

Lithium-ion Battery State of Charge (SOC) Estimation Based on HHO-LSTM

Cao Mingxing*, Gao Guozhang

School of Naval Architecture, Ocean and Energy Power Engineering, Wuhan University of Technology, Wuhan, China

Email address:

599698654@qq.com (Cao Ming-xing), ggzhang815@163.com (Gao Guozhang)

*Corresponding author

To cite this article:

Cao Mingxing, Gao Guozhang. (2025). Lithium-ion Battery State of Charge (SOC) Estimation Based on HHO-LSTM. *Science Discovery*, 13(6), 149-158. <https://doi.org/10.11648/j.sd.20251306.18>

Received: 28 October 2025; **Accepted:** 11 December 2025; **Published:** 24 December 2025

Abstract: State of Charge (SOC) estimation is one of the core contents of the Battery Management System (BMS), and it is crucial for evaluating the safety of using lithium-ion batteries as the power source for ships. Based on the existing problems in SOC estimation methods, this paper proposes a method for SOC estimation that uses the Long Short-Term Memory (LSTM) neural network as the core algorithm, optimizes parameters through the Harris Hawk Optimization Algorithm (HHO), and builds a co-simulation model. Firstly, using four batteries, B0005, B0006, B0007, and B0018, from the National Aeronautics and Space Administration (NASA) as the basic data, the random forest algorithm and Pearson correlation coefficient analysis are used to identify the linear and nonlinear correlations with the target variable respectively, eliminate redundant or irrelevant variables, and screen out features with high correlation with SOC to improve estimation accuracy. Then, the obtained features are input into the LSTM neural network model, and the Harris optimization algorithm is used to optimize the model parameter combination. Finally, digital simulation experiments are conducted. The results show that the HHO-LSTM prediction model improves the estimation accuracy of SOC. This method provides a new idea for SOC estimation and has certain effectiveness and feasibility.

Keywords: State of Charge Estimation, Harris Hawk Optimization Algorithm, Long Short-Term Memory Neural Network, Pearson Correlation Coefficient Analysis

基于HHO-LSTM的锂离子电池SOC估计

曹明兴*, 高国章

武汉理工大学船海与能源动力工程学院, 武汉, 中国

邮箱

599698654@qq.com (曹明兴), ggzhang815@163.com (高国章)

摘要: 电池荷电状态 (State of Charge, SOC) 估计是电池管理系统 (Battery Management System, BMS) 的核心内容之一, 对于评估以锂离子电池作为船舶动力来源的安全性至关重要。基于现有SOC估计方法存在的问题, 本文提出了一种以长短期记忆 (Long Short-Term Memory, LSTM) 神经网络为核心算法, 通过哈里斯鹰优化算法 (Harris Hawk Optimization Algorithm, HHO) 完成参数优化, 搭建出联合仿真模型的方法进行SOC估计。首先, 以美国国家航空航天局 (NASA) 的B0005、B0006、B0007和B0018四块电池作为基础数据, 通过随机森林算法以及皮尔逊相关系数分析, 分别识别与目标变量的线性及非线性相关性, 剔除冗余或无关变量, 筛选出与SOC具有高度相关性的特征, 以提高估计精确度; 然后, 将得到的特征输入到长短期记忆神经网络模型中, 并使用哈里斯优化算法完成对模型参数组合的优化; 最后, 进行了数字仿真实验, 结果表明, HHO-LSTM预测模型提高了对SOC的估计精度, 该方法为SOC估计提出一种新的思路, 且有一定的有效性和可行性。

关键词: 荷电状态估计, 哈里斯鹰优化算法, 长短期记忆神经网络, 皮尔逊相关系数分析

1. 引言

锂离子电池因其具有能量密度高、自放电率低、循环寿命长等优势,在船舶动力中的应用越来越广泛。电池管理系统(BMS)对保障动力电池的安全稳定运行起着至关重要的作用,它为船舶运行的安全性、准确性和可靠性的重要保障[1]。

锂离子电池动力系统是典型的复杂非线性系统,其特性与行为受多因素耦合作用影响显著。这些因素的动态变化会引发电池内部特性与外特性的非线性演变,使得难以准确估算电池荷电状态(SOC)。特别是锂电池在船舶航行的实际工作过程中呈现的强非线性特征,以及工作环境条件(如温度、充放电电流)的动态多变性,进一步加剧了SOC精确估算的技术难度。

锂电池SOC预测方法多样,主要分为传统法、基于等效电路模型法以及基于数据驱动法。开路电压(Open Circuit Voltage Method, OCV)法和安时积分法是常见的传统SOC预测方法。开路电压法依据电池开路电压与SOC的对应关系,通过测量开路电压并查询预先建立的数据来查找对应的估算SOC。但实际应用中,电池难以长时间保持开路状态,且OCV-SOC关系受电池老化、温度等因素影响,并非完全线性,导致该方法预测精度有限,通常仅作为辅助手段[2]。安时积分法预测SOC是通过对电池充放电电流积分计算电量来实现,其原理简单,无需复杂硬件。不过,它对SOC初始值依赖性强,初始值不准确会使误差随时间累积,且易受电池自放电、电流测量误差等因素干扰。为此,研究人员提出结合其他传感器数据校正初始值、采用自适应积分步长、引入温度补偿机制以及利用卡尔曼滤波算法补偿电流测量误差等改进措施来提高其预测精度[3-9]。

卡尔曼滤波法和滑模观测器法是基于等效电路模型的SOC预测的基本的方法。卡尔曼滤波法利用电池等效电路模型描述电池动态特性,通过卡尔曼滤波器估计和更新模型参数,能有效处理系统和测量噪声,提升预测精度。然而,该方法对电池模型准确性要求高,模型参数辨识困难,计算复杂度也较高,在实际应用中可能面临实时性问题。为此,研究人员提出(Extended Kalman Filter, EKF)、无迹卡尔曼滤波(Unscented Kalman Filter, UKF)等改进算法。EKF通过线性化处理将卡尔曼滤波应用于锂电池SOC预测,但处理强非线性问题时线性化近似可能引入较大误差;UKF则通过优化采样策略提高在非线性系统中的预测精度[10-13]。滑模观测器法通过设计滑模面实现对SOC的观测,具有较强的鲁棒性,对系统参数变化和外部干扰不敏感。但滑模运动过程中产生的抖振现象会影响观测精度,且滑模面设计需要经验和技巧。为克服抖振问题,研究人员提出自适应滑模观测器、积分滑模观测器等改进方法,通过调整滑模控制律或引入积分环节来抑制抖振,提高SOC预测精度[14]。

