过程的电压曲线　　(c) 放电过程的电压曲线

图 1　B05 电池不同循环次数下的充放电曲线

Fig. 1 Charge/discharge curves for B05 battery under different cycle times

1.3 HI 筛选及降维

与电池 SOH 关联度较弱的 HI 会降低 SOH 估计的准确性，为了进一步验证并筛选出与电池容量高关联度的 HI，采用 Pearson 相关分析法来定量分析所选 HI 与 SOH 之间的相关性，其相关系数 r 的表达式为

$$r=\frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}} \tag{1}$$

式中：N 为电池总循环次数；x_i 和 y_i 分别为第 i 次循环对应的 HI 和 SOH；$\bar{x}$ 和 $\bar{y}$ 分别为 HI 与 SOH 的平均值。

r 的取值范围为[−1, 1]，当 r 的绝对值越接近于 1 时，表明 HI 与 SOH 的相关性越强。4 块电池 HI 与 SOH 间的相关系数计算结果见表 1。舍弃 r 小于 0.97 的 HI，以消除低相关性的 HI 对估计精度的影响。最终选择的 HI 为 HI2、HI5、HI6、HI7。

表 1　电池 SOH 与 HI 的相关系数计算结果

Tab. 1 Calculation results of correlation coefficient between battery SOH and HI

电池编号	不同健康 HI 的相关系数							
	HI1	HI2	HI3	HI4	HI5	HI6	HI7	HI8
B05	−0.970	−0.990	0.903	−0.986	0.996	0.996	−0.971	−0.975
B06	−0.952	0.986	0.935	−0.986	0.990	0.994	−0.981	−0.922
B07	−0.986	0.981	0.785	−0.982	0.992	0.997	−0.969	−0.959
B18	−0.941	0.983	0.560	−0.807	0.981	0.997	−0.993	−0.719

筛选出的高相关性 HI 之间信息重叠会加大计算和分析的难度，使用自编码器 AE(autoencoder)深度学习算法可以降低数据的维数，并保留原始数据中的大部分有用信息，消除 HI 之间的信息冗余、降低模型复杂度并缩短运行时间。自编码器是 1 种无监督深层神经网络，被广泛应用于数据降维与重构[16-18]。AE 由编码器和解码器组成，其中输入层与隐藏层组成编码器，将多维输入映射到低维空间，隐藏层与输出层组成解码器，将隐藏层中的低维数据重构。

AE 的解码和编码过程可表示为

$$h=s(\boldsymbol{w}_1x+\boldsymbol{b}_1) \tag{2}$$

$$x'=s(\boldsymbol{w}_2h+\boldsymbol{b}_2) \tag{3}$$

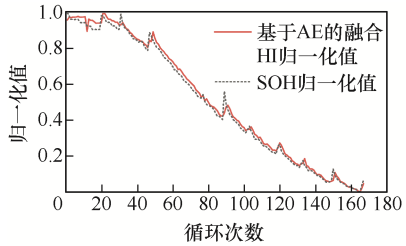
式中：x 和 x' 分别为输入和输出；h 为隐藏层节点；s 为激活函数 sigmoid；$\boldsymbol{w}_1$、$\boldsymbol{w}_2$ 为权重矩阵；$\boldsymbol{b}_1$、$\boldsymbol{b}_2$ 为偏置矩阵。

通过反向传播和梯度下降法对网络进行训练，以损失函数最小化为目标来调整网络参数，使输出数据逼近输入数据，最终得到降维后的输出，损失函数 $L_{\mathrm{AE}}(x,x')$ 为

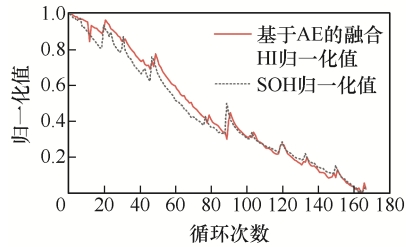
$$L_{\mathrm{AE}}(x,x')=\frac{1}{m}\|x-x'\|^2 \tag{4}$$

式中，m 为训练样本个数。

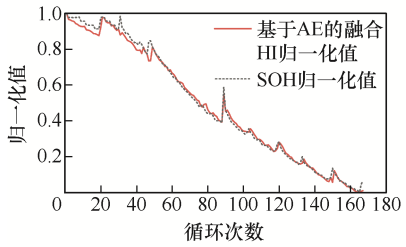
为了定性分析经 AE 降维后的融合 HI 与 SOH 变化的相关性，归一化后的 SOH 与基于 AE 的融合 HI 衰退曲线如图 2 所示。可以看出，基于 AE 的融合 HI 归一化值与 SOH 归一化值变化趋势基本一致，因此经 AE 降维后的融合 HI 与电池 SOH 有高相关性，可以较好地反映电池容量衰退情况。



