batteries

Special Issue Reprint

Machine Learning for Advanced Battery Systems

Edited by
Benben Jiang, Guannan He and Xiang Chen

mdpi.com/journal/batteries

MDPI

Machine Learning for Advanced Battery Systems

Machine Learning for Advanced Battery Systems

Guest Editors

Benben Jiang
Guannan He
Xiang Chen

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Benben Jiang
Department of Automation
Tsinghua University
Beijing
China

Guannan He
College of Engineering
Peking University
Beijing
China

Xiang Chen
Department of Chemical Engineering
Tsinghua University
Beijing
China

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Batteries* (ISSN 2313-0105), freely accessible at: https://www.mdpi.com/journal/batteries/special_issues/15W8XVC62I.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-8425-4 (Hbk)
ISBN 978-3-7258-8426-1 (PDF)
https://doi.org/10.3390/books978-3-7258-8426-1

© 2026 by the authors. Articles in this reprint are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The reprint as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

Junghwan Lee, Longda Wang, Hoseok Jung, Bukyu Lim, Dael Kim, Jiaxin Liu and Jong Lim
Deep Neural Network with Anomaly Detection for Single-Cycle Battery Lifetime Prediction
Reprinted from: *Batteries* **2025**, *11*, 288, https://doi.org/10.3390/batteries11080288 . . . . . **1**

Zichen Du and Renhao Lu
Physics-Informed Neural Networks for Advanced Thermal Management in Electronics and Battery Systems: A Review of Recent Developments and Future Prospects
Reprinted from: *Batteries* **2025**, *11*, 204, https://doi.org/10.3390/batteries11060204 . . . . . **23**

Muhammad Arslan Khan, Yixing Wang and Benben Jiang
Early Prediction of Battery Lifetime Using Centered Isotonic Regression with Quantile-Transformed Features
Reprinted from: *Batteries* **2025**, *11*, 145, https://doi.org/10.3390/batteries11040145 . . . . . **49**

Pierpaolo Dini and Davide Paolini
Exploiting Artificial Neural Networks for the State of Charge Estimation in EV/HV Battery Systems: A Review
Reprinted from: *Batteries* **2025**, *11*, 107, https://doi.org/10.3390/batteries11030107 . . . . . **66**

Büşra Çetinus, Saadin Oyucu, Ahmet Aksöz and Emre Biçer
The Role of Machine Learning in Enhancing Battery Management for Drone Operations: A Focus on SoH Prediction Using Ensemble Learning Techniques
Reprinted from: *Batteries* **2024**, *10*, 371, https://doi.org/10.3390/batteries10100371 . . . . . **106**

Nerijus Paulauskas and Vsevolod Kapustin
Battery Scheduling Optimization and Potential Revenue for Residential Storage Price Arbitrage
Reprinted from: *Batteries* **2024**, *10*, 251, https://doi.org/10.3390/batteries10070251 . . . . . **122**

Yixing Wang and Benben Jiang
Attention Mechanism-Based Neural Network for Prediction of Battery Cycle Life in the Presence of Missing Data
Reprinted from: *Batteries* **2024**, *10*, 229, https://doi.org/10.3390/batteries10070229 . . . . . **134**

Jianlin Li, Yuchen Peng, Qian Wang and Haitao Liu
Status and Prospects of Research on Lithium-Ion Battery Parameter Identification
Reprinted from: *Batteries* **2024**, *10*, 194, https://doi.org/10.3390/batteries10060194 . . . . . **149**

Alireza Rastegarpanah, Mohammed Eesa Asif and Rustam Stolkin
Hybrid Neural Networks for Enhanced Predictions of Remaining Useful Life in Lithium-Ion Batteries
Reprinted from: *Batteries* **2024**, *10*, 106, https://doi.org/10.3390/batteries10030106 . . . . . **171**

Chia-Nan Wang, Fu-Chiang Yang, Nhut T. M. Vo and Van Thanh Tien Nguyen
Enhancing Lithium-Ion Battery Manufacturing Efficiency: A Comparative Analysis Using DEA Malmquist and Epsilon-Based Measures
Reprinted from: *Batteries* **2023**, *9*, 317, https://doi.org/10.3390/batteries9060317 . . . . . **194**

Yongkang Ding, Xinjiang Chen and Jianxiao Wang
Deep Reinforcement Learning-Based Method for Joint Optimization of Mobile Energy Storage Systems and Power Grid with High Renewable Energy Sources
Reprinted from: *Batteries* **2023**, *9*, 219, https://doi.org/10.3390/batteries9040219 . . . . . **214**

batteries

MDPI

Article

Deep Neural Network with Anomaly Detection for Single-Cycle Battery Lifetime Prediction

Junghwan Lee [1,2,*], Longda Wang [3], Hoseok Jung [1], Bukyu Lim [1], Dael Kim [1], Jiaxin Liu [2] and Jong Lim [1,3]

[1] Research and Development Department, Batter Machine Co., Ltd., Seoul 06233, Republic of Korea
[2] Livision AI Technology Co., Ltd., Weilv Building, Wujin District, Changzhou 213100, China
[3] College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China
* Correspondence: rory@battermachine.com

Abstract

Large-scale battery datasets often contain anomalous data due to sensor noise, communication errors, and operational inconsistencies, which degrade the accuracy of data-driven prognostics. However, many existing studies overlook the impact of such anomalies or apply filtering heuristically without rigorous benchmarking, which can potentially introduce biases into training and evaluation pipelines. This study presents a deep learning framework that integrates autoencoder-based anomaly detection with a residual neural network (ResNet) to achieve state-of-the-art prediction of remaining useful life at the cycle level using only a single-cycle input. The framework systematically filters out anomalous samples using multiple variants of convolutional and sequence-to-sequence autoencoders, thereby enhancing data integrity before optimizing and training the ResNet-based models. Benchmarking against existing deep learning approaches demonstrates a significant performance improvement, with the best model achieving a mean absolute percentage error of 2.85% and a root mean square error of 40.87 cycles, surpassing prior studies. These results indicate that autoencoder-based anomaly filtering significantly enhances prediction accuracy, reinforcing the importance of systematic anomaly detection in battery prognostics. The proposed method provides a scalable and interpretable solution for intelligent battery management in electric vehicles and energy storage systems.

Keywords: remaining useful life; li-ion batteries; autoencoder; deep neural networks; anomaly detection; machine learning

1. Introduction

As global adoption of electric vehicles (EVs) continues to accelerate, precise prediction of the remaining useful life (RUL) of lithium-ion batteries has become increasingly critical for optimizing maintenance schedules, mitigating operational risks, and ensuring vehicle safety and operational efficiency. Recent advancements in cloud-based data acquisition capabilities have enabled the extensive collection and real-time transmission of operational data from lithium-ion battery systems in EVs and energy storage systems (ESS) [1,2]. However, such large-scale datasets commonly include anomalous or unqualified samples attributable to measurement inconsistencies, communication errors and noise, complicating accurate battery prognostics [3]. Consequently, developing robust methodologies to effectively manage and mitigate the influence of anomalous data remains one of the challenges in battery prognostics. Additionally, recent studies have introduced novel methodologies, such as Chen et al.'s impedance estimation using support vector regression optimized by a

cuckoo search algorithm [4] and Li et al.'s hybrid neural network employing feature fusion and transfer learning strategies for improved battery state-of-health (SOH) estimation [5].

Autoencoder-based methods have emerged as powerful tools for anomaly filtering, data generation, and robust feature representation. Depending on their structure and training objectives, these models can be broadly categorized into five architectural types: standard autoencoders (AE), denoising autoencoders (DAE), variational autoencoders (VAE), conditional variational autoencoders (CVAE), and adversarial autoencoders (AAE).

Standard autoencoders typically reduce dimensionality or transform features for predictive tasks [6,7]. DAEs are effective in handling sensor noise and missing data, enhancing robustness in real-world battery prognostics [8–10]. VAEs model uncertainty, generate synthetic data, and detect anomalies through reconstruction errors and probabilistic distributions [11–14]. CVAEs utilize supervised signals to guide generative modeling and improve predictive accuracy [15,16]. AAEs enforce feature realism and domain invariance, further improving generalization [14,17,18]. While various autoencoder architectures have been explored, direct comparisons remain challenging due to limited datasets with explicit failure annotations. Therefore, their effectiveness is often indirectly assessed by their contributions to performance improvements in downstream tasks, such as RUL or SOH prediction, after anomaly filtering [6,8,9,16,19,20].

Another crucial factor affecting deep learning-based prediction is the scale of the training and validation datasets. As Iglesia et al. [21] highlight in their systematic review, dataset size is crucial for ensuring robust model generalization. Models trained on relatively small datasets risk overfitting to dataset-specific degradation patterns, thereby reducing their applicability across diverse real-world conditions. In this study, we utilize the largest publicly available lithium-ion battery dataset from Stanford-MIT [22], which comprises 124 battery cells. Although this dataset has been criticized for its near-constant discharge conditions [23], it encompasses 72 distinct charging policies, offering a variety of operational patterns. This diversity helps mitigate concerns of overfitting and provides a more comprehensive basis for developing and evaluating deep learning models in battery RUL prediction.

Hsu et al. [24] reported significant advancements in predicting RUL using DNNs, achieving notably low prediction errors with a mean absolute percentage error (MAPE) below 4% and a root mean square error (RMSE) of fewer than 50 cycles using only a single cycle of input data. This demonstrates substantial potential for high-accuracy, early-life battery prognostics leveraging DNN architectures. However, a critical examination of their data selection and filtering methodology reveals important areas that need clarification and standardization. Although the authors excluded battery samples based on experimental issues such as early test termination, temperature recording failures, and excessive signal noise, explicit quantitative thresholds or statistically supported benchmarks for these exclusions were not defined. Specifically, from the original dataset, only 115 batteries had discharge data deemed sufficient within the initial 100 cycles for the Discharge DNN. Further selection reduced this to 95 batteries with acceptable charging data, and eventually to 81 batteries possessing consistently high-quality discharge and charge data for use in their Full RUL DNN.

Despite detailing reasons for these exclusions, the authors did not provide quantitative standards or statistical validation to ensure the absence of potential biases. Clearly defined standards would enhance the reliability of the findings and help differentiate model innovation from improvements that may result from selective data omission or biases. Furthermore, a systematic machine-learning approach can reveal patterns and enable discoveries that surpass human intuition and traditional methods [25].

Several studies have employed the MIT dataset for early-stage RUL prediction as shown in Table 1. Tang et al. [26] used a CNN-LSTM model with early incremental capacity (IC) curves, achieving 146 RMSE and 13.4% MAPE using 100 cycles. Couture and Lin [27] combined health indicators with CNN-based image features and a fully connected neural network (FCNN), reporting 79.7 RMSE and 11.2% MAPE using 10 cycles from 94 filtered cells. Safavi et al. [28] applied a CNN-XGBoost model on all 124 cells, reaching 106 RMSE and 7.5% MAPE. Gong et al. [29] used evolutionary feature selection (FS), neighborhood component analysis (NCA), and differential evolution (DE) combined with XGBoost/relevance vector machine (RVM), obtaining 43 RMSE and 5.21% MAPE. Fei et al. [30] employed CNN (VGG-16) with gaussian process regression (GPR), achieving 112 RMSE and 8.2% MAPE. Sanz-Gorrachategui et al. [31] used adversarial CNNs with leave-one-out cross-validation (CV), reporting 62 RMSE and 5% MAPE. Yang et al. [32] and Zhang et al. [33] proposed hybrid CNN models that integrate 2D/3D convolutions and attention mechanisms, achieving state-of-the-art performance with as few as 10–60 cycles across all 124 cells. Zhang et al. [33] specifically utilized a hybrid parallel residual (HPR)-CNN architecture.

Table 1. Comparative summary of machine learning methods for battery RUL prediction.

Reference	Model Architecture	Samples	Data Split	Cycles	RMSE (Cycles)	MAPE (%)
[34]	ResNet-based CNN	124	33%/35%/32%	40	128	10.7
[24]	Res-Inception + Attention	81	80%/20%	1	< 50	< 4
[26]	Parallel CNN-LSTM	124	60%/13%/27%	100	146	13.4
[27]	VGG-11 + FCNN (Transfer Learning)	94	80%/20%	10	79.7	11.2
[28]	CNN + XGBoost (Coati Optimization)	124	33%/67%	100	106	7.5
[29]	Evolutionary FS (NCA+DE) + XGBoost/RVM	123	70%/30%	100	43	5.21
[30]	CNN (VGG-16) + Feature Wrapper + GPR	123	70%/30%	100	112	8.2
[31]	CNN (IC curves) + Adversarial Training	124	Leave-one-out CV	10	62	5
[32]	Hybrid CNN (2D+3D) + Attention	124	80%/20%	60	13	1.1
[33]	Hybrid Parallel Residual CNN (HPR-CNN)	124	70%/30%	10	25	3.47

This study proposes a robust ResNet-based deep learning framework integrated with autoencoder-driven anomaly detection to significantly advance battery RUL prediction. By systematically filtering anomalous data, this approach effectively reduces bias and enhances model generalization. A key novelty of this work lies in comprehensively comparing multiple autoencoder architectures to determine their effectiveness in anomaly detection, highlighting how the choice of architecture directly impacts RUL prediction performance. Extensive hyperparameter optimization of both autoencoder models and the ResNet architecture underscores the rigorous and innovative nature of this experimental effort. Given the practical difficulty of acquiring explicitly labeled faulty battery data, autoencoders offer a scalable, unsupervised approach for anomaly detection [35]. The proposed framework achieves state-of-the-art RUL prediction accuracy (MAPE 2.85%, RMSE 40.87 cycles) using only single-cycle battery data. Furthermore, interpretability is enhanced through detailed feature importance analysis using SHAP and tree-based methods, providing a transparent and comparative methodology for anomaly filtering and significantly improving downstream predictive performance.

2. Method

2.1. Overall Framework

The overall framework of this study, illustrated in Figure 1, comprises multiple systematic procedures designed to ensure robust and accurate prediction of RUL. Initially, interpolation techniques are applied to standardize all battery cycle data, resulting in a consistent representation of 1000 data points per cycle. Subsequent feature selection is carefully guided by criteria that ensure independence across battery cycles, explicitly excluding

statistical features that vary across cycles. Instead, the study utilizes raw measurements directly obtainable from practical battery management systems (BMS), including voltage, current, temperature, and integrated capacity during both charging and discharging phases for training autoencoders. K-fold cross-validation is implemented to thoroughly assess the generalization and predictive performance of the proposed models. This validation procedure partitions the dataset into three subsets, iteratively utilizing each subset for validation while training with the remaining subsets. Furthermore, autoencoders detect and remove anomalous data samples based on reconstruction error criteria. Consequently, 25 out of the initial 124 cells are excluded, leaving a dataset comprising 99 cells for subsequent analysis. Even after this filtering step, the remaining dataset size surpasses that utilized in previously published benchmark studies, notably the single-cycle RUL prediction approach by Hsu et al. [24]. Finally, hyperparameter optimization of the ResNet architecture is conducted through a systematic grid search approach.

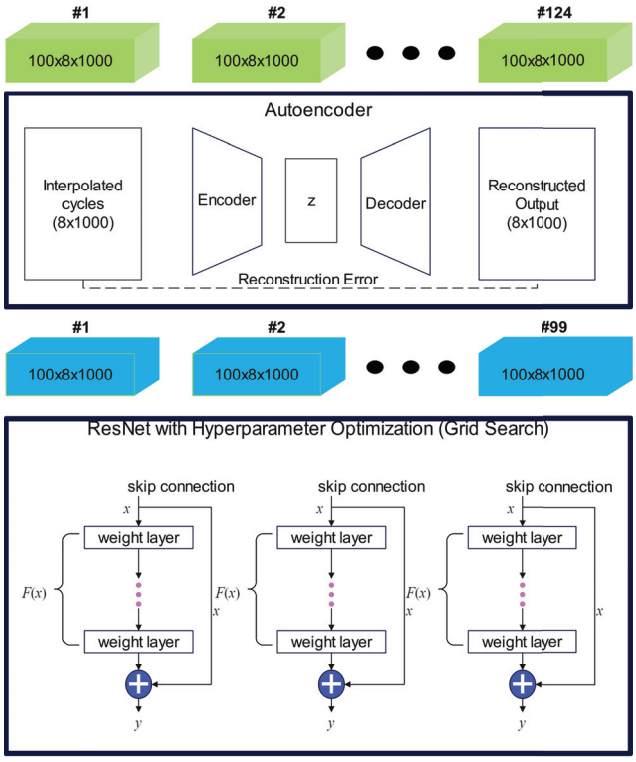


Figure 1. Overall framework of the proposed methodology for lithium-ion battery RUL prediction. The workflow comprises two primary stages. First, raw battery cycle data (green, dimensions: 100 cycles × 8 features × 1000 time steps per cycle) are interpolated and processed by autoencoder architectures for systematic anomaly filtering based on reconstruction error. The eight sequential input features per cycle include voltage, current, temperature, integrated charge capacity (QC), integrated discharge capacity (QD), linearly interpolated discharge capacity (QDlin), linearly interpolated temperature (Tdlin), and internal resistance (IR). The autoencoder identifies anomalous cycles, filtering them out to yield a refined dataset (blue). The refined dataset is subsequently used as input to the ResNet model, optimized through hyperparameter tuning (grid search), specifically adjusting kernel size and incorporating batch normalization layers. The ResNet inputs consist of 14 statistical features per battery cycle: average voltage, maximum voltage, minimum voltage, average temperature, maximum temperature, minimum temperature, integrated charge capacity (QC), integrated discharge capacity (QD), linearly interpolated discharge capacity (QDlin), linearly interpolated temperature (Tdlin), internal resistance (IR), average incremental capacity (dQ/dV), maximum incremental capacity (dQ/dV), and minimum incremental capacity (dQ/dV). The final ResNet model outputs cycle-level RUL predictions for each battery cell.

2.2. Autoencoders

Convolutional autoencoder(CAE) and sequence-to-sequence (Seq2Seq) architectures in Appendix A are developed to systematically identify and remove anomalous or inconsistent battery cycle data based on reconstruction errors. Each CAE variant employs 1D-CNNs, using specialized dimensionality reduction strategies tailored to either raw sequential battery data or statistical features summarizing each cycle. The Seq2Seq autoencoder variants utilize LSTM units, which explicitly capture sequential dependencies within battery cycle data. Autoencoder models are designed to learn latent representations of input data and reconstruct the original input with minimal error.

The transformation of input data through a CAE follows a hierarchical feature extraction process, where each layer refines the representation of the previous one. Mathematically, the latent representation at a given layer is expressed as:

$$\boldsymbol{h}^{(l)} = \sigma\Big(\boldsymbol{W}^{(l)} * \boldsymbol{h}^{(l-1)} + \boldsymbol{b}^{(l)}\Big) \tag{1}$$

where $\boldsymbol{h}^{(l)}$ denotes the latent feature representation at layer l, $\boldsymbol{W}^{(l)}$ and $\boldsymbol{b}^{(l)}$ correspond to the convolutional filter weights and biases, respectively, and $*$ represents the convolution operation. The activation function $\sigma(\cdot)$ introduces nonlinearity to enable complex feature extraction. The reconstructed output is obtained by applying a transposed convolution to the final latent representation:

$$\hat{\boldsymbol{X}} = \hat{\boldsymbol{W}}^{(\text{out})} * \hat{\boldsymbol{h}}^{(\text{final})} + \hat{\boldsymbol{b}}^{(\text{out})} \tag{2}$$

where $\hat{\boldsymbol{h}}^{(\text{final})}$ represents the final encoded feature representation, and $\hat{\boldsymbol{W}}^{(\text{out})}$ and $\hat{\boldsymbol{b}}^{(\text{out})}$ are learnable parameters of the decoder. This structure enables CAEs to learn spatially meaningful latent features, which is advantageous in applications such as anomaly detection and feature compression.

In contrast, the Seq2Seq autoencoder employs RNNs to capture temporal dependencies in sequential data. The encoding process utilizes an LSTM network to transform an input sequence into a fixed-dimensional latent representation:

$$\boldsymbol{h}_t^{(\text{enc})} = \text{LSTM}\big(\boldsymbol{x}_t, \boldsymbol{h}_{t-1}^{(\text{enc})}\big) \tag{3}$$

where $\boldsymbol{x}_t$ denotes the input vector at time step t and $\boldsymbol{h}_t^{(\text{enc})}$ is the encoder LSTM's hidden state at time t. The final hidden state of the encoder serves as a compressed representation of the entire input sequence. This encoded representation is then fed into a decoder LSTM, which sequentially reconstructs the original input:

$$\boldsymbol{h}_t^{(\text{dec})} = \text{LSTM}\big(\hat{\boldsymbol{x}}_{t-1}, \boldsymbol{h}_{t-1}^{(\text{dec})}\big) \tag{4}$$

$$\hat{\boldsymbol{x}}_t = f\big(\boldsymbol{h}_t^{(\text{dec})}\big) \tag{5}$$

Here, $\boldsymbol{h}_t^{(\text{dec})}$ is the decoder's hidden state at time t, $\hat{\boldsymbol{x}}_{t-1}$ is the reconstructed output from the previous time step, and $f(\cdot)$ represents a transformation function that maps the decoder's hidden state to the output space. This Seq2Seq structure effectively learns temporal dependencies, enabling sequence-level reconstructions.

Both CAE and Seq2Seq autoencoders are typically trained by minimizing the reconstruction loss, which is commonly measured using the mean squared error (MSE):

$$\mathcal{L}(\boldsymbol{X}, \hat{\boldsymbol{X}}) = \frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{X}_i - \hat{\boldsymbol{X}}_i\|^2. \tag{6}$$

where N represents the number of samples, and $\boldsymbol{X}_i$ and $\hat{\boldsymbol{X}}_i$ denote the original and reconstructed inputs, respectively. The objective of training is to optimize the model parameters such that the reconstructed output closely matches the input, thereby ensuring an accurate and meaningful representation of the data in the latent space.

The convolutional autoencoder with average pooling (CAE-AvgPool) employs one-dimensional convolutional layers combined with average pooling operations to reduce temporal dimensionality while preserving smooth and stable patterns in battery behavior. In contrast, the convolutional autoencoder with max pooling (CAE-MaxPool) uses max pooling layers that retain only the strongest activations at each pooling step, thereby enhancing sensitivity to transient anomalies or abrupt variations in the battery signal. The deep convolutional autoencoder (CAE-Deep) consists of multiple convolutional layers with progressively decreasing channel dimensions, allowing it to capture hierarchical temporal dependencies across multiple scales. Its decoder mirrors the encoder using symmetric convolutional and upsampling layers to reconstruct the original input dimensions. Additionally, the convolutional autoencoder with statistical feature extraction (CAE-StatisticalFeature) replaces pooling operations with strided convolutions to compress input features while preserving essential statistical characteristics of the battery cycle data. The decoder reconstructs these representations through convolutional and upsampling layers, concluding with a fully connected output layer matched to the original statistical feature dimensions.

The Simple-Seq2Seq encodes and reconstructs the entire battery sequence using a straightforward single-layer LSTM structure, providing a baseline for temporal modeling performance. The Chunked-Seq2Seq divides input sequences into segments or "chunks", which are individually processed via TimeDistributed LSTM layers, explicitly modeling localized temporal characteristics within battery data. Reconstruction is performed correspondingly through segmented decoder layers, enhancing local temporal pattern fidelity. The Hierarchical-Seq2Seq incorporates hierarchical LSTM structures to concurrently capture both fine-grained local patterns and overarching global temporal dependencies.

Table 2 presents the averaged training and validation metrics obtained through 3-fold cross-validation for each autoencoder architecture. The variations in reconstruction errors are influenced by factors such as dimensionality reduction strategies, the complexity of input data patterns, and the architecture's capacity to reconstruct these patterns accurately. For instance, architectures that employ heavily compressed latent representations or model intricate sequential dependencies may exhibit higher reconstruction errors due to the increased difficulty in reconstructing the data. Conversely, autoencoders that process simplified statistical feature inputs often achieve lower reconstruction errors due to the reduced data complexity and dimensionality. It is essential to note that the reconstruction metrics alone are insufficient to directly assess the relative anomaly detection effectiveness of each architecture. A critical evaluation criterion is the ability of the reconstruction error to differentiate between abnormal and normal cells. The Seq2Seq-based autoencoder required approximately 20 s for prediction, whereas the CAE completed predictions in approximately 2 s, indicating that the CAE achieved nearly tenfold faster inference performance. These results were obtained using a dataset comprising 124 samples, each encompassing 100 cycles and 1000 time sequences, conducted on an Intel Core i9-13900KF CPU paired with an NVIDIA GeForce RTX 4090 GPU. It should also be noted that a thorough evaluation

of the training durations was not possible due to the variability introduced by the early stopping criteria.

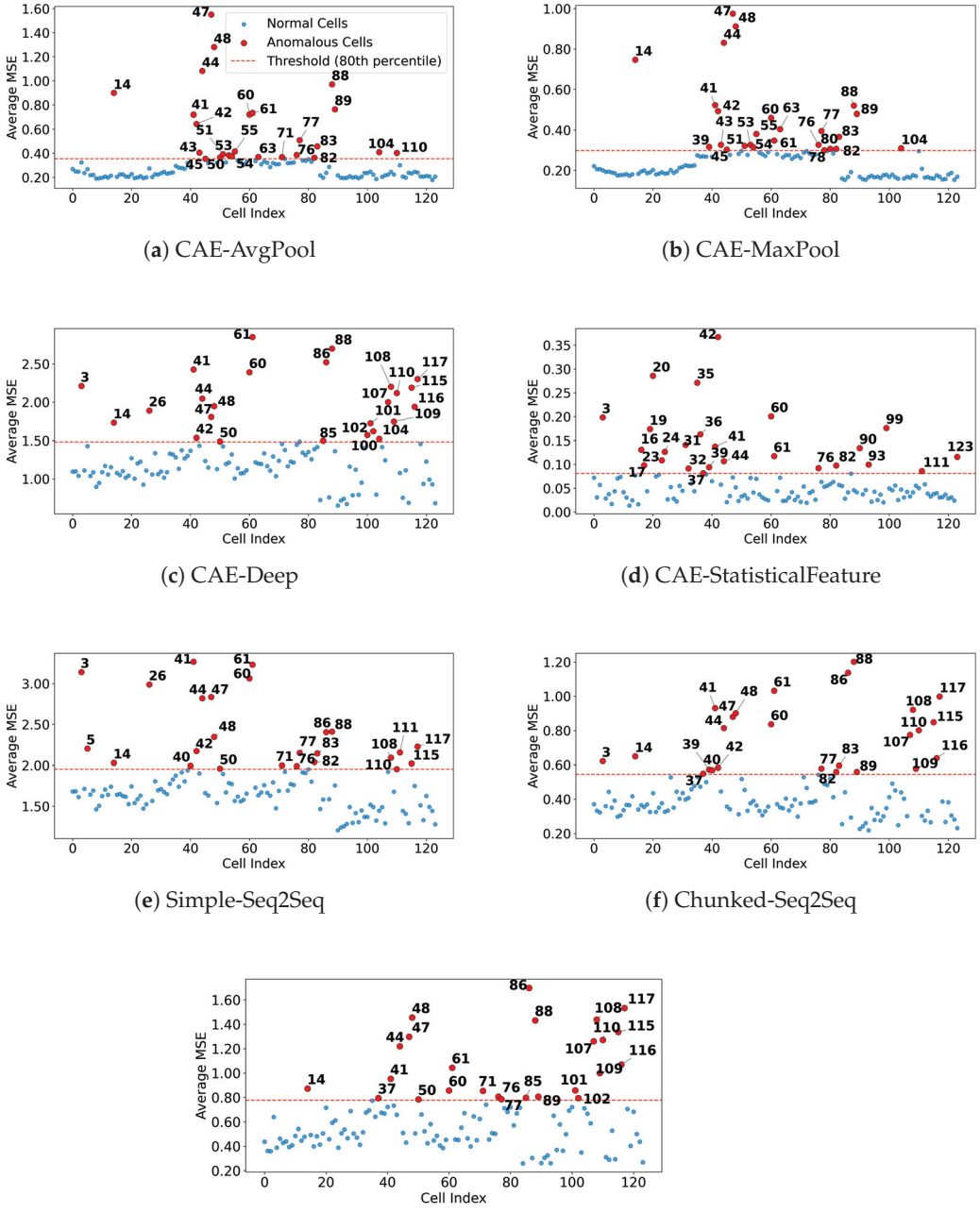
Table 2. Average training and validation metrics across 3-fold cross-validation for autoencoder architectures.

Model	Model Params	Prediction Time (s)	Training (Average)			Validation (Average)		
			MSE (%)	MAE (%)	RMSE (%)	MSE (%)	MAE (%)	RMSE (%)
CAE-AvgPool	4080	1.9	4.10	11.5	19.8	6.30	12.1	21.4
CAE-MaxPool	4080	1.8	3.50	10.3	18.6	4.20	10.5	18.6
CAE-Deep	223,960	1.8	37.4	38.4	61.1	38.1	38.7	61.2
CAE-StatisticalFeature	95,535	1.0	0.05	1.32	2.11	0.11	1.94	3.20
Simple-Seq2Seq	1088	18.9	36.4	37.9	60.3	38.3	38.9	60.9
Chunked-Seq2Seq	1088	22.9	17.8	22.6	42.2	18.2	22.9	42.4
Hierarchical-Seq2Seq	2176	24.3	21.6	25.2	46.4	22.1	25.6	46.7

The training approach employed k-fold cross-validation to objectively evaluate the model's generalization performance. For convolutional autoencoder (CAE) architectures, models were trained consistently with a fixed batch size of 64 and a learning rate of 0.0001. The training and validation loss curves across multiple folds indicated stable convergence behavior. In the case of the CAE-StatisticalFeature, batch sizes varied from 1 to 32, while the learning rate was constant at 0.001. Experimental results showed a clear trend where increasing batch sizes correlated with enhanced training stability and a statistically significant reduction in reconstruction errors. Similar experimentation was performed for the Seq2Seq autoencoder variants, emphasizing batch size effects. The Simple-Seq2Seq architecture exhibited optimal training performance at a batch size of 128, as evidenced by consistently lower validation errors. Training runs with smaller batch sizes demonstrated inconsistent convergence, higher variability in training outcomes, and reduced reconstruction accuracy.

2.3. Anomaly Detection

Reconstruction errors obtained during model training were analyzed using the 80th percentile approach to determine an anomaly detection threshold. This threshold was specifically selected to ensure a fair comparison with prior studies by controlling the proportion of anomalies identified. Despite filtering anomalies at the 80th percentile, the resulting dataset retained more samples compared to previous research, ensuring robust evaluation. Battery cells exceeding this derived threshold were identified as anomalous in Figure 2. Normal cells are represented in blue, whereas cells identified as anomalous are marked in red. Notably, significant variability in anomaly detection is observed across the autoencoder architectures. For instance, the CAE-AvgPool and CAE-MaxPool models predominantly detect anomalies concentrated within mid-range cell indices, suggesting these architectures' sensitivity may be limited to specific data patterns or localized anomalies. In contrast, the CAE-Deep, Simple-Seq2Seq, and Chunked-Seq2Seq architectures demonstrate broader, more consistent anomaly detection capabilities, effectively identifying anomalies across a wider range of cell indices. Overall, pooling-based convolutional architectures exhibit moderate anomaly detection effectiveness, while deeper convolutional and sequence-based models offer superior consistency and robustness. These observations emphasize the importance of carefully selecting appropriate autoencoder architectures tailored to the operational scenarios and anomaly characteristics of specific battery datasets.



(**a**) CAE-AvgPool

(**b**) CAE-MaxPool

(**c**) CAE-Deep

(**d**) CAE-StatisticalFeature

(**e**) Simple-Seq2Seq

(**f**) Chunked-Seq2Seq

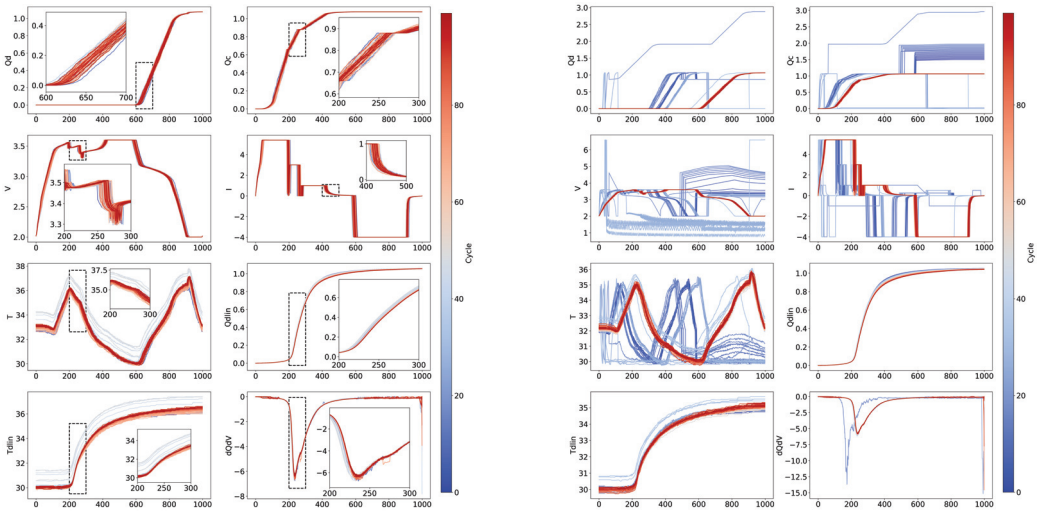
(**g**) Hierarchical-Seq2Seq

Figure 2. Comparison of anomaly detection outcomes based on reconstruction errors from various autoencoder architectures. Each subplot corresponds to a different autoencoder method: (**a**) CAE-AvgPool, (**b**) CAE-MaxPool, (**c**) CAE-Deep, (**d**) CAE-StatisticalFeature, (**e**) Simple-Seq2Seq, (**f**) Chunked-Seq2Seq, and (**g**) Hierarchical-Seq2Seq. Blue points represent battery cells identified as normal, while red points indicate anomalous cells whose average mean squared error (MSE) exceeds the anomaly threshold, defined by the 80th percentile of reconstruction errors (shown as a red dashed line). Numbers next to red points specify the cell indices identified as anomalous. The figure illustrates distinct patterns in anomaly detection effectiveness across different autoencoder architectures, reflecting their varying strategies for dimensionality reduction and their capabilities in accurately reconstructing the battery cycle data.

For visual inspection, the data classified by the autoencoder as normal and anomalous are illustrated in Figure 3. In normal cell data (Figure 3a), the charge capacity (Q_C), discharge capacity (Q_D), and linearly interpolated discharge capacity (Q_{Dlin}) exhibit a gradual decline over successive cycles. Conversely, temperature-related features (T and linearly interpolated temperature, T_{dlin}) show an upward trend, aligning with typical degradation patterns observed in lithium-ion batteries under controlled cycling conditions [36]. These observations are consistent with established degradation mechanisms such as gradual capacity fade driven by incremental loss of lithium inventory and active material degradation [37]. In contrast, anomalous cell data (Figure 3b) exhibit irregular patterns across most features, except for the Q_{Dlin}, which remains relatively stable. These irregularities may result from sensor malfunctions, manufacturing variability, or data acquisition errors. Notably, voltage (V) and current (I) profiles demonstrate abrupt fluctuations indicative of inconsistent charge-discharge dynamics that significantly deviate from typical lithium-ion cell operations. The abnormal temperature data (Figure 3c) specifically highlights deviations predominantly within temperature-related features, even though charge and discharge capacities remain within expected operational ranges. Such temperature anomalies could be attributed to environmental chamber temperature fluctuations previously identified in [24,34]. A particularly significant observation is the classification of the shortest end-of-life (EOL) cell (Figure 3d) as abnormal, despite initially following a capacity degradation trajectory resembling typical aging. This cell reached failure after just 148 cycles, considerably below the typical dataset range (150–2300 cycles). The autoencoder's assignment of a high reconstruction error to this cell likely arises from its distinct degradation trajectory, which deviates from the typical patterns observed across the broader dataset. Overall, the presented results underscore the effectiveness of the autoencoder-based anomaly detection approach in differentiating between normal aging, irregular capacity fade, thermal anomalies, rapid EOL failures, and potential experimental irregularities.