支持向量机(Support Vector Machine, SVM)、神经网络(Artificial Neural Network, ANN)、长短期记忆(Long and Short-Term Memory, LSTM)网络是基于数据驱动的SOC预测的常用的方法。SVM最初用于数据分类,后应用于电池SOC预测,通过拟合电池运行数据与SOC的关系进行预测,具有较好的泛化能力,能处理高维数据。但处理大规模数据集时计算复杂度高,模型参数选择对预测结果

影响大。为优化其性能,研究人员提出粒子群优化PSO(Particle Swarm Optimization)-SVM、遗传算法GA(Genetic Algorithm)-SVM等改进算法,通过优化核函数参数和惩罚因子,提高其在锂电池SOC预测中的准确性和稳定性[15-17]。神经网络在锂电池SOC预测中同样应用广泛,包括浅层神经网络和深度神经网络。浅层神经网络结构简单、训练速度快,但处理复杂电池数据时特征提取能力有限,预测精度较低。卷积神经网络(Convolutional Neural Network, CNN)和LSTM作为深度神经网络具有强大的特征提取和学习能力。CNN通过局部连接和权值共享策略有效提取数据空间特征,但处理时序数据时缺乏记忆功能;LSTM作为循环神经网络(Recurrent Neural Network, RNN)的变体,通过引入遗忘门、输入门和输出门,能有效处理时间序列数据中的长期依赖关系,在锂电池SOC预测中表现较好。不过,LSTM存在训练时间长、易陷入局部最优等问题。研究人员提出引入注意力机制、采用自适应学习率等改进方法来提高其训练效率和预测精度[18-21]。针对LSTM存在的不足,文献[22]提出一种基于随机搜索(Random Search, RS)优化长短期记忆网络的SOC估计方法,以解决现有SOC估计方法的不足,提升估计精度与模型性能,但该方法实验主要在特定的电池型号上进行,未充分验证该方法在其他不同类型、规格电池上的适用性。不同电池的化学特性、充放电曲线和性能表现存在差异,在一种电池上有效的方法在其他电池上可能效果不佳。文献[23]则提出了一种结合LSTM和多元线性回归算法的电动汽车SOC预测方法,针对传统的LSTM在复杂工况下的不足,构建LR-LSTM联合预测模型,交叉验证表明该模型预测可全状态、全气候应用。在不同季节和温度下模型稳定,高温下性能略好,但该方法的LSTM网络超参数靠经验确定,如初始设定的轮次、批次大小等参数。这种方式缺乏系统性和客观性,可能无法找到最优超参数组合,限制模型性能提升,导致预测精度未达最佳。

针对现如今LSTM在SOC估计的应用层面存在的局限性,本文提出了一种利用 Harris 鹰优化算法(Harris Hawks Optimizer, HHO)优化LSTM神经网络模型。HHO作为一种新兴的受自然界中鹰捕猎行为启发的优化算法,实现对复杂优化问题的高效求解,通过该算法寻找LSTM中的最优超参数(隐藏神经元层层、学习率、批次大小等),提升模型的全局搜索能力和收敛特性,增强模型的预测精度和准确性。本研究针对传统的LSTM在SOC估计的应用层面存在的不足,提出了一种新的解决方向,以提高SOC估计精度,确保船舶动力系统安全稳定。

2. 数据描述与预处理

2.1. 数据描述

本文以美国国家航空航天局(NASA)的B0005、B0006、B0007和B0018四块电池放电样本(下文简称为5、6、7和18)作为基础数据,放电过程以2A的恒定电流进行。5、6、7和18号电池的放电终止电压分别为2.7V、2.5V、2.2V和2.5V。随着电池不断地循环老化,每个循环的放电截至点不断提前,额定容量持续衰减,至30%(从2Ahr降至1.4Ahr)后实验停止,5号电池放电电流、温度以及电压参数如图1所示。针对同一循环,随着放电电流从-2A