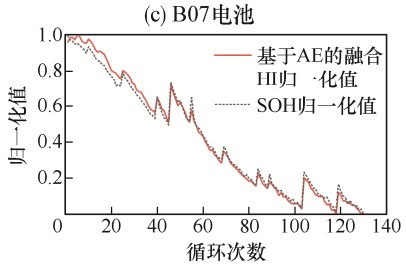
(a) B05电池



(b) B06电池



(c) B07电池



(d) B18电池

图 2 归一化后的 SOH 与融合 HI 衰退曲线

Fig. 2 Decay curves for normalized values of SOH and fused HI

2 基于 IGPR 的锂离子电池 SOH 估计

2.1 GPR 基本原理

GPR 是 1 种非参数回归的机器学习方法，可对任意输入变量和输出变量之间的非线性函数关系进行建模，并提供具有概率意义的预测分布[12]。假设给定 1 个训练样本序列 $D=(\boldsymbol{X}, \boldsymbol{y})$，其中 $\boldsymbol{X}=\left\{\boldsymbol{x}_1^n, \boldsymbol{x}_2^n, \cdots, \boldsymbol{x}_N^n\right\}$，由 n 维 N 个输入向量组成，$\boldsymbol{y}$ 为 n 维输入向量对应的输出向量。当输入对应的输出 $f(x)$形成的随机过程为高斯过程时，其定义为

$$f(x) \sim \mathrm{GP}(m(x), k(x, x')) \tag{5}$$

式中：$m(x)$为均值函数；$k(x, x')$ 为协方差函数。

带噪声的 $\boldsymbol{y}$ 的先验分布可以表示为

$$\boldsymbol{y} \sim N(m(x), \boldsymbol{K}(x, x)+\sigma_{\mathrm{N}}^2 \boldsymbol{I}_{\mathrm{N}}) \tag{6}$$

式中：σ_{N}^2 为噪声方差；$\boldsymbol{I}_{\mathrm{N}}$ 为单位矩阵；$\boldsymbol{K}$ 为 1 个协方差矩阵，通过协方差函数计算得到。

对于测试样本输入 x^*，其对应的估计值记为 y^*，训练样本输出 y 与估计值 y^* 服从联合高斯分布，即

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \boldsymbol{K}(x, x)+\sigma_{\mathrm{N}}^2 \boldsymbol{I}_{\mathrm{N}} & \boldsymbol{K}(x, x^*) \\ \boldsymbol{K}(x, x^*)^{\mathrm{T}} & \boldsymbol{K}(x^*, x^*) \end{bmatrix}\right) \tag{7}$$

由贝叶斯后验分布可得

$$y^*|_{x^*, D} \sim N(\bar{y}^*, \mathrm{cov}(y^*)) \tag{8}$$

$$\bar{y}^* = \boldsymbol{K}(x^*, x)[\boldsymbol{K}(x, x)+\sigma_{\mathrm{N}}^2 \boldsymbol{I}_{\mathrm{N}}]^{-1} y \tag{9}$$

$$\mathrm{cov}\, y^* = \boldsymbol{K}(x^*, x^*) - \boldsymbol{K}(x, x^*)^{\mathrm{T}} \cdot [\boldsymbol{K}(x, x)+ \sigma_{\mathrm{N}}^2 \boldsymbol{I}_{\mathrm{N}}]^{-1} \boldsymbol{K}(x, x^*) \tag{10}$$

式中，$\bar{y}^*$ 和 $\mathrm{cov}(y^*)$ 为 y^* 的均值和方差。

GPR 还能给出输出值的置信区间来反映估计结果的不确定性，估计值的 95%置信区间可表示为

$$\left[\bar{y}^* - 1.96 \times \sqrt{\mathrm{cov}(y^*)}, \bar{y}^* + 1.96 \times \sqrt{\mathrm{cov}(y^*)}\right] \tag{11}$$