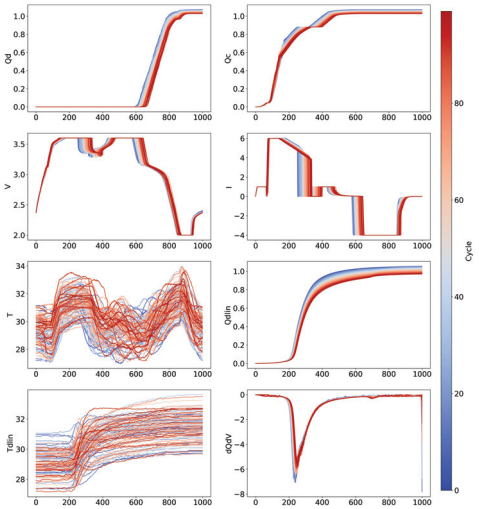
2.4. Preliminary Assessment of Dataset Quality and Feature Analysis

Random Forest (RF) models [38] were utilized to evaluate dataset effectiveness and to conduct feature importance analysis following the removal of anomalous cells from the dataset. The rationale for adopting RF models was twofold. First, RF models offer a computationally efficient approach to preliminarily assess the effectiveness of various anomaly detection methods applied to enhance dataset quality. Due to the significant computational requirements and prolonged training durations associated with DNNs, this preliminary assessment allows for the identification of the most promising anomaly detection methods. By quantifying the degree of improvement in dataset quality early in the analysis pipeline, researchers can make informed decisions to optimize computational resource allocation and streamline subsequent DNN training processes. Second, RF models inherently possess superior interpretability compared to DNNs. This interpretability is critical as it facilitates a comprehensive feature importance analysis, clarifying how specific features contribute to the predictive performance of the model. Consequently, RF-based feature importance analysis enhances the transparency of the modeling process and provides essential validation for the selected anomaly detection methods.

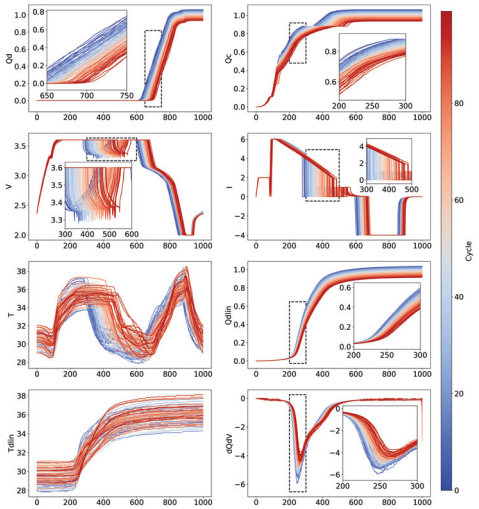


(**a**) Normal Cell Data

(**b**) Abnormal Cell Data



(**c**) Abnormal Temperature



(**d**) Shortest EOL

Figure 3. Comparison of normal and abnormal battery cell data illustrating characteristic measurement patterns across cycles: (**a**) Normal cell data showing consistent, stable trends in charge capacity (QC), discharge capacity (QD), linearly interpolated discharge capacity (QDlin), voltage (V), current (I), temperature (T), and linearly interpolated temperature (Tdlin). (**b**) Abnormal cell data demonstrating irregularities, voltage spikes, and inconsistent patterns indicative of potential sensor or operational issues. (**c**) Abnormal temperature behavior specifically highlighting unusual temperature fluctuations that impact battery performance, consistent with previously reported anomalies [22,24]. (**d**) Battery cells with notably short end-of-life (EOL), exhibiting accelerated degradation or irregular trajectories compared to typical cells; these particular samples were typically excluded in previous studies [29,30]. In all subfigures, the color gradient (blue to red) corresponds to cycle indices, where blue represents earlier cycles and red indicates later cycles, clearly illustrating the evolution of battery measurements over time.

Table 3 summarizes the performance of RF models trained on datasets processed through various anomaly detection methods, evaluated using RMSE and MAPE metrics for both training and validation sets. For RUL prediction, the MAPE metric is particularly critical because RMSE values can be disproportionately influenced by cells with shorter EOL. In contrast, MAPE provides a relative measure of prediction accuracy that is less dependent on the variability in cell lifetimes, making it a more reliable and robust metric when evaluating predictive performance across datasets containing cells with diverse EOLs.

RF models trained on standard datasets, comprising cycles 1–100 and cycles 21–120, yielded relatively lower errors on the training set but exhibited higher errors on the validation set. Observations indicating frequent anomalies within initial cycles informed the decision to exclude cycles 0–20. However, minimal performance improvement after this exclusion implies persistent data complexity and suggests caution regarding potential cycle-selection bias in future analyses. Datasets refined through autoencoder-based anomaly removal methods generally demonstrated improved predictive performance on the validation set, highlighting the effectiveness of these approaches. Among these methods, the CAE-StatisticalFeature approach yielded the lowest validation RMSE and a low validation MAPE. Similarly, the CAE-MaxPool method exhibited strong validation performance, validating its utility in enhancing dataset quality. Conversely, Seq2Seq approaches showed mixed results. Although the Simple-Seq2Seq method had relatively low training errors, validation errors remained relatively high. Chunked-Seq2Seq and Hierarchical-Seq2Seq methods produced higher validation errors, indicating challenges in effectively detecting anomalies within complex sequential data.

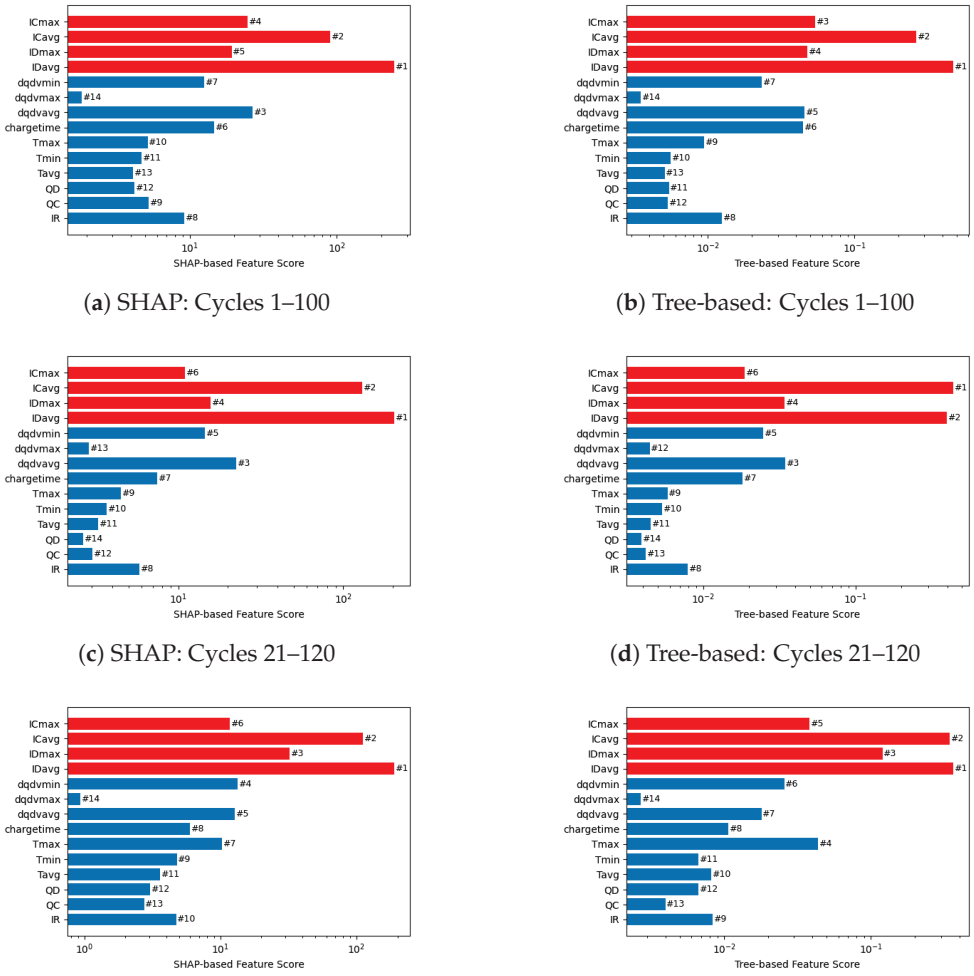
Table 3. Performance comparison of Random Forest models trained on various datasets after anomaly removal.

Dataset for Random Forest	**Training Set**		**Validation Set**	
	RMSE	**MAPE (%)**	**RMSE**	**MAPE (%)**
100 cycles (cycles 1–100)	78.52	5.25	183.25	18.28
100 cycles (cycles 21–120)	71.40	5.24	176.98	19.28
CAE-AvgPool Anomaly Removal	84.43	5.64	123.59	11.14
CAE-MaxPool	79.99	5.17	118.76	10.69
CAE-Deep	89.92	6.51	139.18	15.93
CAE-StatisticalFeature	88.29	5.31	111.00	11.80
Simple-Seq2Seq	85.45	4.51	153.26	14.30
Chunked-Seq2Seq	82.26	5.71	213.94	17.32
Hierarchical-Seq2Seq	86.12	5.95	197.97	13.69

Performance of the Random Forest model trained on datasets processed by different anomaly detection methods.

Figure 4 presents a comparative analysis of feature importance derived from RF models across different datasets, utilizing SHapley Additive exPlanations (SHAP) and tree-based methodologies. All analyzed features originate from Lee et al. [34], with newly introduced features specifically highlighted in red to emphasize their novel contribution. Key features include internal resistance (IR), which indicates electrical resistance within the battery cell, as well as charge and discharge capacities (QC and QD), reflecting the battery's capacity to store and deliver electrical charge, respectively. Temperature-related metrics such as average (T_{avg}), minimum (T_{min}), and maximum (T_{max}) temperatures capture thermal behaviors throughout battery cycles. Additionally, incremental capacity features include average (dQ/dV_{avg}), maximum (dQ/dV_{max}), and minimum (dQ/dV_{min}) values. Current-related parameters, covering average and maximum values during both charge ($I_{C,\text{avg}}$, $I_{C,\text{max}}$) and discharge cycles ($I_{D,\text{avg}}$, $I_{D,\text{max}}$), further contribute insights into battery performance and stress conditions. Features newly introduced and emphasized by Lee et al. are distinctly highlighted in red within the figure. Figure 4a,b illustrate the feature importance scores for the dataset covering cycles 1–100. Both SHAP and tree-based analyses consistently identify newly introduced features from [34] (e.g., $I_{C,\text{max}}$, $I_{C,\text{avg}}$, $I_{D,\text{max}}$, and D_{avg}) as highly significant predictors, validating their importance in battery degradation modeling. The alignment between SHAP and tree-based results strengthens the validity of the identified feature importance, as the convergence of two complementary methods reduces the likelihood of analytical bias. In Figure 4c,d, covering cycles 21–120, the same

newly introduced features remain prominently influential, confirming their robustness across different cycle ranges. However, minor differences in feature rankings between methods and cycle ranges underline the sensitivity of these metrics to specific dataset configurations and reinforce the importance of employing multiple analytical methods. Figure 4e,f present feature importance scores after anomaly removal. Most significantly, the importance of the maximum temperature feature (T_{max}) substantially increases after filtering anomalies, suggesting that removing anomalous cells reveals the critical predictive role of temperature-related features.



(**a**) SHAP: Cycles 1–100

(**b**) Tree-based: Cycles 1–100

(**c**) SHAP: Cycles 21–120

(**d**) Tree-based: Cycles 21–120

(**e**) SHAP: Anomaly-filtered

(**f**) Tree-based: Anomaly-filtered

Figure 4. Comparison of feature importance using SHAP and Tree-based methods across different cycle ranges and anomaly filtering conditions.

2.5. Deep Neural Network Training

The proposed model utilizes a ResNet-34-based architecture, adapted for one-dimensional convolutional processing. Residual connections inherent in ResNet architectures address the vanishing gradient problem, allowing the use of deeper neural networks that can extract complex, hierarchical temporal features. Each input sample is structured as a tensor of dimensions, where the first dimension (14) corresponds to distinct input features, and the second dimension represents a single channel suitable for 1D convolutional operations. The initial processing step involves a convolutional layer with 64 filters, a kernel size of 7, and a stride of 2, activated by the Rectified Linear Unit (ReLU). This convolutional operation is followed by a max-pooling layer with a pooling size of 3 and a stride of 2, reducing the temporal dimension while preserving critical feature representa-

tions. Subsequently, the network is organized into sequential stages comprising residual blocks designed for hierarchical temporal feature extraction. Stage 1 contains three residual blocks, each equipped with 64 filters. Stage 2 employs an initial convolutional layer with a stride of 2 to halve the temporal dimension, followed by three residual blocks comprising 128 filters each. Stage 3 further reduces the temporal dimension through an additional stride-2 convolutional layer, followed by seven residual blocks, each containing 256 filters, to capture higher-level temporal abstractions. Stage 4 comprises three residual blocks with 512 filters, providing deep-level feature extraction capabilities. Finally, a global average pooling layer aggregates temporal features, transforming them into a fixed-length vector that summarizes temporal information across all stages. A fully connected layer subsequently utilizes this aggregated vector to generate the final regression output. Overall, the described ResNet34-based model contains approximately 4.2 million trainable parameters Hyperparameter optimization of the ResNet architecture was performed using a systematic grid search approach. Prior research has extensively identified and evaluated the critical hyperparameters influencing model performance [34]. Building upon these established insights, the current study strategically narrowed its focus to optimize two pivotal hyperparameters: the implementation of batch normalization layers and the selection of optimal batch size.

Table 4 presents the comparative training and validation results of a ResNet model using data preprocessed by various autoencoder techniques. The baseline established by raw, unfiltered data exhibited high validation errors. The results indicate significant variability in ResNet performance depending on the preprocessing method employed. Among the convolutional autoencoder variations, CAE-Deep achieved better generalization than pooling-based CAE methods, as evidenced by lower validation errors. Conversely, pooling-based methods such as CAE-AvgPool and CAE-MaxPool yielded higher validation errors, suggesting limited effectiveness in capturing relevant features for this specific application. Notably, the Simple-Seq2Seq autoencoder consistently provided superior validation results, achieving the lowest validation without a k-fold split, thus indicating its potential as a robust preprocessing method. Interestingly, the analysis revealed increased validation errors across all methods when using the k-fold split, contrary to typical expectations that additional cross-validation splits enhance generalization. This finding underscores the necessity for careful selection and optimization of cross-validation parameters tailored to the specific dataset.

Table 4. Training and validation performance of ResNet using data filtered by various autoencoders.

Autoencoder	k-Fold	Training Performance			Validation Performance		
		MAE	MAPE (%)	RMSE	MAE	MAPE (%)	RMSE
No Filter (Raw Data)	3	25.74	3.56	38.40	85.36	21.86	117.43
CAE-AvgPool	3	45.54	5.77	62.04	123.65	10.14	202.48
CAE-MaxPool	3	15.57	2.11	21.08	93.23	10.39	127.00
CAE-Deep	3	11.57	1.71	17.20	70.19	9.75	110.56
CAE-StatisticalFeature	3	12.21	1.97	15.54	69.31	10.34	94.89
Simple-Seq2Seq	3	8.78	1.27	12.07	61.07	9.12	87.71
Chunked-Seq2Seq	3	18.21	2.57	24.47	60.64	7.56	98.10
Hierarchical-Seq2Seq	3	9.92	1.39	14.10	74.38	11.23	107.82
CAE-AvgPool	0	19.78	2.59	27.67	63.54	8.40	90.48
CAE-MaxPool	0	13.65	1.82	20.66	91.77	12.13	123.97
CAE-Deep	0	21.38	2.90	28.63	46.84	6.92	69.31
CAE-StatisticalFeature	0	14.12	2.03	19.03	89.31	9.86	172.54
Simple-Seq2Seq	0	10.66	1.47	14.62	53.24	7.11	87.65
Chunked-Seq2Seq	0	27.40	3.44	42.23	53.91	7.19	100.57
Hierarchical-Seq2Seq	0	9.09	1.36	11.82	64.71	9.47	96.50

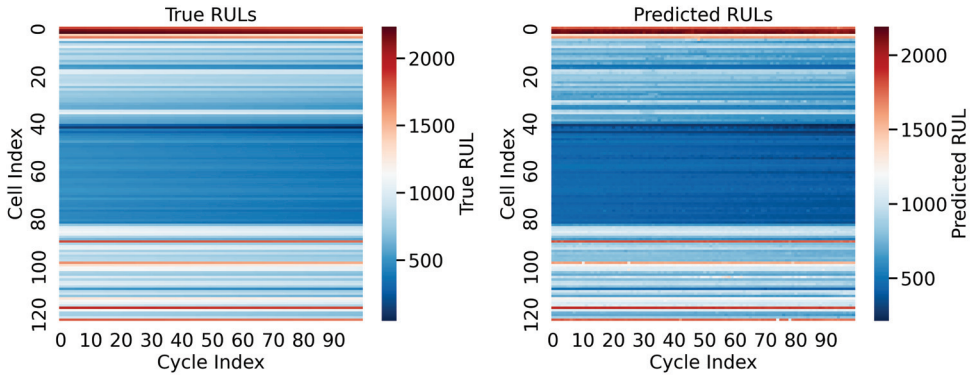
Table 5 presents a comparative analysis of performance metrics for various models. The baseline ResNet model, without any filtering, recorded comparatively higher errors, indicating limited predictive capability relative to the advanced architectures. The ResNet with Simple-Seq2Seq achieved the best performance, demonstrating the lowest error across all metrics. It is inherently challenging to explicitly attribute performance differences among autoencoder architectures due to the subtle and complex interactions among model design, hyperparameters, and data characteristics. However, given the sequential nature of battery degradation data, it is reasonable to infer that the Simple-Seq2Seq architecture's explicit temporal modeling capability is crucial to its superior performance, as demonstrated consistently in our empirical evaluations. The ResNet with CAE-Deep and ResNet with CAE-StatisticalFeature models also showed notable performance improvements compared to basic ResNet models.

Table 5. Performance metrics for different models.

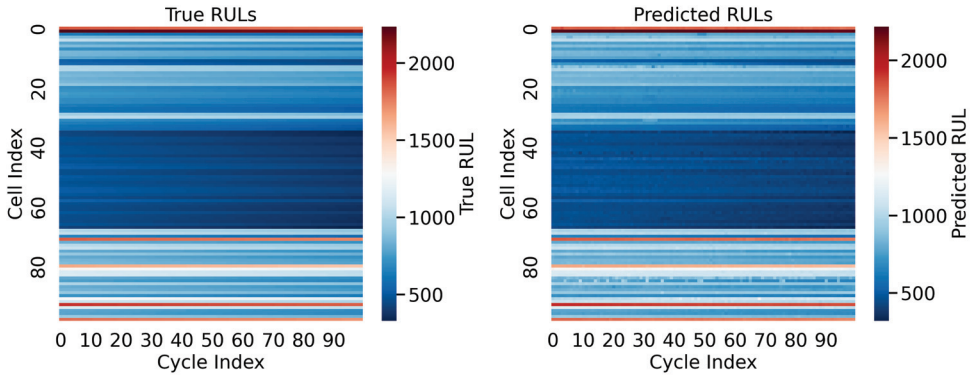
Model	MAPE	RMSE	MAE
Res-Inception + Attention [24]	4.00	50.00	–
ResNet [34]	7.25	62.91	37.76
ResNet with CAE-AvgPool	5.92	69.99	46.20
ResNet with CAE-MaxPool	3.78	60.11	31.25
ResNet with CAE-Deep	3.34	52.02	23.41
ResNet with CAE-StatisticalFeature	3.66	44.85	23.74
ResNet with Simple-Seq2Seq	2.85	40.87	19.34
ResNet with Chunked-Seq2Seq	3.57	49.22	26.78
ResNet with Hierarchical-Seq2Seq	3.38	50.07	22.94

Figure 5 provides an image visualization comparing cycle-by-cycle true and predicted remaining useful life (RUL) across multiple battery cells and cycles. This visualization is intended to provide an intuitive, qualitative overview of model accuracy differences between methods, complementing detailed numerical comparisons. Color intensity corresponds directly to RUL values: darker blue indicates lower RUL (approaching end-of-life), while lighter colors or red indicate higher RUL values (earlier in battery life). The subfigures contrast (a) the baseline ResNet model, which shows noticeable deviations and inconsistencies relative to the true RUL values, with (b) the ResNet integrated with Simple-Seq2Seq, clearly demonstrating improved accuracy and better alignment between predicted and true RUL values across battery cycles and cells.

Figure 6 provides a precise, quantitative evaluation of the models, presenting the distribution of cycle-by-cycle prediction errors between true and predicted RULs across individual cells for different model architectures. The ResNet with Simple-Seq2Seq architecture achieves the highest prediction accuracy and robustness, demonstrated by tighter error distributions and minimal maximum absolute errors within ± 200 cycles across nearly all cells. Additionally, the ResNet with CAE-StatisticalFeature architecture shows strong stability, maintaining relatively small error distributions for most evaluated cells. Architectures such as ResNet with CAE-AvgPool, CAE-Deep, Chunked-Seq2Seq, and Hierarchical-Seq2Seq outperform the baseline ResNet, resulting in overall reduced prediction errors. However, these models occasionally produce noticeable outliers with higher maximum absolute errors for specific cells. These findings suggest that integrating a Simple-Seq2Seq architecture with ResNet significantly enhances predictive accuracy and reliability for RUL estimation.

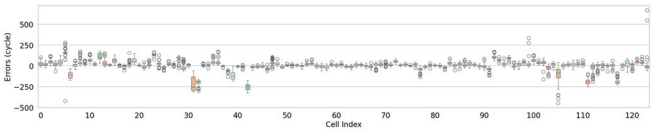


(**a**) ResNet

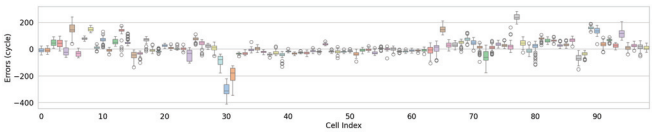


(**b**) ResNet with CAE-AvgPool

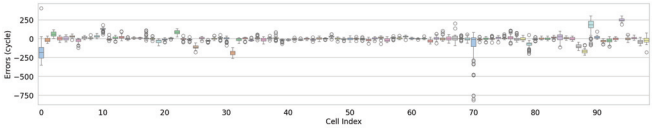
Figure 5. Image visualization comparing cycle-by-cycle true and predicted RUL across battery cells and cycles for (**a**) ResNet and (**b**) ResNet with Simple-Seq2Seq.



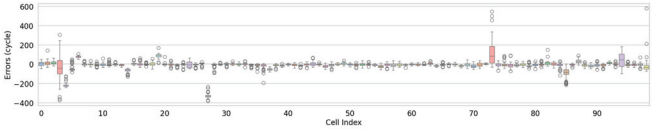
(**a**) ResNet



(**b**) ResNet with CAE-AvgPool

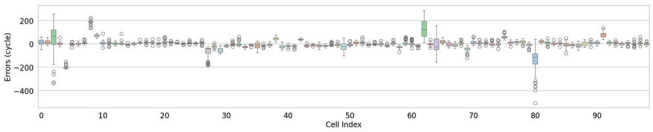


(**c**) ResNet with CAE-MaxPool

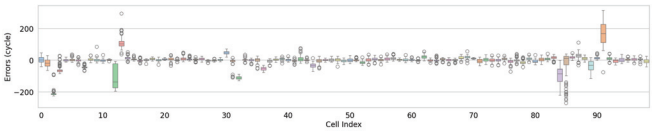


(**d**) ResNet with CAE-Deep

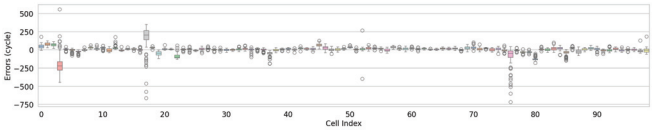
Figure 6. *Cont.*



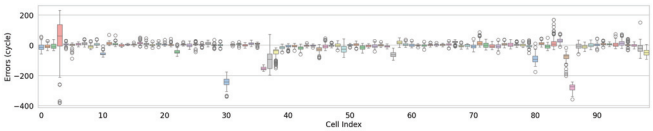
(**e**) ResNet with CAE-StatisticalFeature



(**f**) ResNet with Simple-Seq2Seq



(**g**) ResNet with Chunked-Seq2Seq



(**h**) ResNet with Hierarchical-Seq2Seq

Figure 6. Distribution of cycle-by-cycle prediction errors between true and predicted RUL across individual cells for various model architectures. The x-axis represents the cell index, and the y-axis shows prediction errors (in cycles) across all cycles for each cell.

3. Discussion

The proposed study significantly advances battery RUL prediction through a novel integration of anomaly detection techniques using diverse autoencoder architectures combined with a ResNet-based deep neural network model. This research highlights substantial progress in predictive reliability for EV and ESS applications by achieving state-of-the-art accuracy in cycle-by-cycle battery RUL prediction.

The ResNet model, integrated with the Simple-Seq2Seq autoencoder, was identified as the most accurate and stable model, providing lower prediction errors across multiple metrics and minimal maximum absolute errors. This stability is particularly crucial for safety-critical applications, such as EVs, where the consistent and reliable prediction of battery health impacts operational safety, prevents unexpected failures, and ensures overall system reliability. CAE-AvgPool and CAE-MaxPool exhibited comparatively limited effectiveness, likely due to the potential loss of critical transient information during the pooling process. Conversely, the CAE-StatisticalFeature architecture was robust in filtering anomalies, underlining its potential for computational efficiency and reliable anomaly detection.

The systematic anomaly filtering approach, integrated through autoencoders, reduced potential biases associated with sample selection. This finding represents an indirect yet practical validation of autoencoder applications for real-world anomaly detection, as evidenced by the clear performance improvement when comparing the baseline ResNet performance to that of anomaly-filtered models. The chosen anomaly detection threshold—set at the 80th percentile of reconstruction errors—was intentionally selected to transparently demonstrate the beneficial effect of autoencoder-based filtering rather than to establish optimal anomaly detection performance metrics. Future research could benefit from integrating advanced anomaly detection strategies, such as hybrid models, probabilistic frameworks, and dynamic thresholding approaches [11], to further enhance accuracy.

Despite achieving significant advancements, the persistent gap between training and validation errors indicates the need for more comprehensive datasets. Increasing validation errors with higher k-fold splits further support the necessity for expanded datasets to enhance model generalizability and accuracy. Future studies should therefore continue to explore advanced anomaly detection methodologies, assess their applicability across varied operational conditions and battery chemistries, and leverage larger datasets to ensure improved robustness and enhanced real-world predictive performance.

4. Conclusions

This study proposes a deep learning framework that combines autoencoder-based anomaly detection with a ResNet-based neural network for accurate and reliable cycle-by-cycle RUL prediction of lithium-ion batteries. By integrating anomaly detection into the preprocessing pipeline, the framework effectively addresses the challenges of anomalous battery data, resulting in significantly enhanced prediction stability and accuracy. Among the tested architectures, the Simple-Seq2Seq autoencoder proved most effective, yielding the lowest prediction errors and demonstrating superior generalization performance.

A key strength of the methodology lies in its systematic approach to identifying and filtering anomalous data, which improves dataset quality and model robustness. The experimental results confirm the effectiveness of the proposed framework, achieving an MAPE of 2.85% and an RMSE of 40.87 cycles, outperforming prior single-cycle RUL prediction studies. The results highlight the importance of comprehensive preprocessing and anomaly management in enhancing predictive reliability in real-world applications.

While this study focuses on convolutional and sequence-based autoencoder architectures, other promising variants, such as variational and adversarial autoencoders, were not explored due to the significant time and computational resources required for their training and optimization. Future work will aim to extend the framework by incorporating these advanced architectures, along with adaptive thresholding techniques and hybrid anomaly detection strategies.

Author Contributions: Conceptualization, J.L. (Junghwan Lee); methodology, J.L. (Junghwan Lee); software, J.L. (Junghwan Lee), B.L. and D.K.; validation, J.L. (Junghwan Lee), L.W. and J.L. (Jiaxin Liu); formal analysis, J.L. (Junghwan Lee); investigation, J.L. (Junghwan Lee) and H.J.; resources, D.K. and B.L.; data curation, J.L. (Junghwan Lee), D.K. and B.L.; writing—original draft preparation, J.L. (Junghwan Lee); writing—review and editing, J.L. (Junghwan Lee), L.W. and J.L. (Jong Lim); visualization, J.L. (Junghwan Lee) and J.L. (Jiaxin Liu); supervision, J.L. (Junghwan Lee); project administration, J.L. (Junghwan Lee); funding acquisition, H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was part of the BatterMachine research and development project funded by the Korea Technology & Information Promotion Agency for SMEs of South Korea (Grant No. 00442413).

Data Availability Statement: Codes are available on GitHub: https://github.com/BatterMachineCo/dnn_autoencoders.git GitHub Repository, accessed on 27 July 2025.

Conflicts of Interest: Authors Junghwan Lee, Hoseok Jung, Bukyu Lim, Dael Kim and Jong Lim were employed by the company Research and Development Department, Batter Machine Co., Ltd. Authors Junghwan Lee and Jiaxin Liu were employed by the company Livision AI Technology Co., Ltd. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

The autoencoder architectures evaluated in this study utilize two distinct types of input data:

- Time-Series Input (used by CAE-AvgPool, CAE-MaxPool, CAE-Deep, Simple-Seq2Seq, Chunked-Seq2Seq, Hierarchical-Seq2Seq): Each input sample has dimensions of (None, 1000, 8), representing battery cycle measurements over 1000 time-steps, with the following eight sequential features per step:
    1. Voltage
    2. Current
    3. Temperature
    4. Integrated charge capacity (QC)
    5. Integrated discharge capacity (QD)
    6. Linearly interpolated discharge capacity (QDlin)
    7. Linearly interpolated temperature (Tdlin)
    8. Internal resistance (IR)
- Statistical Feature Input (CAE-StatisticalFeature only): This input consists of a single timestep per cycle with 14 summary statistics derived from each battery cycle, matching the features utilized in the ResNet model for RUL prediction:
    1. Average voltage
    2. Maximum voltage
    3. Minimum voltage
    4. Average temperature
    5. Maximum temperature
    6. Minimum temperature
    7. Integrated charge capacity (QC)
    8. Integrated discharge capacity (QD)
    9. Linearly interpolated discharge capacity (QDlin)
    10. Linearly interpolated temperature (Tdlin)
    11. Internal resistance (IR)
    12. Average incremental capacity (dQ/dV)
    13. Maximum incremental capacity (dQ/dV)
    14. Minimum incremental capacity (dQ/dV)

Table A1. The architecture consists of convolutional layers with progressively increasing filters, batch normalization for stable training, average pooling layers for dimensionality reduction, and upsampling layers for reconstruction.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 1000, 8)	0
Conv1D	(None, 1000, 32)	800
BatchNormalization	(None, 1000, 32)	128
AveragePooling1D	(None, 500, 32)	0
Conv1D	(None, 500, 64)	6208
BatchNormalization	(None, 500, 64)	256
AveragePooling1D	(None, 250, 64)	0
Conv1D	(None, 250, 128)	24,704
BatchNormalization	(None, 250, 128)	512
AveragePooling1D	(None, 125, 128)	0
UpSampling1D	(None, 250, 128)	0
Conv1D	(None, 250, 64)	24,640
BatchNormalization	(None, 250, 64)	256
UpSampling1D	(None, 500, 64)	0
Conv1D	(None, 500, 32)	6176
BatchNormalization	(None, 500, 32)	128
UpSampling1D	(None, 1000, 32)	0
Conv1D	(None, 1000, 8)	776
BatchNormalization	(None, 1000, 8)	32

Table A2. The model architecture utilizes convolutional layers, batch normalization, max pooling for dimensionality reduction, and upsampling for reconstructing the input dimensions.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 1000, 8)	0
Conv1D	(None, 1000, 8)	648
BatchNormalization	(None, 1000, 8)	32
MaxPooling1D	(None, 100, 8)	0
Conv1D	(None, 100, 8)	648
BatchNormalization	(None, 100, 8)	32
MaxPooling1D	(None, 10, 8)	0
Conv1D	(None, 10, 8)	648
BatchNormalization	(None, 10, 8)	32
MaxPooling1D	(None, 1, 8)	0
UpSampling1D	(None, 10, 8)	0
Conv1D	(None, 10, 8)	648
BatchNormalization	(None, 10, 8)	32
UpSampling1D	(None, 100, 8)	0
Conv1D	(None, 100, 8)	648
BatchNormalization	(None, 100, 8)	32
UpSampling1D	(None, 1000, 8)	0
Conv1D	(None, 1000, 8)	648
BatchNormalization	(None, 1000, 8)	32

Table A3. This deep convolutional autoencoder architecture employs multiple convolutional layers with increasing filters to capture hierarchical feature representations, followed by average pooling layers for dimensionality reduction, and corresponding upsampling and convolutional layers for accurate reconstruction.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 1000, 8)	0
Conv1D	(None, 1000, 128)	10,368
BatchNormalization	(None, 1000, 128)	512
AveragePooling1D	(None, 100, 128)	0
Conv1D	(None, 100, 64)	81,984
BatchNormalization	(None, 100, 64)	256
AveragePooling1D	(None, 10, 64)	0
Conv1D	(None, 10, 32)	20,512
BatchNormalization	(None, 10, 32)	128
AveragePooling1D	(None, 1, 32)	0
Conv1D	(None, 1, 16)	528
UpSampling1D	(None, 10, 16)	0
Conv1D	(None, 10, 32)	5152
BatchNormalization	(None, 10, 32)	128
UpSampling1D	(None, 100, 32)	0
Conv1D	(None, 100, 64)	20,544
BatchNormalization	(None, 100, 64)	256
UpSampling1D	(None, 1000, 64)	0
Conv1D	(None, 1000, 128)	82,048
BatchNormalization	(None, 1000, 128)	512
Conv1D	(None, 1000, 8)	1032

Table A4. This architecture employs LSTM layers to capture sequential dependencies, compressing and subsequently reconstructing the input sequences.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 1000, 8)	0
LSTM (Encoder)	(None, 8)	544
RepeatVector	(None, 1000, 8)	0
LSTM (Decoder)	(None, 1000, 8)	544

Table A5. The architecture splits input sequences into chunks for localized sequential modeling, utilizing TimeDistributed LSTM layers to process each chunk individually and reconstruct the original data shape.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 1000, 8)	0
Reshape	(None, 10, 100, 8)	0
TimeDistributed (LSTM Encoder)	(None, 10, 8)	544
TimeDistributed (RepeatVector)	(None, 10, 100, 8)	0
Reshape	(None, 10, 100, 8)	0
TimeDistributed (LSTM Decoder)	(None, 10, 100, 8)	544
Reshape	(None, 1000, 8)	0

Table A6. This hierarchical sequence-to-sequence model captures both local and global temporal dependencies by processing segmented data chunks through TimeDistributed and LSTM layers, enabling detailed reconstruction of the input sequence.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 1000, 8)	0
Reshape	(None, 10, 100, 8)	0
TimeDistributed (LSTM Encoder)	(None, 10, 8)	544
LSTM (Global Encoder)	(None, 8)	544
RepeatVector	(None, 10, 8)	0
LSTM (Global Decoder)	(None, 10, 8)	544
TimeDistributed (RepeatVector)	(None, 10, 100, 8)	0
Reshape	(None, 10, 100, 8)	0
TimeDistributed (LSTM Decoder)	(None, 10, 100, 8)	544
Reshape	(None, 1000, 8)	0

Table A7. The architecture utilizes convolutional layers with batch normalization for feature extraction, pooling layers for dimensionality reduction, and upsampling layers with convolutional operations followed by a dense layer for reconstruction of statistical input features.