归0，表示放电结束，电池端电压上升温度降低；针对不同循环，随着电池额定容量持续衰减，放电时间减少，图1曲线拐点随循环进程不断在时间坐标轴上前移。

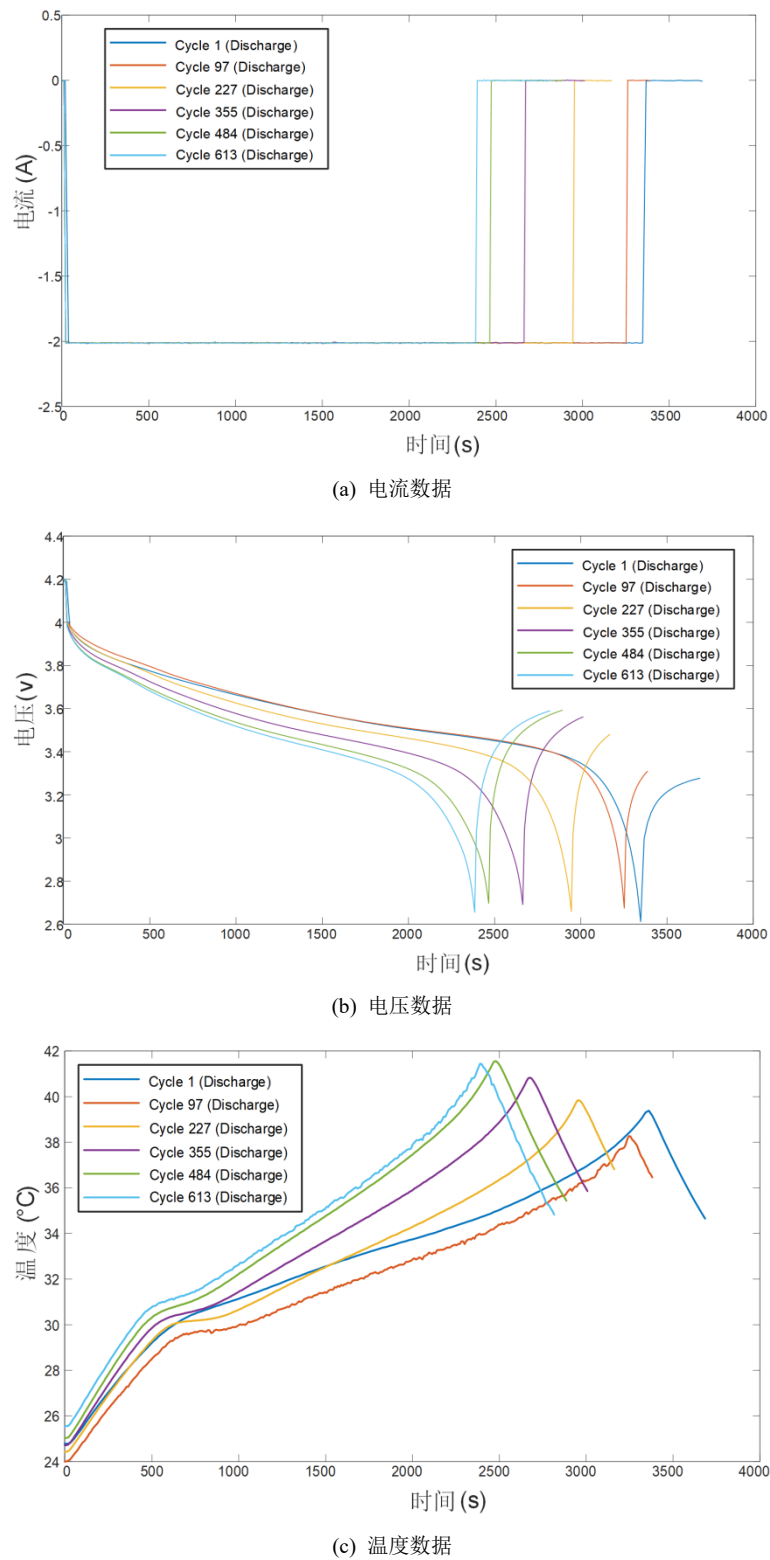


图1. 5号电池放电过程电压、电流、温度数据。

2.2. 数据预处理

本文采用皮尔斯相关性分析在多变量分析中通过计算变量间的相关系数，识别与目标变量线性相关性高的特征，剔除冗余或无关变量，从而降低特征维度，为后续模型的

训练奠定基础。该方法中系数的取值范围在-1~1之间，当系数趋近于1时，两个变量呈强正线性相关；当系数趋近于-1时，两个变量呈强负线性相关；当趋近于0时，变量之间几乎不存在线性相关关系。从图1相关系数矩阵图中可知，放电容量、放电能量以及放电时间与SOC线性相关性高，线性负相关性皆大于97%。

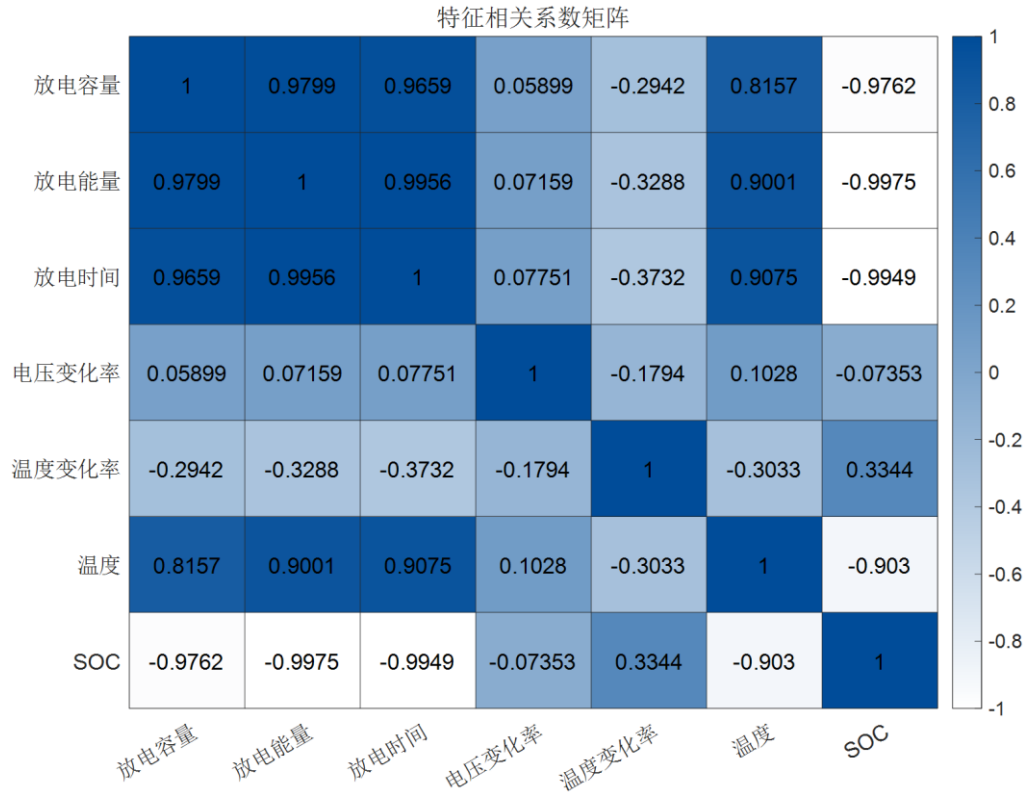


图2 皮尔逊相关系数矩阵图。

皮尔逊相关系数只能衡量线性相关性,如果变量之间存在非线性关系,皮尔逊相关系数可能无法捕捉。针对以上问题,本文加入随机森林模型对特征参数进行重要性计算,捕捉参数间的非线性关系,从而完善特征重要性分析。随机森林算法评估特征重要性的核心逻辑是:通过量化每个特征对模型预测精度的贡献程度,或其在决策树分裂过程中降低数据不纯度的能力,来衡量特征的重要性,结果数值越大,特征对预测的准确性越大。

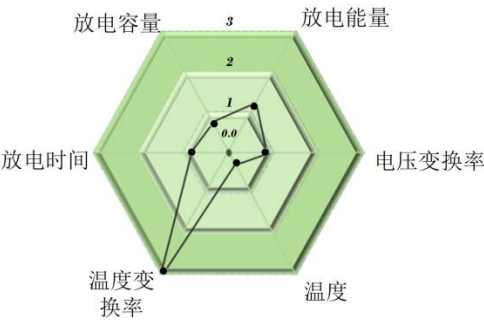
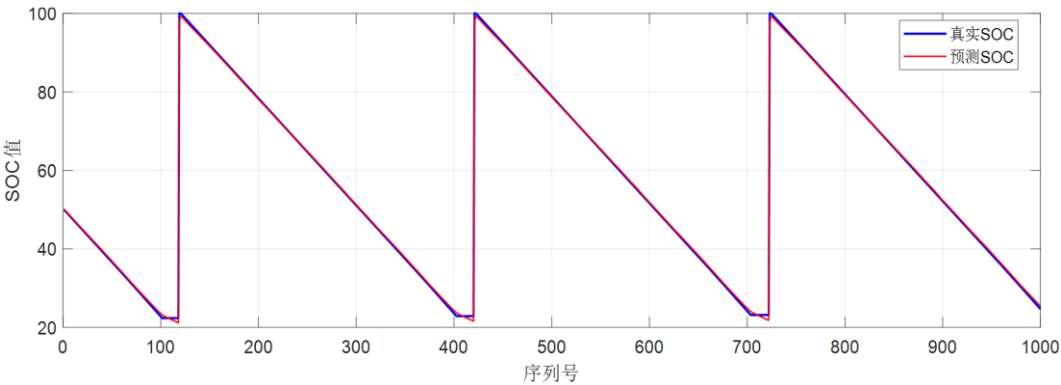


