DOI：10.13234/j.issn.2095-2805.2026.3.242　　**中图分类号**：TM912　　**文献标志码**：A

基于 EMD-LSTM-Attention 的锂离子电池 SOC 预测

朱　博，刘雪芹，王宇珲，李　涛，程观银
(重庆理工大学理学院，重庆 400054)

摘要：锂离子电池荷电状态SOC(state-of-charge)预测是电池管理系统中重要的一环，精确地预测锂离子电池SOC对保障电池安全运行具有重要意义。针对锂离子电池SOC的预测问题，利用经验模态分解EMD(empirical mode decomposition)、长短期记忆LSTM(long short-term memory)循环神经网络与注意力(Attention)机制搭建锂离子电池SOC预测模型。以锂离子电池充放电过程中的电压、电流、表面温度与经过EMD分解后的电压作为模型的特征；模型的LSTM层能够从锂离子电池历史循环数据中捕捉电池时间信息，并总结出长期的依赖关系；模型的注意力层能够捕捉特征输入和SOC输出之间的非线性和动态关系，优化各循环阶段和各个特征之间的权重关系，进一步提高SOC预测精度。使用马里兰大学公共电池数据集进行实验，结果表明与仅使用LSTM的模型相比，所提模型在3种不同工况下的平均预测误差分别降低了49.36%、45.61%、46.93%。

关键词：锂离子电池；荷电状态；经验模态分解；长短期记忆神经网络；注意力机制

SOC Prediction for Lithium-ion Batteries Based on EMD-LSTM-Attention

ZHU Bo, LIU Xueqin, WANG Yuhui, LI Tao, CHENG Guanyin
(School of Science, Chongqing University of Technology, Chongqing 400054, China)

Abstract: The state-of-charge (SOC) prediction for lithium-ion batteries is an important part of a battery management system, and accurately predicting their SOC is of significance for ensuring their safe operation. On this basis, empirical mode decomposition (EMD), long short-term memory (LSTM) recurrent neural network and Attention mechanism are used to build an SOC prediction model for a lithium-ion battery. The voltage, current, surface temperature and the voltage decomposed by EMD in its charging and discharging processes are taken as the characteristics of the model. The LSTM layer of the model can capture the battery time information from the historical cycle data of the lithium-ion battery, and the long-term dependence relationship is also summarized. The Attention layer of the model can capture the nonlinear and dynamic relationship between feature input and SOC output, so as to optimize the weight relationship between each cycle stage and each feature and further improve the SOC prediction accuracy. Experiments were carried out based on the public battery data set of University of Maryland, and results show that compared with that of a model which only uses LSTM, the average prediction errors of the proposed model under three different working conditions were reduced by 49.36%, 45.61% and 46.93%, respectively.

Keywords: Lithium-ion battery; state-of-charge (SOC); empirical mode decomposition (EMD); long short-term memory (LSTM) neural network; Attention mechanism

收稿日期：2023-02-02；**修回日期**：2023-09-14；**录用日期**：2023-12-13；**网络首发日期**：2024-04-20
基金项目：国家自然科学基金资助项目(11774041)；重庆市基础与前沿研究计划资助项目(cstc2015jcyjA50033)；重庆理工大学研究生教育高质量发展资助项目(gzlcx20222079)
This work is supported by National Natural Science Foundation of China under the grant 11774041; Chongqing Research Program of Basic Research and Frontier Technology under the grant cstc2015jcyjA50033; Chongqing University of Technology Graduate Education High Quality Development Program under the grant gzlcx20222079

当前全球能源危机日益严重，社会对环境保护逐渐重视。电动汽车的快速发展，全球电动汽车保有量占比逐步上升，大幅度减少了由燃油汽车排放

的温室气体，有效缓解了当前的能源危机与环境恶化的压力[1]。锂离子电池凭借其制造成本低、能量密度高、自放电率低、使用寿命长等优点，已成为电动汽车最常见的储能载体[2]。锂离子电池荷电状态SOC(state-of-charge)预测是电池管理系统中重要的一环，准确的 SOC 预测可以有效避免过充电或过放电情况发生，对于电动汽车的安全和性能维护至关重要，还能给驾驶人员制定驾驶策略提供有用的信息[3]。