Layer (Type)	Output Shape	Number of Parameters
InputLayer	(None, 14, 1)	0
Conv1D	(None, 7, 128)	512
BatchNormalization	(None, 7, 128)	512
Conv1D	(None, 4, 64)	24,640
BatchNormalization	(None, 4, 64)	256
Conv1D	(None, 2, 32)	6176
BatchNormalization	(None, 2, 32)	128
Conv1D	(None, 2, 16)	1552
BatchNormalization	(None, 2, 16)	64
Conv1D	(None, 2, 1)	17
UpSampling1D	(None, 4, 1)	0
Conv1D	(None, 4, 16)	64
BatchNormalization	(None, 4, 16)	64
UpSampling1D	(None, 8, 16)	0
Conv1D	(None, 8, 32)	1568
BatchNormalization	(None, 8, 32)	128
UpSampling1D	(None, 16, 32)	0
Conv1D	(None, 16, 64)	6208
BatchNormalization	(None, 16, 64)	256
Conv1D	(None, 16, 128)	24,704
Flatten	(None, 2048)	0
Dense	(None, 14)	28,686
Reshape	(None, 14, 1)	0

References

1. Bose, B.; Shaosen, S.; Li, W.; Gao, L.; Wei, K.; Garg, A. Cloud-Battery management system based health-aware battery fast charging architecture using error-correction strategy for electric vehicles. *Sustain. Energy Grids Netw.* **2023**, *36*, 101197. [CrossRef]
2. Lee, J.; Sun, H.; Liu, Y.; Li, X.; Liu, Y.; Kim, M. State-of-Health Estimation and Anomaly Detection in Li-Ion Batteries Based on a Novel Architecture with Machine Learning. *Batteries* **2023**, *9*, 264. [CrossRef]
3. Sulzer, V.; Mohtat, P.; Aitio, A.; Lee, S.; Yeh, Y.T.; Steinbacher, F.; Khan, M.U.; Lee, J.W.; Siegel, J.B.; Stefanopoulou, A.G.; et al. The challenge and opportunity of battery lifetime prediction from field data. *Joule* **2021**, *5*, 1934–1955. [CrossRef]
4. Chen, K.; Zhu, J.; Zhou, S.; Liu, K.; Gao, B.; Gao, G.; Wu, G. Impedance estimation for lithium-ion battery using support vector regression and cuckoo search algorithm. *CSEE J. Power Energy Syst.* **2023**. [CrossRef]
5. Li, Y.; Gao, G.; Chen, K.; He, S.; Liu, K.; Xin, D.; Luo, Y.; Long, Z.; Wu, G. State-of-health prediction of lithium-ion batteries using feature fusion and a hybrid neural network model. *Energy* **2025**, *319*, 135163. [CrossRef]
6. Ren, L.; Dong, J.; Wang, X.; Meng, Z.; Zhao, L.; Deen, M.J. A Data-Driven Auto-CNN-LSTM Prediction Model for Lithium-Ion Battery Remaining Useful Life. *IEEE Trans. Ind. Inform.* **2021**, *17*, 3478–3487. [CrossRef]
7. Zhang, J.; Wang, Y.; Jiang, B.; He, H.; Huang, S.; Wang, C.; Zhang, Y.; Han, X.; Guo, D.; He, G.; et al. Realistic Fault Detection of Li-Ion Battery via Dynamical Deep Learning. *Nat. Commun.* **2023**, *14*, 5940. [CrossRef]
8. Xia, W.; Xu, J.; Liu, B.; Duan, H. A Novel Denoising Autoencoder Hybrid Network for Remaining Useful Life Estimation of Lithium-Ion Batteries. *Energy Sci. Eng.* **2024**, *12*, 3390–3400. [CrossRef]
9. Hong, S.; Kang, M.; Kim, J.; Baek, J. Sequential Application of Denoising Autoencoder and Long-Short Recurrent Convolutional Network for Noise-Robust Remaining-Useful-Life Prediction Framework of Lithium-Ion Batteries. *Comput. Ind. Eng.* **2023**, *179*, 109231. [CrossRef]
10. Jeon, J.; Cheon, H.; Jung, B.; Kim, H. ProADD: Proactive Battery Anomaly Dual Detection Leveraging Denoising Convolutional Autoencoder and Incremental Voltage Analysis. *Appl. Energy* **2024**, *373*, 123757. [CrossRef]
11. Sun, C.; He, Z.; Lin, H.; Cai, L.; Cai, H.; Gao, M. Anomaly Detection of Power Battery Pack Using Gated Recurrent Units Based Variational Autoencoder. *Appl. Soft Comput.* **2023**, *132*, 109903. [CrossRef]
12. He, X.; Sun, B.; Zhang, W.; Su, X.; Ma, S.; Li, H.; Ruan, H. Inconsistency Modeling of Lithium-Ion Battery Pack Based on Variational Auto-Encoder Considering Multi-Parameter Correlation. *Energy* **2023**, *277*, 127409. [CrossRef]
13. Tao, S.; Ma, R.; Zhao, Z.; Ma, G.; Su, L.; Chang, H.; Chen, Y.; Liu, H.; Liang, Z.; Cao, T.; et al. Generative Learning Assisted State-of-Health Estimation for Sustainable Battery Recycling with Random Retirement Conditions. *Nat. Commun.* **2024**, *15*, 10154. [CrossRef] [PubMed]
14. Chan, J.; Han, T.; Pan, E. Variational Autoencoder-Driven Adversarial SVDD for Power Battery Anomaly Detection on Real Industrial Data. *J. Energy Storage* **2024**, *103*, 114267. [CrossRef]
15. Jiao, R.; Peng, K.; Dong, J. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Conditional Variational Autoencoders-Particle Filter. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8831–8843. [CrossRef]
16. Zhang, H.; Guo, K.; Chen, Y.; Sun, J. Remaining Useful-Life Prediction of Lithium Battery Based on Neural-Network Ensemble via Conditional Variational Autoencoder. *Appl. Intell.* **2025**, *55*, 34. [CrossRef]
17. Cai, L.; Li, J.; Xu, X.; Jin, H.; Meng, J.; Wang, B.; Wu, C.; Yang, S. Automatically Constructing a Health Indicator for Lithium-Ion Battery State-of-Health Estimation via Adversarial and Compound Stacked Autoencoder. *J. Energy Storage* **2024**, *84*, 110711. [CrossRef]
18. Cai, L.; Yan, J.; Jin, H.; Meng, J.; Peng, J.; Wang, B.; Liang, W.; Teodorescu, R. A Two-Stage Method with Twin Autoencoders for the Degradation Trajectories Prediction of Lithium-Ion Batteries. *J. Energy Chem.* **2025**, *103*, 759–772. [CrossRef]
19. Guille-Escuret, C.; Rodriguez, P.; Vazquez, D.; Mitliagkas, I.; Monteiro, J. CADet: Fully Self-Supervised Anomaly Detection with Contrastive Learning. In Proceedings of the Advances in Neural Information Processing Systems 36 (NeurIPS 2023), New Orleans, LA, USA, 10–16 December 2023; Curran Associates Inc.: Red Hook, NY, USA, 2023.
20. Raghuram, J.; Chandrasekaran, V.; Jha, S.; Banerjee, S. A General Framework For Detecting Anomalous Inputs to DNN Classifiers. In Proceedings of the 38th International Conference on Machine Learning (ICML 2021), Virtual, 18–24 July 2021; Volume 139, pp. 8764–8775.
21. de la Iglesia, D.H.; Corbacho, C.C.; Dib, J.Z.; Alonso-Secades, V.; López Rivero, A.J. Advanced Machine Learning and Deep Learning Approaches for Estimating the Remaining Life of EV Batteries—A Review. *Batteries* **2025**, *11*, 17. [CrossRef]
22. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D.; et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **2019**, *4*, 383–391. [CrossRef]
23. Lu, J.; Xiong, R.; Tian, J.; Wang, C.; Hsu, C.W.; Tsou, N.T.; Sun, F.; Li, J. Battery degradation prediction against uncertain future conditions with recurrent neural network enabled deep learning. *Energy Storage Mater.* **2022**, *50*, 139–151. [CrossRef]
24. Hsu, C.W.; Xiong, R.; Chen, N.Y.; Li, J.; Tsou, N.T. Deep neural network battery life and voltage prediction by using data of one cycle only. *Appl. Energy* **2022**, *306*, 118134. [CrossRef]

25. Segler, M.H.S.; Preuss, M.; Waller, M.P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **2018**, *555*, 604–610. [CrossRef]
26. Tang, Y.; Yang, K.; Zheng, H.; Zhang, S.; Zhang, Z. Early prediction of lithium-ion battery lifetime via a hybrid deep learning model. *Measurement* **2022**, *199*, 111530. [CrossRef]
27. Couture, J.; Lin, X. Image- and health indicator-based transfer learning hybridization for battery RUL prediction. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105120. [CrossRef]
28. Safavi, V.; Mohammadi Vaniar, A.; Bazmohammadi, N.; Vasquez, J.C.; Keysan, O.; Guerrero, J.M. Early prediction of battery remaining useful life using CNN-XGBoost model and Coati optimization algorithm. *J. Energy Storage* **2024**, *98*, 113176. [CrossRef]
29. Gong, D.; Gao, Y.; Kou, Y.; Wang, Y. Early prediction of cycle life for lithium-ion batteries based on evolutionary computation and machine learning. *J. Energy Storage* **2022**, *51*, 104376. [CrossRef]
30. Fei, Z.; Zhang, Z.; Yang, F.; Tsui, K.L.; Li, L. Early-stage lifetime prediction for lithium-ion batteries: A deep learning framework jointly considering machine-learned and handcrafted data features. *J. Energy Storage* **2022**, *52*, 104936. [CrossRef]
31. Sanz-Gorrachategui, I.; Wang, Y.; Guillén-Asensio, A.; Bono-Nuez, A.; Martín-del Brío, B.; Orlik, P.V.; Pastor-Flores, P. Remaining Useful Life Estimation of Used Li-Ion Cells With Deep Learning Algorithms Without First Life Information. *IEEE Access* **2024**, *12*, 147798–147808. [CrossRef]
32. Yang, X.; Xie, H.; Zhang, L.; Yang, K.; Liu, Y.; Chen, G.; Ma, B.; Liu, X.; Chen, S. Early-stage degradation trajectory prediction for lithium-ion batteries: A generalized method across diverse operational conditions. *J. Power Sources* **2024**, *612*, 234808. [CrossRef]
33. Zhang, Q.; Yang, L.; Guo, W.; Qiang, J.; Peng, C.; Li, Q.; Deng, Z. A deep learning method for lithium-ion battery remaining useful life prediction based on sparse segment data via cloud computing system. *Energy* **2022**, *241*, 122716. [CrossRef]
34. Lee, J.; Sun, H.; Liu, Y.; Li, X. A machine learning framework for remaining useful lifetime prediction of li-ion batteries using diverse neural networks. *Energy AI* **2024**, *15*, 100319. [CrossRef]
35. Cao, R.; Zhang, Z.; Shi, R.; Lu, J.; Zheng, Y.; Sun, Y.; Liu, X.; Yang, S. Model-Constrained Deep Learning for Online Fault Diagnosis in Li-Ion Batteries Over Stochastic Conditions. *Nat. Commun.* **2025**, *16*, 1651. [CrossRef] [PubMed]
36. Attia, P.M.; Grover, A.; Jin, N.; Severson, K.A.; Markov, T.M.; Liao, Y.H.; Chen, M.H.; Cheong, B.; Perkins, N.; Yang, Z.; et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **2020**, *578*, 397–402. [CrossRef] [PubMed]
37. Tebbe, J.; Hartwig, A.; Jamali, A.; Senobar, H.; Wahab, A.; Kabak, M.; Kemper, H.; Khayyam, H. Innovations and prognostics in battery degradation and longevity for energy storage systems. *J. Energy Storage* **2025**, *114*, 115724. [CrossRef]
38. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Review

Physics-Informed Neural Networks for Advanced Thermal Management in Electronics and Battery Systems: A Review of Recent Developments and Future Prospects

Zichen Du [1,*,†] and Renhao Lu [2,†]

[1] Ansys, Inc., Austin, TX 78758, USA
[2] Nancy E. and Peter C. Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY 14853, USA; rl839@cornell.edu
* Correspondence: joey.du@ansys.com
† These authors contributed equally to this work.

Abstract: The growing complexities, power densities, and cooling demands of modern electronic systems and batteries—such as three-dimensional integrated circuit chip packaging, printed circuit board assemblies, and electronics enclosures—have pushed the urgency for efficient and dynamic thermal management strategies. Traditional numerical methods like computational fluid dynamics (CFD) and the finite element method (FEM) are computationally impractical for large-scale or real-time thermal analysis, especially when dealing with complex geometries, temperature-dependent material properties, and rapidly changing boundary conditions. These approaches typically require extensive meshing and repeated simulations for each new scenario, making them inefficient for design exploration or optimization tasks. Physics-informed neural networks (PINNs) emerge as a powerful alternative approach that incorporates physical principles such as mass and energy conservation equations into deep learning models. This approach delivers rapid and adaptable resolutions to the partial differential equations that govern heat transfer and fluid dynamics. This review examines the basic principle of PINN and its role in thermal management for electronics and batteries, from the small unit scale to the system scale. We highlight recent advancements in PINNs, particularly their superior performance compared to traditional CFD methods. For example, studies have shown that PINNs can be up to 300,000 times faster than conventional CFD solvers, with temperature prediction differences of less than 0.1 K in chip thermal models. Beyond speed, we explore the potential of PINNs in enabling efficient design space exploration and predicting outcomes for previously unseen scenarios. However, challenges such as training convergence in fine-grained or large-scale applications remain. Notably, research combining PINNs with LSTM networks for battery thermal management at a 2.0 C charging rate has achieved impressive results—an R^2 of 0.9863, a mean absolute error (MAE) of 0.2875 °C, and a root mean square error (RMSE) of 0.3306 °C—demonstrating high predictive accuracy. Finally, we propose future research directions that emphasize the integration of PINNs with advanced hardware and hybrid modeling techniques to advance thermal management solutions for next-generation electronics and battery systems.

Keywords: physics-informed neural networks; thermal management; partial differential equation (PDE); electronics; battery; energy system; heat transfer; fluid dynamics; computational fluid dynamics (CFD)

1. Introduction

Electronics and battery systems, which stand out as the two primary contributors to power consumption and heat generation in modern devices, drive the need for advanced thermal management solutions [1,2]. Effective thermal management is essential for maintaining the long-term performance and reliability of electronics and battery systems, particularly as modern devices continue to shrink in size while increasing in power, presenting significant challenges. Consequently, evaluating the thermal performance of these systems during the early design phase is critical, as pre-screening design parameters and manufacturing processes can demand substantial resources [3]. Achieving rapid and precise thermal predictions for various design parameter combinations is thus a key priority. Historically, the finite element method (FEM) and computational fluid dynamics (CFD) have been employed, leveraging computational resources to conduct numerical analyses of temperature profiles and fluid flow based on the fundamental equations governing mass, momentum, and energy conservation [4–7]. The finite volume method or finite element method, which are common numerical techniques, address partial differential equations (PDEs) by dividing the domain into discrete control volumes or elements and integrating the governing equations across them [8]. This process requires a deep understanding of the underlying physics and the creation of control volumes or meshes, which is critical for obtaining accurate results while managing computational resource demands. However, this approach is often slow and inefficient, particularly considering the evolution of electronic systems from individual chips to packages and full systems and from single battery cells to battery packs and large-scale battery systems featuring billions of transistors and high thermal design power. The next-generation power density system can reach over 100 W/cm^2 [9–11]; thus, CFD is becoming increasingly impractical for addressing the thermal simulation needs of today's advanced and rapidly evolving technologies.

As machine learning becomes more popular, benefiting from advancements in algorithms and GPU-based parallel solvers, its integration with thermal management has emerged as a growing trend [12,13]. Typically, machine learning has found widespread application in areas like generative AI [14,15], image processing [16–19], and content creation [20–22]. However, traditional ML and DL methods face several important limitations in scientific and engineering contexts. First, they typically require large volumes of labeled data to train accurate models—a resource that can be expensive or impractical to obtain, especially for training intricate physical systems. Often, generating these datasets relies on extensive, resource-intensive simulations such as CFD, further limiting scalability [23,24]. Second, conventional ML models are highly sensitive to imperfect data, such as noise, sparsity, or bias in training datasets, which can significantly degrade their predictive performance and robustness. This issue is especially problematic in real-world engineering applications, where experimental or simulation data may be limited or of variable quality [25]. Third, most ML models generally function as "black boxes", producing outputs without offering a clear physical interpretation of the internal parameters [26,27]. This lack of transparency makes it difficult to ensure that model predictions are physically realistic or to trust their extrapolation to unseen scenarios. As a result, purely data-driven models may fail to generalize when faced with new boundary conditions, rare events, or out-of-distribution cases that are common in thermal management.

Physics-informed neural networks (PINNs), on the other hand, overcome this limitation by embedding governing physical principles, typically in the form of PDEs, directly into the training process [28]. The physics-informed technique can participate in stages of initialization, loss function calculation, and architecture design or be hybridized with conventional DL models. This approach enables PINNs to deliver accurate predictions even with limited data, making them well-suited for thermal management tasks governed

by established physical laws, such as Fourier's law of heat conduction or the Navier–Stokes equations for fluid dynamics [29–32]. Unlike conventional neural networks, PINNs are trained not only on observational data but also on the underlying physics, which serves as a regularization mechanism, narrowing the range of possible solutions and improving their ability to generalize. This makes PINNs particularly advantageous in situations in which traditional numerical methods are too slow or computationally demanding, such as real-time thermal regulation or solving inverse problems like parameter estimation.

Recent research has demonstrated the versatility of PINNs across various thermal management domains. For instance, in building thermal modeling, PINNs have been used to develop control-oriented models that combine the interpretability of physical laws with the expressive power of neural networks, as seen in studies like Gokhale et al.'s work on control-oriented thermal models for buildings [33]. Similarly, Wang et al.'s research on heat transfer in porous media using PINNs demonstrated their ability to accurately predict temperature and heat flux fields without labeled data, achieving computation accelerations five orders of magnitude greater than those of numerical methods [34]. More examples are discussed in Sections 2 and 3. These efforts suggest that PINNs could revolutionize thermal management in electronics, although the field is still emerging, with fewer direct studies compared to other areas.

In this review paper, we will discuss the applications of PINNs in electronics and battery systems. We will begin by explaining the foundational working principles of the PINN and its variants in Section 2. In Section 3, we will explore PINN research conducted in electronics thermal management at different scales, from chips to board to systems. In Section 4, following the same scale category, we describe the PINN research conducted in battery thermal management, from single cells to battery packs to battery systems. Within these sections, based on this understanding of the challenges related to different-scale thermal management in electronics and batteries, we will then evaluate the integration of PINNs with other machine learning techniques and explore variations of the PINN framework. Lastly, we will outline potential future opportunities and prospects for leveraging PINNs in these domains.

2. PINNs and Variations

PINNs represent a significant paradigm shift in scientific machine learning, particularly in solving PDEs governing heat transfer, fluid dynamics, and multiphysics problems. Unlike conventional deep learning models, which require extensive labeled datasets, PINNs embed physical laws—such as energy conservation and material properties—directly into the training process. This hybrid approach enables PINNs to generate accurate, physically consistent solutions, even with limited experimental or simulated data, making them highly valuable for thermal modeling in electronics and battery systems.

2.1. Mathematical Formulation of PINNs

PINNs approximate the solutions of PDEs, which can be generally expressed as follows:

$$\mathcal{N}[\boldsymbol{u}(\mathbf{x}, t)] = 0,\ \mathbf{x} \in \boldsymbol{\Omega},\ t \in [0,\ T] \tag{1}$$

where $\mathcal{N}$ represents a differential operator that encodes the governing physics; $\boldsymbol{\Omega}$ is the computational domain; and $\boldsymbol{u}(\mathbf{x}, t)$ is the solution of the PDE, with $\mathbf{x}$ as the coordinate and t as the time of the physical field.

Heat transfer in electronics and battery systems is primarily governed by conduction and convection, with radiation typically being negligible due to small surface areas, moderate temperatures, and low-emissivity materials [35]. In heat conduction simulation, $\mathcal{N}$ is defined according to Fourier's law as follows [36]:

$$\mathcal{N} = \nabla\cdot(k\nabla T) + \dot{q} - \rho c\frac{\partial T}{\partial t} \tag{2}$$

where T is the temperature, k is the thermal conductivity, ρ is the density, c is the specific heat capacity, and $\dot{q}$ is the energy generated per unit volume.

In convection simulation, 3D incompressible flow dynamics are usually considered, where $\mathcal{N}$ consist of two components based on Navier–Stokes equations, as follows [29]:

$$\begin{gathered}\mathcal{N}_1 = \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p - \frac{1}{Re}\nabla^2\boldsymbol{u} \\ \mathcal{N}_2 = \nabla\boldsymbol{u}\end{gathered} \tag{3}$$

where $\boldsymbol{u}$, t, *and* p are the fluid velocity, time, and pressure, respectively, and Re is the Reynolds number.

Based on this equation, a force convection problem, which is a common heat management scenario, can be achieved by coupling the heat conduction and the fluid dynamics. Hence, the PDEs will be defined as follows [37]:

$$\begin{gathered}\mathcal{N}_1 = \frac{\partial T}{\partial t} + (\boldsymbol{u}\cdot\nabla)T - \frac{1}{Pe}\nabla^2 T \\ \mathcal{N}_2 = \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p - \frac{1}{Re}\nabla^2\boldsymbol{u} + RiT \\ \mathcal{N}_3 = \nabla\boldsymbol{u}\end{gathered} \tag{4}$$

where Pe and Ri denote the Peclet and Richardson numbers, respectively.

Meanwhile, PDE problems usually include boundary and initial conditions, as follows:

$$\begin{gathered}\boldsymbol{u}(\mathbf{x}, 0) = \boldsymbol{h}(\mathbf{x}),\ \mathbf{x} \in \boldsymbol{\Omega} \\ \boldsymbol{u}(\mathbf{x},\ t) = \boldsymbol{g}(\mathbf{x}, t), \mathbf{x} \in \partial\boldsymbol{\Omega},\ t \in [0, T]\end{gathered} \tag{5}$$

where $\boldsymbol{\Omega}$ and $\partial\boldsymbol{\Omega}$ are the computational domain and boundary, respectively, $\boldsymbol{h}(\mathbf{x})$ is the initial condition, and $\boldsymbol{g}(\mathbf{x}, t)$ represents the boundary conditions. The backbone of the PINN is usually a deep neural network (DNN), which takes $\mathbf{x}$ and t as inputs and outputs the approximated solution of the PDE.

To calculate the solution, the PINN minimizes the non-negative residual error associated with Equations (1) and (5).

$$\begin{gathered}\mathcal{L}_N = \int_{[0,T]\times\boldsymbol{\Omega}} \left(\mathcal{N}[\boldsymbol{u}(\mathbf{x}, t)]\right)^2 dt d\mathbf{x} \\ \mathcal{L}_I = \int_{\boldsymbol{\Omega}} \left(\boldsymbol{u}(\mathbf{x}, 0) - \boldsymbol{h}(\mathbf{x})\right)^2 d\mathbf{x} \\ \mathcal{L}_B = \int_{[0,T]\times\partial\boldsymbol{\Omega}} \left(\boldsymbol{u}(\mathbf{x},\ t) - \boldsymbol{g}(\mathbf{x}, t)\right)^2 dt d\mathbf{x}\end{gathered} \tag{6}$$

PINNs leverage automatic differentiation (AD) from deep learning frameworks to compute the derivatives required for the residual calculation. Unlike finite difference or finite element methods, AD avoids numerical discretization errors. Since evaluating the full integral in residual loss can be computationally expensive, PINNs often employ Monte Carlo sampling [38]. A subset of points $\left(\mathbf{x}^i, t^i\right)$ is randomly sampled from the computational domain, and the mini-batch training algorithm is used to iteratively update the neural network parameters [39].

$$\begin{gathered}\mathcal{L}_N = \frac{1}{N_N}\sum\left(\mathcal{N}\left[\boldsymbol{u}\left(\mathbf{x}^{\boldsymbol{i}}, t^i\right)\right]\right)^2 \\ \mathcal{L}_I = \frac{1}{N_I}\sum\left(\boldsymbol{u}\left(\mathbf{x}^i, 0\right) - \boldsymbol{h}\left(\mathbf{x}^i\right)\right)^2 \\ \mathcal{L}_B = \frac{1}{N_B}\sum\left(\boldsymbol{u}\left(\mathbf{x}^i, t^i\right) - \boldsymbol{g}\left(\mathbf{x}^i, t^i\right)\right)^2\end{gathered} \tag{7}$$

If some ground truth data are available in the computational domain, an additional data-based loss term can also be included.

$$\mathcal{L}_{Data} = \frac{1}{N_{Data}} \sum (\boldsymbol{u}(\mathbf{x}^{\boldsymbol{i}}, t^i) - \boldsymbol{u}_{Data}(\mathbf{x}^{\boldsymbol{i}}, t^i))^2 \tag{8}$$

Hence, the loss function of the PINN can be defined as the combination of the following residuals:

$$\mathcal{L} = \mathcal{L}_N + \mathcal{L}_I + \mathcal{L}_B + \mathcal{L}_{Data} \tag{9}$$

This approach allows PINNs to generalize beyond the training data and maintain consistency with underlying physics, making them particularly effective for thermal management problems with limited experimental data or unknown boundary conditions [29,37]. For a comprehensive overview of PINN methodologies—including neural network architectures, strategies for integrating physical laws, and applications across fields such as fluid dynamics, materials science, energy, and medicine—see the recent review by Farea et al. [40].

2.2. PINN Variants

While the standard PINN framework offers a powerful tool for solving PDE-governed problems with limited data, it still faces notable challenges when applied to complex real-world scenarios—such as those encountered in electronics and battery thermal management. These challenges include (1) difficulty in accurately enforcing boundary conditions, which can degrade solution fidelity; (2) computational inefficiency when scaling to high-dimensional or multiscale problems; (3) poor convergence and susceptibility to local minima, particularly in stiff systems or extrapolation tasks; and (4) limited ability to capture complex physical phenomena like turbulence or multiphysics interactions [41]. To address these challenges, several variants of PINNs have been developed, each focusing on different limitations, which are reviewed in detail in the following subsections.

2.2.1. Balancing Residual and Boundary Losses

One of the most critical challenges in training PINNs is the imbalance between the PDE residual loss and the boundary condition loss, which can significantly hinder convergence and accuracy. This issue often arises because the residual loss, which stems from the PDE constraints across the entire domain, can dominate the boundary loss by several orders of magnitude, leading to the poor satisfaction of boundary conditions. Wang et al. analyzed this phenomenon as a gradient pathology, showing that conventional training dynamics result in vanishing boundary loss gradients, thereby biasing the model toward interior solutions that violate boundary constraints. To address this issue, they proposed a learning rate annealing algorithm and a novel PINN architecture to rebalance gradient flows during training, which led to 50–100$\times$ improvements in predictive accuracy across various benchmark problems [42]. Building on this, Yao et al. introduced MultiAdam, a scale-invariant optimizer that adaptively rescales gradients using parameter-wise second-moment statistics. Unlike manual or static reweighting, MultiAdam automatically harmonizes the contributions of loss terms at different scales, maintaining consistent convergence across complex PDE domains. This method improved solution accuracy by 1–2 orders of magnitude across diverse physics scenarios, demonstrating robust performance in multiscale PINN training [43]. Together, these approaches mark a significant step forward in developing reliable and physically consistent PINNs, particularly for thermal management tasks that demand precision at boundary interfaces.

2.2.2. Adaptive Sampling Strategies

Sampling strategies play a pivotal role in the training dynamics and accuracy of PINNs, as the locations of residual collocation directly influence how well the solution captures the complex features of the PDE. Uniform sampling, while simple and widely adopted, often results in poor performance in regions with steep gradients or localized phenomena. To address this issue, several adaptive sampling strategies have been proposed. Nabian et al. introduced an importance sampling approach that selects collocation points proportional to the loss value, effectively focusing training on regions with greater residuals and accelerating convergence without additional hyperparameters [39]. Wu et al. conducted a comprehensive study and proposed RAD and RAR-D, which dynamically adjust point distributions based on residual magnitudes and outperform traditional strategies with fewer collocation points [44]. Tang et al. advanced this solution further with the DAS-PINN, a generative model-based method that learns the residual distribution, yielding strong results for high-dimensional or irregular PDEs [45]. More recently, Yu et al. proposed MCMC-PINNs, which use a modified Markov Chain Monte Carlo method to sample collocation points according to a canonical distribution based on PDE residuals. This method adapts the proposal distribution to domain geometry and ensures more thorough exploration of complex solution landscapes while maintaining convergence guarantees [46]. Together, these innovations in adaptive sampling significantly enhance both the efficiency and precision of PINNs, particularly in multiscale or unbounded domain problems.

2.2.3. Variational Formulations in PINNs

While traditional PINNs enforce PDEs through their strong form—minimizing residuals at discrete collocation points (Equations (6) and (7))—this approach often suffers from instability, especially when high-order derivatives are involved or when the solution lacks smoothness. Variational formulations offer a powerful alternative by expressing the PDE in its weak form, where the equation is satisfied in an integrated sense against test functions. This allows for smoother losses, reduced differentiation order (via integration by parts), and improved stability, especially in complex or high-dimensional settings. The evolution of variational PINNs reflects growing recognition of these benefits. The earliest prominent example is the Deep Ritz Method by E and Yu, which recasts the PDE solution as the minimizer of a variational energy functional. This method constructs the trial solution using a deep neural network and optimizes an integral form of the PDE's energy without enforcing pointwise residuals, thus avoiding high-order derivative computations and enabling efficient training, even in high dimensions [47]. Following this, the VPINN framework by Kharazmi et al. generalized the approach by embedding PINNs in a Petrov-Galerkin setting. In VPINNs, the trial space consists of neural networks, while the test space is constructed using classical polynomial bases (e.g., Legendre polynomials). This variational residual reduces differential order via integration by parts and replaces dense collocation with sparse quadrature, leading to improved numerical stability and efficiency [48]. Building on VPINNs, the hp-VPINNs introduced adaptive domain decomposition and hierarchical polynomial refinement, enabling localized learning and better handling of sharp gradients or singularities in the solution [49]. Around the same time, VarNet proposed a variational training strategy that is fully discretization-free and operates over space–time volumes rather than isolated points. By training on integral residuals and using adaptive sampling driven by residual feedback, VarNet enables smoother, more sample-efficient learning and is particularly well-suited for parametric and control applications [50]. Collectively, variational approaches provide a powerful and more physically grounded alternative to classic PINNs, making them especially well-suited for problems with irregular domains, lower solution regularity, or high computational complexity.

2.2.4. Domain Decomposition PINNs

As PINNs are extended to large-scale or multiphysics systems, they often suffer from high computational costs and poor convergence, especially when trying to capture complex or discontinuous physical phenomena [41]. To address these limitations, domain decomposition techniques have been incorporated into the PINN framework, giving rise to variations such as Conservative PINNs (cPINNs) [51] and eXtended PINNs (XPINNs) [52]. These architectures divide the computational domain into smaller subdomains, within which localized PINNs are trained. This strategy not only enhances scalability and parallelizability but also enables tailored network architectures for different subregions of the problem domain.

The cPINN framework focuses on conservation laws, enforcing continuity in both the solution and flux across the boundaries of decomposed subdomains. Each subdomain employs a separate neural network, with interface conditions ensuring physical consistency by stitching local solutions together [51]. This includes enforcing average solution continuity and flux conservation at shared interfaces, which is a critical step for solving hyperbolic PDEs like the Euler equations (Figure 1). Based on this approach, XPINNs further generalize the domain decomposition concept beyond conservation laws, supporting arbitrary space–time decompositions for any type of PDEs. This includes both convex and non-convex geometries, time-dependent or time-independent problems, and even cases with moving interfaces [52]. Each subdomain is governed by its own neural network and optimized independently. XPINNs introduced interface conditions such as residual continuity and average solution enforcement, allowing seamless stitching across irregular domains. In order to fully leverage the advantages of the cPINN and XPINN, a parallelized implementation has also been proposed based on a hybrid MPI + X programming model (where X can be CPUs or GPUs). This enables the efficient training of PINNs on distributed hardware. The parallel framework supports both weak and strong scaling and significantly reduces training time by exploiting localized computation within subdomains.

The physics-informed neural network (PINN) has emerged as a powerful modeling framework for electronics thermal management (ETM), offering a compelling alternative to conventional numerical solvers. By embedding physical laws—such as Fourier's law and the Navier–Stokes equations (Equations (2)–(4))—directly into the training process, PINNs eliminate the need for large datasets and deliver fast, accurate solutions to partial differential equations, even in data-sparse or complex settings. As summarized in Table 1, the adaptability of PINNs is further strengthened by a rich ecosystem of variants designed to overcome key training and scalability challenges. For instance, advanced optimizers like MultiAdam and reweighting strategies mitigate the imbalance between PDE residual and boundary losses, improving solution fidelity. Adaptive sampling methods, such as importance sampling, DAS-PINNs, and MCMC-PINNs, dynamically allocate collocation points where they are most needed, boosting convergence and efficiency in multiscale and sharp-gradient regions. Variational formulations—including the Deep Ritz Method, VPINNs, and VarNet—reformulate PINNs in a weak form, reducing the order of derivatives and enhancing stability for irregular geometries. Domain decomposition approaches like cPINNs and XPINNs enable parallelization and localized learning, making PINNs scalable to large or heterogeneous ETM problems. Among these variants, domain decomposition methods (such as cPINNs and XPINNs) are particularly promising for electronics thermal management, where modeling large-scale, multi-material systems and capturing localized hot spots are critical. Their main advantage lies in improved scalability and parallelizability; however, their practical implementation can be complex, especially when managing interface conditions between subdomains. Loss balancing strategies are also valuable for both electronics and battery applications, as they improve the accuracy of

boundary and interface temperature predictions, which is essential for device reliability and safety. The main challenge with these approaches is that optimal weighting can be problem-dependent and may require extensive hyperparameter tuning. In battery systems, adaptive sampling approaches show strong potential, as they can efficiently target regions with rapid thermal changes or sparse data, which are common issues in battery packs and large-format cells. Ultimately, the choice of variant depends on the complexity and scale of the system being modeled, the available data, and computational resources.

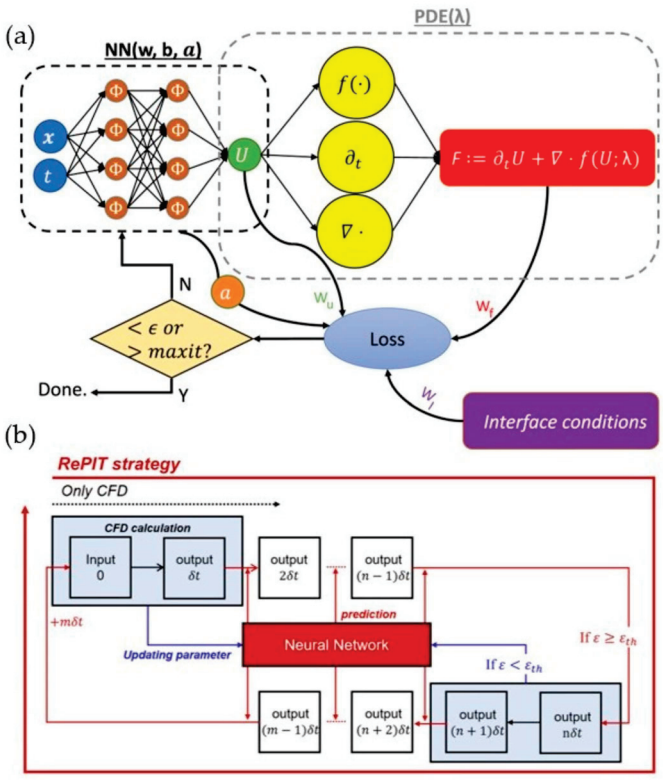


Figure 1. Schematics of PINN and its variations. (**a**) Schematic of the conservative physics-informed neural network (cPINN) architecture [51]. The addition of interface loss distinguishes cPINNs from traditional PINNs, enabling improved performance on conservation laws and problems with sharp gradients. (**b**) The residual-based physics-informed transfer learning (RePIT) strategy as a representative hybrid PINN approach [53]. The workflow alternates between conventional CFD computation and neural network-based predictions. The figure has been republished with permission from each indicated reference (ref. [51] for (**a**), ref. [53] for (**b**)).

Table 1. Summary of major PINN variants and their core innovations.