从 GPR 的推导过程可知，估计值由核函数确定，因此核函数的选择和核函数中超参数的优化是 GPR 的关键。

2.2 基于核函数组合选择与 AVOA 优化的 GPR 模型

由于电池容量衰退过程中伴随有明显的容量回升现象，用单一核函数的 GPR 模型不足以很好地描述这一特征，本文选择平方指数核函数(SE)与线性核函数(LIN)组合的方式来捕捉电池容量的整体下降趋势和回升现象，以实现更准确的 SOH 估

计，组合核函数表达式为

$$k(x,x')=\sigma_{f_1}^2\exp\left[-\frac{(x-x')^2}{2l^2}\right]+\sigma_{f_2}^2(x-c)(x'-c) \tag{12}$$

式中，σ_{f_1}、σ_{f_2}、l、c 为组合核函数中的超参数，同时引入 σ_{N}，将这些超参数集合记为 $\theta=[\sigma_{f_1}^2,\sigma_{f_2}^2,l,c,\sigma_{\mathrm{N}}]$。

超参数可以通过最小化负对数似然函数和共轭梯度法 CGM(conjugate gradient method)求得。CGM 是 1 种局部优化算法，存在计算结果依赖于初始值的人为设定、容易陷入局部最优等缺点，因此本文使用 AVOA 代替 CGM 来优化 GPR 模型。AVOA 通过计算种群的适应度值来确定最优和次优秃鹫作为引领者，而剩余的秃鹫则根据饱食率的大小来确定当前迭代中跟随的引领者，算法每迭代一次，最优秃鹫就更新一次，直到达到设定的最大迭代次数，输出最优的值[19-20]。适应度值的计算公式为

$$\begin{cases}R(i)=\begin{cases}B_1 & P_i=L_1\\ B_2 & P_i=L_2\end{cases}\\ P_i=f_i/\sum_{i=1}^{n}f_i\end{cases} \tag{13}$$

式中：B_1、B_2 分别为最优、次优秃鹫的位置；$R(i)$ 为其他秃鹫的位置；i 为当前迭代次数；L_1、L_2 为[0, 1]之间的 2 个随机数，且 $L_1+L_2=1$；P_i 为选择最佳秃鹫的概率；f_i 为秃鹫适应度值。

饱食率的计算公式为

$$F=(2r_1+1)\cdot z\cdot\left(1-\frac{i}{i_{\max}}\right)+h\cdot\left[\sin^w\left(\frac{\pi}{2}\cdot\frac{i}{i_{\max}}\right)+\cos\left(\frac{\pi}{2}\cdot\frac{i}{i_{\max}}\right)-1\right] \tag{14}$$

式中：F 为秃鹫饱食率；h 为[−2, 2]之间的随机数；z 为[−1, 1]之间的随机数；r_1 为[0, 1]之间的随机数；w 为自定义参数；$i_{\max}$ 为最大迭代次数。AVOA 根据饱食率 $|F|$ 的大小来确定秃鹫位置的更新策略。

2.3 基于 IGPR 的锂离子电池 SOH 估计流程

基于改进高斯过程回归 IGPR(improved Gaussian process regression)的 SOH 估计模型的核心思想是利用 AVOA 对 GPR 中的组合核函数进行超参数寻优，其估计流程如图 3 所示。

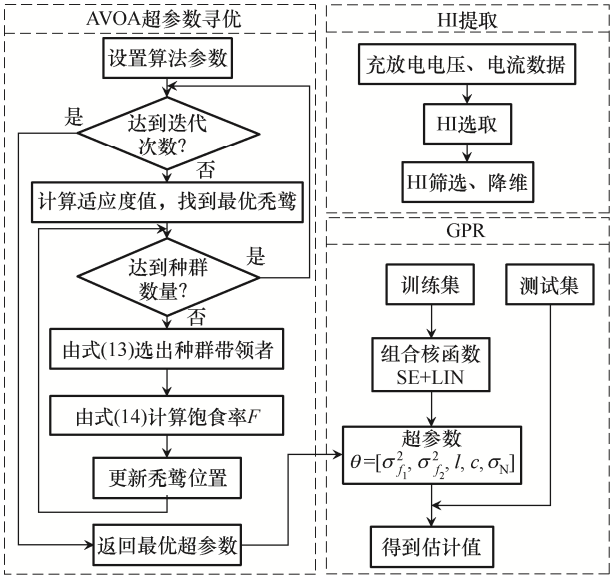


图 3　基于 IGPR 的电池 SOH 估计流程

Fig. 3 Battery SOH estimation process based on IGPR

根据图 3，IGPR 估计模型的主要步骤如下。

步骤 1 HI 提取。从锂离子电池充放电数据曲线中选取与电压、电流、时间相关的 8 个 HI，然后利用 Pearson 相关系数法评估 8 个 HI 与电池 SOH 的相关性，筛选出相关系数较高的 HI，并使用 AE 算法对其进行降维。