但是电池的SOC不能被直接测量[4]，只能通过传感器测量得到间接参数(电流、电压、温度等)，然后利用已有间接参数来预测电池的 SOC。目前已有多种针对锂离子电池 SOC 预测的方法。安时积分法和开路电压法被广泛使用，但由于抗噪声能力较弱[5]，且电池使用后需要静置至稳态才能计算 SOC 值，不适用于在线预测；基于模型的方法旨在通过建立复杂的数学方程来模拟电池行为，这种方法通常需要特定领域的知识和复杂的程序[6-7]，且该方法高度依赖于电池模型的准确度，由于电池的复杂动态特性，以及温度和老化的影响，特别是一些电池的开路电压对 SOC 变化不敏感，因此在实际应用中难以获得准确的电池模型[8-9]。

近年来，随着硬件性能的提高，数据驱动的方法取得了长足的进步。深度学习算法可用于探究采样数据与SOC 之间的非线性映射[10-11]。本文提出了带有注意力机制的经验模态分解与长短期记忆神经网络 EMD-LSTM-Attention(empirical mode decomposition-long short-term memory neural network-Attention mechanism)联合预测模型。该模型不需要考虑电池复杂的内部原理和反应机理，能够在有限的领域知识内模拟电池单元的复杂非线性动态特性。模型加入了注意力层，来增强其专注于长输入序列的关键特征的能力。注意力层通过权重分配来突出重要特征对神经网络输出的影响。注意力层能够约束 LSTM 相邻输出 SOC 之间的变化，使网络输出平滑，进一步提高 SOC 预测性能。

针对模型输入，锂离子电池初始电压是 1 个关于时间的持续动态信号，因此可以应用信号处理的时频分析法对电压值进行处理。对于平稳信号，可以使用傅里叶变换进行处理，例如快速傅里叶变换与离散傅里叶变换[12]。但在锂离子电池的实际使用过程中，电压变化是非平稳的，因此电压是 1 个非平稳的动态信号。对于这种非平稳信号频谱的分析，目前主流方式有窗口傅里叶变换和小波分解，但此类变换方式都存在明显不足，易受到人为因素影响，如窗口宽度与基函数的设置等[13-14]，因此针对非平稳信号构建无人工参数的模态分解方法十分必要。经验模态分解算法在处理锂离子电池电压这类非平稳信号时具有良好的效果。本文引入额外的 EMD 电压作为 1 个特征，以增强网络映射电池非线性特性的能力，减少输出 SOC 波动，进一步提高 SOC 预测精度。

1　EMD-LSTM-Attention 模型

本文采用 EMD-LSTM-Attention 联合预测模型，其结构如图 1 所示。首先 EMD 算法部分对锂离子电池数据进行经验模态分解，将分解后的时间序列数据输入带有注意力机制的 LSTM 层，然后由带有注意力机制的 LSTM 层提取数据序列的时间特征。整个模型结合了 3 种算法的优势，EMD 算法部分对数据集中的非平稳时间序列数据进行自适应分解，然后得到平稳的时间序列数据，LSTM 部分有效解决了长距离依赖序列的问题，带有注意力机制的 LSTM 层给予不同时间戳的数据不同的权重，使模型能够优先关注到对当前时间 SOC 影响较大的充放电数据。实验结果表明，整个模型对数据的捕捉学习能力得到了大幅提高，预测精度得到进一步提升。

1.1　经验模态分解算法

EMD 算法由 Huang 等[15]于 1998 年提出。EMD 算法能够根据数据本身的时间尺度特性进行信号分解，不需要人为设置分解窗口宽度与基函数，具有较强的自适应性。由于省去了设置基函数这一人为过程，EMD 使信号分解更简单、更高效。EMD 可用于序列信号分解，是 1 种较好的非线性与非平