Category	Key Methods	Highlights
Loss Balancing	MultiAdam [43], Gradient Reweight [42]	Adaptive optimizer, rebalancing loss terms
Sampling Strategies	Importance sampling [39], RAD [42], DAS-PINNs [45], MCMC-PINNs [46]	Adaptive and probabilistic point selection
Variational Form	Deep Ritz [47], VPINNs [48,49], VarNet [50]	Weak form enforcement, lower derivative order
Domain Decomposition	cPINNs [51], XPINNs [52]	Local networks, interface stitching

While many machine learning approaches are fundamentally limited by the size and quality of available datasets, a key advantage of PINNs is their ability to leverage

governing physical laws, thereby reducing reliance on large experimental or simulated datasets. In PINN frameworks, the physical equations themselves guide model learning, enabling accurate solutions even when data are sparse or partially missing. However, the effectiveness of PINNs can still be influenced by data-related factors such as measurement noise, limited boundary condition data, or incomplete coverage of operating regimes. Thus, while PINNs are more robust to data scarcity than purely data-driven methods, careful attention to the available data—however limited—remains important, particularly in complex or ill-posed real-world scenarios.

3. Applications of PINNs in Electronics Thermal Management

For general heat transfer physics, there are three modes of heat transfer—conduction, convection, and radiation—depending on the heat transfer medium [54]. Conduction heat transfer is governed by Fourier's Law, which states that heat flux is linearly dependent on the temperature gradient when thermal conductivity is constant in a steady-state scenario, or on density and heat capacity. Temperature-dependent thermal conductivity, density, or heat capacity will introduce non-linear physical characteristics into this scenario [55,56]. Convection is governed by Newton's Law, and the heat transfer coefficient depends on fluid properties, such as temperature, fluid velocity, or flow regime (laminar or turbulent) [57]. Radiation is governed by the Stefan–Boltzman Law and is non-linear due to the fourth-power dependence on temperature [58]. With power densities with varying intensities and different cooling requirement scales, thermal management systems employed in electronics can generally be categorized into passive cooling and active cooling [59]. Passive cooling refers to the method that dissipates heat without any active energy inputs; techniques in this category include heat sink, thermal pads/interfaces, heat pipes, and vapor chambers, whereas active cooling requires additional power input, such as a fan, liquid cooling, or jet impingement cooling. Besides these traditional cooling mechanisms, two-phase liquid cooling is increasingly standard in data centers, while immersion cooling is gaining traction for its efficiency [60–62].

The advancement of cutting-edge information and digital technologies—including 5G, artificial intelligence, cloud computing, autonomous vehicles, and data centers—has greatly increased the need for functional electronics, driving demand in their power densities and reliabilities [63,64]. Maintaining a safe operating temperature is essential for the proper functioning of each unit, making thermal management systems very critical. Depending on the physical size of the system, thermal management can be classified into three levels: chip, board, and system [35]. The thermal management challenges vary across the three levels. As depicted in Figure 2, at the chip level (Figure 2a), densely packed power tiles represent each functional unit, with conduction as the dominant heat transfer mode and boundary conditions set by the heat transfer coefficient. Material properties may be isotropic and temperature-dependent, varying based on design accuracy requirements. At the board level (Figure 2b), when it involves a lot of components, such as electronic control module, cooling solutions, PCB, etc., heat transfer becomes more complex as various mediums—such as air or dielectric liquids—come into play. In addition, multiple components with varying power inputs are present, which can include factors like direct current internal resistance losses or electromagnetic power losses involving multiphysics considerations. Despite this increased complexity and the larger number of design parameters, board-level challenges are not necessarily greater than those at the chip level, largely because most passive electrical and heat-generating components exhibit isotropic characteristics. Finally, at the system level (Figure 2c), the intricate geometry of components not only increases the volume of data but also lengthens both the training and forecasting periods.

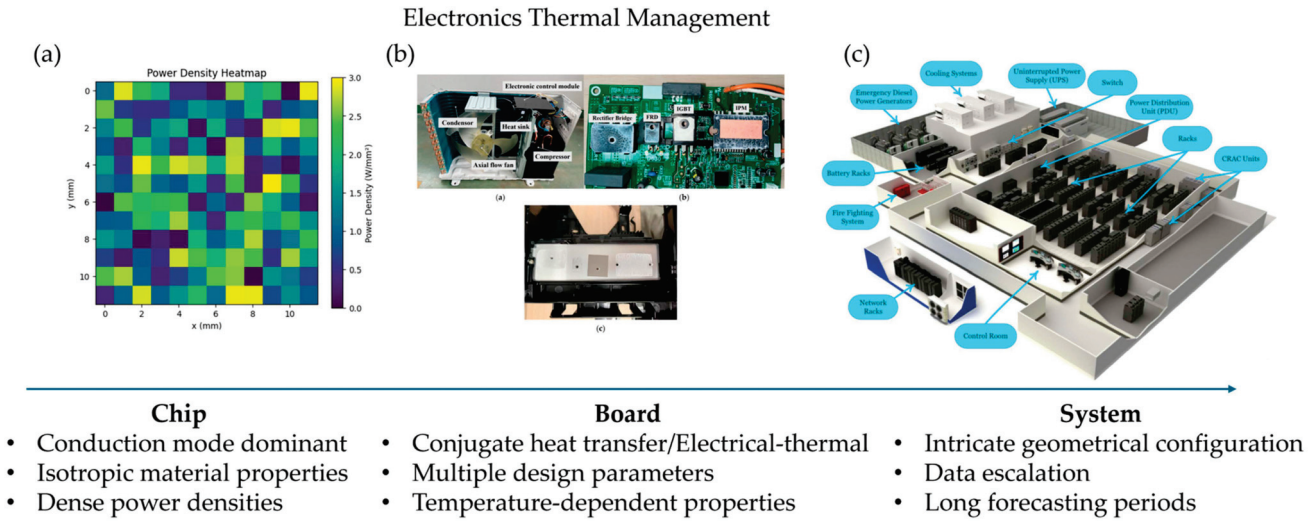


Figure 2. Electronics thermal management challenges on (**a**) Chip, (**b**) Board, and (**c**) System. Image sources: (**b**) [65], (**c**) [66], Figure (**b**) was reproduced from [65] with permission under the terms of the Creative Commons. Figure (**c**) has been republished with permission from each indicated reference [66].

Following this scaling methodology, this section will analyze selective PINN applications at the chip, board, and system levels of electronics thermal management and aims to bridge the connections between different scales.

3.1. Chip Thermal Management

For chip thermal management, heat is generated at the nanoscale with intensive density and is mainly conducted through the die before it reaches the chip surface, is cooled through external multilayer package and thermal interface materials, and is eventually dissipated through external conduction or convection [35]. Therefore, conduction is the dominant heat transfer mode at the chip level, while external conduction or convection can be simplified and simulated with heat transfer coefficient boundary condition settings. Radiation is negligible due to the small surface area and low emissivity (polished surface) [35]. According to Fourier's Law, conduction heat transfer is linear when thermal conductivity is constant. There are two main challenges associated with chip thermal management, namely, temperature-dependent material properties (thermal conductivity, density, and heat capacity) and high-dimensional power densities due to the hierarchical structure or advanced 3D IC structure [67].

Chip technology has become very dense with power tiles, requiring a very fine resolution for each power tile to be captured correctly, which is not possible with the current finite element analysis (FEA) approach, as the discretization of the full chip level is extremely time-consuming and resource-intensive. The high-dimensional and nonlinear PDEs have also been designed to accelerate using machine learning approaches. Traditional machine learning approaches to chip thermal management require large data inputs. For example, Sadiqbatcha et al. built the long short-term memory (LSTM) model to train experimental infrared thermal image results to estimate representative spatial features of 2D heatmaps with similar accuracy and high efficiency [68]. Chen et al. [69] adopted graph convolution networks (GCN) with global features, skip connections, edge-based attention, and principle neighborhood aggregation to efficiently estimate thermal maps of 2.5D chiplet-based systems while demonstrating strong generalization to unseen datasets.

On the other hand, PINN demonstrates good potential for overcoming the data-intensive limitations of traditional machine learning. Liu et al. applied the physics-aware

operator learning method (DeepONet) to a 21 × 21 × 11 mesh grid-based single-cuboid geometry with a 2D power map with constant HTC boundary conditions [70]. In the study, firstly, two families of design configurations, namely, boundary conditions for each individual surface and the locations and intensity of external or internal heat sources, were encoded as input functions for different "branch nets" to be fed into the framework. While all the sampled coordinates were fed into another sub-network, namely, "trunk net", the k branch nets, and one trunk net were combined via a Hadamard (elementwise) product and summed to represent the predicted temperature field. The framework was then trained using multi-input DeepONet (Figure 3) [71]. The total loss was minimized using gradient descent based on an automatic differentiation algorithm. The result was 300,000 times faster than the commercial Celsius 3D solver, with a max/min temperature difference of less than 0.1 K. However, this method reveals limitations in scalability and generalization, as it does not include orthotropic or temperature-dependent material properties.

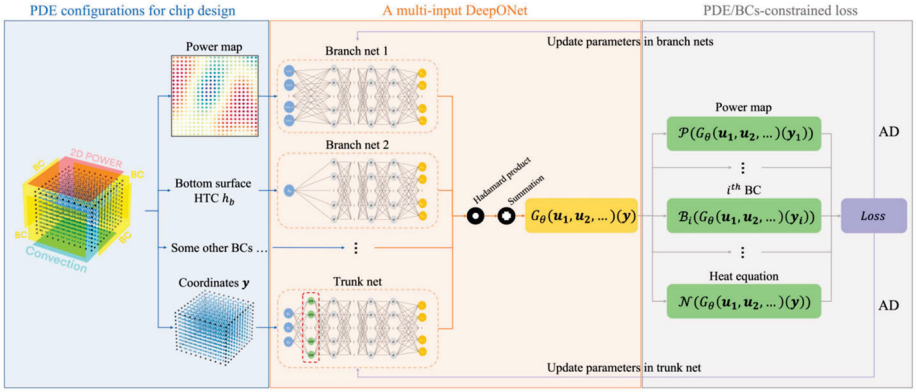


Figure 3. DeepOHeat framework. Figure has been republished from [71] with permission under the terms of the Creative Commons.

Another study conducted by Chen et al. involved an enhanced PINN approach designed for a rapid and accurate full-chip thermal analysis of Very Large-Scale Integration (VLSI) chips, namely, ThermPINN [72]. Standard PINNs, while innovative, suffer from slow training convergence. In Chen's work, the temperature distribution equation can be expressed with the separation of variables into two variables along the x and y directions to form two cosine vectors with the coefficient C_{pq}. The authors also considered that both thermal conductivity and leakage power vary with temperature and used appropriate models to depict their relationship to temperature across a specific temperature range. Firstly, spatial variables are separated into cosine vectors, forming a Q × P matrix whose inner product with the C_{pq} matrix yields the temperature value with a discrete cosine neural network. Second, the effective convection coefficient, m, and the ambient temperature, T_0, are parameterized to connect with C_{pq} with multi-layer perception (MLP). A back-propagation algorithm is used to learn the MLP parameters. The thermal equation is then coupled with a loss function, and an unsupervised learning method is used to train the networks. The author also applied a plain PINN as a benchmark, where the position, the effective convection coefficient, and the ambient temperature are directly parametrized to train the MLP.

The findings indicate that the ThermPINN's accuracy was slightly below that of the plain PINN, but in the same order of significance. However, it demonstrated strong potential for faster training, making it more practical for EDA applications, with both PINNs outperforming traditional FEM solutions in terms of speed. The model parameterizes key variables—the ambient temperature and effective convection coefficient—enabling efficient design space exploration and uncertainty quantification (UQ).

3.2. Board Thermal Management

In board thermal management, the electrical components hosted on a printed circuit board—such as resistors, capacitors, memory cards, integrated circuits (ICs), heatsinks, and solder joints—vary widely by application and industry, with some being passive and others producing high-density, dynamic thermal power under different operating conditions [73]. Thermal management poses a significant challenge, although techniques like air cooling, cold plates, jet impingement, and immersion cooling are well-established, addressing the diverse safe temperature limits, space constraints, and power demands across components and product generations, which complicates redesign and simulation efforts. Simplified simulations of board thermal management using black box reduced-order models (e.g., linear time invariant [74], linear parameter varying [75], and singular value decomposition [76]) are possible but often lack field-specific accuracy or require extensive training data, whereas PINNs offer a promising approach for prescreening and estimating untested scenarios.

On the board, there can be power chips with different layers soldering with baseplates or printed circuit boards (PCBs). In a study conducted by Yang et al., four SiC MOSFET chips are mounted on DBC AIN substrates with integrated deionized water pin-fin cooling channels on top. Five parameters, including initial temperature, heat flux, chip total powers, baseplate thickness, and pin-fin heights, are explored in terms of design space [77]. For the SiC power module, due to the presence of layers with different material properties, representing separate physical domains, the authors applied a total of nine Fourier neural networks with soft coupling constraints in PINNs for rapid design exploration. This approach outperforms traditional FEM solvers for faster thermal profile prediction. The seven Fourier networks represent and fit the thermal equations for each physical domain, whereas two Fourier networks address fluid flow and convective heat transfer PDEs, resulting in a total of nine Fourier NNs as training inputs. Random sampling points are then generated uniformly across the domain and boundaries to evaluate a loss function, which integrates PDE residuals, boundary condition errors, and interface continuity penalties, which are balanced using adaptive weights. The model undergoes iterative training via backpropagation with the Adam optimizer and an exponential decay learning rate, minimizing the loss below a threshold (10^{-5}) for synchronized convergence across networks. Once trained, the PINN model enables the rapid inference of thermal fields for any input parameter combination without retraining, offering an efficient alternative to traditional numerical methods. The author suggests that in terms of scalability, PINNs delivered performance similar to that of COMSOL for 100 simulations. However, when expanding to a significantly larger set of simulations—such as 10,000 cases—the PINN-based approach demonstrated substantially higher simulation efficiency, particularly for exploring expansive design spaces. However, PINNs were trained on GPUs while benchmark COMSOL simulations are based on CPUs, which is not a fair comparison due to the parallel computation capabilities of GPUs. As the authors introduced geometrical parameters, mesh-dependent accuracy can be a potential risk when scaling.

Another application involves a cooling component on the board, the thermoelectric cooler (TEC), which consists of a series of N-type and P-type semiconductor materials [78]. When an electric current passes through these materials, heat is absorbed from one side (the cold side) and released on the opposite side (the hot side), which is also called the Peltier effect. Simulating and designing TECs can be challenging due to their complex physics and computational demands: (1) TECs operate based on intertwined thermal and electrical phenomena, governed by nonlinear PDEs and highly influenced by temperature-dependent material properties, including thermal conductivity, the Seeback coefficient, and electrical conductivity; (2) 3D FEA simulation that requires fine spatial discretization results

in large systems of equations and high demand for computer resources; and (3) identifying optimized parameters escalates the complexity. The study conducted by Chen et.al introduced a surrogate model that reduces 3D TEC geometry to a 1D problem that incorporates key parameters like current density and thermal boundary conditions [79]. The implicit physics-constrained neural network (IPCNN) framework proposed by the authors employs a two-stage training process. Firstly, temperature-dependent material properties are approximated using an extreme learning machine (ELM), followed by a PINN enforcing TEC-specific PDEs in the second stage. This method achieves an 8.5× speed increase over traditional COMSOL simulations and enhances stability compared to conventional PINNs. Additionally, a hybrid finite element neural network (FENN) method integrates the surrogate model into COMSOL, yielding a 5.1× speedup and 5.4× memory reduction for VLSI chip thermal analysis. The author demonstrated that this IPCNN approach showed much better convergence to a loss of 10^{-6} versus 10^{-1} for traditional PINNs in smaller length ranges. It avoids the slow convergence and large errors of one-step PINN training by separating the modeling of material properties and PDE solutions, thereby reducing the optimization search space. However, a decrease in accuracy was observed as the parameter ranges widened (e.g., length from 0.05 to 1.2 mm), suggesting scalability limitations, in addition to the increased complexity of implementing a two-stage process compared to a single-step PINN. The scalability of broader parameter ranges based on this study can potentially be achieved by employing multiple neural networks for subregions, as suggested in the text. Further refinement could also involve adaptive learning rates or advanced network architectures to maintain accuracy across diverse conditions. Integrating more physics constraints or exploring alternative approximators beyond ELM could also boost performance. Overall, this study underscores IPCNN's potential as a transformative tool in TEC modeling, balancing efficiency and precision while identifying pathways for future enhancements in board thermal management.

Another study conducted by Farrag et al. involved the soldering reflow process (SRP) [80], which occurs during the PCB manufacturing process. In SRP, solder paste is melted and then solidified to connect electronic components to PCBs. Precise temperature control is critical to ensure the PCB's quality. Due to the escalating complexity of the PCB and a greater number of different electrical components, monitoring and accommodating the SRP process is extremely difficult. The SRP-PINN model proposed by Farrag et al. leverages PINNs to predict temperature distribution across PCBs, ensuring solder joints meet manufacturer-specified thermal profiles for quality assurance. Unlike traditional CFD approaches, which are computationally intensive, the SRP-PINN integrates PDEs into a DNN to achieve accurate predictions with limited experimental data. The study uses a 1D heat transfer model along the PCB length, trains the PINN with sparse data from one recipe, and demonstrates its generalizability across different PCB designs and soldering recipes. The experimental results, which were conducted using a Heller 1707 W reflow oven and SAC305 solder paste, show the model's effectiveness, achieving 98% accuracy compared to 97% for a hybrid physics–ML benchmark, with potential applications in real-time manufacturing optimization.

3.3. System Thermal Management

When it comes to system scale, in addition to the above-mentioned chip- and board-level challenges, intricate geometrical configurations can pose a challenge due to the complexities of CAD geometries, such as in electronics enclosures and large-scale cooling components [81–84]. To precisely capture the fluid flow and temperature profile, setting non-conformal mesh at different regions to balance calculation accuracy and computer resources requires a significant amount of human instruction and experience, let alone the

time to start a simulation [85,86]. A PINN can be very useful in resolving repetitive work and save running time when only several parameters need to be modified for a new design.

In the data center field, energy consumption is long-term and intensive, and heat generation is significant due to the computing requirement. The primary cooling source is the heating, ventilation, and air conditioning (HVAC) system, while the heat is generated from racks, power supplies, and power generators. Chen et al. demonstrated notable performance in a six-month case study on data center thermal modeling using an adaptive physically consistent neural network (A-PCNN) (Figure 4) [66]. The approach leveraged an NN with Softplus activation functions, replacing traditional preset and fixed coefficients to reduce trial-and-error costs and increase flexibility. Specifically, it reduced the mean absolute error by 17.3% for a 15 min forecast and by 79.2% over a 7-day period.

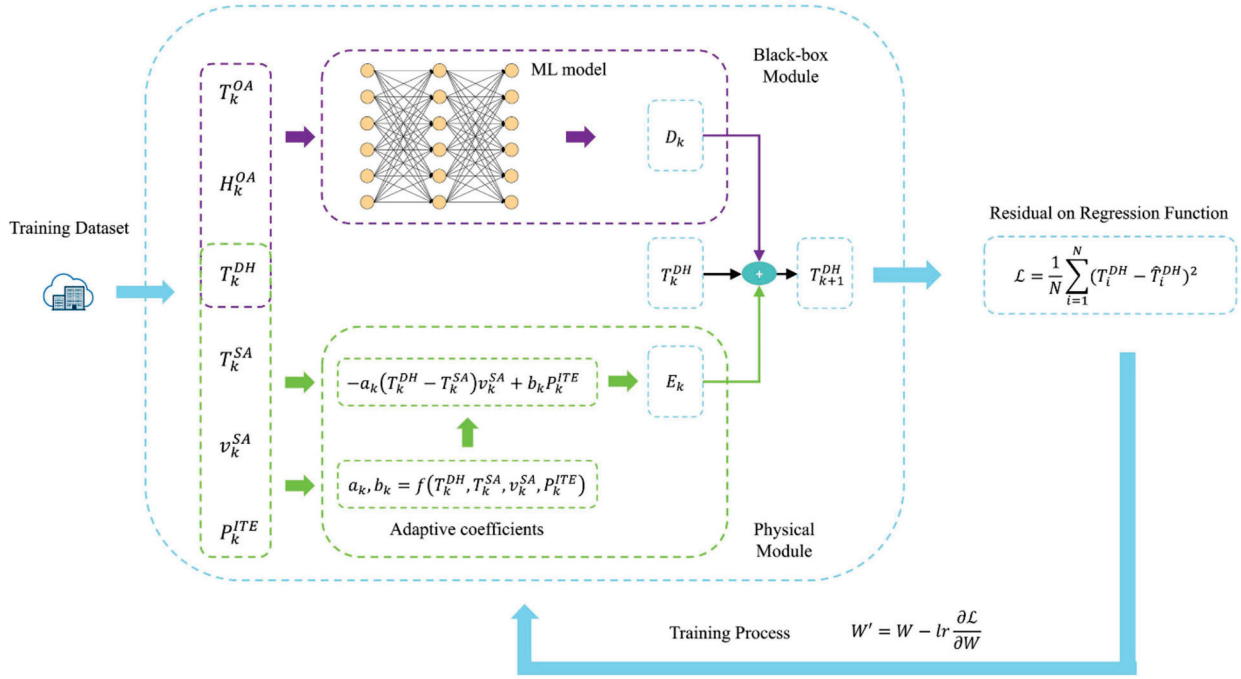


Figure 4. Adaptive physically consistent neural network framework. The figure has been republished with permission from [66].

Tanaka et al. developed an alternative modeling approach for system-level thermal systems by employing data compression through proper orthogonal decomposition combined with physics-informed machine learning (PIML). To evaluate the effective PIML technique, two distinct thermal mathematical frameworks were developed to explore how the training dataset size and model complexity influenced prediction precision. Furthermore, the study contrasted the prediction accuracy and computational training expenses between traditional data-driven machine learning and the PIML approach. The results demonstrated that this method accurately forecasted the temperature across both models in diverse heat input scenarios while also reducing the overall training cost by 26% to 81% relative to data-driven machine learning techniques, showcasing the advantages of the physics-informed strategies [87].

Zhang et al. investigated the application of PINNs in simulating fluid flow and heat transfer in manifold microchannel (MMC) heat sinks designed for cooling high-power Insulated Gate Bipolar Transistors (IGBTs), which are critical components in power electronics [88]. The study developed a PINN model with two sub-networks—one for flow dynamics and another for thermal behavior—each employing a DNN with a sine activation function to capture high-order derivatives and mitigate vanishing gradient

issues. Compared to traditional CFD simulations, PINNs show similar trends, such as increased pressure drops and decreased temperatures with higher inlet velocities, although discrepancies occur in regions with rapid flow changes and maximum temperature predictions. The mesh-free nature of PINNs and their ability to embed physical laws into the loss function enable the efficient simulation of complex geometries with fewer data than purely data-driven approaches. Additionally, the paper explores PINNs' potential in solving inverse problems, like estimating kinematic viscosity and thermal diffusivity, highlighting their versatility. Despite computational expense and sensitivity to geometry and hyperparameters, PINNs emerge as a promising alternative for thermal management in engineering applications.

All the system thermal management benefits mentioned above stem from the implementation of the PINN framework. With emerging technologies and fast turn-around requests in high-tech consumer electronics, data centers, and electric vehicles, PINNs have more potential when computational requests escalate.

The key methods for electronics thermal management at different scales, along with highlights for each scale, are summarized Table 2 below:

Table 2. Summary of applications of PINNs in electronics thermal management.

Scale	Key Methods	Highlights
Chip	DeeOHeat [71], ThermPINN [72]	Temperature-dependent properties, DeepONet, variables separation
Board	PINN with SiC design exploration [77], IPCNN [78], SRP-PINN [80]	Design space exploration, Multiphysics
System	A-PCNN [66], PIML [87], PINN with MMC and IGBTs [88]	CAD complexities, data size

4. Applications of PINNs in Battery Thermal Management

Battery systems involve multiple physical phenomena, such as electrochemical reactions, heat transfer, and mechanical stress, making simulation and analysis complex [89,90]. Uses for PINNs include state estimation [91], degradation [92,93], and aging estimation [93], thus proving the interest and potential of PINNs in battery systems. Finegan et al. outlined a perspective, discussed the challenges related to scarce data for analysis, and thus recommended physics-based learning for predicting battery failure more accurately and reliably [94]. Physics integration can include physics-based datasets, physics-informed training constraints, and physics-guided algorithm structures, which are a hybrid approaches that merge data-driven and physics-based methods.

Thermal management in batteries is critical, as thermal runaway (TR) is one of the primary safety concerns in batteries [95,96]. It occurs when an increase in temperature accelerates internal chemical reactions, generating more heat and potentially causing an uncontrollable chain reaction that may result in a fire or explosion. TR in a single LIB can rapidly propagate from the root cell to all adjacent cells, thus resulting in catastrophic accidents in large-scale battery packs and systems. Therefore, efficient battery thermal management systems at different scales are critical for ensuring the safe temperature of each cell and mitigating the TR phenomenon overall.

Battery thermal management simulation presents a range of challenges across different levels—from the cell to the pack to the system (Figure 5). At the cell level, the need to capture non-linear thermal behavior and electrochemical–thermal multiphysics interactions make modeling complex, especially for the accurate real-time prediction of thermal runaway events. Moving to the pack level, cell-to-cell variability, non-uniform temperature

distribution, and the impact of cell arrangement and manufacturing deviations complicate the prediction of thermal behavior and require detailed spatial resolution. At the system level, the challenges expand to modeling under wide operational conditions, ensuring effective heat dissipation coordination across components and maintaining a high level of model sophistication to balance accuracy and computational efficiency. The challenges are not confined to a single scale; instead, they can play a role across different scales. In addition, unlike electronics thermal management, the heat source includes power loss from electrical components, traces, or frequency-varying processes. The heat source in batteries is mainly introduced through multiple endothermic and exothermic electrochemical processes and dynamically changes according to the charging state, which makes capturing the exact source term and modeling battery thermal behavior even more challenging.

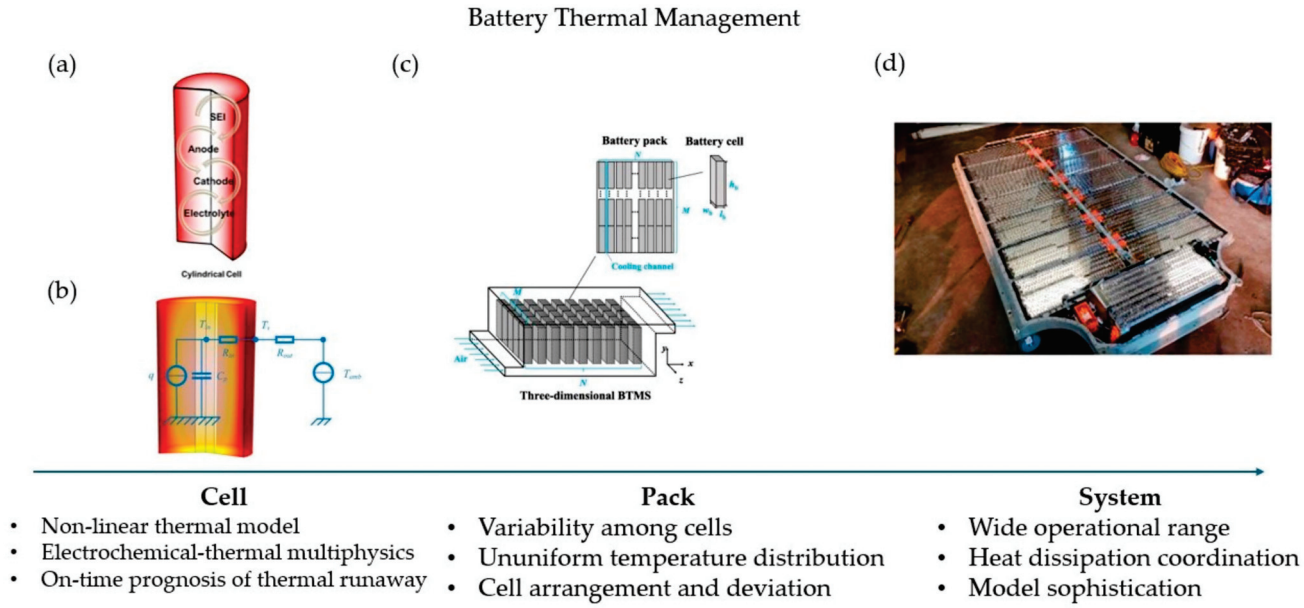


Figure 5. Battery thermal management challenges. Figure (**a**) was reproduced from [97] with permission under the terms of the Creative Commons. Figures (**b**–**d**) have been republished with permission from each indicated reference (ref. [98] for (**b**), ref. [99] for (**c**), ref. [100] for (**d**)).

The interconnections among the challenges related to battery thermal management simulation across the cell, pack, and system levels are deeply intertwined, as issues at one scale directly influence the others, creating a complex multiphysics problem. The complexity of precise cell thermal management propagates to the pack level, where cell-to-cell variability and non-uniform temperature distribution, which are driven by manufacturing deviations and cell arrangement, lead to uneven heat generation and dissipation, further complicating thermal predictions. At the system level, these challenges culminate in the need to model the entire battery pack under diverse operating conditions, where the cumulative effects of cell- and pack-level thermal inconsistencies must be addressed through coordinated heat dissipation strategies, all while balancing computational efficiency with model accuracy. Unlike electronics thermal management, where heat sources are primarily power losses from electrical components, BTMS must account for the dynamic electrochemical heat sources, making the accurate capture of source terms a cross-scale challenge that links the cell, pack, and system levels in a unified modeling framework. Therefore, this section will divide the PINN research into battery cells, battery packs, and battery systems, following the same escalating order as Section 3.

4.1. Battery Cell Thermal Management

For battery thermal profile prediction, the source term is generally simplified with a uniform heat generation source, or thermal resistance network [101,102]. Many researchers have simplified PDEs to create lumped parameter models. Research can, however, be conducted by considering the realistic 3D battery thermal model at different charging states.

Deng et al. applied a PINN to integrate the electric–thermal mechanism of the battery and data information through a weight adaptive function [103]. This was achieved by integrating transient thermal equations, with heat generation uniformly calculated from the relationship between current and voltage. The thermal equation is relatively easy to train due to the constant material property, and the battery heat generation rate is simply determined by joule heating without considering the various reactions and their exothermic rates in the formula. Wang et al. proposed a battery informed neural network (BINN) that selects features (voltage, current, experiment time, etc.) to obtain the cell surface temperature, open circuit voltage, and internal resistance, which reflect aging [98].

Kim et al. applied a multiphysics-informed neural network (MPINN) in thermal runaway analysis using a simple cylindrical lithium-ion cell, without any charge or discharge processes [97]. The MPINN addresses this by embedding physical laws—such as the energy balance equation and the Arrhenius law—into the neural network, enabling it to estimate time- and space-dependent temperature and concentration profiles more effectively than purely data-driven models like standard artificial neural networks (ANNs). The study achieves significant advancements by demonstrating that the MPINN outperforms ANNs in accuracy across various data availability scenarios (Figure 6). With fully labeled data, the MPINN reduces the mean absolute error (MAE) and the root mean squared error (RMSE) compared to ANNs (e.g., MAE of 0.46 vs. 1.17 for temperature). In semi-supervised settings with limited labeled data, the MPINN's errors remain low (MAE of 0.08 vs. 47.46 for ANN), and it can even predict TR without labeled data for positive electrode decomposition, leveraging its physics-informed framework. This capability positions the MPINN as a promising surrogate model for real-time TR prediction and battery safety optimization, which is validated through comparisons with high-fidelity COMSOL simulations. Looking forward, the paper suggests promising improvements, including expanding the MPINN to model additional TR mechanisms like negative electrode and solid electrolyte interphase (SEI) layer decomposition for a more comprehensive representation. Reducing the computational cost of training, currently a bottleneck, would enhance its practicality. Additionally, applying the MPINN to complex battery geometries or multi-cell systems could broaden its real-world applicability, thereby advancing battery management systems and safety designs. These enhancements could solidify the MPINN's role in improving LIB reliability and safety.

4.2. Battery Pack Thermal Management

Cho et al. explores the use of a PINN to tackle significant challenges associated with managing lithium-ion battery packs, particularly the accurate prediction of temperature distributions [104]. The PINN approach addresses these issues by integrating physical laws, such as energy balance equations, directly into the neural network's loss function. This hybrid method combines the strengths of physics-based and data-driven models, reducing the need for large datasets and extensive parameter tuning while maintaining accuracy, especially in scenarios with limited data or unknown initial conditions. The paper demonstrates notable success in applying the PINN method to improve temperature prediction accuracy within lithium-ion battery packs (Figure 7). By embedding physical constraints into the neural network, the PINN achieves a root mean square error (RMSE) of 0.57 °C for the Direct Current Fast Charge (DCFC) test profile and 0.52 °C for the Grade Load (GL)

100 test profile. These results mark a significant improvement over traditional methods, offering a more efficient and reliable way to monitor battery temperatures. Enhanced accuracy is particularly valuable for preventing thermal runaway and extending battery life, which are key concerns in battery management systems. The study also refines the PINN's performance by incorporating additional inputs, such as chamber temperature, and optimizing the neural network architecture. This hybrid approach not only outperforms conventional models in terms of precision but also reduces computational overhead, making it a practical solution for real-time temperature monitoring in battery packs.

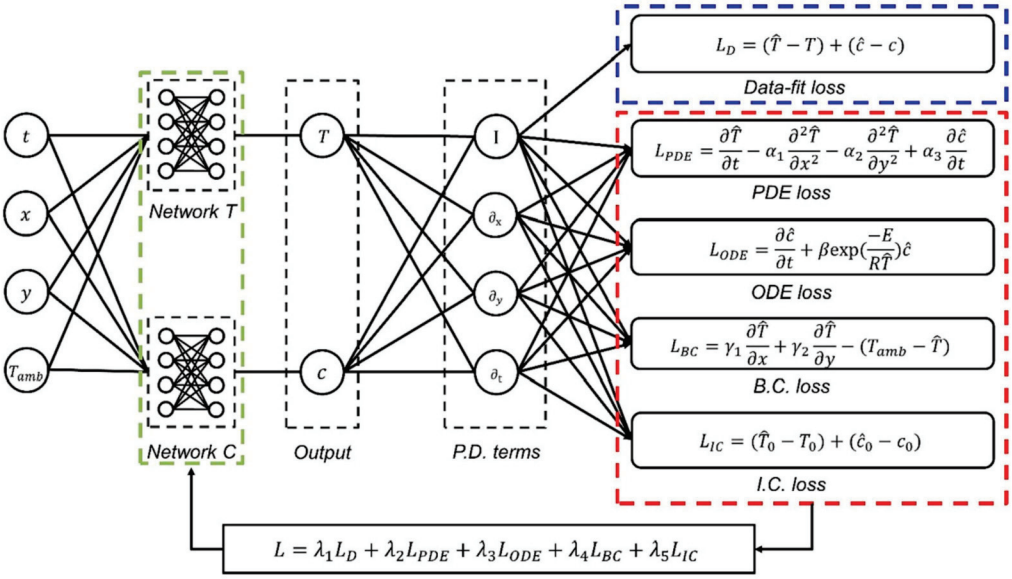


Figure 6. Multiphysics-informed neural network framework. The figure has been republished from [97] with permission under the terms of the Creative Commons.

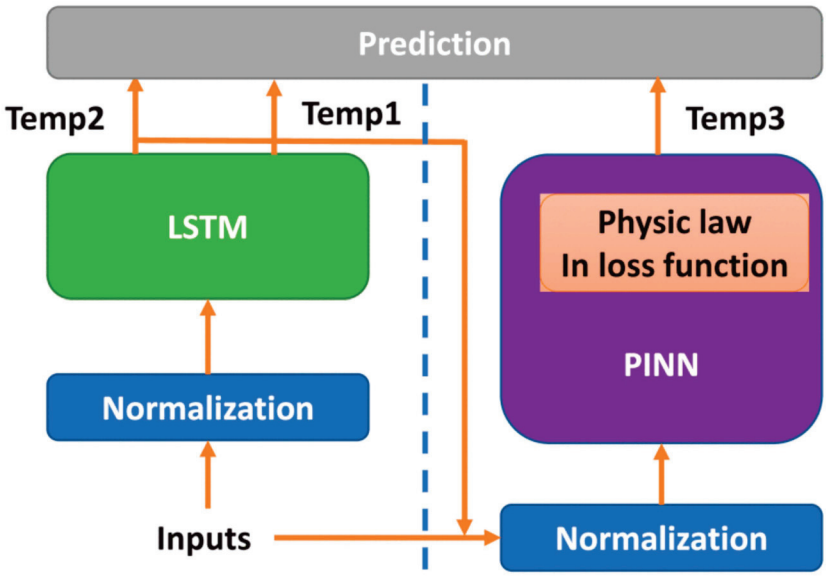


Figure 7. LSTM-PINN hybrid framework. The figure has been republished from [104] with permission under the terms of the Creative Commons.