步骤 2 划分训练集、测试集。将样本总体分为训练集和测试集，其中模型的输入为融合 HI，输出为电池 SOH。

步骤 3 AVOA 寻优。设置 AVOA 中的参数、寻优的目标函数，以训练样本的估计值与真实值的均方根误差为目标函数，将超参数求解转换为对目标函数求最小值，求出最优超参数值。

步骤 4 IGPR 模型测试。将 AVOA 寻到的超参数值代入 GPR，最后将测试集基于训练好的模型进行 SOH 估计。

3　模型验证及分析

为更好地评价模型估计性能，使用平均百分比

误差 MAPE(mean absolute percentage error)、均方根误差 RMSE(root mean squared error)、平均绝对误差 MAE(mean absolute error)作为模型的性能评价指标。MAPE、RMSE、MAE 越小则表明模型的估计性能越好，其计算式分别为

$$\delta_{\mathrm{MAPE}}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{S}_{\mathrm{OH},i}-S_{\mathrm{OH},i}}{S_{\mathrm{OH},i}}\right|\times 100\% \tag{15}$$

$$\delta_{\mathrm{RMSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{S}_{\mathrm{OH},i}-S_{\mathrm{OH},i})^{2}}\times 100\% \tag{16}$$

$$\delta_{\mathrm{MAE}}=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{S}_{\mathrm{OH},i}-S_{\mathrm{OH},i}\right|\times 100\% \tag{17}$$

式中，$\hat{S}_{\mathrm{OH},i}$、$S_{\mathrm{OH},i}$分别为 SOH 的估计值、真实值。

设置秃鹫种群规模 N=50、最大迭代次数 $i_{\max}$=200、寻优上下限为[−20, 20]，各阶段更新策略参数 P_1、P_2、P_3、w、L_1、L_2分别为 0.6、0.4、0.6、2.5、0.8、0.2。

3.1 不同超参数寻优方法的 SOH 估计结果对比

为验证 AVOA-GPR 方法相比 CGM-GPR 方法的优越性，使用 SE 核函数，将 B05 电池前 50%循环数据作为训练集，剩余循环数据作为测试集，对比 AVOA-GPR 与 CGM-GPR 这 2 种超参数寻优方法的 SOH 估计结果，CGM-GPR 中的初始超参数记为 θ，设置 2 组不同超参数初始值为[0, 0]、[0, 5]。图 4 为 2 种方法的 SOH 估计结果，2 种方法的误差评估见表 2。

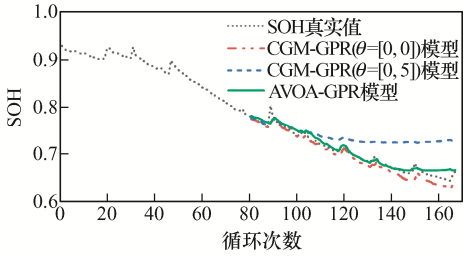


图 4 CGM-GPR 与 AVOA-GPR 的 SOH 估计结果

Fig. 4 SOH estimation results based on CGM-GPR and AVOA-GPR

由图 4 和表 2 可知，AVOA-GPR 模型的估计结果相比 CGM-GPR 模型能够更好地贴合 SOH 真实值并且具有更高的估计精度。当 CGM-GPR 模型的超参数初始值设置为[0, 0]时，其估计结果与 SOH 真实值较接近，而超参数初始值变为[0, 5]时，其估计结果将远远偏离 SOH 真实值，估计精度大大下降，这是由于 CGM-GPR 模型受初始值人为设定的影响易陷入局部最优，从而导致实际应用效果不太理想。综上可知，本文使用的 AVOA-GPR 在超参数寻优中更具优势。

表 2 2 种方法的误差评估

Tab. 2 Error evaluation for two methods

方法	MAPE/%	RMSE/%	MAE/%
AVOA-GRP	0.92	0.88	0.64
CGM-GRP (θ=[0, 0])	1.33	1.12	0.91
CGM-GRP (θ=[0, 5])	5.08	4.28	3.43

3.2 不同核函数的 SOH 估计结果对比

为验证 SE+LIN 组合核函数的有效性，将 B05 电池前 50%循环数据作为训练集，剩余循环数据作为测试集，对比 SE-GPR、LIN-GPR、GPR、SE+LIN-GPR 模型的 SOH 估计结果，所有核函数中的超参数均由 AVOA 寻优。图 5 为基于不同核函数的 SOH 估计结果，其误差评估见表 3。