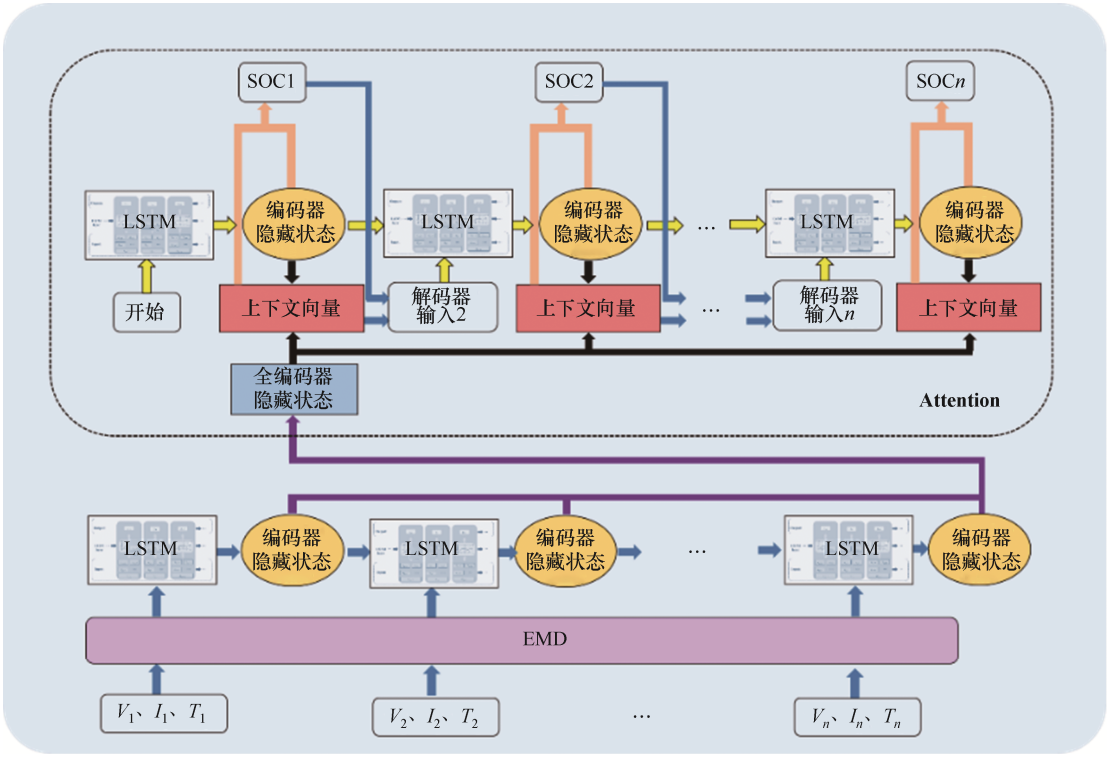


图 1 EMD-LSTM-Attention 模型结构

Fig. 1 Structure of EMD-LSTM-Attention model

稳信号频谱分析工具，在非平稳信号处理上具有显著优势。

1.2 长短期记忆神经网络

LSTM 是由递归神经网络 RNN(recurrent neural network)改进后的 1 种神经网络。与其他 RNN 一样，LSTM 能够有效处理时间序列数据。与传统 RNN 相比，LSTM 能够在训练过程中大幅减少梯度消失现象，有效解决长距离依赖序列的问题，并表现出更高的预测性能。其具体计算过程为

$$f_t=\sigma(\boldsymbol{w}_f[h_{t-1},x_t]+\boldsymbol{b}_f) \tag{1}$$

$$i_t=\sigma(\boldsymbol{w}_i[h_{t-1},x_t]+\boldsymbol{b}_i) \tag{2}$$

$$\tilde{c}_t=\tanh(\boldsymbol{w}_c[h_{t-1},x_t]+\boldsymbol{b}_c) \tag{3}$$

$$c_t=f_tc_{t-1}+i_t\tilde{c}_t \tag{4}$$

$$o_t=\sigma(\boldsymbol{w}_o[h_{t-1},x_t]+\boldsymbol{b}_o) \tag{5}$$

$$h_t=o_t\cdot\tanh(c_t) \tag{6}$$

式中：x_t 为当前时间步的输入；h_{t-1} 与 h_t 分别为上一时间步与当前时间步的隐藏状态；c_{t-1} 与 c_t 分别为上一时间步与当前时间步的记忆状态；$\tilde{c}_t$ 为当前时间步的单元状态；f_t、i_t、o_t 分别为遗忘门、输入门、输出门，f_t 决定了上一时间步的记忆状态 c_{t-1} 遗忘的程度，其取值区间为[0, 1]，f_t=0 代表遗忘 c_{t-1} 中的所有信息，f_t=1 代表 c_{t-1} 中的所有信息完全不变，i_t 控制当前状态信息的更新程度，结合 h_{t-1} 与 x_t 将 c_{t-1} 更新为 c_t，o_t 控制哪些信息用以输出，得到当前时间步的隐藏状态 h_t；σ(·) 与 tanh(·) 为传输函数；$\boldsymbol{w}$ 和 $\boldsymbol{b}$ 分别为上述门的权重矩阵和偏置向量。