Looking ahead, the paper suggests several avenues to further enhance the PINN approach for battery pack applications. One promising improvement is the integration of more comprehensive physical models into the neural network, such as those accounting for varying thermal properties or complex heat transfer mechanisms within the battery

pack. This could lead to even greater prediction accuracy by capturing a broader range of physical phenomena. Another potential advancement lies in optimizing the neural network architecture, possibly by exploring hybrid models or alternative network types to better handle the nonlinear dynamics of battery systems. Additionally, expanding the dataset to encompass a wider variety of operating conditions and battery types could improve the model's generalizability, making it adaptable to diverse real-world scenarios and battery configurations. These enhancements could solidify PINN's role as a cornerstone in battery management systems, driving further improvements in safety, efficiency, and performance for lithium-ion battery applications.

4.3. Battery System Thermal Management

When battery packs are embedded in large systems, such as electronics and electric vehicles, simulation becomes more challenging.

Shen et al. applied PINNs to address the challenge of accurately estimating temperature distributions in large-format lithium-ion blade batteries, which is a critical aspect of thermal management in electric vehicle battery packs [105]. Effective temperature control is vital because it directly impacts battery performance, safety, and lifespan, with excessive heat potentially leading to thermal runaway. Traditional physics-based models, while detailed, are computationally demanding and require extensive manual parameter tuning, making them impractical for real-time use. Conversely, data-driven models depend on large datasets, which may not always be available, especially under diverse operating conditions. The paper introduces a PINN model that integrates a simplified multi-node thermal model into an LSTM neural network, combining physical laws with machine learning (Figure 8). This hybrid approach reduces the need for extensive data and calibration by embedding heat transfer equations into the network's loss function, enabling real-time temperature predictions with improved accuracy and interpretability, particularly for large-format batteries with non-uniform heat generation.

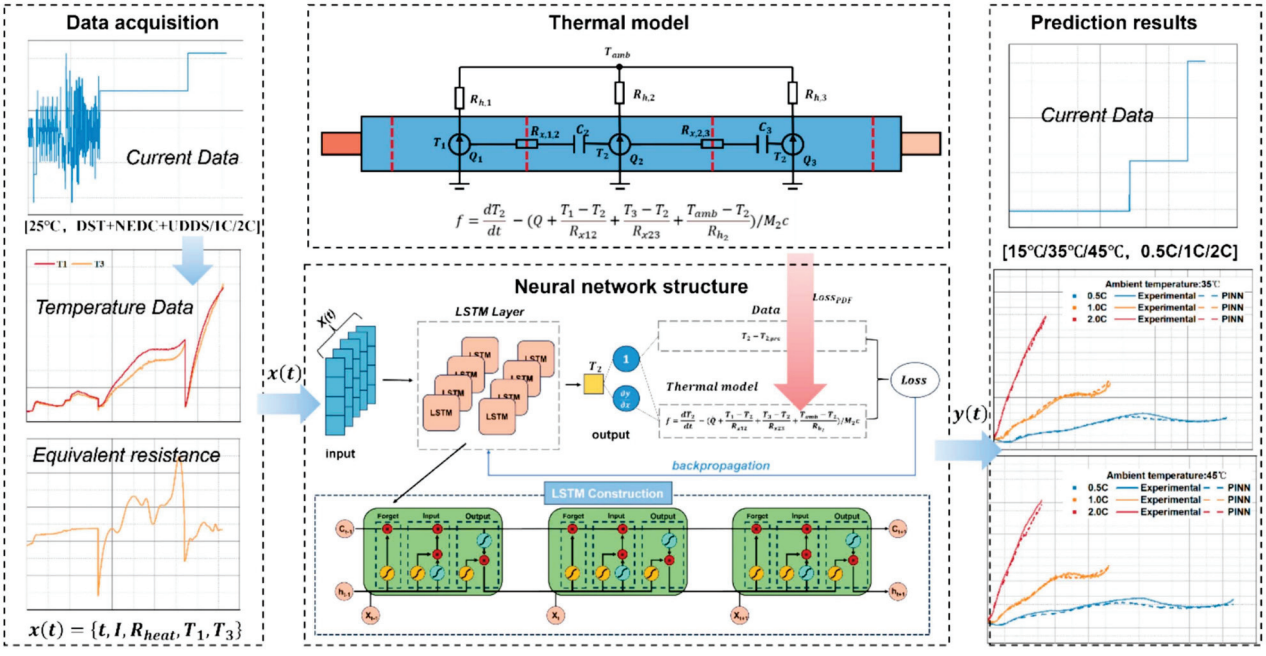


Figure 8. PINN-based real-time battery temperature estimation framework. The figure has been republished with permission from [105].

In this study, experimental data were gathered using a test bench equipped with a Nebula Charging and Discharging testing system, a GDBELL high- and low-temperature

test chamber, three K-type thermocouples, and an upper computer. A series of battery thermal characteristic experiments, including discharging, cooling, charging, resting, and discharging, were performed under varying current rate and environmental conditions to capture the real-time temperature profiles of the blade battery. To develop the framework, a 1D thermal model was established along the length direction of the large-format blade battery, as heat transfer in the width and thickness directions was negligible. The model incorporates thermal resistance between nodes, heat generation, Joule heating, and convective heat transfer, and is simplified as follows:

$$f = \frac{dT_2}{dt} - \left(Q + \frac{T_1 - T_2}{R_{x12}} + \frac{T_3 - T_2}{R_{x23}} + \frac{T_{amb} - T_2}{R_{h2}} \right) / M_2 c \tag{10}$$

Here, Tx denotes the temperature at different nodes, while R_h and R_x represent the thermal resistance between nodes, including convective heat transfer resistance (R_h) and equivalent internal resistance (R_x). The R_h values were determined using experimentally measured data via the recursive least squares (RLS) method, whereas R_x values were derived from open circuit voltage (OCV) testing combined with the Joule heating formula. The transient temperature data, equivalent resistance curves, experimental data, and thermal model were integrated into the PINN model, which includes an LSTM layer to handle long-sequence time-series data, thereby preserving and updating critical temperature information. The model is trained by minimizing the loss function, ultimately outputting the transient central temperature of the battery as a temperature curve.

The paper demonstrates significant achievements through the development and testing of the PINN model for battery temperature estimation. By incorporating a one-dimensional thermal model with three nodes—accounting for heat generation, transfer, and dissipation—the PINN leverages LSTM to capture the time-series nature of temperature changes. Tested under various charging conditions (0.5 C, 1.0 C, and 2.0 C), the model outperforms traditional approaches like backpropagation neural networks (BP-NN) and standalone LSTMs. For instance, at a 2.0 C charging rate, it achieves an R^2 of 0.9863, a mean absolute error (MAE) of 0.2875 °C, and a root mean square error (RMSE) of 0.3306 °C, showcasing high predictive accuracy. A notable advantage is its ability to learn parameters, such as equivalent internal resistance, automatically via neural network training, thereby eliminating manual calibration. These results, validated through a realistic experimental setup with a battery test bench, highlight the PINN's effectiveness in providing precise, real-time temperature estimates and enhancing battery management systems' ability to prevent thermal issues. The paper outlines several promising directions for enhancing the PINN model's capabilities. One key improvement is extending temperature estimation to the battery module level, predicting the overall temperature field across multiple cells in a pack. This would offer a more comprehensive thermal profile for practical applications but would require addressing increased complexity in modeling inter-cell interactions. Another avenue is refining the thermal model by incorporating additional physical phenomena, such as detailed heat transfer mechanisms or variable thermal properties, to further improve accuracy. Expanding the dataset to include a broader range of operating conditions (e.g., different charge rates, ambient temperatures, and battery aging states) could enhance the model's generalizability and robustness. Additionally, optimizing the neural network architecture—potentially by exploring advanced hybrid designs—could improve its ability to handle complex, nonlinear thermal dynamics. These advancements would strengthen the PINN's role in real-time thermal management, supporting safer and more efficient battery pack operations in electric vehicles.

The applications of PINNs in battery thermal management discussed in this section are summarized in Table 3. However, this area is not well understood due to the complexity

of electrochemical–thermal multiphysics, which makes it difficult to identify the thermal source and corresponding PDEs, especially with different SOC. Data-driven experimental identification is the current approach being used, which deviates from the intention of PINNs because they are not data-driven.

Table 3. Summary of applications of PINNs in battery thermal management.

Scale	Key Methods	Highlights
Cell	PINN with electric–thermal mechanism [103], BINN [98], MPINN [97]	Electrochemical–thermal multiphysics, real-time TR prediction
Pack	LSTM-PINN [104]	Data size, temperature uniformity, cell arrangement
System	PINN-LSTM [105]	Large format, wide operational range, 1D simplification

5. Conclusions

In this review paper, we first discussed the working principles of PINNs and their variants, as well as their general applications in the physics simulation world in Section 2. Based on this understanding, we then discussed the challenges associated with both electronics and battery thermal management at three scale levels and how PINNs can tackle these challenges accordingly. Subseqently, we reviewed study that has been conducted at the chip level, board level and system level of electronics thermal management in Section 3. PINNs are being used to solve different analysis challenges at different scales. In Section 4, we then reviewed research conducted on battery thermal management with the same scale categorizing standards, from the battery level to the pack level to large-format systems. As the two most heat-intensive components in the high-tech, automobile, and industrial fields, the thermal management of electronics and batteries should be considered together. Gaps clearly exist in the connections between these two components. In the current research, PINNs have solved the challenges of data scarcity and prediction efficiency, and the critical idea is to figure out the correct PDEs to use and adjust the training algorithms to various operating conditions. Despite the promising advantages that PINNs offer over traditional machine learning, as well as their flexibility to combine with different techniques to enhance performance, the application of PINNs is not without challenges. They can be computationally intensive, sensitive to the correct formulation of governing equations and neural network architecture, and may face convergence issues or require extensive hyperparameter tuning when dealing with stiff or highly nonlinear systems, making their deployment in real-time or large-scale industrial settings more challenging. Therefore, this review points out both the pros and cons of applying PINNs, which will benefit from inspiration in using advanced machine learning approaches to address system-level electronics and battery system thermal management.

Author Contributions: Z.D.: Conceptualization, investigation, formal analysis, methodology, project administration, resources, supervision, validation, visualization, Writing—original draft, Writing—review & editing. R.L.: Investigation, literature review, formal analysis, methodology, visualization, Writing—original draft, validation, resources, Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: Author Zichen Du is employed by the Ansys, Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Pistoia, G. *Battery Operated Devices and Systems*; Elsevier: Amsterdam, The Netherlands, 2009; ISBN 978-0-444-53214-5.
2. Wang, Q.; Ping, P.; Zhao, X.; Chu, G.; Sun, J.; Chen, C. Thermal Runaway Caused Fire and Explosion of Lithium Ion Battery. *J. Power Sources* **2012**, *208*, 210–224. [CrossRef]
3. De Bock, H.P.; Huitink, D.; Shamberger, P.; Lundh, J.S.; Choi, S.; Niedbalski, N.; Boteler, L. A System to Package Perspective on Transient Thermal Management of Electronics. *J. Electron. Packag.* **2020**, *142*, 041111. [CrossRef]
4. Falcone, M.; Palka Bayard De Volo, E.; Hellany, A.; Rossi, C.; Pulvirenti, B. Lithium-Ion Battery Thermal Management Systems: A Survey and New CFD Results. *Batteries* **2021**, *7*, 86. [CrossRef]
5. Li, X.; He, F.; Ma, L. Thermal Management of Cylindrical Batteries Investigated Using Wind Tunnel Testing and Computational Fluid Dynamics Simulation. *J. Power Sources* **2013**, *238*, 395–402. [CrossRef]
6. Kim, G.-H.; Pesaran, A. Battery Thermal Management Design Modeling. *World Electr. Veh. J.* **2007**, *1*, 126–133. [CrossRef]
7. Chen, W.; Hou, S.; Shi, J.; Han, P.; Liu, B.; Wu, B.; Lin, X. Numerical Analysis of Novel Air-Based Li-Ion Battery Thermal Management. *Batteries* **2022**, *8*, 128. [CrossRef]
8. Eymard, R.; Gallouët, T.; Herbin, R. Finite Volume Methods. In *Handbook of Numerical Analysis*; Elsevier: Amsterdam, The Netherlands, 2000; Volume 7; pp. 713–1018; ISBN 978-0-444-50350-3.
9. Birbarah, P.; Gebrael, T.; Foulkes, T.; Stillwell, A.; Moore, A.; Pilawa-Podgurski, R.; Miljkovic, N. Water Immersion Cooling of High Power Density Electronics. *Int. J. Heat Mass Transf.* **2020**, *147*, 118918. [CrossRef]
10. American Society of Mechanical Engineers (Ed.) *Proceedings of the ASME International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems-2018: Heterogeneous Integration: Microsystems with Diverse Functionality: Servers of the Future, IoT, and Edge to Cloud: Structural and Physical Health Monitoring: Power Electronics, Energy Conversion, and Storage: Autonomous, Hybrid, and Electric Vehicles: Presented at ASME 2018 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, 27–30 August 2018, San Francisco, CA, USA*; The American Society of Mechanical Engineers: New York, NY, USA, 2019; ISBN 978-0-7918-5192-0.
11. Shuai, S.; Du, Z.; Ma, B.; Shan, L.; Dogruoz, B.; Agonafer, D. Numerical Investigation of Shape Effect on Microdroplet Evaporation. In Proceedings of the ASME 2018 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, San Francisco, CA, USA, 27–30 August 2018; American Society of Mechanical Engineers: San Francisco, CA, USA, 2018; p. V001T04A010.
12. Al Miaari, A.; Ali, H.M. Batteries Temperature Prediction and Thermal Management Using Machine Learning: An Overview. *Energy Rep.* **2023**, *10*, 2277–2305. [CrossRef]
13. Abhijith, M.S.; Soman, K.P. Machine Learning Methods for Modeling Nanofluid Flows: A Comprehensive Review with Emphasis on Compact Heat Transfer Devices for Electronic Device Cooling. *J. Therm. Anal. Calorim.* **2024**, *149*, 5843–5869. [CrossRef]
14. Floridi, L.; Chiriatti, M. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds Mach.* **2020**, *30*, 681–694. [CrossRef]
15. OpenAI; Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F.L.; Almeida, D.; Altenschmidt, J.; Altman, S.; et al. GPT-4 Technical Report. *arXiv* **2023**, arXiv:2303.08774.
16. Decencière, E.; Cazuguel, G.; Zhang, X.; Thibault, G.; Klein, J.-C.; Meyer, F.; Marcotegui, B.; Quellec, G.; Lamard, M.; Danno, R.; et al. TeleOphta: Machine Learning and Image Processing Methods for Teleophthalmology. *IRBM* **2013**, *34*, 196–203. [CrossRef]
17. Munawar, H.S.; Hammad, A.W.A.; Waller, S.T. A Review on Flood Management Technologies Related to Image Processing and Machine Learning. *Autom. Constr.* **2021**, *132*, 103916. [CrossRef]
18. Lu, R. Complex Wavelet Mutual Information Loss: A Multi-Scale Loss Function for Semantic Segmentation. *arXiv* **2025**, arXiv:250200563.
19. Lu, R. Steerable Pyramid Weighted Loss: Multi-Scale Adaptive Weighting for Semantic Segmentation. *arXiv* **2025**, arXiv:250306604.
20. Summerville, A.; Snodgrass, S.; Guzdial, M.; Holmgard, C.; Hoover, A.K.; Isaksen, A.; Nealen, A.; Togelius, J. Procedural Content Generation via Machine Learning (PCGML). *IEEE Trans. Games* **2018**, *10*, 257–270. [CrossRef]
21. Justesen, N.; Bontrager, P.; Togelius, J.; Risi, S. Deep Learning for Video Game Playing. *IEEE Trans. Games* **2020**, *12*, 1–20. [CrossRef]
22. Xu, M.; Maddage, N.C.; Xu, C.; Kankanhalli, M.; Tian, Q. Creating Audio Keywords for Event Detection in Soccer Video. In Proceedings of the 2003 International Conference on Multimedia and Expo. ICME 03. Proceedings (Cat. No.03TH8698), Baltimore, MD, USA, 6–9 July 2003; IEEE: Piscataway, NJ, USA; p. II–281.
23. Huang, B.; Wang, J. Applications of Physics-Informed Neural Networks in Power Systems—A Review. *IEEE Trans. Power Syst.* **2023**, *38*, 572–588. [CrossRef]

24. Li, A.; Yuen, A.C.Y.; Wang, W.; Chen, T.B.Y.; Lai, C.S.; Yang, W.; Wu, W.; Chan, Q.N.; Kook, S.; Yeoh, G.H. Integration of Computational Fluid Dynamics and Artificial Neural Network for Optimization Design of Battery Thermal Management System. *Batteries* **2022**, *8*, 69. [CrossRef]
25. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-Informed Machine Learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [CrossRef]
26. Li, J.; Lopez, S.A. A Look Inside the Black Box of Machine Learning Photodynamics Simulations. *Acc. Chem. Res.* **2022**, *55*, 1972–1984. [CrossRef] [PubMed]
27. Liu, H.-H.; Zhang, J.; Liang, F.; Temizel, C.; Basri, M.A.; Mesdour, R. Incorporation of Physics into Machine Learning for Production Prediction from Unconventional Reservoirs: A Brief Review of the Gray-Box Approach. *SPE Reserv. Eval. Eng.* **2021**, *24*, 847–858. [CrossRef]
28. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics Informed Deep Learning (Part I): Data-Driven Solutions of Nonlinear Partial Differential Equations. *arXiv* **2017**, arXiv:1711.10561.
29. Cai, S.; Mao, Z.; Wang, Z.; Yin, M.; Karniadakis, G.E. Physics-Informed Neural Networks (PINNs) for Fluid Mechanics: A Review. *Acta Mech. Sin.* **2021**, *37*, 1727–1738. [CrossRef]
30. Wang, H.; Cao, Y.; Huang, Z.; Liu, Y.; Hu, P.; Luo, X.; Song, Z.; Zhao, W.; Liu, J.; Sun, J.; et al. Recent Advances on Machine Learning for Computational Fluid Dynamics: A Survey. *arXiv* **2024**, arXiv:2408.12171.
31. Arzani, A.; Wang, J.-X.; D'Souza, R.M. Uncovering Near-Wall Blood Flow from Sparse Data with Physics-Informed Neural Networks. *Phys. Fluids* **2021**, *33*, 071905. [CrossRef]
32. Zhou, W.; Miwa, S.; Okamoto, K. Advancing Fluid Dynamics Simulations: A Comprehensive Approach to Optimizing Physics-Informed Neural Networks. *Phys. Fluids* **2024**, *36*, 013615. [CrossRef]
33. Gokhale, G.; Claessens, B.; Develder, C. Physics Informed Neural Networks for Control Oriented Thermal Modeling of Buildings. *Appl. Energy* **2022**, *314*, 118852. [CrossRef]
34. Xu, J.; Wei, H.; Bao, H. Physics-Informed Neural Networks for Studying Heat Transfer in Porous Media. *Int. J. Heat Mass Transf.* **2023**, *217*, 124671. [CrossRef]
35. Li, Z.; Luo, H.; Jiang, Y.; Liu, H.; Xu, L.; Cao, K.; Wu, H.; Gao, P.; Liu, H. Comprehensive Review and Future Prospects on Chip-Scale Thermal Management: Core of Data Center's Thermal Management. *Appl. Therm. Eng.* **2024**, *251*, 123612. [CrossRef]
36. Xia, Y.; Meng, Y. Physics-Informed Neural Network (PINN) for Solving Frictional Contact Temperature and Inversely Evaluating Relevant Input Parameters. *Lubricants* **2024**, *12*, 62. [CrossRef]
37. Cai, S.; Wang, Z.; Wang, S.; Perdikaris, P.; Karniadakis, G.E. Physics-Informed Neural Networks for Heat Transfer Problems. *J. Heat Transf.* **2021**, *143*, 060801. [CrossRef]
38. Daw, A.; Bu, J.; Wang, S.; Perdikaris, P.; Karpatne, A. Mitigating Propagation Failures in Physics-Informed Neural Networks Using Retain-Resample-Release (R3) Sampling. *arXiv* **2022**, arXiv:2207.02338.
39. Nabian, M.A.; Gladstone, R.J.; Meidani, H. Efficient Training of Physics-informed Neural Networks via Importance Sampling. *Comput. Civ. Infrastruct. Eng.* **2021**, *36*, 962–977. [CrossRef]
40. Farea, A.; Yli-Harja, O.; Emmert-Streib, F. Understanding Physics-Informed Neural Networks: Techniques, Applications, Trends, and Challenges. *AI* **2024**, *5*, 1534–1557. [CrossRef]
41. Cuomo, S.; Di Cola, V.S.; Giampaolo, F.; Rozza, G.; Raissi, M.; Piccialli, F. Scientific Machine Learning Through Physics–Informed Neural Networks: Where We Are and What's Next. *J. Sci. Comput.* **2022**, *92*, 88. [CrossRef]
42. Wang, S.; Teng, Y.; Perdikaris, P. Understanding and Mitigating Gradient Flow Pathologies in Physics-Informed Neural Networks. *SIAM J. Sci. Comput.* **2021**, *43*, A3055–A3081. [CrossRef]
43. Yao, J.; Su, C.; Hao, Z.; Liu, S.; Su, H.; Zhu, J. Multiadam: Parameter-Wise Scale-Invariant Optimizer for Multiscale Training of Physics-Informed Neural Networks. In Proceedings of the International Conference on Machine Learning, PMLR, Honolulu, HI, USA, 23–29 July 2023; pp. 39702–39721.
44. Wu, C.; Zhu, M.; Tan, Q.; Kartha, Y.; Lu, L. A Comprehensive Study of Non-Adaptive and Residual-Based Adaptive Sampling for Physics-Informed Neural Networks. *Comput. Methods Appl. Mech. Eng.* **2023**, *403*, 115671. [CrossRef]
45. Tang, K.; Wan, X.; Yang, C. DAS-PINNs: A Deep Adaptive Sampling Method for Solving High-Dimensional Partial Differential Equations. *J. Comput. Phys.* **2023**, *476*, 111868. [CrossRef]
46. Yu, T.; Yong, H.; Liu, L.; Wang, H.; Chen, H. MCMC-PINNs: A Modified Markov Chain Monte-Carlo Method for Sampling Collocation Points of PINNs Adaptively. *Authorea Preprint* **2023**. [CrossRef]
47. Yu, B.; E, W. The Deep Ritz Method: A Deep Learning-Based Numerical Algorithm for Solving Variational Problems. *Commun. Math. Stat.* **2018**, *6*, 1–12.
48. Kharazmi, E.; Zhang, Z.; Karniadakis, G.E. Variational Physics-Informed Neural Networks for Solving Partial Differential Equations. *arXiv* **2019**, arXiv:191200873.
49. Kharazmi, E.; Zhang, Z.; Karniadakis, G.E. Hp-VPINNs: Variational Physics-Informed Neural Networks with Domain Decomposition. *Comput. Methods Appl. Mech. Eng.* **2021**, *374*, 113547. [CrossRef]

50. Khodayi-Mehr, R.; Zavlanos, M. VarNet: Variational Neural Networks for the Solution of Partial Differential Equations. In Proceedings of the Learning for Dynamics and Control, PMLR, Berkeley, CA, USA, 10–11 June 2020; pp. 298–307.
51. Jagtap, A.D.; Kharazmi, E.; Karniadakis, G.E. Conservative Physics-Informed Neural Networks on Discrete Domains for Conservation Laws: Applications to Forward and Inverse Problems. *Comput. Methods Appl. Mech. Eng.* **2020**, *365*, 113028. [CrossRef]
52. Jagtap, A.D.; Karniadakis, G.E. Extended Physics-Informed Neural Networks (XPINNs): A Generalized Space-Time Domain Decomposition Based Deep Learning Framework for Nonlinear Partial Differential Equations. *Commun. Comput. Phys.* **2020**, *28*, 2002–2041. [CrossRef]
53. Jeon, J.; Lee, J.; Vinuesa, R.; Kim, S.J. Residual-Based Physics-Informed Transfer Learning: A Hybrid Method for Accelerating Long-Term CFD Simulations via Deep Learning. *Int. J. Heat Mass Transf.* **2024**, *220*, 124900. [CrossRef]
54. Incropera, F.P.; DeWitt, D.P.; Bergman, T.L.; Lavine, A.S. (Eds.) *Fundamentals of Heat and Mass Transfer*, 6th ed.; Wiley: Hoboken, NJ, USA, 2007; ISBN 978-0-471-45728-2.
55. Liaw, S.P.; Yeh, R.H. Fins with Temperature Dependent Surface Heat Flux—I. Single Heat Transfer Mode. *Int. J. Heat Mass Transf.* **1994**, *37*, 1509–1515. [CrossRef]
56. Das, S.K.; Putra, N.; Thiesen, P.; Roetzel, W. Temperature Dependence of Thermal Conductivity Enhancement for Nanofluids. *J. Heat Transf.* **2003**, *125*, 567–574. [CrossRef]
57. Bejan, A. *Convection Heat Transfer*, 4th ed.; Wiley: Hoboken, NJ, USA, 2013; ISBN 978-0-470-90037-6.
58. Kaviany, M. *Heat Transfer Physics*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2014; ISBN 978-1-107-04178-3.
59. Bernardo, C.; Johann, W.K. Energietechnische Gesellschaft, Ed.; *Proceedings/CIPS 2012, 7th International Conference on Integrated Power Electronics Systems: 6–8 March 2012, Nuremberg, Germany*; Incl. CD-ROM; ETG-Fachbericht; VDE-Verl: Berlin, Germany, 2012; ISBN 978-3-8007-3414-6.
60. Ohadi, M.M.; Dessiatoun, S.V.; Choo, K.; Pecht, M.; Lawler, J.V. A Comparison Analysis of Air, Liquid, and Two-Phase Cooling of Data Centers. In Proceedings of the 2012 28th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), San Jose, CA, USA, 18–22 March 2012; IEEE: Piscataway, NJ, USA; pp. 58–63.
61. Gong, Y.; Zhou, F.; Ma, G.; Liu, S. Advancements on Mechanically Driven Two-Phase Cooling Loop Systems for Data Center Free Cooling. *Int. J. Refrig.* **2022**, *138*, 84–96. [CrossRef]
62. Yuan, X.; Zhou, X.; Pan, Y.; Kosonen, R.; Cai, H.; Gao, Y.; Wang, Y. Phase Change Cooling in Data Centers: A Review. *Energy Build.* **2021**, *236*, 110764. [CrossRef]
63. Abro, G.E.M.; Zulkifli, S.A.B.M.; Kumar, K.; El Ouanjli, N.; Asirvadam, V.S.; Mossa, M.A. Comprehensive Review of Recent Advancements in Battery Technology, Propulsion, Power Interfaces, and Vehicle Network Systems for Intelligent Autonomous and Connected Electric Vehicles. *Energies* **2023**, *16*, 2925. [CrossRef]
64. Zhang, Y.; Udrea, F.; Wang, H. Multidimensional Device Architectures for Efficient Power Electronics. *Nat. Electron.* **2022**, *5*, 723–734. [CrossRef]
65. Shan, L.; Bu, C.; Su, Y.; Wu, J.; Wang, Y.; Shen, L.; Xie, J. Towards Feasible Thermal Management Design of Electronic Control Module for Variable Frequency Air Conditioner Function in Extremely High Ambient Temperatures. *Electronics* **2025**, *14*, 1595. [CrossRef]
66. Chen, D.; Chui, C.-K.; Lee, P.S. Adaptive Physically Consistent Neural Networks for Data Center Thermal Dynamics Modeling. *Appl. Energy* **2025**, *377*, 124637. [CrossRef]
67. Ding, B.; Zhang, Z.-H.; Gong, L.; Xu, M.-H.; Huang, Z.-Q. A Novel Thermal Management Scheme for 3D-IC Chips with Multi-Cores and High Power Density. *Appl. Therm. Eng.* **2020**, *168*, 114832. [CrossRef]
68. Sadiqbatcha, S.I.; Zhang, J.; Amrouch, H.; Tan, S.X.-D. Real-Time Full-Chip Thermal Tracking: A Post-Silicon, Machine Learning Perspective. *IEEE Trans. Comput.* **2021**, *71*, 1411–1424. [CrossRef]
69. Chen, L.; Jin, W.; Tan, S.X.-D. Fast Thermal Analysis for Chiplet Design Based on Graph Convolution Networks. In Proceedings of the 2022 27th Asia and South Pacific Design Automation Conference (ASP-DAC), Taipei, Taiwan, 17 January 2022; IEEE: Piscataway, NJ, USA; pp. 485–492.
70. Liu, Z.; Li, Y.; Hu, J.; Yu, X.; Shiau, S.; Ai, X.; Zeng, Z.; Zhang, Z. DeepOHeat: Operator Learning-Based Ultra-Fast Thermal Simulation in 3D-IC Design. In Proceedings of the 2023 60th ACM/IEEE Design Automation Conference (DAC), San Francisco, CA, USA, 9 July 2023; IEEE: Piscataway, NJ, USA; pp. 1–6.
71. Jin, P.; Meng, S.; Lu, L. MIONet: Learning Multiple-Input Operators via Tensor Product. *SIAM J. Sci. Comput.* **2022**, *44*, A3490–A3514. [CrossRef]
72. Chen, L.; Lu, J.; Jin, W.; Tan, S.X.-D. Fast Full-Chip Parametric Thermal Analysis Based on Enhanced Physics Enforced Neural Networks. In Proceedings of the 2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD), San Francisco, CA, USA, 28 October 2023; IEEE: Piscataway, NJ, USA; pp. 1–8.
73. Garimella, S.V.; Persoons, T.; Weibel, J.A.; Gektin, V. Electronics Thermal Management in Information and Communications Technologies: Challenges and Future Directions. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2017**, *7*, 1191–1205. [CrossRef]

74. Asgari, S.; Hu, X.; Tsuk, M.; Kaushik, S. Application of POD plus LTI ROM to Battery Thermal Modeling: SISO Case. *SAE Int. J. Commer. Veh.* **2014**, *7*, 278–285. [CrossRef]
75. Hu, X.; Asgari, S.; Lin, S.; Stanton, S.; Lian, W. A Linear Parameter-Varying Model for HEV/EV Battery Thermal Modeling. In Proceedings of the 2012 IEEE Energy Conversion Congress and Exposition (ECCE), Raleigh, NC, USA, 15–20 September 2012; IEEE: Piscataway, NJ, USA; pp. 1643–1649.
76. Hu, X.; Asgari, S.; Yavuz, I.; Stanton, S.; Hsu, C.-C.; Shi, Z.; Wang, B.; Chu, H.-K. A Transient Reduced Order Model for Battery Thermal Management Based on Singular Value Decomposition. In Proceedings of the 2014 IEEE Energy Conversion Congress and Exposition (ECCE), Pittsburgh, PA, USA, 14–18 September 2014; IEEE: Piscataway, NJ, USA; pp. 3971–3976.
77. Yang, Y.; Wang, Z.; Liao, Y.; Kong, W.; Shi, X.; Hu, R.; Yao, Y. A Parameterized Thermal Simulation Method Based on Physics-Informed Neural Networks for Fast Power Module Thermal Design. *IEEE Trans. Power Electron.* **2025**, *40*, 9200–9210. [CrossRef]
78. Hamid Elsheikh, M.; Shnawah, D.A.; Sabri, M.F.M.; Said, S.B.M.; Haji Hassan, M.; Ali Bashir, M.B.; Mohamad, M. A Review on Thermoelectric Renewable Energy: Principle Parameters That Affect Their Performance. *Renew. Sustain. Energy Rev.* **2014**, *30*, 337–355. [CrossRef]
79. Chen, L.; Jin, W.; Zhang, J.; Tan, S.X.-D. Thermoelectric Cooler Modeling and Optimization via Surrogate Modeling Using Implicit Physics-Constrained Neural Networks. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **2023**, *42*, 4090–4101. [CrossRef]
80. Farrag, A.; Kataoka, J.; Yoon, S.W.; Won, D.; Jin, Y. SRP-PINN: A Physics-Informed Neural Network Model for Simulating Thermal Profile of Soldering Reflow Process. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2024**, *14*, 1098–1105. [CrossRef]
81. Liu, H.; Wen, M.; Yang, H.; Yue, Z.; Yao, M. A Review of Thermal Management System and Control Strategy for Automotive Engines. *J. Energy Eng.* **2021**, *147*, 03121001. [CrossRef]
82. Du, D.; Darkwa, J.; Kokogiannakis, G. Thermal Management Systems for Photovoltaics (PV) Installations: A Critical Review. *Sol. Energy* **2013**, *97*, 238–254. [CrossRef]
83. Nadjahi, C.; Louahlia, H.; Lemasson, S. A Review of Thermal Management and Innovative Cooling Strategies for Data Center. *Sustain. Comput. Inform. Syst.* **2018**, *19*, 14–28. [CrossRef]
84. Zhang, K.; Zhang, Y.; Liu, J.; Niu, X. Recent Advancements on Thermal Management and Evaluation for Data Centers. *Appl. Therm. Eng.* **2018**, *142*, 215–231. [CrossRef]
85. Pogorelskiy, S.; Kocsis, I. BIM and Computational Fluid Dynamics Analysis for Thermal Management Improvement in Data Centres. *Buildings* **2023**, *13*, 2636. [CrossRef]
86. Schmidt, R.R.; Cruz, E.E.; Iyengar, M. Challenges of Data Center Thermal Management. *IBM J. Res. Dev.* **2005**, *49*, 709–723. [CrossRef]
87. Tanaka, H.; Nagai, H. Thermal Surrogate Model for Spacecraft Systems Using Physics-Informed Machine Learning with POD Data Reduction. *Int. J. Heat Mass Transf.* **2023**, *213*, 124336. [CrossRef]
88. Zhang, X.; Tu, C.; Yan, Y. Physics-Informed Neural Network Simulation of Conjugate Heat Transfer in Manifold Microchannel Heat Sinks for High-Power IGBT Cooling. *Int. Commun. Heat Mass Transf.* **2024**, *159*, 108036. [CrossRef]
89. Jordan, S.M.; Schreiber, C.O.; Parhizi, M.; Shah, K. A New Multiphysics Modeling Framework to Simulate Coupled Electrochemical-Thermal-Electrical Phenomena in Li-Ion Battery Packs. *Appl. Energy* **2024**, *360*, 122746. [CrossRef]
90. Grazioli, D.; Magri, M.; Salvadori, A. Computational Modeling of Li-Ion Batteries. *Comput. Mech.* **2016**, *58*, 889–909. [CrossRef]
91. Wang, F.; Zhai, Z.; Zhao, Z.; Di, Y.; Chen, X. Physics-Informed Neural Network for Lithium-Ion Battery Degradation Stable Modeling and Prognosis. *Nat. Commun.* **2024**, *15*, 4332. [CrossRef]
92. Wen, P.; Ye, Z.-S.; Li, Y.; Chen, S.; Xie, P.; Zhao, S. Physics-Informed Neural Networks for Prognostics and Health Management of Lithium-Ion Batteries. *IEEE Trans. Intell. Veh.* **2024**, *9*, 2276–2289. [CrossRef]
93. Navidi, S.; Thelen, A.; Li, T.; Hu, C. Physics-Informed Machine Learning for Battery Degradation Diagnostics: A Comparison of State-of-the-Art Methods. *Energy Storage Mater.* **2024**, *68*, 103343. [CrossRef]
94. Finegan, D.P.; Zhu, J.; Feng, X.; Keyser, M.; Ulmefors, M.; Li, W.; Bazant, M.Z.; Cooper, S.J. The Application of Data-Driven Methods and Physics-Based Learning for Improving Battery Safety. *Joule* **2021**, *5*, 316–329. [CrossRef]
95. Feng, X.; Ouyang, M.; Liu, X.; Lu, L.; Xia, Y.; He, X. Thermal Runaway Mechanism of Lithium Ion Battery for Electric Vehicles: A Review. *Energy Storage Mater.* **2018**, *10*, 246–267. [CrossRef]
96. Xu, B.; Lee, J.; Kwon, D.; Kong, L.; Pecht, M. Mitigation Strategies for Li-Ion Battery Thermal Runaway: A Review. *Renew. Sustain. Energy Rev.* **2021**, *150*, 111437. [CrossRef]
97. Kim, S.W.; Kwak, E.; Kim, J.-H.; Oh, K.-Y.; Lee, S. Modeling and Prediction of Lithium-Ion Battery Thermal Runaway via Multiphysics-Informed Neural Network. *J. Energy Storage* **2023**, *60*, 106654. [CrossRef]
98. Wang, Y.; Xiong, C.; Wang, Y.; Xu, P.; Ju, C.; Shi, J.; Yang, G.; Chu, J. Temperature State Prediction for Lithium-Ion Batteries Based on Improved Physics Informed Neural Networks. *J. Energy Storage* **2023**, *73*, 108863. [CrossRef]
99. Chen, K.; Song, M.; Wei, W.; Wang, S. Design of the Structure of Battery Pack in Parallel Air-Cooled Battery Thermal Management System for Cooling Efficiency Improvement. *Int. J. Heat Mass Transf.* **2019**, *132*, 309–321. [CrossRef]

100. Liu, H.; Wei, Z.; He, W.; Zhao, J. Thermal Issues about Li-Ion Batteries and Recent Progress in Battery Thermal Management Systems: A Review. *Energy Convers. Manag.* **2017**, *150*, 304–330. [CrossRef]
101. Gümüşsu, E.; Ekici, Ö.; Köksal, M. 3-D CFD Modeling and Experimental Testing of Thermal Behavior of a Li-Ion Battery. *Appl. Therm. Eng.* **2017**, *120*, 484–495. [CrossRef]
102. Esmaeili, J.; Jannesari, H. Developing Heat Source Term Including Heat Generation at Rest Condition for Lithium-Ion Battery Pack by up Scaling Information from Cell Scale. *Energy Convers. Manag.* **2017**, *139*, 194–205. [CrossRef]
103. Deng, H.-P.; He, Y.-B.; Wang, B.-C.; Li, H.-X. Physics-Dominated Neural Network for Spatiotemporal Modeling of Battery Thermal Process. *IEEE Trans. Ind. Inform.* **2024**, *20*, 452–460. [CrossRef]
104. Cho, G.; Zhu, D.; Campbell, J.J.; Wang, M. An LSTM-PINN Hybrid Method to Estimate Lithium-Ion Battery Pack Temperature. *IEEE Access* **2022**, *10*, 100594–100604. [CrossRef]
105. Shen, K.; Xu, W.; Lai, X.; Li, D.; Meng, X.; Zheng, Y.; Feng, X. Physics-Informed Machine Learning Estimation of the Temperature of Large-Format Lithium-Ion Batteries under Various Operating Conditions. *Appl. Therm. Eng.* **2025**, *269*, 126200. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

Early Prediction of Battery Lifetime Using Centered Isotonic Regression with Quantile-Transformed Features

Muhammad Arslan Khan, Yixing Wang and Benben Jiang *

Department of Automation, Tsinghua University, Beijing 100084, China;
arslan22@mails.tsinghua.edu.cn (M.A.K.); yx-wang21@mails.tsinghua.edu.cn (Y.W.)
* Correspondence: bbjiang@tsinghua.edu.cn

Abstract: The rapid development of lithium-ion (Li-ion) batteries has raised requirements for cycle life prediction to ensure safety and reliability. However, the intricate and nonlinear behavior of the battery degradation process poses significant challenges in accurately predicting its cycle life at an early stage. This work addresses the battery lifetime prediction problem by leveraging machine learning (ML) methods. First, we apply quantile transformation (QT) to both features and battery cycle lives to improve their correlation. Next, we adopt a centered isotonic regression (CIR) method in our work, a variant of isotonic regression (IR) that centers the feature values and then reduces overfitting to improve model performance. Finally, we perform convex regression (CR) to capture specific patterns in the battery dataset where monotonicity is absent and achieve an overall prediction for battery cycle lives. To validate our proposed method, we have done a comprehensive comparison among several different benchmarks, including elastic net, gradient boosting regression tree, decision tree, support vector machine, random forest, and Gaussian process regression. In contrast to existing methods, our CIR model has shown the best performance, with an average percentage error of 9.8% and a root mean square error of 149 cycles. These experiment results demonstrate the capability and potential of our proposed CIR model in the problem of battery lifetime prediction.