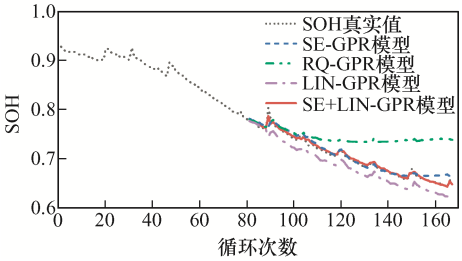


图 5 基于不同核函数的 SOH 估计结果

Fig. 5 SOH estimation results based on different kernel functions

表 3 不同核函数的误差评估

Tab. 3 Error evaluation for different kernel functions

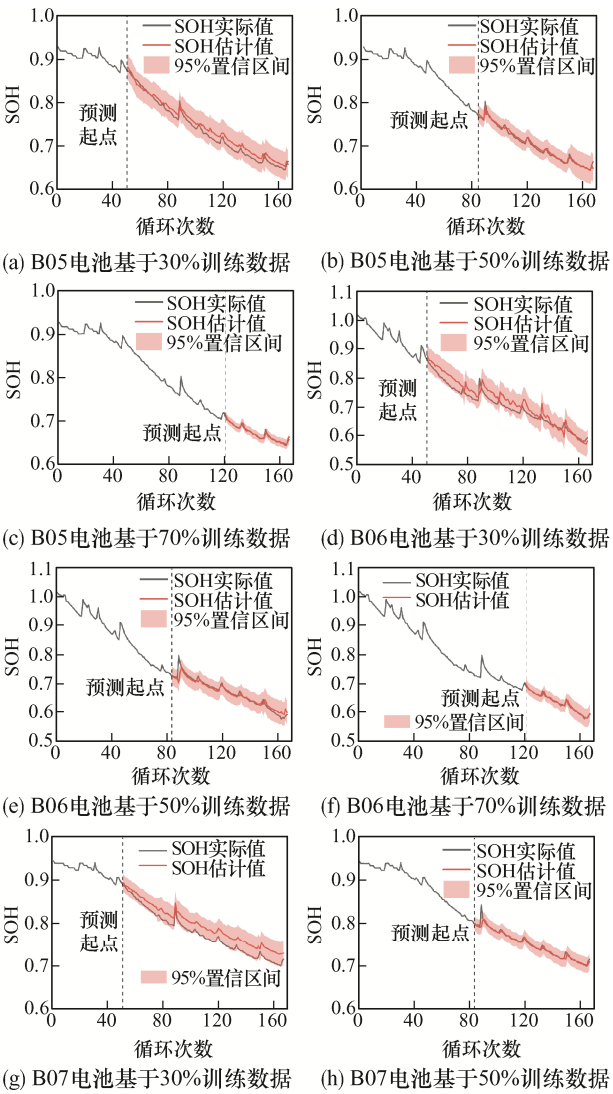
核函数	MAPE/%	RMSE/%	MAE/%
SE	0.92	0.88	0.64
RQ	6.17	5.12	4.16
LIN	2.70	2.04	1.89
SE+LIN	0.48	0.46	0.34

由图 5 和表 3 可知，基于不同核函数的模型 SOH 估计结果差异很大。LIN-GPR 模型和 RQ-GPR 模型的 SOH 估计结果均明显偏离 SOH 真实值，其

中 RQ-GPR 模型的估计结果最差，MAPE、RMSE、MAE 分别达到了 6.17%、5.12%、4.16%，而 SE-GPR 模型的估计结果明显好于前两者，但其在最后 10 个循环的估计结果并不理想。SE+LIN-GPR 模型的 SOH 估计结果最好，估计精度明显优于其他 3 种单一核函数。综上可知，本文提出的 SE+LIN 组合核函数相较于单一核函数能够更准确地捕捉到电池的容量回升和衰退趋势，可以实现更高精度的 SOH 估计。

3.3　单个电池不同训练数据量的 SOH 估计结果对比

为验证所提出的改进 GPR 模型的准确性，采用不同的训练量对电池进行测试，分别选取 4 块电池前 30%、50%、70%的循环数据作为训练集，剩余循环数据作为测试集。图 6 所示为 4 块电池基于不同训练量的 SOH 估计结果及 95%置信区间，该置信区间用来表示估计结果的不确定性，越窄代表估计结果的可信度越高。误差评估见表 4。