1.3 注意力机制

LSTM 虽然能够有效减少梯度消失和梯度爆炸对长时间序列的影响，但是并不能完全消除影响。因此在模型中引入注意力(Attention)层，使模型既能发挥 LSTM 的长时间捕捉的优势，又能够利用注意力机制对权重进行分配。在训练过程中，大量的特征数据将被输入到深度学习模型中。带有注意力机制的模型可以有效地访问输入序列中所有时间步长的状态，并且训练模型从数据中筛选出对输出有显著影响的特征，并增加这些特征的权重，为最相关的特征指定更高的权重，同时降低不重要特征的权重。其具体计算过程为

$$s(\boldsymbol{h}_{\mathrm{m}}, \boldsymbol{h}_t) = \boldsymbol{w}_{\mathrm{s}}\boldsymbol{h}_t^{\mathrm{T}} + \boldsymbol{b}_{\mathrm{s}} \tag{7}$$

$$s(\boldsymbol{h}_{\mathrm{m}}, \boldsymbol{h}_t) = \boldsymbol{h}_{\mathrm{m}}\boldsymbol{h}_t^{\mathrm{T}} \tag{8}$$

$$\alpha_{\mathrm{i}} = \frac{\exp(s(\boldsymbol{h}_{\mathrm{m}}, \boldsymbol{h}_t))}{\sum_{j=1}^{N}\exp(s(\boldsymbol{h}_{\mathrm{m}}, \boldsymbol{h}_t))} \tag{9}$$

$$h_t^* = \sum_{i=1}^{N}\alpha_{\mathrm{i}}\boldsymbol{h}_{\mathrm{m}} \tag{10}$$

$$\hat{y}_t = \boldsymbol{w}h_t^* + \boldsymbol{b} \tag{11}$$

式中：$\boldsymbol{h}_{\mathrm{m}}$ 为特征向量；$\boldsymbol{h}_t$ 为当前时间步的隐藏状态向量；$s(\boldsymbol{h}_{\mathrm{m}}, \boldsymbol{h}_t)$ 为采用内积运算的注意力打分函数；$\boldsymbol{w}_{\mathrm{s}}$ 与 $\boldsymbol{b}_{\mathrm{s}}$ 分别为全连接层的权重矩阵与偏置向量；α_{i} 为注意力权重；h_t^* 为注意力层输出；$\hat{y}_t$ 为长序列 SOC 的预测值；$\boldsymbol{w}$ 与 $\boldsymbol{b}$ 分别为预测运算中的权重矩阵与偏置向量。

2　锂离子电池数据集

采用马里兰大学高级生命周期工程中心 CALCE(center for advanced life cycle engineering)数据集[16]对所提出的 SOC 预测模型进行验证。CALCE 数据集包含了A123型 $LiFePO_4$ 电池在 3 个环境温度(0 ℃、25 ℃、40 ℃)下的充放电数据，包括充放电电压、工作电流、电池温度等。每个温度下均针对 3 种测试工况进行测试，3 种工况分别为动态压力测试 DST(dynamic stress test)工况、高速公路驾驶工况(US06)和联邦城市驾驶 FUDS (federal urban driving schedule)工况。3 种工况分别模拟了不同的驾驶环境，能够较好地模拟真实驾驶情况，适合用于锂离子电池 SOC 预测模型的验证。

3　结果与分析

首先使用第 2 节所提 A123 型电池数据集中 3 种温度数据中的 2 种作为训练集，剩余 1 种作为测试集进行训练与验证。图 2~图 4 分别为 DST、US06、FUDS 工况下，使用 0 ℃与 40 ℃数据作为训练集，25 ℃数据作为测试集的预测结果。根据实验结果可以看出，模型在不同工况下都有较好的预测效果。