Keywords: data analytics; data-driven prediction; lithium-ion battery; battery lifetime prediction; feature transformation

1. Introduction

Many applications, including electric vehicles, computers, and smartphones, have adopted lithium-ion (Li-ion) batteries due to their relative low cost, long cycle life, and high energy and power density [1–3]. Li-ion batteries are very important for reducing the operational costs of battery storage systems in smart grids and microgrids [4,5]. However, Li-ion batteries degrade with time through internal electrochemical reactions as well as other environmental factors. These problems not only increase operation costs by forcing the replacement of Li-ion batteries but also lead to serious accidents, including explosions and even bodily injuries. Therefore, optimizing operations and ensuring safety and accuracy in the systems will be important. In addition, early-stage lifetime prediction will save much time for the manufacture, production, and optimization of batteries. In high-throughput battery optimization processes, such as multistep rapid charging, early lifetime prognostics minimize the number of testing cycles required for each cell, thereby accelerating the overall optimization process [6]. In most cases, the approaches for battery health and lifetime prognostics are classified into two groups: model-based and data-driven methods.

The model-based methods pay attention to the underlying aging mechanisms that define batteries by lists of equations describing the underlying chemical and physical reactions. These methods employ modern filtering algorithms, including the particle filter [7] and Kalman filter [8], to continuously modify model parameters in order to predict the battery's remaining useful life (RUL). These methods highlight that only a few developed empirical models are built based on assuming simple mathematical relations between the capacity of the battery and the aging cycles in a way that does not consider any internal electrochemical mechanisms. The commonly developed empirical models widely involve methods such as linear [9], polynomial [10], exponential [11], logarithmic [12], and hybrid [13] methods. However, there are a few drawbacks to model-based methods. Firstly, the efficiency of lifetime predictions is heavily based on the strength of the underlying degradation models. Furthermore, these methods generally require extensive data along the degradation trajectory to achieve accurate predictions. It is difficult to predict the battery lifetime accurately at an early stage of degradation due to the inherently nonlinear behavior of capacity degradation and the negligible changes during the initial battery cycles [6].

The data-driven methods, which do not require explicit battery models, regard the battery system as a black box and then derive the RUL and state of health (SOH) from extracted features. The selection of input features generally influence the effectiveness of data-driven approaches [14]. Elastic net (EN) regression is applied to predict RUL using early-cycle charging and discharging curves. After that, several features derived from voltage, temperature, capacity, and internal resistance are used as inputs for the machine learning (ML) model to predict battery lifetime [6]. Liu et al. [15] introduced the Box-Cox transformation (BCT) to model the relationship between extracted features and Li-ion battery capacities. These features were then used to predict the RUL using a relevance vector machine (RVM). Nuhic et al. [16] implemented the support vector machine (SVM) to develop embedded diagnostics and prognostics for estimating the SOH and RUL of Li-ion batteries. Zhang et al. [17] achieved the battery lifetime prediction by combining feature construction with physical features, specifically electrochemical model parameters. Paulson et al. [18] employed sequential feature selection to identify suitable features from a total of 396 extracted features. Yang et al. [19] extracted 72 features from voltage, temperature, and capacity data, and utilized various feature combinations with a gradient boosting regression tree (GBRT) to predict battery lifetime. Fei et al. [20] used four different feature selection methods to select important features from 42 manually extracted ones and then predicted the battery's RUL using various ML techniques. Zhang et al. [21] employed neural networks to predict RUL using features extracted from discharge current, terminal voltage, discharge capacity, and IR. Fei et al. [22] automatically extracted features using convolutional neural networks (CNNs) and combined them with manually crafted features based on domain knowledge to enhance prediction accuracy. Gong et al. [23] extracted 20 capacity degradation-related features and utilized various feature selection methods to improve the accuracy of battery lifetime predictions. He et al. [24] employed CNNs to predict RUL by generating graphical features from the capacity-voltage curves. Zhao et al. [25] implemented a transfer learning technique and developed a model called convolutional neural network with attention (CNN-AT) that integrates an attention mechanism during feature extraction to improve prediction accuracy. However, considering the nonlinear degradation behavior of Li-ion batteries, most of these features have a nonlinear correlation with the battery's cycle life. Applying appropriate nonlinear transformations to the original features would facilitate the feature selection. Zhang et al. [26] developed a linear model between transformed capacity and battery cycle life using BCT, which can process the capacity data to predict the RUL.

In this paper, we consider a diverse set of 20 features from raw data to capture key aspects of battery aging dynamics [6]. We transform the complete feature set into a lower-dimensional subset to address potential irrelevance and redundancy, thereby enhancing predictive accuracy. We apply quantile transformation (QT) to all 20 features and select 11 features to train the centered isotonic regression (CIR) model. The CIR method, a variant of isotonic regression (IR), centers feature values and mitigates the overfitting to improve model performance. Finally, we employ convex regression (CR) to capture specific patterns in the battery dataset where monotonicity is insufficient, thereby improving the overall prediction of battery cycle life. These methods take into consideration the nonlinear degradation performance of Li-ion batteries and improve the accuracy of the lifetime predictions.

Based on the above discussion, we outline our major contributions to advance the state of the art in battery lifetime prediction, summarized below:

- We demonstrate the effectiveness of applying QT to both features and battery cycle life. This technique reduces the impact of outliers and non-normality in battery degradation data, resulting in a more stable and informative feature representation for modeling. The improved feature distribution enhances prediction accuracy.
- We introduce CIR, a variant of IR approach, to capture the monotonic relationship between transformed features and battery cycle life. By centering feature values, CIR mitigates overfitting and improves the model's ability to identify underlying trends, leading to more accurate predictions, particularly in the early stages of battery life.
- To address cases where the monotonicity assumption of CIR is insufficient, we employ CR. It enables the model to capture specific patterns in battery dataset, further enhancing battery cycle life forecasts and overall predictive accuracy.

To further evaluate the impact of different ML models on the early-stage prediction of Li-ion battery lifetime, a comprehensive analysis was performed using the feature subset selected by the QT method. The ML models evaluated in this paper are EN, gradient boosting egression tree (GBRT), decision tree (DT), SVM, RF, and Gaussian process regression (GPR). With an average percentage error (APE) value of 9.8% and a root mean square error (RMSE) of 149 cycles, the results show that our proposed CIR model works better than other ML algorithms.

The rest of the paper is organized as follows. Section 2 describes the fundamental principles of QT, IR, CIR, CR, and the overall framework for Li-ion battery lifetime prediction. Section 3 explains the dataset used and the feature extraction process. Section 4 explains feature selection and compares the results for the battery RUL prediction. Finally, Section 5 summarizes this paper.

2. Methods

This section focuses on various ML-based approaches for RUL battery prediction: QT, IR, CIR, and CR. The effectiveness of models is evaluated based on the prediction accuracy and reliability.

2.1. Quantile Transformation

Quantile Transformation (QT) is a preprocessing technique developed to conduct nonlinear transformations on data, one of the widely used methods in the ML domain, which differs from other conventional methods of linear scaling [27,28]. In particular, this is useful for situations where transformation to a uniform or normal distribution significantly increases interpretability and comparability. We have implemented QT to transform the input data into a uniform distribution.

Each feature undergoes a separate transformation process for the QT method. First, it estimates the cumulative distribution function (CDF) of the original data. Next, the QT utilizes the estimated CDF to convert the input data into a uniform distribution, ensuring that all features have a uniform distribution between 0 and 1 [29]. Applying this transformation process separately to each feature maintains the ranking of the feature values relative to each other and reduces sensitivity to outliers. We can express this feature transformation mathematically as follows:

$$X_{tf} = {F_G}^{-1}\left(F_d(X)\right) \tag{1}$$

where $F_d(X)$ is the CDF of the original feature, and ${F_G}^{-1}$ is the inverse CDF of the target distribution, which is uniform in this case. In other words, the transformation of the feature will make sure that the input data is redistributed according to the quantiles of the uniform distribution and makes the features more comparable but less influenced by the original characteristics of the distribution.

We selected the QT since it is most appropriate for the nature of battery degradation data. This type of battery data is usually non-normal and has outliers, and these can have a negative impact on ML model performance. QT addresses this problem by transforming the features and battery cycle lives into a uniform distribution. This consistency maximizes comparability of characteristics and reduces the effect of outliers without loss of rank order of the data, which is crucial in monitoring the relative pattern of battery degradation. Even though other transformation techniques like log transformation or Box-Cox can handle skewness, they often result in distributional assumptions that may not be justified. Standardization and min-max scaling address feature scaling but do not necessarily address non-normality or outliers. QT provides a broader, more robust solution for battery dataset, addressing non-normality and outlier effects while preserving rank information and enabling feature comparability, which leads to improved model performance.

2.2. Isotonic Regression

Isotonic regression (IR) also known as monotonic regression is a nonparametric estimation technique with an imposed monotonicity constraint on the estimated function [30,31]. To estimate a monotone increasing univariate function $y = F(x)$ based on sequence of observations $y = y_1, y_2, \ldots, y_m$ at m distinct data points $x_1 < x_2 < \ldots < x_m$, IR does not require a parametric form for $F(x)$. The goal of IR is to find a function $\hat{F}$ such that the sum of squared errors $\sum_i \left(\hat{F}(x_i) - y_i\right)^2$ is minimal under the constraint that $\hat{F}$ is monotonic. If the original data do not violate the monotonicity condition, IR returns the data unchanged. If there are violations, the IR estimate adjusts the values by replacing the observations that violate the monotonicity constraint with sequences of identical values [32].

In most IR cases, a usual problem lies in the fact that no explicit function estimates are available between data points, as IR gives only the values at data points, which is important in regression tasks that need to understand the $F(x)$ behavior over a continuous range. In real-world applications, estimating a continuous and strictly increasing curve is often more realistic and valuable. For this purpose, linear interpolation among data points can be employed to extend estimates across the entire range of x. This results in a piecewise linear curve that provides a better approximation of $F(x)$ while preserving monotonicity. The approach remains nonparametric and relatively simple.

The Pooled-Adjacent-Violators Algorithm (PAVA) plays a significant role in IR: the monotonicity constraint satisfaction is estimated by this function. The approach in this work will be performed through ordered data monotonicity violation correction by pooling adjacent violators and replacing them with weighted averages, which normally gives simple and necessary adjustments. PAVA combines flexibility in nonparametric estimation with a basic and efficient robust monotonic fit method. PAVA maintains monotonicity,

allowing for reliable estimation and inverse estimation tasks without assuming an explicit functional form. The steps for PAVA can be stated as Algorithm 1.

Algorithm 1: Pooled-Adjacent-Violators Algorithm (PAVA)

Input: (y, n)

- y = Sequence of values $\{y_1, y_2, \ldots, y_m\}$
- n = Sequence of weights $\{n_1, n_2, \ldots, n_m\}$

Output: $(\hat{F})$

- $\hat{F}$ = Sequence of adjusted values $\{\hat{F}_1, \hat{F}_2, \ldots, \hat{F}_m\}$

Procedure:

Initialization:

- For each $i = 1, \ldots, m$:
- $\hat{F}_i \leftarrow y_i$
- $C \leftarrow \emptyset$ where C represents current set of contiguous violations

while loop:

- $V = \{i : i < n, \hat{F}_i > \hat{F}_{i+1}\} \neq \emptyset$, where V denotes violations
- $h \leftarrow \min(V)$; first violation identification
- Note: V identifies all violations; C is current contiguous violation. If contiguity breaks then if-statement resets C
- **if** $h \epsilon C$ **then**
- $C \leftarrow C \cup \{h + 1\}$
- **else**
- $C \leftarrow \{h, h + 1\}$
- **end if**
- **For** $i \in C$:
- $\hat{F}_i \leftarrow \frac{\sum_{k \in C} n_k y_k}{\sum_{k \in C} n_k y_k}$

end while loop

return $\hat{F} = \hat{F}_1, \hat{F}_2, \ldots, \hat{F}_m$.

end Procedure

The algorithm basically starts from the initial point $(x_1,\ y_1)$ and finds the first evidence of non-monotonicity, then replaces the violating values with weighted average. This process continues, pooling values as needed until all monotonicity violations are resolved. The final output is the minimal adjustment required to achieve monotonicity. Piecewise-linear interpolation between the points $(x_1, \hat{F}_1), \ldots, (x_m, \hat{F}_m)$ is used to estimate function F over the continuous range $[x_1, x_m]$. The estimates can be used to find inverse values by identifying the corresponding x-value.

In IR, the estimates remain the same as the original observations when there are no violations of monotonicity. However, if monotonicity is violated, the IR estimate may include parts that remain constant to ensure the monotonicity constraint is met. Also, the data points do not influence the PAVA process, as they are not included in its input. These constraints can pose challenges in applying IR to regression tasks, especially since a strictly monotonic relationship between dependent and independent variables is typically expected, and flat stretches in the estimate are often seen as undesirable. In order to avoid the flat stretches in the estimate without violation of monotonicity, the CIR approach is further established.

2.3. Centered Isotonic Regression

Centered isotonic regression (CIR), a developed approach to IR, presents an interesting variant of IR for handling monotonicity violations at interior data points [32]. In CIR, we resolve the violations by collapsing the IR estimates onto a single point. The x-coordinate of this point is the weighted average of the involved data points, using the same weighting method as the point estimates.

The CIR point estimates, denoted as $\widetilde{F}$, retain the distinctive values of the IR estimates but with fewer repetitions and possibly with different associated x values. This simplification makes the CIR method more efficient than PAVA, as it pools both x and y values, thereby reducing the need for in-depth data analysis.

CIR works as the shrinkage estimator along the x-axis. When a violation involves only two points, the location of $\left(\widetilde{x}_i, \widetilde{F}_i\right)$ can be found by finding the midpoint between the original data points and the IR estimates along the x-axis. The process of the CIR algorithm is detailed in Algorithm 2.

Algorithm 2: Centered Isotonic Regression (CIR) Algorithm

Input: (x, y, n)

- $x =$ Sequence of input variables $\{x_1, x_2, \ldots, x_m\}$
- $y =$ Sequence of output variables $\{y_1, y_2, \ldots, y_m\}$
- $n =$ Sequence of weights $\{n_1, n_2, \ldots, n_m\}$

Output: $\left(\widetilde{x}, \widetilde{F}, \widetilde{n}\right)$

- $\widetilde{x} =$ Adjusted sequence of input variables
- $\widetilde{F} =$ Adjusted sequence of output variables
- $\widetilde{n} =$ Adjusted sequence of weights

Procedure:

Initialization:

- For each $i = 1, \ldots, m$, set $\widetilde{x}_i \leftarrow x_i, \widetilde{F}_i \leftarrow y_i, \widetilde{n}_i \leftarrow n_i$
- Set $\widetilde{m} \leftarrow m$

while loop:

- $V \equiv \left\{ i : i < m, \left(\widetilde{F}_i > \widetilde{F}_{i+1} t\right) \vee \left(\widetilde{F}_i = \widetilde{F}_{i+1} \wedge \widetilde{F}_i \in (0,1)\right)\right\} \neq \emptyset :$
- $h \leftarrow \min(V)$
- Update $\widetilde{F}_h \leftarrow \frac{\widetilde{n}_h \widetilde{F}_h + \widetilde{n}_{h+1} \widetilde{F}_{h+1}}{\widetilde{n}_h + \widetilde{n}_{h+1}}$
- Update $\widetilde{x}_h \leftarrow \frac{\widetilde{n}_h x_h + \widetilde{n}_{h+1} x_{h+1}}{\widetilde{n}_h + \widetilde{n}_{h+1}}$
- Update $\widetilde{n}_h \leftarrow \widetilde{n}_h + \widetilde{n}_{h+1}$
- Remove $h + 1$
- Update $\widetilde{m} \leftarrow m - 1$

end while loop:

Boundary Conditions:

- **if** $\widetilde{x}_1 > x_1$, **then**:
- Add point $1 : \left(x_1, \widetilde{F}_1, 0\right)$
- Update $\widetilde{m} \leftarrow m + 1$
- **if** $\widetilde{x}_{\widetilde{m}} < x_m$, **then**:
- Add point $\widetilde{m} + 1 : \left(x_m, \widetilde{F}_{\widetilde{m}}, 0\right)$
- Update $\widetilde{m} \leftarrow \widetilde{m} + 1$

return $\widetilde{x}, \widetilde{F}, \widetilde{n}$.

end Procedure

CIR treats the identical values in y series as violations of strict monotonicity, except for sequences of 0 s and 1 s, which are considered as limits. When shrinkage occurs, shown on the x-axis, CIR may produce fewer points than initially provided. The resulting x-values, known as shrinkage points and denoted as $\widetilde{x}$, differ from the original data points. Employing piecewise-linear addition at shrinkage points enables unique reverse estimates for any given value in $\left(\widetilde{F}_1, \widetilde{F}_{\widetilde{m}}\right)$ because the CIR estimate is continuously increasing, but potentially at the boundaries.

The final steps of the algorithm are triggered when monotonicity violations occur at the boundaries of the input range. The main loop substitutes data points involved in violations with a single shrinkage point, which can modify x_1 or remove x_m, thereby restricting the range. The if-statements at the end introduce flat stretches to restore the range to $[x_1, x_m]$. Along these added stretches, the CIR estimate matches the IR estimate. The CIR implementation used in this work follows the PAVA directly. As a non-parametric algorithm, PAVA does not require hyperparameter tuning. It works by iteratively pooling adjacent violators of monotonicity until the monotonicity condition is met. This is a completely data-driven process with no tuning parameters involved. Therefore, our CIR model has no hyperparameters to optimize, ensuring reproducibility without the need for hyperparameter optimization.

2.4. Convex Regression

Convex regression (CR) is a technique used to fit a convex function to a set of data points [33,34]. A function $f(x)$ is considered convex if, for any two points x_1 and x_2 in its domain, the line part connecting $f(x_1)$ and $f(x_2)$ lies on or above the graph of the function. Mathematically, a function $f(x)$ is convex if:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2), \quad \forall \lambda \epsilon [0,1] \tag{2}$$

where λ denotes the weight of x_1 and x_2 in above equation. CR is particularly useful when the relationship between variables is expected to be convex. The constraints in CR frameworks often enforce this convex relationship. The goal of CR is to ensure that the estimated function remains convex while closely fitting the observed data. Unlike linear regression, which is based on a certain linearity assumption between independent and dependent variables, in CR there is no assumption about the parametric form. Instead, it counts on the convexity constraint to drive its fitting, which gives rise to a more flexible and precise model for convex data compared with linear models.

The implementations of CR may vary. Our implementation of IR extends CIR in a way that ensures the fitted function is not only monotonic but characterized by a global minimum or maximum at a specific point, thereby capturing the convexity of the data. The CR process identifies pivot points, dividing the data into sections where the function switches between convex parts. Defining pivot points is critical to maintaining the convexity of the general function. To find the global minima and maxima of the data, smoothing is performed using moving averages. Once the pivot points are identified, the data are split into two parts at the pivot point. For each part, CIR is performed, where the pivot ensures the global convexity. The convex function $f_j(x)$ in part j satisfies:

$$f_j(x) = a_j x + b_j \tag{3}$$

where a_j and b_j are coefficients to be determined during the fit process in that part, and the overall function has to be convex for the entire dataset.

In order to make predictions, the position of a new data point relative to the pivot points is assessed by the fitted model, which utilizes the relative position in order to decide whether the lower part, upper part, or an interpolation between parts is applied. In datasets with numerous features, CR using CIR can be done separately for each feature in order to preserve convexity across all dimensions and support more complex relationships between variables. In fact, CR proves very useful in situations where there is intrinsic convexity between variables, such as in optimization problems where the objective functions often turn out to be convex. By using appropriate CIR techniques, CR can capture data patterns while keeping the function of the expected convex structure.

2.5. Data-Driven Framework for Battery Lifetime Prediction

The framework initiates with the collection of cycle data from LFP/graphite cells, where discharge capacity is recorded over multiple cycles, each reflecting varying stages of the cell's lifecycle. After that, some features are extracted: current, voltage, temperature, and capacity, leading to a multi-dimensional feature space within which important information regarding the battery operation is captured, as shown in Figure 1. These features are transformed using QT, mitigating the influence of outliers with a uniform distribution, improving both stability and the performance of predictive models. After the QT transformation, the most significant features are selected for battery RUL prediction using proposed regression techniques.

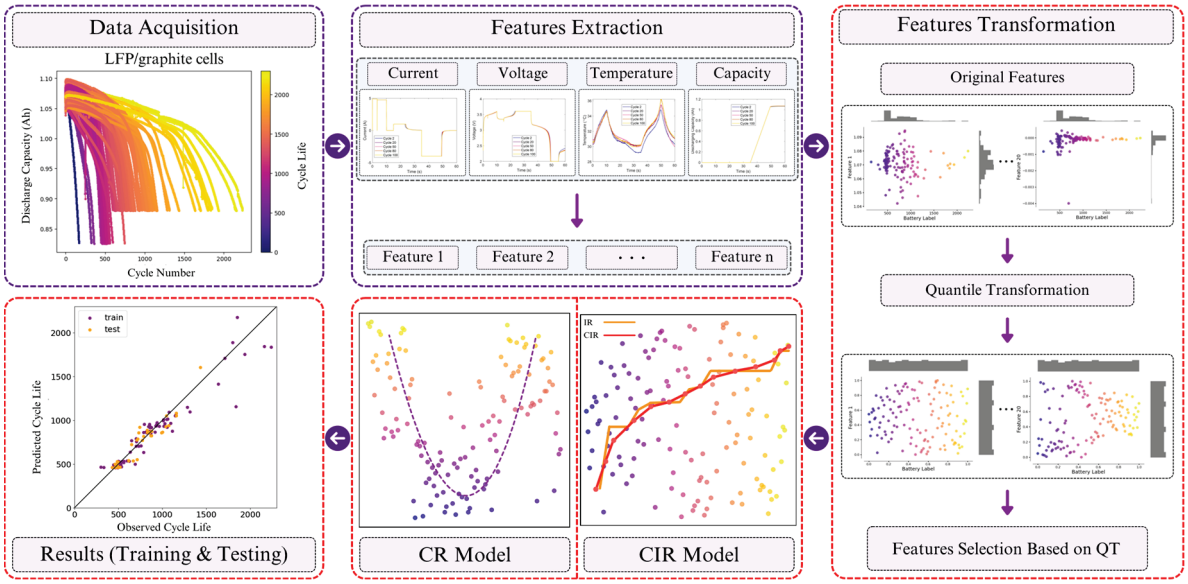


Figure 1. The flowchart of our proposed method for predicting battery cycle life.

The next step involves developing two predictive models: CIR and CR. The CIR model effectively captures the relationship between features and battery cycle lives by reducing overfitting, leading to more accurate and organized predictions. The CR is applied to fit a convex function to the data, effectively capturing specific patterns observed in the dataset where monotonicity is insufficient. The model performance is evaluated by comparing predicted and observed battery cycle lives. The proposed model demonstrates significantly better performance in predicting battery lifetime.

3. Dataset and Feature Analysis

3.1. Dataset Description

In this study, we employ the MIT battery dataset provided by Severson et al. [6], which contains 124 publicly available commercial LFP/graphite A123 APR18650M1A batteries with a 1.1 Ah nominal capacity and 3.3 V nominal voltage. These batteries experienced diverse fast charging conditions ranging from 3.6 C to 6 C with constant current-constant voltage (CC-CV). All batteries followed a uniform discharge protocol. Charging from a 0% to a 80% state of charge (SOC) used single or dual step policies with fixed C-rates over defined SOC intervals (e.g., 6 C from 0 to 50% SOC, then 4 C to 80%). Batteries then charged from 80% to 100% SOC through a 1C CC-CV protocol, terminating at 3.6 V (cutoff: C/50), and discharged at 4 C to 2.0 V (cutoff: C/50). The dataset comprises 124 cells, with each cell involving cycle life, charge policy, and structural summary, incorporating parameters such as cycle number, discharging capacity, charging capacity, IR, maximum temperature,

average temperature, minimum temperature, and charging duration, along with cycle-related features. These batches are classified by considering the sequential order of the battery cell production under varying environmental conditions. Particularly, the dataset comprises scalar variables, including dynamic features such as voltage, current, capacity, temperature, and IR, all subject to continual changes over time. A cycle is deemed complete upon a full charging and later discharging. The lifecycle of all batteries spans from 148 to 2233 under 72 distinctive charging conditions in the controlled laboratory environment.

Analyzing the discharge capacity over the first 1000 cycles shows minimal capacity fade in the first 100 cycles, followed by an accelerated decline towards the end of battery life, a common trend in batteries. The intersection of the capacity fade curves emphasizes the weak correlation between the initial capacity and lifetime, an expected observation because of the less degradation during the early cycles. Generally, in the dataset under investigation, 81% of the cells exhibit an increase in capacity after 100 cycles, underscoring the significance of factors beyond mere capacity measurements. The small increases in capacity after a slow cycle or rest period could be caused by the charge stored in the area of the negative electrode, which is outside the positive electrode's bounds.

Recognizing the limited predictive capacity of the correlations mentioned above relies on capacity fade curves, and we choose an alternative data-driven strategy that covers a complete set of cycling data. This strategy includes full voltage profiles of each cycle and other measures that include cell internal resistance and temperature. These changes in data trends allow for broader and deeper analysis, highlighting the importance of predictive models. The feature extraction from raw data will be done using the approach outlined by Severson et al. [6]. The process of feature extraction is crucial for enhancing the accuracy of cycle life predictions.

3.2. Features Extraction

Our work focuses on the first 100 cycles, taking into account the methodology proposed by Severson et al. [6]. In the complex aging dynamics of Li-ion batteries, several critical parameters effectively capture voltage, charge and discharge capacity, and temperature, as shown in Figure 2 [35].

The feature extraction is accomplished by obtaining 20 features, which are similar to the features in [6], to enable the models to accurately predict battery lifetime, while the details of the 20 features are listed in Table A1 in the Appendix A. Initially, a range of different degradation data from the first 100 cycles is taken into consideration, where most cells do not reveal a significant capacity fade. The 20 features are analyzed in terms of sources and processes of extraction, hence giving a diversified overview with respect to battery aging dynamics. Specifically, geometric analyses and summary statistics of discharge voltage curves and derivatives provide the discharge capacity features, which are the most significant observations for battery degradation. This transformation $\Delta Q(V)$ conveys all the information in the discharge voltage curve and its derivatives, representing the most interesting point in effective degradation evaluation.

Additionally, features F-1 to F-20 are extracted from the discharge voltage curves, with F-1 representing the discharge capacity at cycle 2 and F-2 denoting the difference between the maximum discharge capacity and the capacity at cycle 2. F-3 corresponds to the discharge capacity at cycle 100, F-4 signifies the temperature integral over time from cycle 2-100, and F-5 manifests the average charge time across the first five cycles.

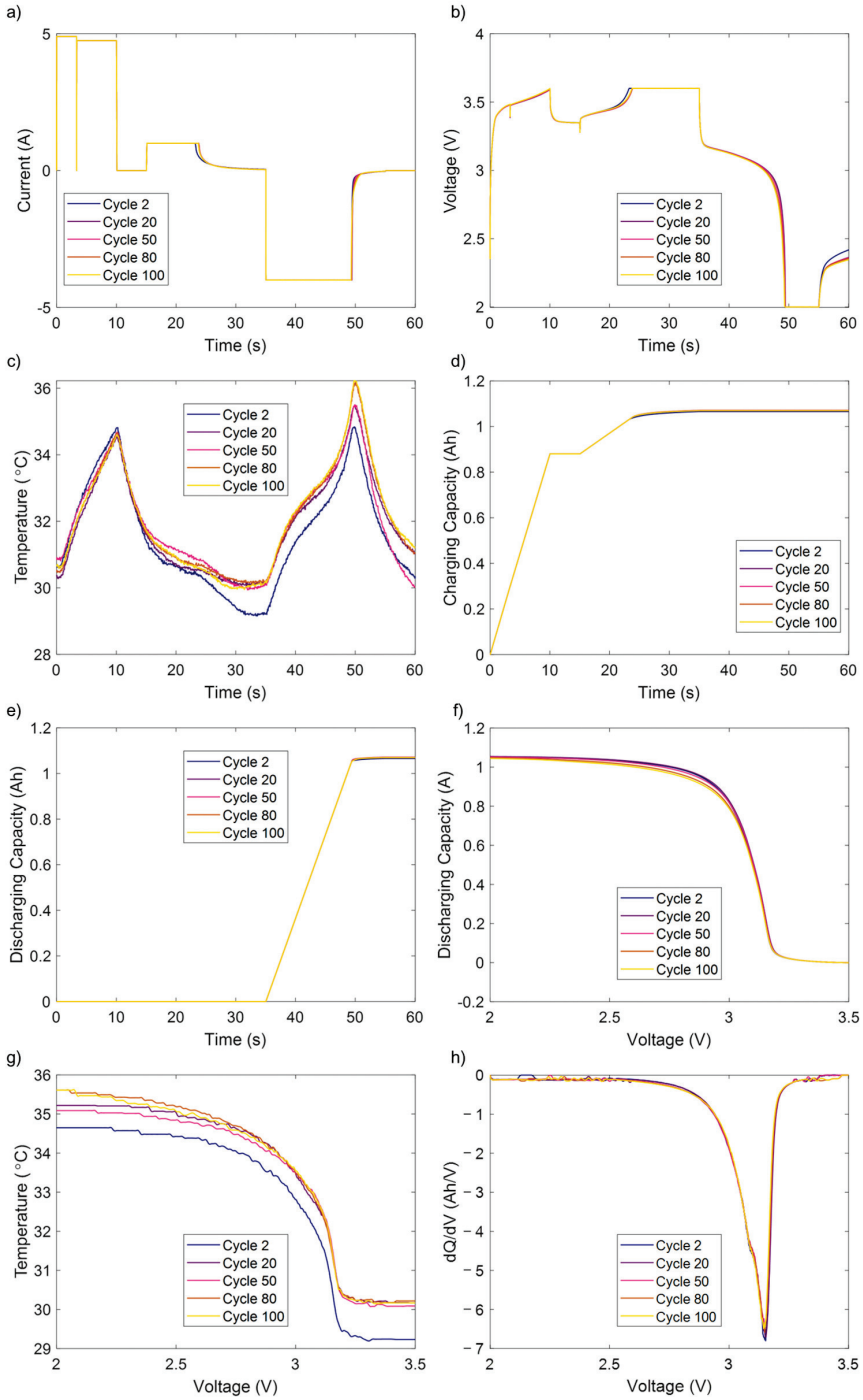


Figure 2. Data visualization of a random LFP/graphite cell: (**a**) current vs. time; (**b**) voltage vs. time; (**c**) temperature vs. time; (**d**) charging capacity vs. time; (**e**) discharging capacity vs. time; (**f**) discharging capacity vs. voltage; (**g**) temperature vs. voltage. (**h**) incremental capacity (IC) curves vs. voltage.

The next set of features, F-6 to F-11, are defined by summary statistics of the $Q_{diff}\ (V)$ curves, which characterize the difference in capacity between cycles 2–100; it is a critical feature in our analysis, including the minimum, mean, variance, skewness, and kurtosis, with F-11 comprising the value at 2 V for Q_{diff} . The features F-12 to F-17 emphasize capturing temperature variations and the behavior of the battery's capacity degradation over time. F-12 and F-13 denote the maximum and minimum temperatures recorded from cycles 2 to 100, respectively. F-14 and F-15 describe the slope and intercept of the capacity fade curve from cycles 2 to 100, while F-16 and F-17 capture the slope and intercept of the

capacity fade curve from cycles 91 to 100. For IR, F-18 is the minimum IR from cycle 2 to cycle 100, a measure of efficiency while in operation. F-19 actually gives the IR value of cycle 2; this is at the very start of cycling and provides a starting point from which all changes can be monitored. Finally, F-20 captures the difference in IR between cycle 100 and cycle 2, giving an idea of degradation and how that impacts the lifetime of the battery. Subsequently, the extracted features undergo feature selection, which aims to find the most critical indicators of battery degradation and performance.