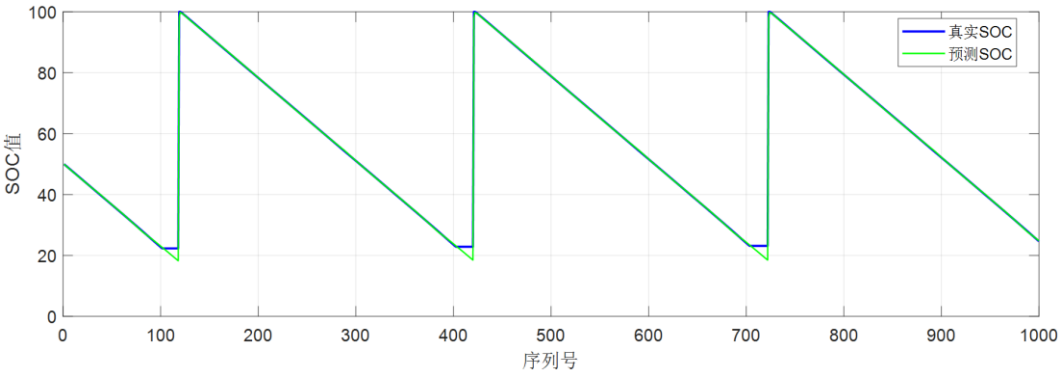
图3 随机森林算法重要性。

从图3的分析结果,温度变换率、放电容量、放电能量以及放电时间与SOC具有强非线性相关性。结合两种重要性分析结果,本文选取放电容量、放电时间以及放电能量(下文简称为三特征参数)作为HHO-LSTM神经网络模型的输出参数。

为验证本文选取特征参数的方法的有效性,选取选取两类特征参数集进行对比,一类为三特征参数集;另一类为本文提出的时间特征参数集,涵盖电压、电流、温度、放电时间。将这两类特征参数集分别输入 LSTM 预测模型,以均方根误差(RMSE)和平均绝对误差(MAE)作为模型预测性能的评价指标,完成SOC的预测对比实验,预测曲线如图4所示。根据表1的预测误差对比显示,基于三特征参数集的 LSTM 模型,其 RMSE 为 0.0179, MAE 为 0.00125;而基于本文时间特征参数集的 LSTM 模型, RMSE 显著降低至 0.00201, MAE 降至 0.00141。对比可知,本文选取的时间特征参数集在 LSTM 模型中表现出更优的预测精度,有效验证了所提特征参数选取方法的有效性,能够为电池状态预测等相关任务提供更具价值的特征支撑。



(a) 三特征模型预测的SOC曲线



(b) 时间特征模型预测的SOC曲线

图4 不同特征参数选取SOC预测对比图。

表1 不同特征参数选取的预测误差对比。

特征参数集	包含数据类型	预测模型	RMSE	MAE
三特征参数	放电容量、放电时间、放电能量	LSTM	0.0179	0.0125
时间特征参数	电压、电流、温度、放电时间	LSTM	0.0201	0.0141

3. 基于HHO-LSTM的SOC估算方法

3.1. HHO算法原理

HHO算法是一种元启发式算法,它利用种群(哈里斯鹰)作为候选解,在迭代过程中逐步引导目标(兔)接近最优解[24]。HHO算法模仿哈里斯鹰在自然界中独特的协作捕食行为,通过探索阶段、探索与开发过渡阶段和开发阶段,根据多种策略提高算法局部搜索能力,有效的平衡全局搜索和局部搜索,相较于传统算法如模拟退火算法、灰狼优化算法等,能够根据搜索进程自动调整搜索策略,在全局范围内寻找最优解的同时,也能对局部区域进行精细搜索,提高解的质量。

3.1.1. 全局探索阶段

哈里斯鹰在整个搜索空间中寻找目标,利用两种策略来发现目标,迭代时用概率 q 来更新位置,计算式为:

$$X(t+1)=\begin{cases} X_{rand}(t)-r_1|X_{rand}(t)-2r_2X(t)| & , q \geq 0.5 \\ (X_{rabbit}(t)-X_m(t))-r_3(lb+r_4(ub-lb)) & , q < 0.5 \end{cases} \quad (1)$$

式中: $X(t+1)$ 为第 $t+1$ 次迭代中鹰的位置, X_{rabbit} 为当下捕食到最优个体的位置; ub 和 lb 分别为搜索过程变

量的上下限; X_{rand} 为当前种群内部随机选择的鹰的位置; $r_1 \sim r_4$ 以及 q 皆为 $[0,1]$ 的均匀随机数; $r_1 \sim r_4$ 为当前种群的平均位置,具体表示为:

$$X_m(t)=\frac{1}{N}\sum_{i=1}^N X_i(t) \quad (2)$$

其中: N 表示哈里斯鹰种群规模。

3.1.2. 探索 and 开发过渡阶段

HHO算法计算猎物的逃逸能量,猎物随逃跑过程而导致逃逸能量的大幅度降低,通过对逃逸能量的比较完成从全局搜索阶段到探索和开发阶段的选择,猎物的逃逸能量计算公式为:

$$E=2E_0(1-\frac{t}{T}), \quad (3)$$

$$E_0=2\beta-1, \quad (4)$$

式中: T 为最大迭代次数; E_0 为逃逸能量初始状态, E_0 是在 $(-1,1)$ 选取的随机数; β 为逃逸概率,是在 $(0,1)$ 选取的随机数; t 为当前进化代数; E 为猎物的逃逸能量。

当逃逸能量满足 $|E| \geq 1$ 时, 算法进入探索阶段; 当逃逸能量 $|E| < 1$ 时, 算法则进入开发阶段。

3.1.3. 开发阶段

在此阶段, 猎物通常试图逃过哈里斯鹰的追捕, 因此结合逃逸能量 E 和逃逸概率 β 形成了4种不同的鹰群围猎猎物的攻击策略。

1) 软包围阶段

此时 $|E| \geq 0.5$, $\beta \geq 0.5$, 此时猎物具有较高的逃逸能量但最终逃逸概率低, 哈里斯鹰对猎物缓慢的接近, 更新位置调整策略如下:

$$\begin{aligned} X(t+1) &= \Delta X(t) - E |JX_{rabbit}(t) - X(t)|, \\ J &= 2(1-r_5) \\ \Delta X(t) &= X_{rabbit}(t) - X(t), \end{aligned} \quad (5)$$