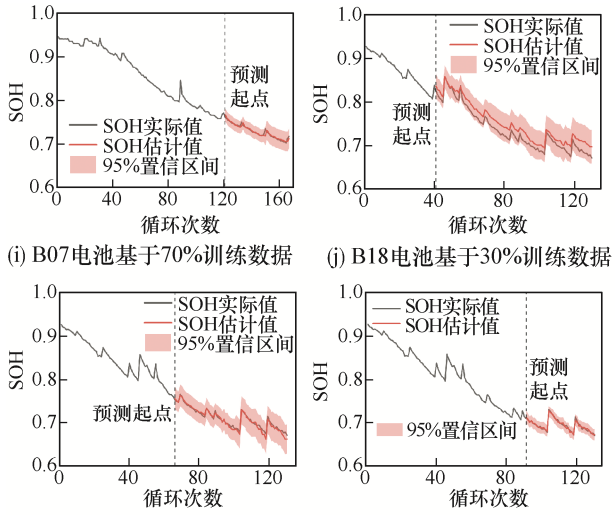


(a) B05电池基于30%训练数据　(b) B05电池基于50%训练数据

(c) B05电池基于70%训练数据　(d) B06电池基于30%训练数据

(e) B06电池基于50%训练数据　(f) B06电池基于70%训练数据

(g) B07电池基于30%训练数据　(h) B07电池基于50%训练数据



(i) B07电池基于70%训练数据　(j) B18电池基于30%训练数据

(k) B18电池基于50%训练数据　(l) B18电池基于70%训练数据

图 6　4 块电池基于不同训练量的 SOH 估计结果

Fig. 6 SOH estimation results for four batteries based on different training amounts

表 4　4 块电池的误差评估

Tab. 4 Error evaluation for four batteries

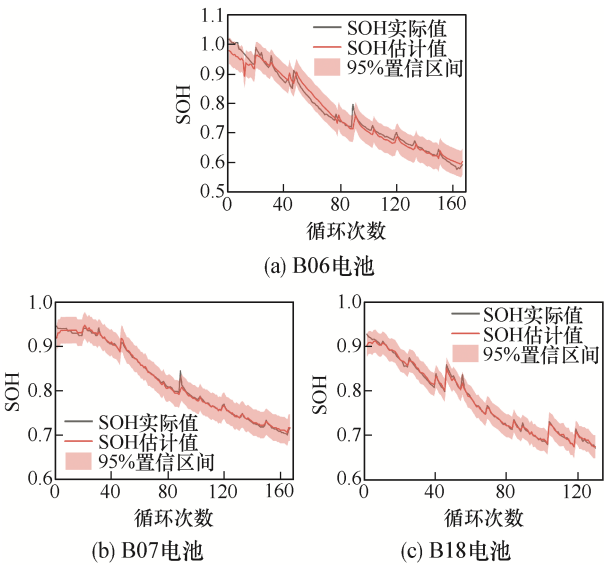
电池编号	训练量/%	MAPE/%	RMSE/%	MAE/%
B05	30	1.47	1.27	1.35
	50	0.48	0.46	0.34
	70	0.15	0.21	0.09
B06	30	3.71	2.71	2.73
	50	1.06	1.31	0.71
	70	0.23	0.28	0.15
B07	30	3.51	2.42	2.68
	50	0.37	0.64	0.28
	70	0.18	0.27	0.13
B18	30	2.23	1.71	1.61
	50	0.39	0.42	0.27
	70	0.15	0.20	0.10

由图 6 和表 4 可知，随着训练量的增加，SOH 估计的准确率不断提高，置信区间也逐渐变窄，表明估计模型有较高的可靠性。在训练量仅有 30%的情况下，模型的估计趋势与 SOH 真实值很相近，其中 B05 电池的误差均在 1.50%以下。当训练量增加到 50%时，模型已经具有较好的效果，能够实现

准确的 SOH 估计。B06 电池的误差略高于其他 3 块电池，这是由于 B06 电池的融合 HI 的相关性略低于其他电池(如图 3 所示)，然而这一误差依然在可接受范围内，仍然能证明本文方法具有较高的精度。因此本文提出的 IGPR 模型在不同的训练集下均具有较强的准确性。

3.4 多电池数据的 SOH 估计结果

为评估模型的泛化能力，选择 B05 电池的全部数据作为训练集，进行模型训练，使用 B06、B07、B18 电池作为测试集验证模型的估计精度，估计结果如图 7 所示，误差评估见表 5。