4. Experimental Results

4.1. Feature Selection Using Quantile Transformation

In this section, the QT applied to the 20 features is a major preprocessing step that balances the varying distributions of the features, causing them to fall into a uniform distribution. It mitigates the effect of highly skewed distributions and outliers, making data consistent and easy to handle, which is mainly helpful for models sensitive to the distribution and size of the input data. The QT improves feature comparability and its correlations with the battery cycle life by distributing the data uniformly. The major result of this transformation is to equalize the distribution of each feature. This uniformity ensures that no single feature dominates the model due to scale or distribution, allowing for a more balanced input into the ML algorithms intended to deploy.

The features that most likely predict battery lifetime will be those that provide good, consistent trends with the battery's cycle life. A feature that clearly exhibits a linear or non-linear trend and has a strong correlation with the battery cycle life may be an important predictor in the model. The modeling method should rank these features due to their high impact on the accuracy of the model.

Alternatively, features that exhibit a weak or unclear pattern with the battery cycle life may contribute less to the prediction. However, we cannot fully ignore such features, and their potential predictive power may require further transformation or feature engineering. One way to capture more complicated relationships that are not immediately evident in raw form is by generating polynomial features or interaction terms.

By considering each of the 20 features individually, we have noticed that features, such as F-1 to F-5, may show linear correlations with the battery cycle life. Even if these linear correlations are easy to model, redundancy should be avoided, especially if the features have a strong correlation with the battery cycle life. In such conditions, decreasing the number of features might enhance model performance by avoiding overfitting.

The F-6 offers a unique case because of its convex pattern. The non-linear connection implies a more complex relation between the feature and battery cycle life, maybe indicating a situation where performance increases to a certain extent before declining. It's necessary to include this complexity in model as it is able to reflect a significant component of the underlying battery behavior that a simpler linear model might neglect. F-7 to F-10 exhibit a combination of linear and non-linear patterns, and their role in the model will depend on how strongly these features correlate with the battery cycle life. If these features show significant patterns, then they could be key predictors.

Similarly, F-11 to F-15 display varying degrees of correlation with the battery cycle life. The strongly correlated features are probably important indicators of battery lifetime, but weakly correlated features could need more adjustments to make them more helpful in the model. Lastly, F-16 to F-20 could display a range of patterns, from strong linear relationships to more complex, non-linear interactions, and understanding these relationships is vital for constructing a robust predictive model.

In addition to ordering these features, the QT method revealed the underlying linear and non-linear patterns of association between the features and the battery cycle life. The

selection of modeling approaches and feature engineering methodologies is guided by these patterns, which guarantee that the model can accurately capture the complex patterns in the battery dataset. We used the correlation matrix generated with Pearson correlation, to guide feature selection after QT. We identified the 11 most highly correlated features with battery cycle life, while minimizing the inclusion of strongly inter-correlated features to avoid redundancy.

Based on the aforementioned discussion, we select the following 11 features: F-2, F-4, F-5, F-6, F-8, F-14, F-15, F-17, F-18, F-19, and F-20 for our regression CIR model. Currently, the features that strongly correlated or displayed significant patterns indicating predictive value with the battery cycle life are the ones that are shown in this selection. We propose a ML based method for the early prediction of Li-ion battery lifetime. This framework is composed of three main parts: feature transformation, feature selection, and ML based predictive modeling. Initially, a diverse set of 20 features is extracted from the raw degradation data to capture various aspects of battery aging dynamics. Subsequently, to address potential feature redundancy and irrelevance, the full feature set is transformed to produce a lower-dimensional subset, thereby enhancing the model's predictive accuracy. This study uses the selected 11 features to evaluate the performance of various ML algorithms in terms of battery lifetime prediction.

This paper utilized 123 batteries from the MIT dataset, except for that with the shortest life, whose degradation pattern differed from other cells [24]. Stratified random sampling considers the intrinsic variability between the long-lived and short-lived cells in a given dataset for balanced representation. We divide the data so that 2/3 (83 samples) form the training set, and 1/3 (40 samples) constitute the testing set. This maintains the same ratio of long-lived to short-lived cells in the subsets [19]. We repeat the stratified random sampling procedure twenty times [20], to make the experiment less variable and average the results across these repetitions to make them reliable and robust.

This work uses two crucial statistical metrics, APE and RMSE, to compare the performance of our proposed method with other benchmarks. APE calculates the average percentage difference between the predicted and observed cycle lives, providing a valuable understanding of the overall accuracy of the predictive models. A lower APE value reflects a higher degree of accuracy and alignment between predicted and observed battery lifetimes, which means that the proposed method efficiently improves lifetime predictions. The APE is given as

$$\text{APE}(\%) = \frac{1}{N}\sum_{i=1}^{N} \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \tag{4}$$

where y_i represents the observed cycle life, $\hat{y}_i$ denotes the predicted cycle life, and the total number of battery samples is N. In contrast, RMSE provides a total measure of the magnitude of the differences between predicted and actual cycle lives, capturing variance and bias in the predictive models. The RMSE is stated as

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{5}$$

RMSE figures out the differences between the predicted and observed cycle lives to show how well the prediction model works overall; lower RMSE values mean that the model is more accurate and reliable at predicting battery lifetimes.

4.2. Results Comparison with Benchmarks

This section employs the CIR approach to predict battery cycle life and compares it with several ML techniques, including EN [6], GBRT [19], DT [19], SVM [19], RF [20], and

GPR [20]. In the SVM model, we adopt the medium Gaussian kernel function. In the GPR model, we choose the exponential kernel function. As discussed in Section 4.1, we split the dataset into two parts: 2/3 for training and 1/3 for testing. To ensure its validity, we executed stratified random sampling and repeated for 20 times. Performance metrics give the average results over multiple iterations as an evaluation of predictive accuracy for each model.

The metric results are summarized in Table 1, which compares the predictive performance of the CIR method with that obtained from other models. Utilizing a total of 11 selected features as input, CIR performed better than alternative methods for the two criteria of APE and RMSE. This advantage is because CIR accounts for monotonicity in the data, which usually occurs in battery lifetime predictions. Figure 3 shows the relationship between observed and predicted battery lifetime. The x-axis represents the observed values, the y-axis represents the predictions, and the blue and red dots indicate the training and testing results, respectively.

Table 1. Model metrics for different prediction methods.

Model	APE (%)		RMSE (Cycles)	
	Train	Test	Train	Test
CIR Model	7.3	9.8	120	149
EN	8.9	10.8	172	179
GBRT	7.7	10.8	161	181
DT	6.5	11.3	107	157
SVM	9.7	11.5	210	214
RF	8.5	11.4	206	217
GPR	8.6	10.4	164	180

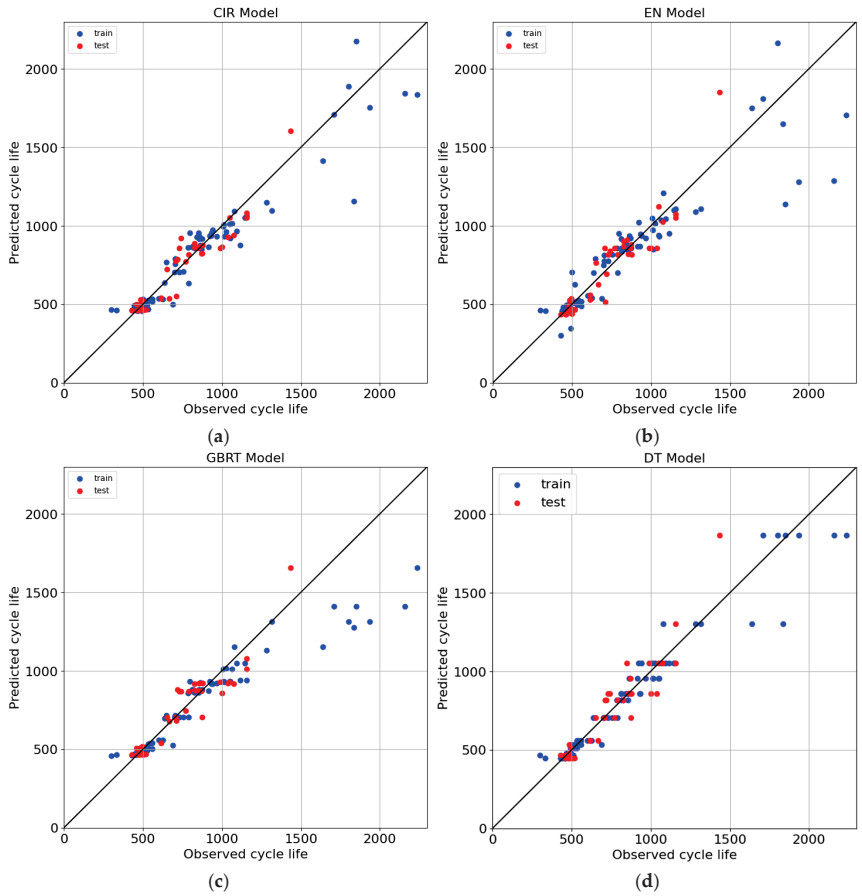


Figure 3. *Cont.*

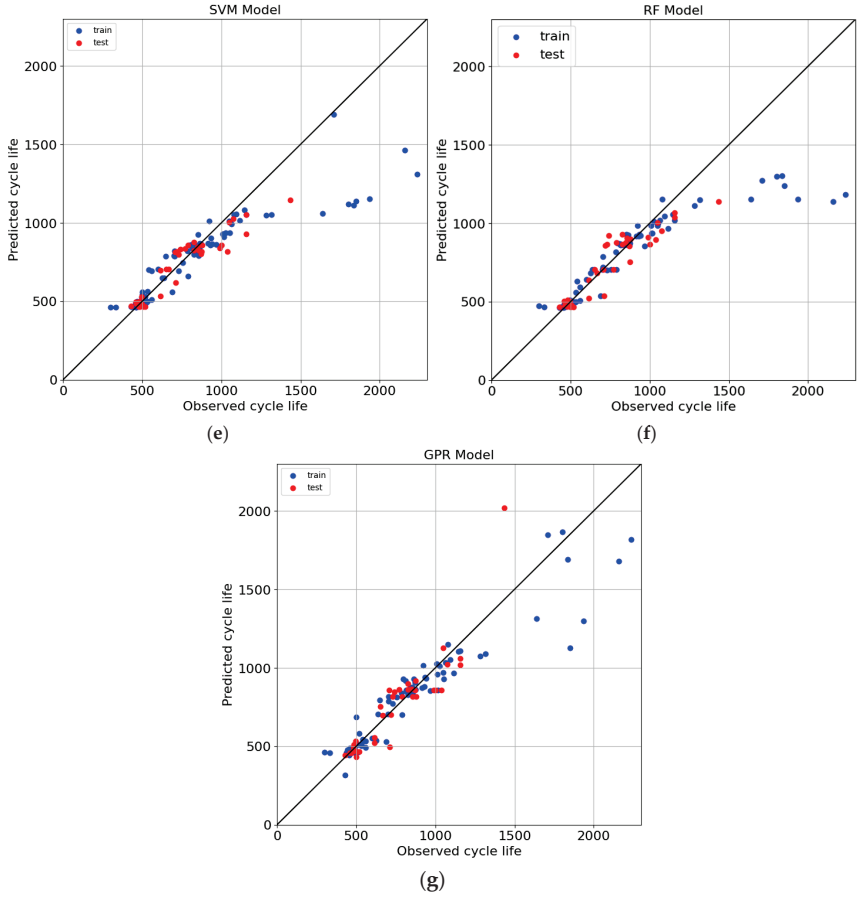


Figure 3. Battery cycle life prediction results of (**a**) CIR; (**b**) EN; (**c**) GBRT; (**d**) DT; (**e**) SVM; (**f**) RF; and (**g**) GPR.

The evaluation provided APE values of 9.8% and RMSE of 149 cycles, respectively, for the CIR method. These findings highlight the accuracy of the CIR method compared to traditional models including, EN and SVM, which show higher error rates. The EN model showed an APE of 10.8%, with RMSE values of 179 cycles. While EN effectively balances the limitations of regression techniques, it struggles to capture complex patterns, which are reflected in its relatively high error rates, signifying less accurate predictions.

Similarly, GBRT achieved an APE of 10.8%, with RMSE values of 181 cycles. Despite its ability to manage non-linear patterns and reduce bias, GBRT is prone to overfitting, particularly with smaller datasets, as indicated by its elevated RMSE. The DT model obtained an APE of 11.3%, with RMSE values of 157 cycles. While DT is easy to interpret, they often overfit, as evidenced by the significant increase in RMSE, indicating weaker generalization.

The SVM demonstrated an APE of 11.5%, with RMSE values of 214 cycles. While SVM excels in handling high-dimensional data, these models can be computationally expensive and sensitive to kernel selection. The high RMSE values suggest challenges in fitting the model to the complex patterns present in the data. The RF model obtained an APE of 11.4%, with RMSE values of 217 cycles. However, RF is robust and effective at minimizing overfitting; it cannot directly account for patterns in the data.

Finally, the GPR achieved an APE of 10.4%, with RMSE value of 180 cycles. While GPR offers a probabilistic approach to regression, it faces scalability challenges and high computational costs, particularly with large datasets, leading to risks of overfitting, as suggested by its performance metrics.

5. Conclusions

This paper demonstrated the effectiveness of a data-driven methodology for early Li-ion battery lifetime prediction using QT, CIR, and CR. A key outcome of this work is the collaborative benefit of combining these methods to address challenges related to non-normality, outliers, and complex degradation patterns in battery data. QT enhances correlations between features and battery cycle lives, facilitating the discovery of stronger and more generalized relationships. CIR, a modified IR variant incorporating centering, mitigates overfitting while improving the model's ability to capture hidden monotonic trends. Moreover, CR enables the accurate modeling of specific degradation patterns, further enhancing predictive accuracy.

The results highlight the significant potential of data-driven approaches for early-stage battery lifetime prediction. The achieved APE of 9.8% and RMSE of 149 cycles on the MIT dataset compare favorably with existing ML methods, demonstrating the practical applicability of the proposed methodology for battery optimization and management. A key insight from this study is the importance of accounting for the intrinsic characteristics of battery degradation data when selecting and integrating modeling techniques. While QT, CIR, and CR have been explored in various applications, their combination for early battery lifetime prediction yields notable performance improvements.

Future work includes extending this approach to other battery chemistries and datasets, exploring advanced feature engineering and selection techniques incorporating physics-informed features to enhance predictive accuracy. Additionally, optimizing computational efficiency will be crucial for scaling the method to real-world battery management systems.

Author Contributions: Conceptualization, M.A.K., Y.W. and B.J.; methodology, M.A.K. and Y.W.; software, M.A.K. and Y.W.; validation, M.A.K. and Y.W.; formal analysis, M.A.K. and Y.W.; investigation, M.A.K. and Y.W.; resources, M.A.K., Y.W. and B.J.; data curation, M.A.K. and Y.W.; writing—original draft preparation, M.A.K. and Y.W.; writing—review and editing, M.A.K., Y.W. and B.J.; visualization, M.A.K. and Y.W.; supervision, B.J.; project administration, B.J.; funding acquisition, B.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Natural Science Foundation of China under grant number 62273197, and the Beijing Natural Science Foundation under grant number L233027.

Data Availability Statement: The authors have not generated any new experimental data in this research, and the data utilized in the research is available at https://data.matr.io/1 (accessed on 20 January 2025, under the "data-driven prediction of battery cycle life before capacity degradation" project.

Acknowledgments: We would like to thank Xiaofei Zhao at Tsinghua University for helpful discussion.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Features extraction from the MIT dataset for battery lifetime prediction [6].

Feature No.	Feature Details	Feature Representation
F-1	Discharge capacity at cycle 2	QC_2
F-2	Difference between maximum discharge capacity and cycle 2	$\mathrm{Max}(QC_1, \ldots, QC_{100}) - QC_2$
F-3	Discharge capacity at cycle 100	QC_{100}
F-4	Integral of temperature over time, cycles 2–100	$\mathrm{Int}T_{2-100}$
F-5	Average charge time first 5 cycles	$\mathrm{Avg}t_{1-5}$
F-6	Minimum of voltage discharge curve	$\mathrm{Min}Q_{diff}\ (V)$

Table A1. *Cont.*

Feature No.	Feature Details	Feature Representation
F-7	Mean of voltage discharge curve	$\mathrm{Mean}Q_{diff}\ (V)$
F-8	Variance of voltage discharge curve	$\mathrm{Var}Q_{diff}\ (V)$
F-9	Skewness of voltage discharge curve	$\mathrm{Skew}Q_{diff}\ (V)$
F-10	Kurtosis of voltage discharge curve	$\mathrm{Kurt}Q_{diff}\ (V)$
F-11	Value at 2V of voltage discharge curve	$Q_{diff}\mid_{V=2\mathrm{V}}$
F-12	Maximum temperature cycles 2–100	$\mathrm{Max}T_{2-100}(V)$
F-13	Minimum temperature cycles 2–100	$\mathrm{Min}T_{2-100}(V)$
F-14	Slope of the capacity fade curve, cycles 2–100	$\mathrm{Slope}CF_{2-100}$
F-15	Intercept of the capacity fade curve, cycles 2–100	$\mathrm{Intp}CF_{2-100}$
F-16	Slope of the capacity fade curve, cycles 91–100	$\mathrm{Slope}CF_{91-100}$
F-17	Intercept of the capacity fade curve, cycles 91–100	$\mathrm{Intp}CF_{91-100}$
F-18	Minimum internal resistance, cycles 2-100	$\mathrm{Min}IR_{2-100}$
F-19	Internal resistance cycle 2	IR_2
F-20	Internal resistance difference between cycle 100 and cycle 2	$IR_{100} - IR_2$

References

1. Bhadriraju, B.; Kwon, J.S., II; Khan, F. An Adaptive Data-Driven Approach for Two-Timescale Dynamics Prediction and Remaining Useful Life Estimation of Li-Ion Batteries. *Comput. Chem. Eng.* **2023**, *175*, 108275. [CrossRef]
2. Cho, S.; Jeong, H.; Han, C.; Jin, S.; Lim, J.H.; Oh, J. State-of-Charge Estimation for Lithium-Ion Batteries under Various Operating Conditions Using an Equivalent Circuit Model. *Comput. Chem. Eng.* **2012**, *41*, 1–9. [CrossRef]
3. Cui, X.; Shen, W.; Zhang, Y.; Hu, C.; Zheng, J. Novel Active LiFePO4 Battery Balancing Method Based on Chargeable and Dischargeable Capacity. *Comput. Chem. Eng.* **2017**, *97*, 27–35. [CrossRef]
4. Homan, B.; ten Kortenaar, M.V.; Hurink, J.L.; Smit, G.J.M. A Realistic Model for Battery State of Charge Prediction in Energy Management Simulation Tools. *Energy* **2019**, *171*, 205–217. [CrossRef]
5. Wang, S.; Guo, D.; Han, X.; Lu, L.; Sun, K.; Li, W.; Sauer, D.U.; Ouyang, M. Impact of Battery Degradation Models on Energy Management of a Grid-Connected DC Microgrid. *Energy* **2020**, *207*, 118228. [CrossRef]
6. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D.; et al. Data-Driven Prediction of Battery Cycle Life before Capacity Degradation. *Nat. Energy* **2019**, *4*, 383–391. [CrossRef]
7. Wang, D.; Yang, F.; Tsui, K.L.; Zhou, Q.; Bae, S.J. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Spherical Cubature Particle Filter. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 1282–1291. [CrossRef]
8. Zheng, X.; Fang, H. An Integrated Unscented Kalman Filter and Relevance Vector Regression Approach for Lithium-Ion Battery Remaining Useful Life and Short-Term Capacity Prediction. *Reliab. Eng. Syst. Saf.* **2015**, *144*, 74–82. [CrossRef]
9. Burgess, W.L. Valve Regulated Lead Acid Battery Float Service Life Estimation Using a Kalman Filter. *J. Power Sources* **2009**, *191*, 16–21. [CrossRef]
10. Micea, M.V.; Ungurean, L.; Cârstoiu, G.N.; Groza, V. Online State-of-Health Assessment for Battery Management Systems. *IEEE Trans. Instrum. Meas.* **2011**, *60*, 1997–2006. [CrossRef]
11. He, W.; Williard, N.; Osterman, M.; Pecht, M. Prognostics of Lithium-Ion Batteries Based on Dempster-Shafer Theory and the Bayesian Monte Carlo Method. *J. Power Sources* **2011**, *196*, 10314–10321. [CrossRef]
12. Yang, F.; Wang, D.; Xing, Y.; Tsui, K.L. Prognostics of Li(NiMnCo)O2-Based Lithium-Ion Batteries Using a Novel Battery Degradation Model. *Microelectron. Reliab.* **2017**, *70*, 70–78. [CrossRef]
13. Hu, C.; Ye, H.; Jain, G.; Schmidt, C. Remaining Useful Life Assessment of Lithium-Ion Batteries in Implantable Medical Devices. *J. Power Sources* **2018**, *375*, 118–130. [CrossRef]
14. Chen, X.; Liu, X.; Shen, X.; Zhang, Q. Applying Machine Learning to Rechargeable Batteries: From the Microscale to the Macroscale. *Angew. Chemie* **2021**, *133*, 24558–24570. [CrossRef]
15. Liu, D.; Zhou, J.; Liao, H.; Peng, Y.; Peng, X. A Health Indicator Extraction and Optimization Framework for Lithium-Ion Battery Degradation Modeling and Prognostics. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 915–928. [CrossRef]
16. Nuhic, A.; Terzimehic, T.; Soczka-Guth, T.; Buchholz, M.; Dietmayer, K. Health Diagnosis and Remaining Useful Life Prognostics of Lithium-Ion Batteries Using Data-Driven Methods. *J. Power Sources* **2013**, *239*, 680–688. [CrossRef]
17. Zhang, Y.; Feng, X.; Zhao, M.; Xiong, R. In-Situ Battery Life Prognostics amid Mixed Operation Conditions Using Physics-Driven Machine Learning. *J. Power Sources* **2023**, *577*, 233246. [CrossRef]
18. Paulson, N.H.; Kubal, J.; Ward, L.; Saxena, S.; Lu, W.; Babinec, S.J. Feature Engineering for Machine Learning Enabled Early Prediction of Battery Lifetime. *J. Power Sources* **2022**, *527*, 231127. [CrossRef]
19. Yang, F.; Wang, D.; Xu, F.; Huang, Z.; Tsui, K.L. Lifespan Prediction of Lithium-Ion Batteries Based on Various Extracted Features and Gradient Boosting Regression Tree Model. *J. Power Sources* **2020**, *476*, 228654. [CrossRef]

20. Fei, Z.; Yang, F.; Tsui, K.L.; Li, L.; Zhang, Z. Early Prediction of Battery Lifetime via a Machine Learning Based Framework. *Energy* **2021**, *225*, 120205. [CrossRef]
21. Zhang, Y.; Peng, Z.; Guan, Y.; Wu, L. Prognostics of Battery Cycle Life in the Early-Cycle Stage Based on Hybrid Model. *Energy* **2021**, *221*, 119901. [CrossRef]
22. Fei, Z.; Zhang, Z.; Yang, F.; Tsui, K.L.; Li, L. Early-Stage Lifetime Prediction for Lithium-Ion Batteries: A Deep Learning Framework Jointly Considering Machine-Learned and Handcrafted Data Features. *J. Energy Storage* **2022**, *52*, 104936. [CrossRef]
23. Gong, D.; Gao, Y.; Kou, Y.; Wang, Y. Early Prediction of Cycle Life for Lithium-Ion Batteries Based on Evolutionary Computation and Machine Learning. *J. Energy Storage* **2022**, *51*, 104376. [CrossRef]
24. He, N.; Wang, Q.; Lu, Z.; Chai, Y.; Yang, F. Early Prediction of Battery Lifetime Based on Graphical Features and Convolutional Neural Networks. *Appl. Energy* **2024**, *353*, 122048. [CrossRef]
25. Zhao, W.; Ding, W.; Zhang, S.; Zhang, Z. Yong. A deep learning approach incorporating attention mechanism and transfer learning for lithium-ion battery lifespan prediction. *J. Energy Storage* **2024**, *75*, 109647. [CrossRef]
26. Zhang, Y.; Xiong, R.; He, H.; Pecht, M.G. Lithium-Ion Battery Remaining Useful Life Prediction With Box–Cox Transformation and Monte Carlo Simulation. *IEEE Trans. Ind. Electron.* **2019**, *66*, 1585–1597. [CrossRef]
27. Mahedy Hasan, S.M.; Uddin, M.P.; Al Mamun, M.; Sharif, M.I.; Ulhaq, A.; Krishnamoorthy, G. A Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders. *IEEE Access* **2023**, *11*, 15038–15057. [CrossRef]
28. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2012**, *12*, 2825–2830.
29. Müller, M.; Huber, F.; Arnaud, M.; Kraemer, A.I.; Altimiras, E.R.; Michaux, J.; Taillandier-Coindard, M.; Chiffelle, J.; Murgues, B.; Gehret, T.; et al. Machine Learning Methods and Harmonized Datasets Improve Immunogenic Neoantigen Prediction. *Immunity* **2023**, *56*, 2650–2663.e6. [CrossRef]
30. Barlow, R.E. *Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression*; John Wiley & Sons Ltd.: Oxford, UK, 1972.
31. Robertson, T. *Order Restricted Statistical Inference*; John Wiley & Sons Ltd.: Oxford, UK, 1988.
32. Oron, A.P.; Flournoy, N. Centered Isotonic Regression: Point and Interval Estimation for Dose–Response Studies. *Stat. Biopharm. Res.* **2017**, *9*, 258–267. [CrossRef]
33. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
34. Simonetto, A. Smooth Strongly Convex Regression. *arXiv* **2021**. [CrossRef]
35. Wang, Y.; Jiang, B. Attention Mechanism-Based Neural Network for Prediction of Battery Cycle Life in the Presence of Missing Data. *Batteries* **2024**, *10*, 229. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Review

Exploiting Artificial Neural Networks for the State of Charge Estimation in EV/HV Battery Systems: A Review

Pierpaolo Dini * and Davide Paolini

Department of Information Engineering, University of Pisa, Via G. Caruso n.16, 56122 Pisa, Italy; davide.paolini@ing.unipi.it
* Correspondence: pierpaolo.dini@unipi.it

Abstract: Artificial Neural Networks (ANNs) improve battery management in electric vehicles (EVs) by enhancing the safety, durability, and reliability of electrochemical batteries, particularly through improvements in the State of Charge (SOC) estimation. EV batteries operate under demanding conditions, which can affect performance and, in extreme cases, lead to critical failures such as thermal runaway—an exothermic chain reaction that may result in overheating, fires, and even explosions. Addressing these risks requires advanced diagnostic and management strategies, and machine learning presents a powerful solution due to its ability to adapt across multiple facets of battery management. The versatility of ML enables its application to material discovery, model development, quality control, real-time monitoring, charge optimization, and fault detection, positioning it as an essential technology for modern battery management systems. Specifically, ANN models excel at detecting subtle, complex patterns that reflect battery health and performance, crucial for accurate SOC estimation. The effectiveness of ML applications in this domain, however, is highly dependent on the selection of quality datasets, relevant features, and suitable algorithms. Advanced techniques such as active learning are being explored to enhance ANN model performance by improving the models' responsiveness to diverse and nuanced battery behavior. This compact survey consolidates recent advances in machine learning for SOC estimation, analyzing the current state of the field and highlighting the challenges and opportunities that remain. By structuring insights from the extensive literature, this paper aims to establish ANNs as a foundational tool in next-generation battery management systems, ultimately supporting safer and more efficient EVs through real-time fault detection, accurate SOC estimation, and robust safety protocols. Future research directions include refining dataset quality, optimizing algorithm selection, and enhancing diagnostic precision, thereby broadening ANNs' role in ensuring reliable battery management in electric vehicles.

Keywords: battery system; data-driven strategy; Smart Industry 5.0; electric vehicles; Artifical Neural Networks (ANNs); SOC monitoring

1. Introduction

With the rise of neural networks and machine learning techniques, there is an increasing demand to apply these technologies to various engineering challenges. These applications include luxury yacht energy management [1,2], energy efficiency in buildings [3,4], industrial robotics [5,6] and electric vehicles [7,8]. AI techniques enable achieving outstanding results while circumventing the difficulties associated with developing accurate physical models of these problems [9]. The transportation industry is investing in

and confronting a range of challenges to improve efficiency, boost performance, enhance connectivity, expand autonomy, and reduce emissions [10,11]. Major automotive markets, especially in Europe, have set ambitious CO_2 reduction targets, positioning electric power trains as one of the most effective solutions. Increasing economic incentives will be allocated to facilitate the transition toward the full adoption of electric vehicles across the region [12], with the goal of achieving this milestone by 2034 [13], as shown in Figure 1.

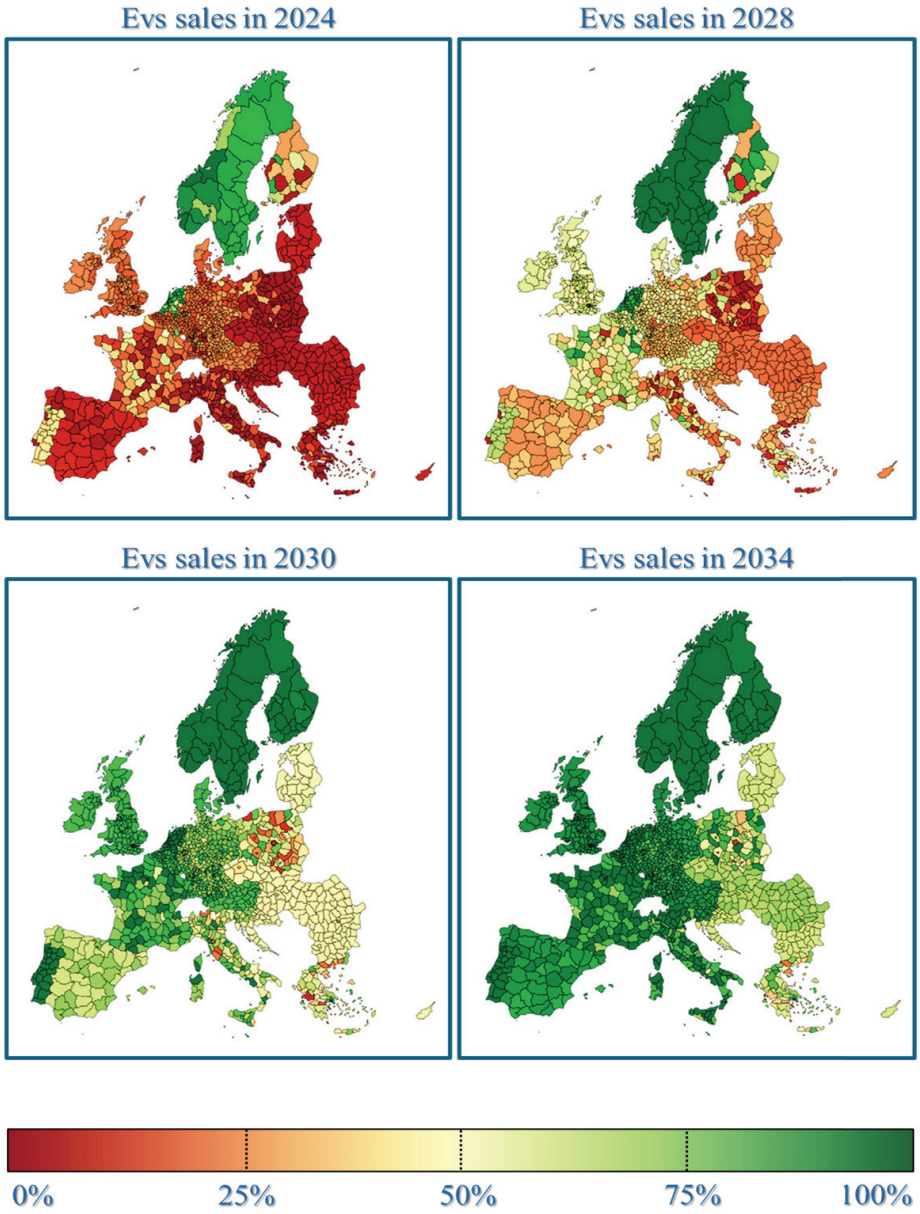


Figure 1. European map of EV sales clustered by how evolved countries are in the transition to net-zero-emission CO_2 passenger car sales [13].

However, one of the critical challenges in this domain is achieving an ideal balance between cost-effectiveness and performance [14]. Batteries, as one of the highest-cost components in EVs [15], play a crucial role in determining both the efficiency and overall economics of these vehicles. An accurate estimation of the battery's State of Charge (SOC) is essential for reducing design costs [16], optimizing vehicle efficiency, and improving overall system reliability [17]. This makes advanced battery management system (BMS) softwares indispensable (e.g., MATLAB/Simulink framework, or Modelica framework), for the precise estimation of SOC [18]. Currently, SOC estimation methods for EV/HV battery

systems are broadly classified into direct and indirect approaches [19,20], as depicted in Figure 2.

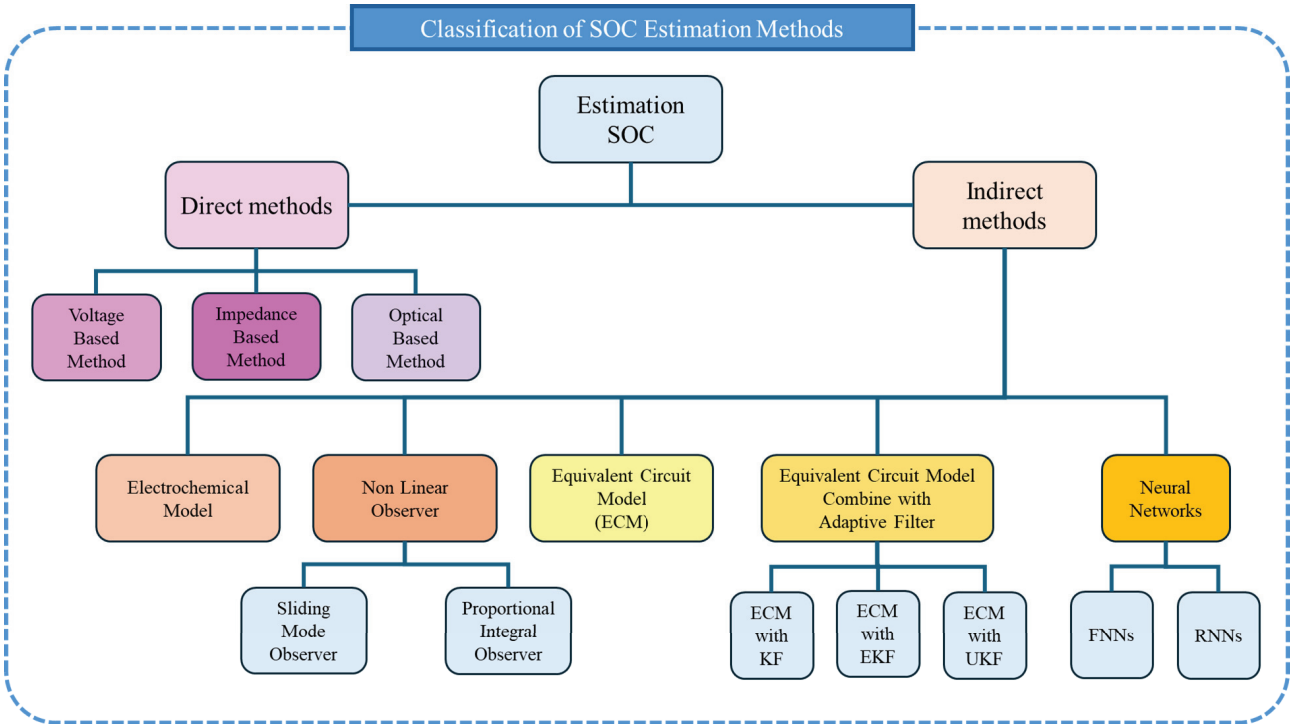


Figure 2. Summary diagram of the major SOC estimators used in the literature.