式中: $\Delta X(t)$ 为第 t 次迭代时哈里斯鹰与目标猎物之间的距离; J 为猎物逃跑时的随机跳跃的距离; r_5 为 $[0,2]$ 之间的随机数。

2) 硬包围阶段

此时 $|E| < 0.5$, $\beta \geq 0.5$, 猎物的逃逸能量极低, 鹰群采取具有侵略性的进攻方式, 更新位置调整策略如下:

$$X(t+1) = X_{rabbit}(t) - E |\Delta X(t)| \quad (6)$$

3) 渐进式快速俯冲的软包围阶段

此时 $|E| \geq 0.5$, $\beta < 0.5$, 猎物逃逸能量高并且具有极高的逃逸成功概率, 需要更新软包围阶段的策略, 获取更加高效的围猎猎物的策略, 具体如下:

$$\begin{cases} Y = X_{rabbit}(t) - E |JX_{rabbit}(t) - X(t)| \\ Z = Y + C \times Levy(D) \\ X(t+1) = \begin{cases} Y & F(Y) < F(x(t)) \\ Z & F(Z) < F(x(t)) \end{cases} \end{cases} \quad (7)$$

式中: D 为问题维度, C 的大小为 $1 \times D$, $Levy$ 的飞行函数如下:

$$Levy(D) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\alpha}}} \quad (8)$$

式中: u 和 v 为 $(0,1)$ 选取的随机数; α 为 $Levy$ 分布的形态参数; σ 是与 α 有关的参数尺度。

4) 渐进式快速俯冲的硬包围阶段

此时 $|E| < 0.5$, $\beta < 0.5$, 猎物逃逸能量不足, 哈里斯鹰采取更加迅猛的方式缩短最优位置与平均位置之间的距离, 位置更新策略如下:

$$\begin{aligned} Y &= X_{rabbit}(t) - E |JX_{rabbit}(t) - X_m(t)| \\ X(t+1) &= \begin{cases} Y, & \text{if } F(Y) < F(x(t)) \\ Z, & \text{if } F(Z) < F(x(t)) \end{cases} \end{aligned} \quad (9)$$

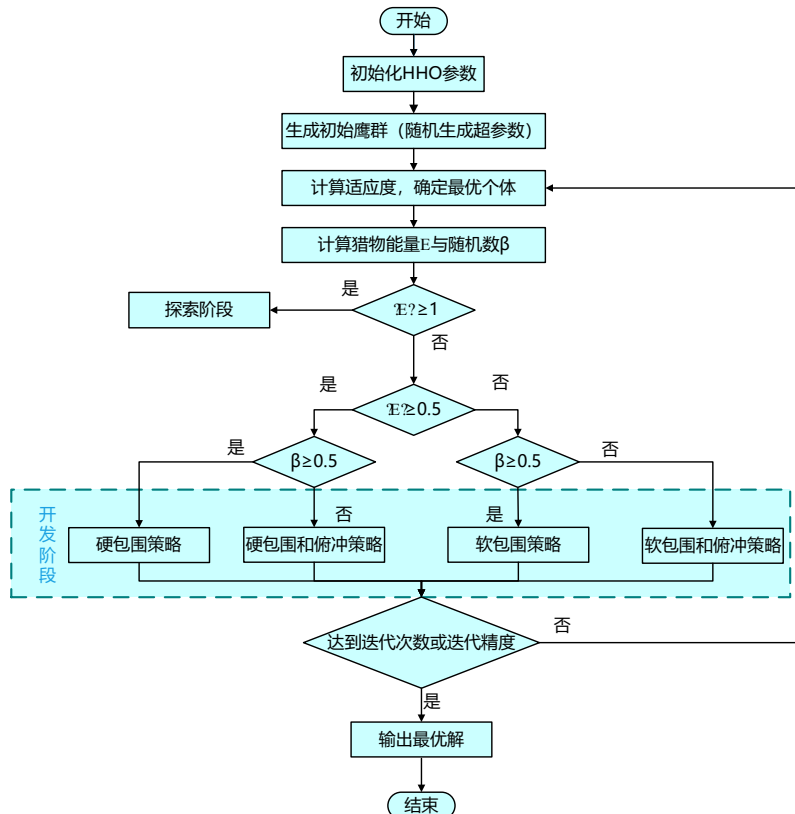


图5 HHO优化算法流程图。

3.2. LSTM神经网络原理

LSTM编码器的核心思想是引入门控机制: 输入门 i_t , 遗忘门 f_t , 和输出门 o_t , 通过门控单元对输入和隐藏状态的控制, 来决定信息的流动和保留[25]。

$$\begin{aligned} i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \end{aligned} \quad (10)$$

其中, σ 表示 *sigmoid* 函数, $W_i x$, $W_f x$ 和 $W_o x$ 分别表示从当前输入 x_t 到各个门的权重矩阵; U_{ix} , U_{fx} 和 U_{ox} 表示从上一步的隐藏状态 h_{t-1} 到各个门的权重矩阵; b_{ix} , b_{fx} 和 b_{ox} 表示每个门的偏置向量。在门控信号计算后, 细胞状态 C_t 通过结合遗忘门和输入门的输出进行更新, 隐藏状态 h_t 是通过输出门和当前细胞状态的非线性变换产生[26]。

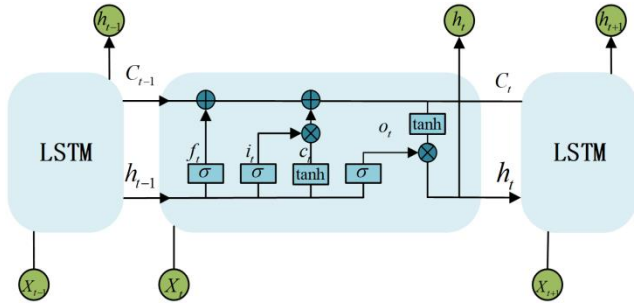


图6. LSTM结构原理图。

$$\begin{aligned} \tilde{C}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \end{aligned} \quad (11)$$

$$\begin{aligned} \tilde{C}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \end{aligned} \quad (11)$$