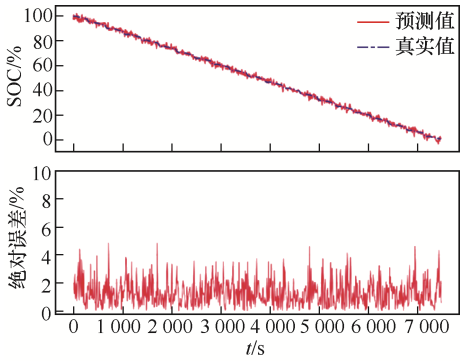


图 2　DST 工况 25 ℃ SOC 预测结果与误差

Fig. 2 SOC prediction results and errors under DST condition and 25 ℃

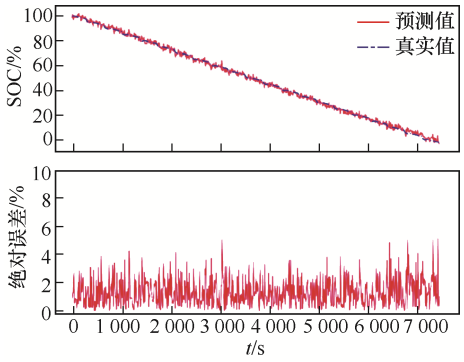


图 3　US06 工况 25 ℃ SOC 预测结果与误差

Fig. 3 SOC prediction results and errors under US06 condition and 25 ℃

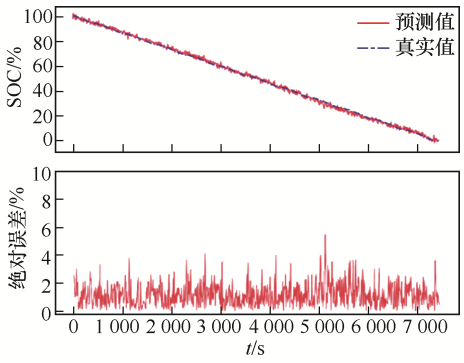


图 4　FUDS 工况 25 ℃ SOC 预测结果与误差

Fig. 4 SOC prediction results and errors under FUDS condition and 25 ℃

分别以 DST、US06、FUDS 工况数据作为测试集的预测结果均方根误差 RMSE(root mean square error)、平均绝对误差 MAE(mean absolute error)及 R^2 见表 1。可见，DST 工况下的最低 RMSE

为 1.556 9%，US06 工况下的最低 RMSE 为 1.354 9%，FUDS 工况下的最低 RMSE 为 1.613 3%。

表 1 同工况不同温度下的预测误差

Tab. 1 Prediction errors under the same condition and different temperatures

训练集	测试集	RMSE/%	MAE/%	R^2/%
DST(25 ℃)+DST(40 ℃)	DST(0 ℃)	1.624 3	1.276 4	98.65
DST(0 ℃)+DST(40 ℃)	DST(25 ℃)	1.590 1	1.286 3	98.70
DST(0 ℃)+DST(25 ℃)	DST(40 ℃)	1.556 9	1.234 7	98.75
US06(25 ℃)+US06(40 ℃)	US06(0 ℃)	1.552 9	1.231 4	98.56
US06(0 ℃)+US06(40 ℃)	US06(25 ℃)	1.438 9	1.160 2	98.65
US06(0 ℃)+US06(25 ℃)	US06(40 ℃)	1.354 9	1.075 8	98.78
FUDS(25 ℃)+FUDS(40 ℃)	FUDS(0 ℃)	1.720 8	1.208 8	98.59
FUDS(0 ℃)+FUDS(40 ℃)	FUDS(25 ℃)	1.655 9	1.326 7	98.73
FUDS(0 ℃)+FUDS(25 ℃)	FUDS(40 ℃)	1.613 3	1.548 9	98.85

本文还验证了 25 ℃下，使用其中 2 种工况的数据作为训练集，剩余 1 种工况数据作为测试集情况下模型的预测性能。同时对比了 25 ℃下使用不同工况数据作为训练集与测试集，采用所提模型与仅使用 LSTM 网络进行 SOC 预测的结果，其中 LSTM 网络参数设置与本模型的 LSTM 层参数保持一致，实验结果如图 5~图 7 所示。可以看出，本文提出的 EMD-LSTM-Attention 模型在各个工况下都具有较好的预测效果。

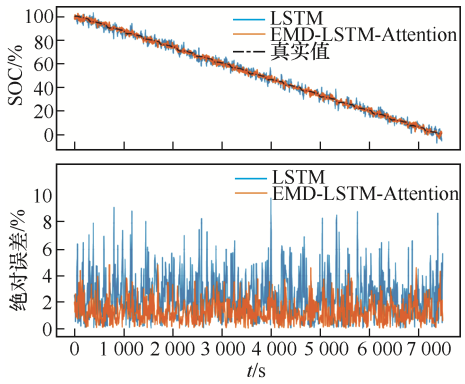


图 5 DST 工况 SOC 预测结果与误差

Fig. 5 SOC prediction results and errors under DST condition

LSTM 与所提模型在 25 ℃不同工况的预测误差对比见表 2。US06 数据与 FUDS 数据作为训练集，DST 数据作为测试集的情况下，所提模型的 RMSE 为 1.406 5%，MAE 为 1.127 4%。US06 数据作为测试集的情况下，所提模型的 RMSE 为 1.346 7%，MAE 为 1.074 4%。FUDS 数据作为测试集的情况下，所提模型的 RMSE 为 1.368 5%，