(a) B06电池

(b) B07电池

(c) B18电池

图 7 多个电池数据的 SOH 估计结果

Fig. 7 SOH estimation results based on data of multiple batteries

表 5 多电池估计的误差评估

Tab. 5 Error evaluation for estimation of multiple batteries

电池编号	MAPE/%	RMSE/%	MAE/%
B06	2.98	2.68	2.87
B07	0.77	0.74	0.64
B18	0.65	0.59	0.51

由图 7 和表 5 可知，B07、B18 的估计值与真实值基本贴合，估计误差均低于 1.00%且真实值也基本落于置信区间内。B06 电池在循环前期的估计值有较大的偏移，但是其误差在 3.00%以内，仍然具有较高的精度。综上，本文提出的 IGPR 模型在多电池实验中仍然有较高的可靠性，具有较强的泛化能力。

3.5 IGPR 模型与其他机器学习模型的对比

为进一步验证 IGPR 模型相较于其他机器学习模型的优越性，将 B05 电池前 50%循环数据作为训练集，剩余循环数据作为测试集，对比其他文献所用模型与本文所建立的 IGPR 模型的 SOH 估计性能，比较结果见表 6。

表 6 不同模型的误差评估

Tab. 6 Error evaluation for different models

模型	MAPE/%	RMSE/%
IGPR	0.48	0.46
GPR[12]	—	0.76
IALO-SVR[9]	0.71	0.53
PSO-RVM[14]	—	1.02

由表 6 可知，IGPR 模型相比于 GPR、IALO-SVR、PSO-RVM 模型有更高的估计精度，表明本文所提出的 IGPR 模型相较其他方法在 SOH 估计的准确性方面，具有较为优异的性能。

4 结论

本文提出 1 种基于健康因子提取和改进 GPR 模型的锂离子电池 SOH 估计方法，并使用 NASA 锂离子电池数据集进行模型验证，结论如下。

(1)结合 SE 和 LIN 核函数提出了 1 种新的组合核函数，能够有效描述电池的容量回升现象和长期衰退趋势。

(2)使用非洲秃鹫优化算法能有效解决传统的共轭梯度法受初值影响易陷入局部最优的问题，减少因人为调整超参数初始值引起的误差，以提高估计精度。

(3) IGPR 模型基于 NASA 电池数据集的 SOH 估计精度较好，证明所提估计模型具有良好的准确性和可靠性。

参考文献：

[1] Sarmah S B, Kalita P, Garg A, et al. A review of state of

health estimation of energy storage systems: Challenges and possible solutions for futuristic applications of Li-ion battery packs in electric vehicles [J]. Journal of Electrochemical Energy Conversion and Storage, 2019, 16(4): 040801.

[2] Fan Yaxiang, Xiao Fei, Li Chaoran, et al. A novel deep learning framework for state of health estimation of lithium-ion battery [J]. Journal of Energy Storage, 2020, 32: 101741.

[3] Tian Huixin, Qin Pengliang, Li Kun, et al. A review of the state of health for lithium-ion batteries: Research status and suggestions [J]. Journal of Cleaner Production, 2020, 261: 120813.

[4] Shu Xing, Shen Shiquan, Shen Jiangwei, et al. State of health prediction of lithium-ion batteries based on machine learning: Advances and perspectives [J]. iScience, 2021, 24(11): 103265.

[5] Gong Dongliang, Gao Ying, Kou Yalin, et al. State of health estimation for lithium-ion battery based on energy features [J]. Energy, 2022, 257: 124812.

[6] Hosseininasab S, Lin Changwei, 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.

[7] Amir S, Gulzar M, Tarar M O, et al. Dynamic equivalent circuit model to estimate state-of-health of lithium-ion batteries [J]. IEEE Access, 2022, 10: 18279-18288.

[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(3): 1182-1195.
Xiong Qing, Di Zhenguo, Ji Shengchang. Review on health state estimation and life prediction of lithium-ion batteries [J]. High Voltage Engineering, 2024, 50(3): 1182-1195. (in Chinese)

[9] Li Qianglong, Li Dezhi, Zhao Kun, et al. State of health estimation of lithium-ion battery based on improved ant lion optimization and support vector regression [J]. Journal of Energy Storage, 2022, 50: 104215.