池化后的特征 X_{pooled} 作为LSTM的输入, 结合其隐藏状态的更新公式, 具体公式如下:

$$\begin{aligned} h_t &= o_t \cdot \tanh(C_t) \\ h_t &= \text{LSTM}(x_{pooled}, h_{t-1}) \end{aligned} \quad (12)$$

3.3. HHO-LSTM模型预测算法

HHO-LSTM综合预测模型主要由LSTM网络和HHO优化算法组成, 其主题的模型结构如图N所示。由于传统的LSTM网络应用于SOC预测时存在明显的缺陷, 例如: LSTM的学习率、隐藏层单元数、Dropout率等超参数对预测结果影响极大, 人工调参(如网格搜索、随机搜索)效率低且易陷入局部最优; 传统的LSTM网络在面对锂电池放电数据存在非线性、时变噪声(如温度波动、电流波动)等状况时, 往往难以快速适配数据特征; 并且训练模型使用的锂电池循环数据在样本有限的情况下, 使用LSTM网络易出现过拟合的现象, 且泛化能力差。因此选用HHO优化算法, 初始化鹰群位置, 将LSTM网络中所需要的学习率、隐藏层单元数、Dropout率、L2正则化系数等关键超参数作为围捕的猎物位置, 利用HHO优化算法输出的最优超参数, 完成对构建的LSTM神经网络的训练。使用HHO算法完成对LSTM网络超参数的优化, 首先需要明确定义超参数的取值范围, 学习率将在(0.01,0.0001)的范围内训练; 隐藏层单元数在[32,128]的范围内完成训练; Dropout的训练范围在[0.1,0.5]; L2正则化系数的训练范围在 $[10^{-5}, 10^{-3}]$ 。HHO优化算法采用均方根误差作为损失函数, 确保优化目标的统一, 通过迭代鹰的位置输出LSTM网络的最优超参数组合。

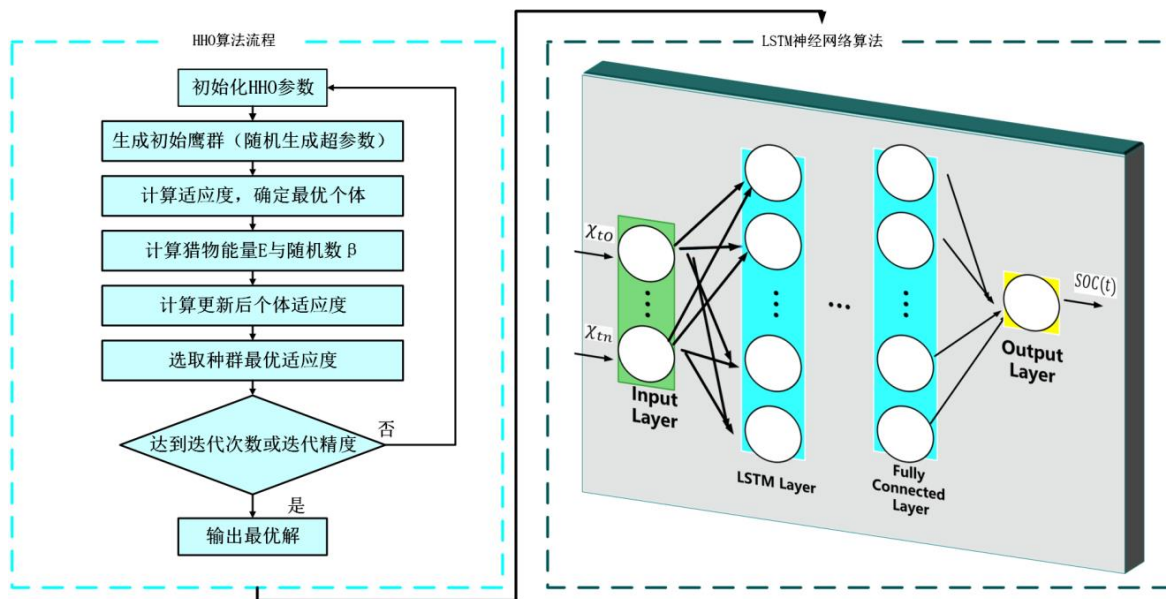


图7 HHO-LSTM流程原理图。

4. 仿真实验与结果分析

4.1. 评价指标

为了更直观的不同模型之间的性能指标，采用了三种评估指标：均方根误差（ E_{RMSE} ）、平均绝对误差（ E_{MAE} ）和决定系数（ R^2 ），基本公式如下：

$$E_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (13)$$

$$E_{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (14)$$

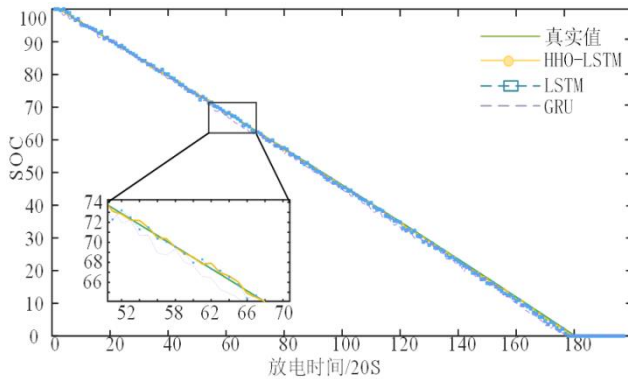
$$R^2 = 1 - \frac{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2} \quad (15)$$

式中： y_i 表示第*i*个样本的真实值； $\hat{y}_i$ 表示第*i*个样本的预测值； N 表示样本总数； $\bar{y}$ 表示数据样本中所有SOC的均值。其中 E_{MAE} 和 E_{RMSE} 都能不同层面的反映预测误差的大小，数值越小预测越准确； R^2 表示模型的与实际过程的相关程度， R^2 越高预测精度越高。各种模型根据评价指标，结果如图9以及表2所示。

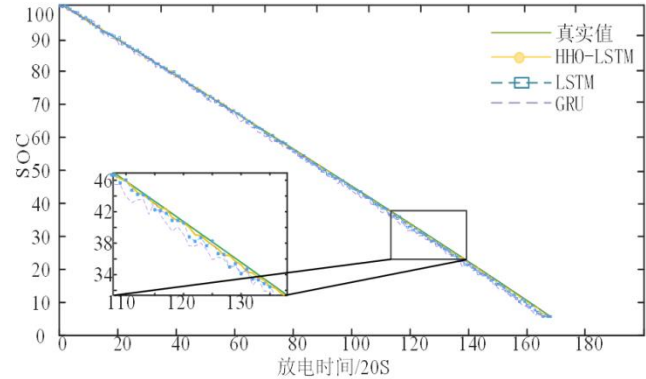
4.2. 模型预测结果与分析

为探究不同深度学习模型在锂电池SOC预测任务中的性能差异，本文选取5、6、7和18号四组电池放电数据展开实验，对比分析HHO-LSTM、LSTM与GRU（Gated Recurrent Unit）三种模型的SOC预测效果，通过评价指标评估预测效果，如图8。