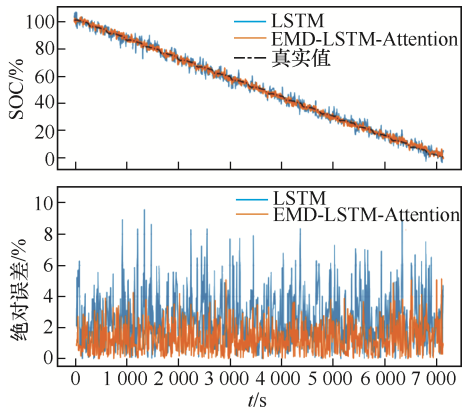


图 6 US06 工况 SOC 预测结果与误差

Fig. 6 SOC prediction results and errors under US06 condition

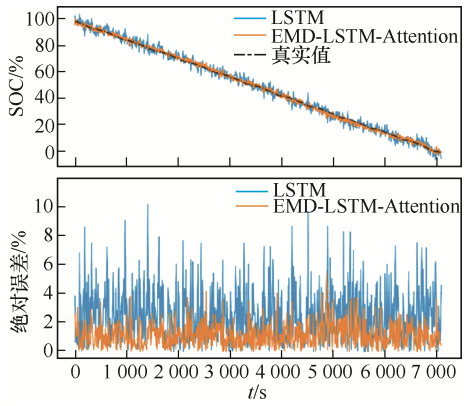


图 7 FUDS 工况 SOC 预测结果与误差

Fig. 7 SOC prediction results and errors under FUDS condition

MAE 为 1.048 6%。相比传统的单一 LSTM 模型，本文所提模型的预测精度更高，在不同工况下的 RMSE 平均值为 1.373 9%，MAE 平均值为 1.083 5%。所提模型在 3 种工况数据集的 SOC 预测值的 RMSE 对比 LSTM 模型分别降低了 39.02%、39.82%、31.18%，平均降低 36.67%；MAE 分别降低了 49.36%、45.61%、46.93%，平均降低 47.30%。

表 2　同温度不同工况下的预测误差

Tab. 2 Prediction errors under different working conditions and the same temperature

训练集	测试集	模型	RMSE/%	MAE/%	R^2/%
US06+FUDS	DST	LSTM	2.306 4	2.226 3	96.43
		EMD-LSTM-Attention	1.406 5	1.127 4	98.52
DST+FUDS	US06	LSTM	2.237 6	1.975 2	95.86
		EMD-LSTM-Attention	1.346 7	1.074 4	98.62
DST+US06	FUDS	LSTM	1.988 4	1.975 8	96.97
		EMD-LSTM-Attention	1.368 5	1.048 6	98.65

4　结语

锂离子电池 SOC 预测是电池管理系统中重要的一环，高效准确地预测锂离子电池 SOC 对保障电池安全运行具有重要的意义。为了进一步提高锂离子电池 SOC 预测准确率，提出了基于 EMD-LSTM-Attention 的锂离子电池 SOC 预测模型。模型利用了经验模态分解算法在非平稳信号处理上的优势，将锂离子电池电压数据分解，并将分解后的电压曲线作为 1 个新的特征量。将电压、电流、温度、分解后的电压作为 SOC 预测模型的输入，以 SOC 值作为输出，直接预测锂离子电池 SOC 值。模型的 LSTM 层能够对历史电池数据进行学习，得到输入与输出的非线性对应关系。注意力层的加入能够使其模型注意到对当前 SOC 影响较大的历史数据，重新分配权重进行参数学习，进一步提高 SOC 的预测精度。使用 CALCE 数据集验证了所提出的 EMD-LSTM-Attention 深度学习模型，实验结果表明，本文提出的 EMD-LSTM-Attention 模型在不同工况下的平均绝对误差为 1.083 5%，均方根误差为 1.373 9%，相较于传统的 LSTM 模型具有更高的预测精度。控制变量实验结果表明，所提 EMD-LSTM-Attention 模型在不同放电工况与不同温度下都具有较高的预测精度，具有良好的泛化能力。实验结果验证了模型的有效性，具有一定的应用价值。

参考文献：

[1] 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.