[10] 杨胜杰, 罗冰洋, 王菁, 等. 基于容量增量曲线峰值区间特征参数的锂离子电池健康状态估算[J]. 电工技术学报, 2021, 36(11): 2277-2287.
Yang Shengjie, Luo Bingyang, Wang Jing, et al. State of health estimation for lithium-ion batteries based on peak region feature parameters of incremental capacity curve [J]. Transactions of China Electrotechnical Society, 2021, 36(11): 2277-2287. (in Chinese)

[11] 韩云飞, 谢佳, 蔡涛, 等. 结合高斯过程回归与特征选择的锂离子电池容量估计方法[J]. 储能科学与技术, 2021, 10(4): 1432-1438.
Han Yunfei, Xie Jia, Cai Tao, et al. Capacity estimation of lithium-ion batteries based on Gaussian process regression and feature selection [J]. Energy Storage Science and Technology, 2021, 10(4): 1432-1438. (in Chinese)

[12] Yang Duo, Zhang Xu, Pan Rui, et al. A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve [J]. Journal of Power Sources, 2018, 384: 387-395.

[13] 张梦迪, 刘洋, 陈健, 等. 基于 ISSA-GPR 的锂离子电池健康状态估计[J]. 电源学报, 2026, 24(2): 233-243.
Zhang Mengdi, Liu Yang, Chen Jian, et al. Estimation of state-of-health of lithium-ion batteries based on ISSA-GPR [J]. Journal of Power Supply, 2026, 24(2): 233-243. (in Chinese)

[14] Guo Peiyao, Cheng Ze, Yang Lei. A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction [J]. Journal of Power Sources, 2019, 412: 442-450.

[15] Ma Yan, Shan Ce, Gao Jinwu, et al. A novel method for state of health estimation of lithium-ion batteries based on improved LSTM and health indicators extraction [J]. Energy, 2022, 251: 123973.

[16] Meng Qinxue, Catchpoole D, Skillicom D, et al. Relational autoencoder for feature extraction [C]// 2017 Inter-national Joint Conference on Neural Networks (IJCNN). Anchorage, AK, USA, 2017: 364-371.

[17] 尹杰, 刘博, 孙国兵, 等. 基于迁移学习和降噪自编码器−长短时间记忆的锂离子电池剩余寿命预测[J]. 电工技术学报, 2024, 39(1): 289-302.
Yin Jie, Liu Bo, Sun guobing, et al. Transfer learning denoising autoencoder-long short termmemory for remaining useful life prediction of Li-ion batteries [J]. Tran-

sactions of China Electrotechnical Society, 2024, 39(1): 289-302. (in Chinese)

[18] 彭自然, 王顺豪, 肖伸平, 等. 基于 KA Informer 的电动汽车动力电池荷电状态和健康状态估算[J]. 电工技术学报, 2025, 40(19): 6378-6394.
Peng Ziran, Wang Shunhao, Xiao Shenping, et al. State of charge and state of health estimation of electric vehicle power battery based on KA informer model [J]. Transactions of China Electrotechnical Society, 2025, 40(19): 6378-6394. (in Chinese)

[19] Abdollahzadeh B, Gharehchopogh F S, Mirjalili S. African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems [J]. Computers & Industrial Engineering, 2021, 158: 107408.

[20] 申江卫, 马文赛, 肖仁鑫, 等. 基于优化高斯过程回归算法的锂离子电池可用容量估算[J]. 中国公路学报, 2022, 35(8): 31-43.
Shen Jiangwei, Ma Wensai, Xiao Renxin, et al. Available capacity estimation of lithium-ion bat-teries based on optimized Gaussian process regr- ession [J]. China Journal of Highway and Tran- sport, 2022, 35(8): 31-43. (in Chinese)



邓鸿栃

作者简介：

邓鸿栃(1998—)，男，硕士。研究方向：锂离子电池健康状态估计。E-mail：dhongli123@163.com。

吴松荣(1977—)，男，通信作者，博士，副教授。研究方向：电力电子技术和机车/动车组牵引传动。E-mail：srwu88@163.com。

刘齐(1999—)，男，硕士。研究方向：电池 SOC 估计。E-mail：1203954416@qq.com。

张翰文(1999—)，男，硕士。研究方向：图像识别。E-mail：925513303@qq.com。

柳博(2000—)，男，硕士。研究方向：电力电子技术。E-mail：691637020@qq.com。

付聪(2000—)，男，硕士。研究方向：电力电子技术。E-mail：2606880283@qq.com。

(责任编辑：郭凯凯)