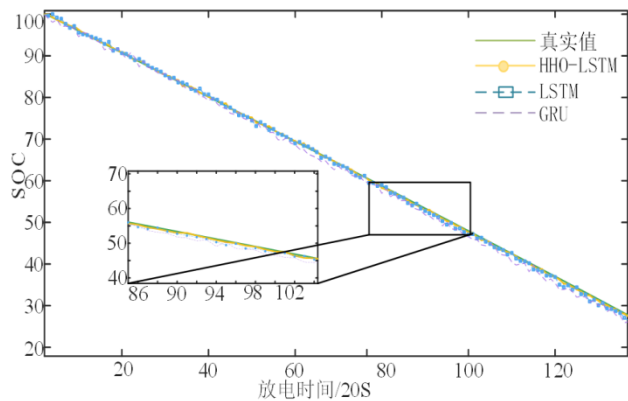
从实际新能源船舶航行过程出发，放电过程是电池发挥效能的核心阶段，此时对SOC的实时预测直接关系到设备的续航判断和安全控制。新能源船舶储能系统放电时，通过放电数据预测SOC可防止过放导致的电池寿命衰减。此外，放电数据的采集成本低、易获取，更适合在实时监测系统中使用，满足工程化应用需求。从实验结果来看，在四组电池放电过程的SOC预测中，HHO-LSTM模型展现出显著更优的预测精度。以SOC随放电时间变化的趋势为例，真实值呈现出平稳且规律的下降态势，HHO-LSTM模型的预测曲线几乎与真实值曲线重合，偏差极小；LSTM模型的预测曲线虽整体能跟随真实值的变化趋势，但与真实值存在一定程度的偏离，误差明显大于HHO-LSTM；GRU网络的预测曲线偏离真实值的程度更为显著，在局部区间（如各子图放大区域所示），与真实值的差距尤为突出。这种性能差异源于模型的结构与优化机制。LSTM通过门控机制一定程度上缓解了长序列依赖问题，但在参数优化上缺乏针对性；GRU简化了LSTM的门控结构，虽提升了计算效率，却进一步损失了对复杂序列特征的捕捉能力，导致预测误差较大。



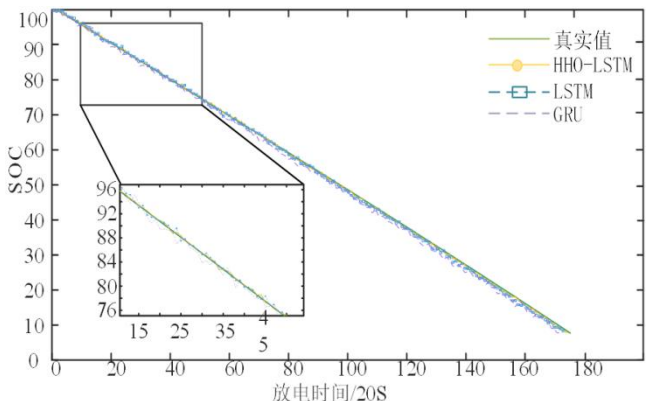
(a) b005电池SOC预测结果



(b) b006电池SOC预测结果



(c) b007电池SOC预测结果



(d) b018电池SOC预测结果

图8. SOC预测对比图。

表2 不同模型的预测误差对比。

模型	E_{RMSE}	E_{MAE}	R^2
GRU	2.3 1%	1.9 5%	0.9 8 0
LSTM	1.8 9%	1.5 0%	0.9 8 5
HHO-LSTM	0.9 1%	0.7 5%	0.9 9 5

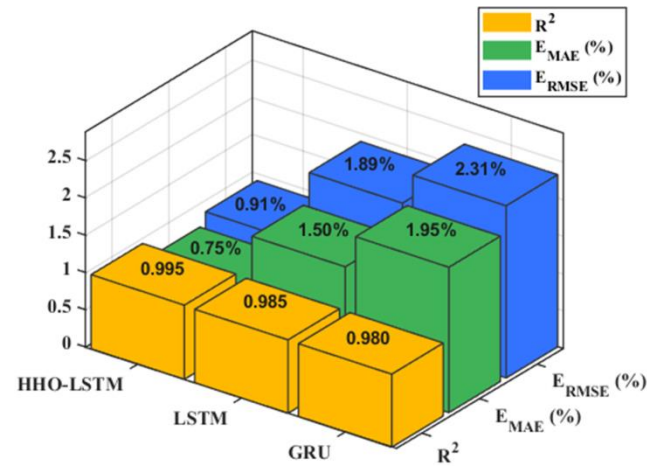


图9 评价指标柱状图。

由表2可知,GRU算法、LSTM算法和HHO-LSTM算法,在使用同一数据集所得到的 E_{RMSE} 分别为2.31%、1.89%、0.91%; E_{MAE} 分别为1.95%、1.50%、0.75%; R^2 分别为0.980、0.985、0.995。从图表的数据结果分析,HHO-LSTM模型在经过哈里斯鹰算法对超参数的优化后, E_{RMSE} 、 E_{MAE} 上显著低于GRU和LSTM模型,将误差限制在0.1%以内;同时 R^2 更接近于1,说明其预测值与真实值的拟合程度更高。在四组电池数据的验证中HHO-LSTM算法均表现出比LSTM和GRU更卓越的预测性能。由此可见,本文加入HHO优化算法,不仅保证了LSTM作为经典的序列建模算法,能捕捉电池数据的时序依赖的优点,而且避免出现超参数设置不合理时预测精度有限的问题,能够在实际电池SOC预测中,帮助LSTM更精准地拟合电池真实SOC变化,为新能源船舶储能系统等领域的电池管理提供可靠的状态监测。根据表2以及图9清晰且直观地揭示了模型性能的优化成效,验证了本文所提HHO-LSTM模型在SOC预测中的可行性。

5. 结论

为了提高锂电池SOC估计的准确性,本文提出了一种基于改进的HHO-LSTM网络的锂电池SOC估计算法,经过实验验证及理论分析。主要结论如下:

- (1) 采用随机森林算法与皮尔逊相关系数进行特征参数重要性分析以筛选特征参数,捕捉特征参数与SOC的线性和非线性关系,弥补单一时间特征在时序关联性捕捉上的不足,所选特征参数在锂电池SOC预测任务中体现出更优的性能,为后续精准的SOC预测打下数据基础。
- (2) 提出HHO-LSTM模型,使用哈里斯鹰优化算法对模型超参进行寻优,选用隐藏层单元数、学习率、Dropout率等超参数进行组合,确定参数,避免传统LSTM模型、GRU模型易陷入局部最优的局限。

(3) 构建的 HHO - LSTM 模型用于锂电池SOC预测,并利用 NASA 数据集进行测试验证。通过与 LSTM 算法、GRU 算法的实验结果对比可知,本文模型为锂电池SOC的准确预测提供了一种可行且有效的方法。

HHO - LSTM 模型充分发挥了哈里斯鹰优化算法对长短期记忆网络参数的优化能力,有效捕捉锂电池SOC的时序变化规律,在预测精度等方面展现出良好的性能,可直接支撑电池管理系统(BMS)的充放电控制与剩余寿命预警,对提升锂电池使用安全性、延长使用周期,以及推动新能源船舶设备储能效率优化具有一定的应用价值。

参考文献

[1] I Dongliang, MAO Qian, et al. Scient metric research and critic analysis of battery state of-charge estimation [J]. Journal of Energy Storage, 2023, 58:106283

[2] 叶丽华, 薛定邦, 彭鼎涵, 等.基于分数阶模型与IEKF的锂离子电池 SOC估计研究 [J]. 农业装备与车辆工程, 2023, 61 (6): 7-10

[3] 周韦润, 姜文刚.基于遗传算法优化扩展卡尔曼滤波的锂电池SOC估计 [J]. 重庆理工大学学报 (自然科学版), 2019, 33 (9): 33 - 39.

[4] 王位.基于平均电极反应动力学的锂离子动力电池模型降阶与SOC估计研究 [D]. 江苏大学2017.

[5] 杜嘉诚. 基于滑模观测器的锂离子电池荷电状态与健康状态估计算法研究 [D].哈尔滨工业大学, 2018.

[6] 骆秀江, 张兵, 黄细霞, 等.基于SVM的锂电池SOC估算 [J]. 电源技术, 2016, 40(02): 287-290.

[7] 徐帅, 刘雨辰, 周飞.基于 RNN 的锂离子电池SOC估算研究进展 [J]. 电源技术, 2021, 45 (2): 263 - 269.

[8] Zhang Z, Xu M, Ma L, et al. A state - of - charge estimation method based on bidirectional LSTM networks for lithium - ion batteries [C]. IEEE International Conference, 2020.

[9] Yang, Yalong, Chen, Siyuan, Chen, Tao, et al. State of Health Assessment of Lithium-ion Batteries Based on Deep Gaussian Process Regression Considering Heterogeneous Features [J]. Journal of Energy Storage, 2023, 61.

[10] Wei, Yupeng, Wu, Dazhong. Prediction of state of health and remaining useful life of lithium-ion battery using graph convolutional network with dual attention mechanisms [J]. Reliability Engineering & System Safety, 2023, 230.

[11] 李超然, 肖飞, 樊亚翔, 等.基于深度学习的锂离子电池SOC和SOH联合估算 [J]. 中国电机工程学报, 2021, 41 (2): 681 - 691.

[12] 郑青根, 杨祥国, 刘冬, 等.改进灰狼优化最小二乘法支持向量机的锂电池剩余寿命预测 [J]. 重庆大学学报, 2023, 46 (11): 78-89.

[13] Fan Y, Xiao F, Li C, et al. A novel deep learning framework for state of health estimation of lithium - ion battery [J]. Journal of Energy Storage, 2020, 32: 101741.

- [14] Liu S, Dong X, Yu X, et al. A method for state of charge and state of health estimation of lithium - ion battery based on adaptive unscented Kalman filter [J]. *Energy Reports*, 2022, 8:426 - 436.
- [15] 常宇佳. 基于ICA与SVR的锂离子电池健康状态估计研究 [D]. 哈尔滨理工大学, 2021.
- [16] Hong S, Kang M, Kim J, et al. Sequential application of denoising autoencoder and long - short recurrent convolutional network for noise - robust remaining - useful - life prediction framework of lithium - ion batteries [J]. *Computers & Industrial Engineering*, 2023, 179:109231.
- [17] Qiyan Yan. SOC Prediction of Power Battery Based on SVM [C]. 2020:2425-2429.
- [18] Li H, Yao Y, Hou J, et al. GPR - Bi - LSTM power battery health state estimation and remaining life prediction based on ICEEMDAN algorithm [J]. 2022.
- [19] Gao M, Bao Z, Zhu C, et al. HFCM - LSTM: A novel hybrid framework for state - of - health estimation of lithium - ion battery [J]. *Energy Reports*, 2023, 9:2577 - 2590.
- [20] Ren Z, Du C. A review of machine learning state - of - charge and state - of - health estimation algorithms for lithium - ion batteries [J]. *Energy Reports*, 2023, 9: 2993 - 3021.
- [21] Hossain Lipu M S, Ansari S, Miah M S, et al. Deep learning enabled state of charge, state of health and remaining useful life estimation for smart battery management system: methods, implementations, issues and prospects [J]. *Journal of Energy Storage*, 2022, 55:105752.
- [22] Chai, Xuqing, Li, Shihao, Liang, Fengwei. A novel battery SOC estimation method based on random search optimized LSTM neural network [J]. *ENERGY*, 2024, 306. <https://doi.org/10.1016/j.energy.2024.132583>
- [23] Hong, Jichao, Wang, Zhenpo, Chen, Wen, et al. Online joint-prediction of multi-forward-step battery SOC using LSTM neural networks and multiple linear regression for real-world electric vehicles [J]. *JOURNAL OF ENERGY STORAGE*, 2020, 30. <https://doi.org/10.1016/j.est.2020.101459>
- [24] 王晓峰, 王军霞, 彭庆媛, 等. 融合多策略的哈里斯鹰优化算法求解Steiner树问题 [J/OL]. *郑州大学学报(理学版)*, 1-8 [2025-10-05]. <https://doi.org/10.13705/j.issn.1671-6841.2024174>
- [25] SHERSTINSKY A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network [J]. *Physica D: Nonlinear Phenomena*, 2020, 404: 132306.
- [26] 于天剑, 曾笑颜, 冯恩来, 等. 基于LSTM-Transformer多通道特征融合的锂电池SOC-SOH联合估计 [J/OL]. *铁道科学与工程学报*, 1-13, 2025-10-06.