[2] 何文轩, 耿磊, 姚芳. 电动汽车动力锂离子电池可靠性关键技术综述[J]. 电源学报, 2024, 22(2): 183-196.
He Wenxuan, Geng Lei, Yao Fang. Overview of key reliability technologies for power lithium-ion batteries used in electric vehicles [J]. Journal of Power Supply, 2024, 22(2): 183-196. (in Chinese)

[3] Meng Jinhao, Ricco M, Luo Guangzhao, et al. An overview and comparison of online implementable SOC estimation methods for lithium-ion battery [J]. IEEE Transactions on Industry Applications, 2018, 54(2): 1583-1591.

[4] 许傲然, 戴菁, 谷彩莲, 等. 基于 WOA 优化扩展卡尔曼算法的锂离子电池 SOC 估算研究[J]. 电源学报, 2025, 23(2): 232-239.
Xu Aoran, Dai Jing, Gu Cailian, et al. Research on SOC estimation of lithium-ion battery based on WOA optimized extended Kalman algorithm [J]. Journal of Power Supply, 2025, 23(2): 232-239. (in Chinese)

[5] Kim I S. Nonlinear state of charge estimator for hybrid

electric vehicle battery [J]. IEEE Transactions on Power Electronics, 2008, 23(4): 2027-2034.

[6] Shen Ping, Ouyang Minggao, Han Xuebing, et al. Error analysis of the model-based state-of-charge observer for lithium-ion batteries [J]. IEEE Transactions on Vehicular Technology, 2018, 67(9): 8055-8064.

[7] Ouyang Quan, Jian Chen, Jian Zheng. State-of-charge observer design for batteries with online model parameter identification: A robust approach [J]. IEEE Transactions on Power Electronics, 2020, 35(6): 5820-5831.

[8] Chen Cheng, Xiong Rui, Yang Ruixin, et al. State-of-charge estimation of lithium-ion battery using an improved neural network model and extended Kalman filter [J]. Journal of Cleaner Production, 2019, 234: 1153-1164.

[9] Shu Xing, Li Guang, Zhang Yuanjian, et al. Stage of charge estimation of lithium-ion battery packs based on improved cubature Kalman filter with long short-term memory model [J]. IEEE Transactions on Transportation Electrification, 2021, 7(3): 1271-1284.

[10] Chemali E, Kollmeyer P J, Preindl M, et al. State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach [J]. Journal of Power Sources, 2018, 400: 242-255.

[11] Ren Xiaoqing, Liu Shulin, Yu Xiaodong, et al. A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM [J]. Energy, 2021, 234: 121236.

[12] Zhang Yituo, Li Chaolin, Jiang Yiqi, et al. Accurate prediction of water quality in urban drainage network with integrated EMD-LSTM model [J]. Journal of Cleaner Production, 2022, 354: 131724.

[13] Chen Junxi, Cao Liangben, Yu Kaihan. EMD-LSTM based deep learning inbound and outbound passenger flow prediction [C]// 2021 2nd International Conference on Artificial Intelligence and Information Systems. Chongqing, China, 2021: 1-6.

[14] Dai Shaowu, Chen Qiangqiang, Liu Zhihao, et al. Time series prediction based on EMD-LSTM model [J]. Journal of Shenzhen University Science and Engineering, 2020, 37(3): 265-270.

[15] Huang N E, Shen Zheng, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-995.

[16] CALCE. Battery data [EB/OL]. [2023-01-15]. https://web.calce.umd.edu/batteries/data.htm



朱博

作者简介：

朱博(1997—)，男，硕士研究生。研究方向：锂离子电池荷电状态预测。E-mail：347947453@qq.com。

刘雪芹(1971—)，男，通信作者，博士，教授。研究方向：新能源材料与器件，人工智能。E-mail：xqliu@cqut.edu.cn。

王宇珲(1996—)，男，硕士。研究方向：时间序列预测。E-mail：756794766@qq.com。

李涛(1996—)，男，硕士。研究方向：时间序列预测。E-mail：15523043058@163.com。

程观银(1997—)，男，硕士。研究方向：新能源材料与器件。E-mail：1804807967@qq.com。

(责任编辑：郭凯凯)