Traditional methods for SOC estimation, such as open-circuit voltage (OCV) measurement [21] and coulomb counting [22], are commonly employed due to their simplicity. However, these methods are often limited by sensor inaccuracies and the inherent uncertainty associated with battery behavior under various operational conditions [23]. To address these limitations, more advanced techniques have been developed, combining equivalent circuit models (ECMs) with Kalman Filters to enhance SOC estimation accuracy [24]. These hybrid methods typically require extensive battery testing to model and calibrate parameters, which can be resource-intensive [25]. In recent years, machine learning has introduced transformative potential to SOC estimation by enabling data-driven approaches that leverage the vast amounts of data generated by modern battery systems. Through representation learning and deep learning, ANN techniques can capture complex patterns in battery data that are challenging to model with conventional physics-based methods [26]. Voltage, current, and temperature signals, among other factors, can be analyzed using ANNs to provide indirect but highly accurate SOC estimates, an approach that aligns well with the requirements of real-time EV applications [27]. Despite the advantages of ANN-based SOC estimation, challenges remain due to the nonlinear nature of batteries, where factors such as degradation, temperature variations, and dynamic load conditions can introduce significant variability in SOC predictions [28,29]. Ensuring accurate SOC estimation is essential not only for optimizing vehicle performance but also for improving passenger safety, enhancing driving comfort, and reducing costs associated with battery overdesign or oversizing [30]. This paper provides a comprehensive overview of recent advancements in ANN-based SOC estimation methods for electric vehicles. By analyzing a range of machine learning algorithms, we aim to highlight the strengths and limitations of these techniques, discussing their applicability and potential to transform EV battery management. For those interested in foundational machine learning concepts, additional resources are available that provide a deeper understanding of these techniques and their

underlying principles. As battery technology continues to evolve rapidly, the availability of extensive datasets for battery performance optimization has grown substantially. These data, encompassing time-series measurements of voltage, current, and temperature, allow for more shaded SOC estimation models, which are critical for the real-time management of energy in EV batteries [31].

2. Methods Analyzed

In this overview, we examine and categorize the main machine learning methods for estimating the battery State of Charge (SOC) in Electric Vehicles (EVs). Before diving into specific approaches, we first address the rationale for selecting machine learning techniques over traditional methods based on Kalman Filters. As illustrated below in Figure 3, Kalman Filters rely on detailed knowledge of the battery's model, requiring several engineering steps to achieve accurate SOC estimation.

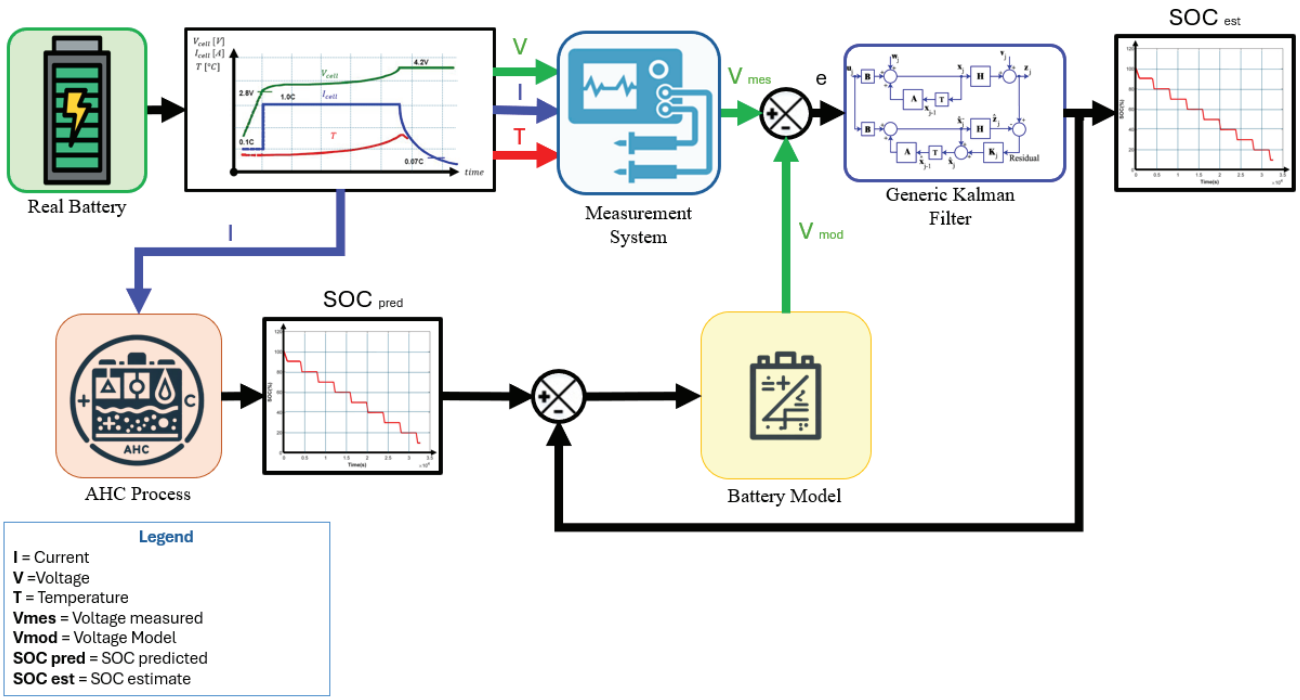


Figure 3. Generic model for estimated SOC with KF.

In contrast, machine learning offers a more streamlined and direct path to reliable results, using datasets built from real battery measurements or Digital Twin (DT) models, as show in Figures 4 and 5. The use of DT models to train neural networks has demonstrated considerable success not only in the electric vehicle industry but also across various applications, such as energy efficiency monitoring in luxury yachts, hotel energy management, autonomous navigation for automotive vehicles, and other robotic systems [32–35]. A key advantage of this approach is its ability to produce a well-trained neural network from DT model outputs, even in early development stages when real-world data are limited. This flexibility makes machine learning particularly valuable in initial project phases, allowing for robust SOC estimation from minimal measurement data and broadening its applicability across different sectors [36].

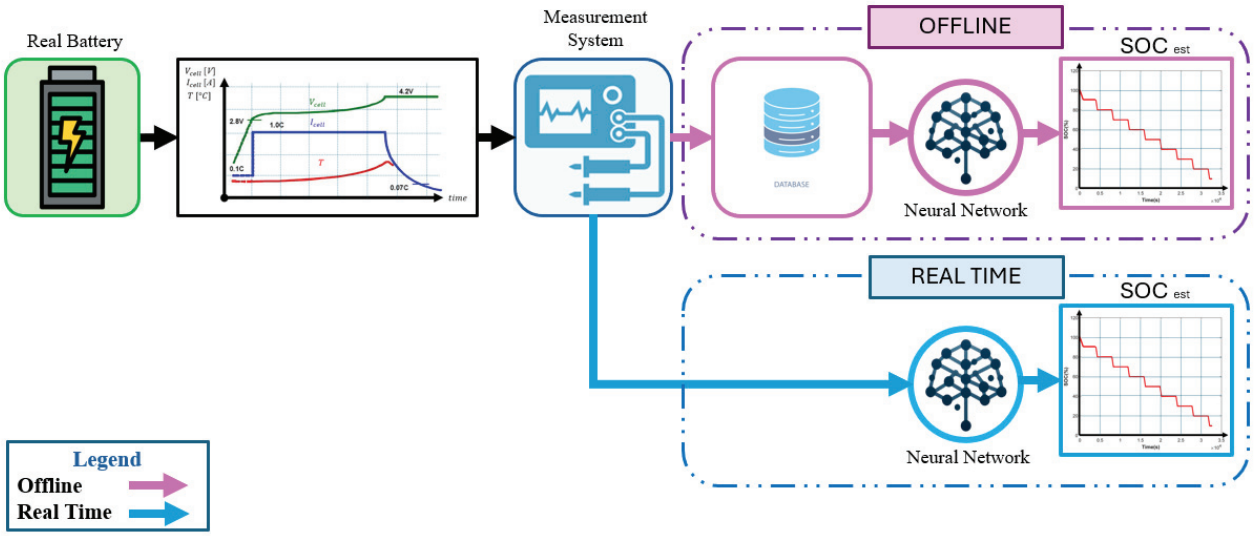


Figure 4. Generic model for estimated SOC with ML with dataset collected by real measurement of battery.

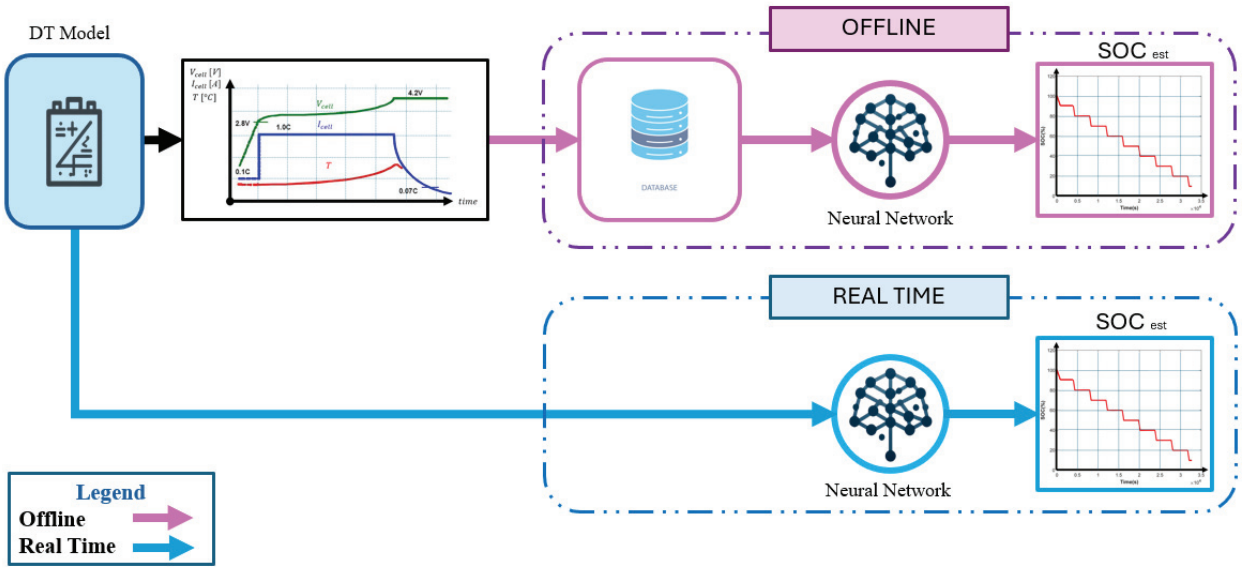


Figure 5. Generic model for estimated SOC with ML with dataset collected by DT model.

In this section of the article, we provide an in-depth analysis of recent machine learning methods developed for addressing the challenge with the State of Charge (SOC) estimation. We present a concise overview of these approaches, highlighting the conceptual frameworks that underpin how each method functions as proposed in the literature. To facilitate a clear understanding of the techniques, we categorize them into two main types:

- Feedforward Neural Networks (FNNs);
- Recurrent Neural Networks (RNNs).

This division enables us to compare their unique architectures, advantages, and limitations in capturing battery dynamics, as well as their suitability for real-time SOC estimation in electric vehicle applications. In this study, we chose to focus exclusively on data-driven approaches based on Feedforward Neural Networks (FNNs) and Recurrent Neural Networks (RNNs) due to their proven effectiveness and suitability for the State of Charge (SOC) estimation. These methods stand out for their ability to model complex nonlinear

relationships between inputs and SOC without relying on detailed and often computationally intensive battery models. This capability makes them particularly well-suited for handling the dynamic and nonlinear behavior inherent in battery systems. Moreover, FNNs and RNNs are highly advantageous for embedded applications in electric vehicles. Their computational efficiency and relatively low memory requirements allow seamless integration into the resource-constrained environments of vehicle systems. This efficiency not only enables real-time SOC estimation but also ensures that these methods can operate effectively under the hardware limitations typical of embedded systems. The rapid advancements and growing interest in FNNs and RNNs within the field further motivated their selection for analysis. These neural network architectures have demonstrated a strong potential to enhance the performance, adaptability, and robustness of battery management systems. By focusing on these approaches, our study aims to contribute to the understanding and development of SOC estimation techniques that are both practical and scalable for real-world applications. Each method's structure and operational principles are examined, with attention given to how they handle data patterns and temporal dependencies inherent in SOC estimation. This categorization will also help to clarify the decision-making process in selecting appropriate neural network types, considering the specific demands of battery state monitoring and predictive accuracy. To ensure a fair and consistent comparison of SOC estimation methods, such as the machine learning-based approaches detailed in this paper, it is crucial to evaluate algorithms under uniform conditions. This involves using comparable datasets, aligning the number of trainable parameters, and applying consistent training and testing methodologies. By following these guidelines, meaningful insights into the relative strengths and limitations of various approaches can be obtained. Specifically, we recommend adhering to the following three principles when evaluating ANNs' SOC estimation techniques:

1. Consistent Quality Metrics: Utilize identical or comparable quality indices across studies to ensure that the evaluation criteria are uniform and allow for objective comparisons. These metrics could include Mean Absolute Error (MAE), root mean squared error (RMSE), or accuracy percentages, as referenced in the literature;
2. Dataset Transparency: Clearly state the dataset used for training and testing, specifying whether it is a standard publicly available dataset or a custom dataset obtained from specific battery configurations. This helps in understanding the results;
3. Algorithm Complexity: Analyze the computational complexity of the proposed methods, highlighting their advantages and drawbacks. This includes evaluating processing time, memory requirements, and the feasibility of real-time implementation on embedded systems;
4. Different Approach: we chose articles that presented a particular method that differed in the interesting way in which they addressed the problem of SOC estimation.

Integrating these principles into the evaluation process provides a structured framework for assessing SOC estimation methodologies. This approach is particularly relevant when comparing the diverse machine learning models discussed earlier, including FNNs, LSTMs, BiLSTMs, and GRUs. By aligning the evaluation standards, the trade-offs between accuracy, computational efficiency, and practical applicability can be better understood, enabling the selection of the most suitable method for real-world deployment.

3. Feedforward Neural Networks (FNNs) for Battery SOC Estimation

Feedforward Neural Networks (FNNs) are powerful tools for implementing complex non-linear mappings in systems with multiple inputs and outputs [37]. Their flexibility and simplicity make them particularly valuable in fields such as battery state estimation, where precise, data-driven models are essential for accurate predictions [38]. In this context, we

focus on a straightforward FNN structure featuring a single hidden layer, an approach commonly used in state estimation due to its simplicity and reliability [39]. In designing an FNN, the choice of activation function $f_{\text{activ}}(x)$ plays a crucial role in defining the network's output behavior. A popular option is the hyperbolic tangent function, defined as

$$f_{\text{activ}}(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tag{1}$$

which bounds the output within the range $[-1, 1]$, creating smooth and stable transitions that often facilitate efficient training. Another interesting choice of activation function is the sigmoid function, especially in binary classification problems, as it maps input values to a range of $(0, 1)$. While similar to tanh, it does not provide zero-centered outputs, which may slow down training in certain networks. Its mathematical expression is

$$f_{\text{activ}}(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Alternatively, the Rectified Linear Unit (ReLU) activation function is widely adopted for its computational efficiency and handling of sparse activations. ReLU is defined as

$$f_{\text{activ}}(x) = \text{ReLU}(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0. \end{cases} \tag{3}$$

The ReLU function's ability to set all negative inputs to zero introduces a piecewise linear structure that enables faster learning and helps to mitigate the vanishing gradient problem, a common issue in neural networks. The choice of activation function can impact both the network's ability to learn from data and the computational efficiency during training [40,41]. For instance, tanh and sigmoid activations are bounded, which can be beneficial for controlling output ranges, but they may suffer from saturation effects, leading to slower gradient updates. On the other hand, ReLU mitigates the vanishing gradient problem, especially in deep networks, but can cause "dead neurons" (neurons that stop updating if they consistently output zero) [42]. The training of an FNN involves optimizing its weights W and biases b to minimize the loss function, commonly defined as the sum of squared errors:

$$L = \frac{1}{2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \tag{4}$$

where y_i is the actual target value, $\hat{y}_i$ is the predicted output from the network, and n is the number of training samples. Backpropagation, a widely used algorithm for this purpose, calculates the partial derivatives of the loss function relative to each weight and bias, iteratively updating them to reduce L. The number of training iterations, or epochs, is often used to establish a stopping condition, though additional criteria, such as improvement in error reduction, can also determine when training should conclude [43]. A generic architecture of FNNs is shown in Figure 6.

For applications like the State of Charge (SOC) estimation in electric vehicle batteries, hybrid approaches that integrate an Equivalent Circuit Model (ECM) with an FNN have proven effective [20]. Traditionally, battery SOC estimation relies on lookup tables (LUTs) to correlate SOC with open-circuit voltage (OCV). However, using an FNN to model this correlation offers greater flexibility and potentially higher accuracy [44]. FNN uses OCV as an input, processes it through a hidden layer with m neurons, and outputs the SOC estimate. The ECM complements this model by estimating OCV from the battery's terminal voltage, V_{terminal}, thus enhancing the robustness of the SOC prediction. This method is schematically represented in Figure 7.

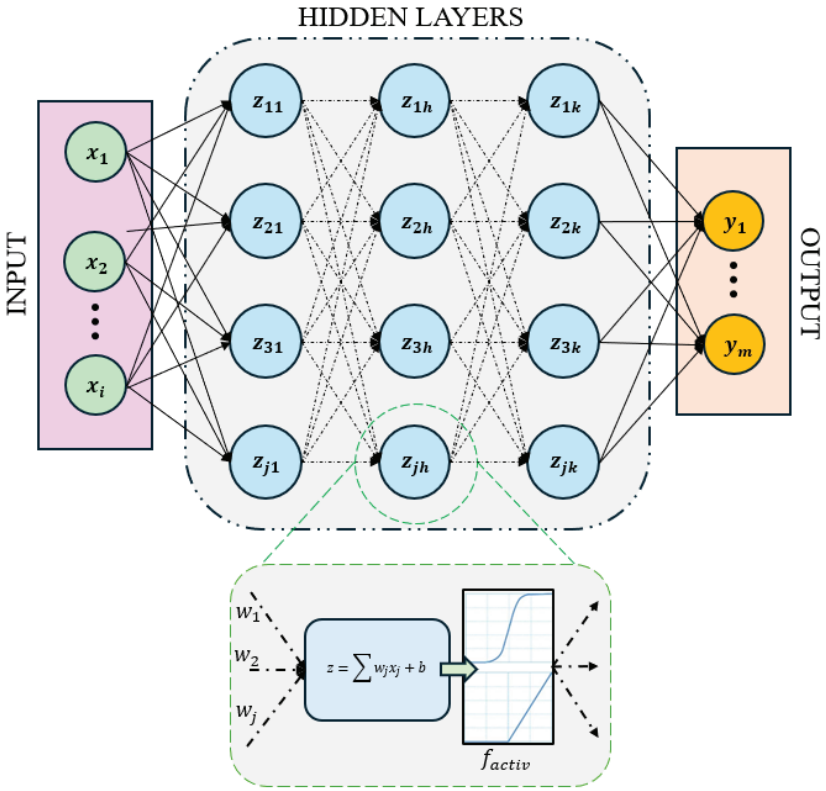


Figure 6. Generic structure of FNNs.

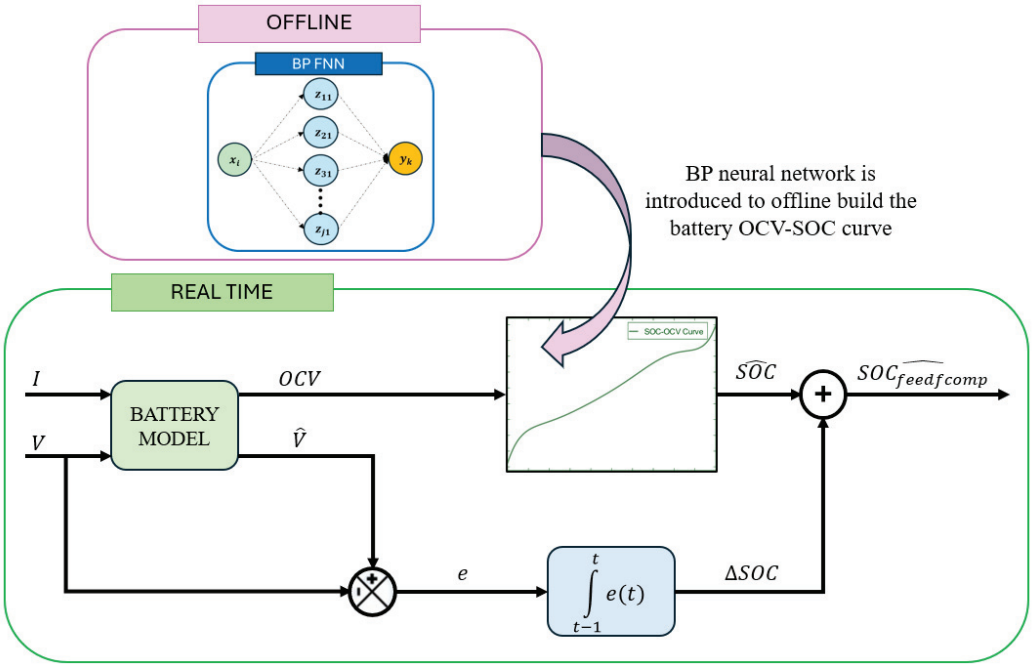


Figure 7. BP neural network introduced to offline build the battery OCV–SOC curve for SOC estimation.

Further improvements in SOC estimation accuracy have been achieved by integrating the FNN with a Kalman Filter [45], specifically an Unscented Kalman Filter (UKF), for Lithium Iron Phosphate (LFP) batteries [46]. In this hybrid model, the FNN generates initial SOC estimates, which the UKF then refines. The FNN operates as a supervised machine learning algorithm, comparing its predictions to known SOC values to compute estimation errors. For training, datasets collected from automotive homologation driving cycles, such

as US06, FUDS, and the Dynamic Stress Test (DST) specified by the U.S. Advanced Battery Consortium, provide diverse operating conditions. Despite the relatively narrow training dataset, the UKF effectively reduces error, increasing the model's accuracy even when applied to new data. This approach can be illustrated by the following Figure 8.

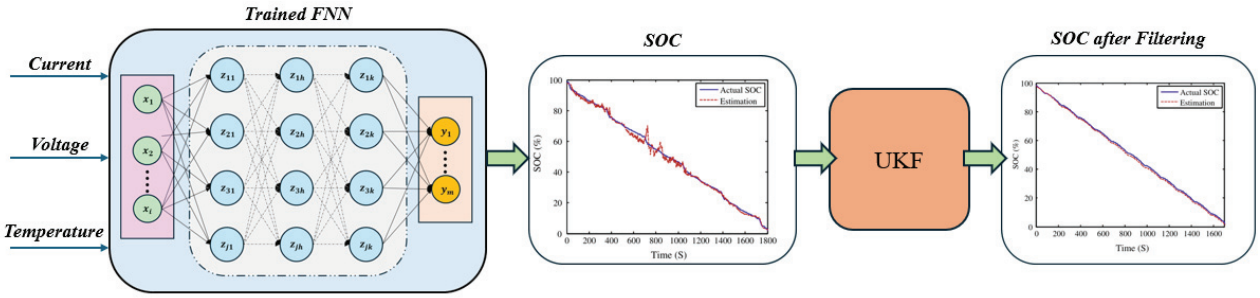


Figure 8. SOC estimation scheme based on FNN and Unscented Kalman Filter (UKF).

Recent studies have demonstrated the potential of Feedforward Neural Networks (FNNs) for estimating the State of Charge (SOC) in hybrid energy storage systems without relying on Kalman Filters [47–50]. One prominent work [51] explored the application of an FNN in a 12V hybrid energy storage configuration that combines a Lithium Iron Phosphate (LFP) battery and a lead–acid battery, both used to power a belt starter–generator system. This system allows the electric machine to operate as either a motor or generator, following a control strategy that cycles the Li-ion battery within a specific SOC range. In this setup, an FNN was developed to estimate the SOC of both batteries simultaneously, leveraging a single network structure. This method can be summarized as in Figure 9.

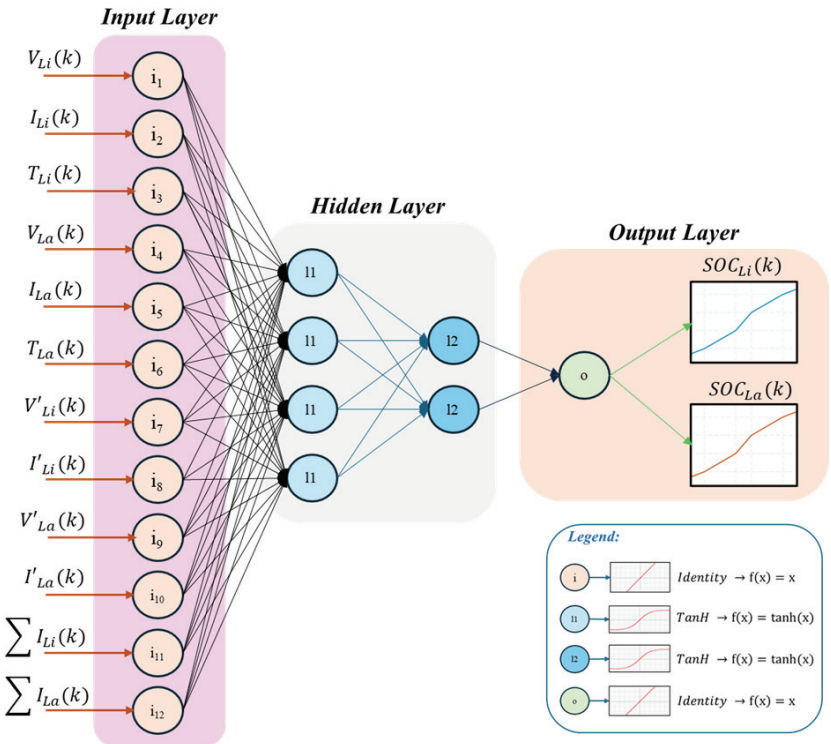


Figure 9. Architecture of BPNN with 4-layer network and 12 inputs, 6 from each battery, for estimating SOC from LFP and LA batteries.

Despite this approach, further research is needed to compare this dual-output FNN with separate single-output models, in order to clarify the advantages of a shared neural

network structure for SOC estimation [52,53]. To address the complexities of accurately predicting SOC—given the influence of various battery parameters like SOC, State of Health (SOH), current, and temperature—some researchers [48,54] have explored using battery polarization states as inputs to the FNN model, as shown in Figure 10.

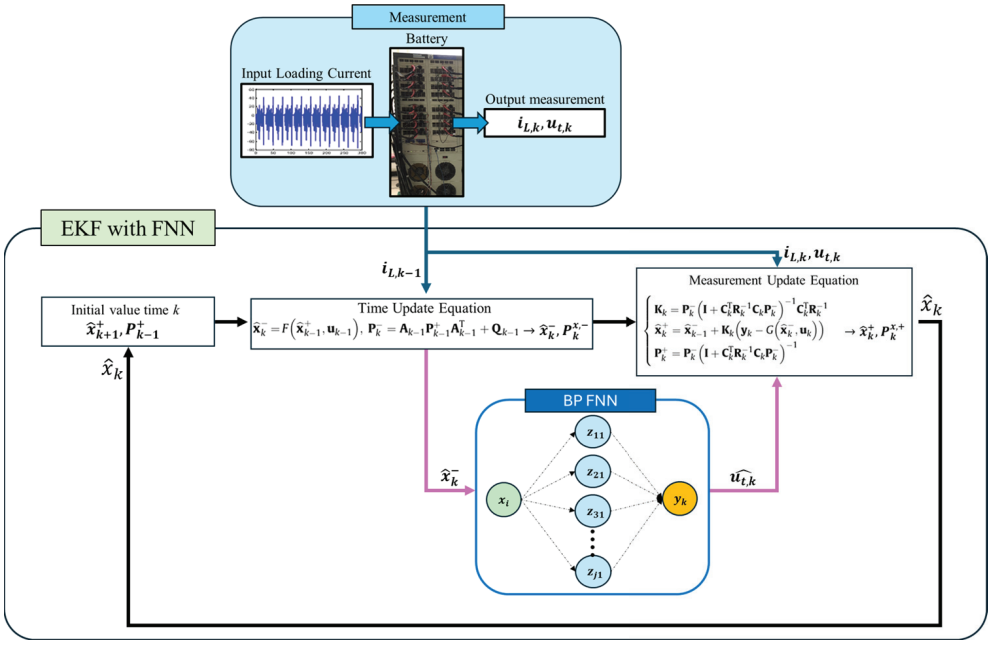


Figure 10. Structure of state estimation algorithm SOC with combined FNN and EKF.

Different polarization state time constants, denoted as τ, can vary (e.g., from 0 to 1000) to account for distinct dynamic behaviors. Four polarization states with specific τ values were chosen, alongside the battery current, voltage, and temperature as inputs to the FNN, as shown in Figure 11.

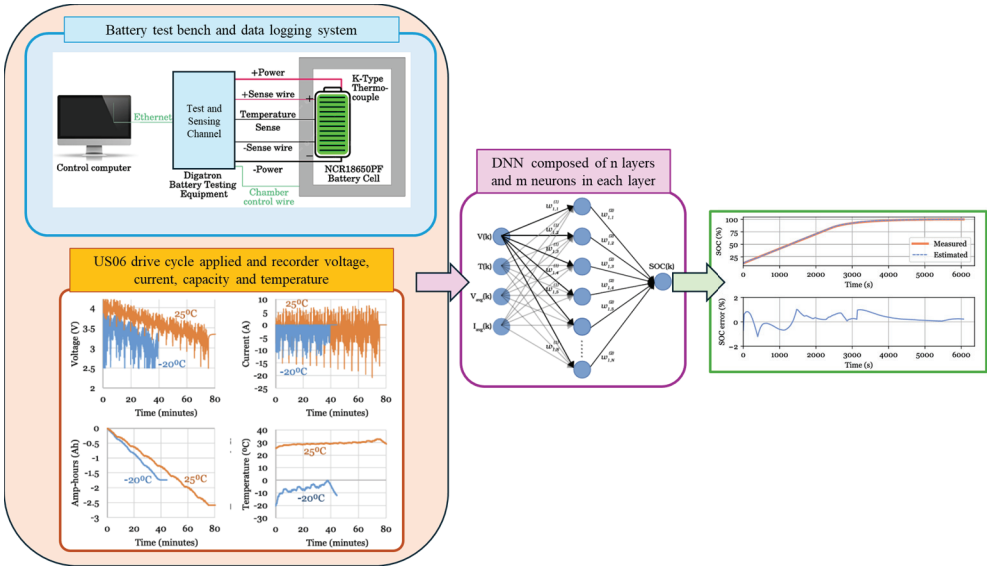


Figure 11. Experimental data collection for training and validation of neural network for SOC estimation with polarization and temperature.

The network was trained and validated on data from various driving cycles—EUDC, HL07, HWFET, and NEDC—each tested under ambient temperatures ranging from −10 °C to 25 °C. This data division allocated 80% for training and the remaining 20% for validation and testing, with additional testing conducted on a Hardware-in-the-Loop (HIL) setup to

evaluate real-world applicability. This study further demonstrated the FNN's capability of accurately estimating SOC across a range of temperatures, including low-temperature conditions down to −20 °C. Although FNNs do not inherently capture time-dependent information, temporal dependencies can be indirectly modeled by creating new input features based on moving averages of the battery's terminal voltage and current. This approach enabled results comparable to those of Recurrent Neural Networks (RNNs), which are specifically designed for sequential data [55–57]. Through a systematic evaluation of FNN architectures—including variations in neuron counts, hidden layers, and averaging window lengths (100 and 400 timesteps)—the study found that, at 25 °C, the optimal results were obtained using a 400-timestep rolling window. As expected, performance degraded at −20 °C due to the adverse effects of low temperatures on Li-ion batteries. The authors hypothesized that model accuracy in these conditions could be improved with a larger dataset or a more complex FNN capable of capturing the intricate dynamics at low temperatures. Another noteworthy finding is the significant improvement in model robustness and accuracy, up to 41%, achievable by using augmented datasets. Additionally, internal resistance data—measured in a laboratory setting along with battery voltage, current, and temperature—were incorporated to train and test the FNN for SOC estimation. Although the direct measurement of internal resistance in vehicles poses practical challenges, it is feasible to integrate a model-based estimate of internal resistance onboard to provide real-time input for both SOC and SOH estimation. In refining FNN structures, optimization algorithms have shown great potential in streamlining the setup process by identifying the optimal FNN structure, thus minimizing reliance on extensive engineering experience to set parameters effectively [58–60]. While initial optimization efforts typically focus on neuron count and learning rates, additional parameters—such as the number of hidden layers and the initialization of weight distributions—could further enhance accuracy. However, considering these factors also introduces added complexity and an increased offline computational burden when searching for optimal structures. An interesting extension of FNN design is presented in [61], where a unique model architecture composed of three parallel FNNs was developed, as summarized in Figure 12.

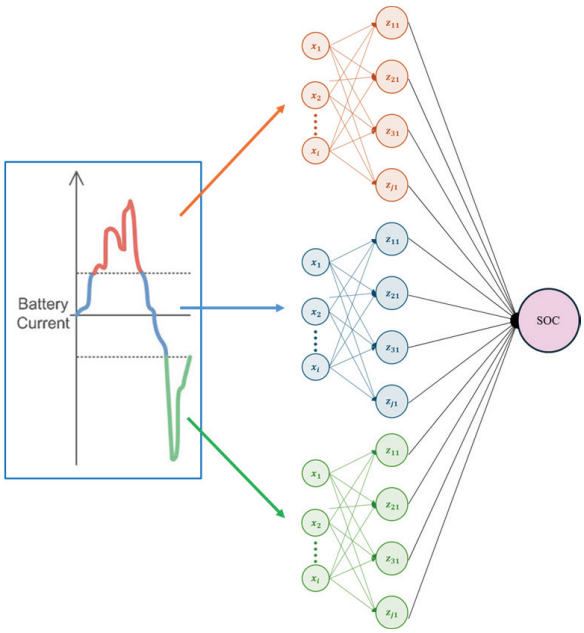


Figure 12. Schematic of three-classificatios Feedforward Neural Networks for battery SOC estimation based on battery current.

Each FNN was individually trained on distinct operational data corresponding to the three primary operating modes of idling, charging, and discharging. Although the accuracy of this model was limited by the available data (with training based on a pulsed profile and validation on the US06 dataset), the study addressed the influence of randomly initialized network weights by performing multiple training iterations. By averaging the resulting error across these training runs, the authors obtained a final error measurement, accounting for variations caused by initial weight values. This insight underscores the impact of initialization in neural network training, a factor often underexplored in the SOC estimation literature despite its importance in determining convergence to desirable local minima. In another study, the authors in [62] present an advanced method for estimating battery State of Charge (SOC) by integrating a Backpropagation Neural Network (BPNN) with the Backtracking Search Algorithm (BSA). The methodology begins by collecting three key input parameters—voltage, current, and temperature—from DST and FUDS drive cycles conducted at three different temperatures (0 °C, 25 °C, and 45 °C). These raw measurements are initially processed through a moving-average low-pass FIR filter to smooth the data, and then normalized to prepare for subsequent analysis. Once preprocessed, the normalized inputs are fed into the BPNN, which leverages its capacity to model complex, nonlinear relationships between the measured parameters and the battery's SOC to produce an initial estimate. However, given that neural networks are prone to becoming trapped in local minima due to suboptimal initial weights and biases, the approach enhances performance by incorporating the BSA. In the optimization stage, BSA fine-tunes the network by determining the optimal number of neurons in the hidden layer and the ideal learning rate, all based on the minimization of the root mean squared error (RMSE). The entire procedure is structured into four distinct stages. In Stage I, raw data are collected, filtered to reduce noise, and normalized. Stage II uses these preprocessed data within the BPNN to generate an initial SOC estimation. Stage III applies BSA to optimize the hidden layer configuration and learning rate, thereby refining the network's performance. Finally, Stage IV focuses on training and validating the optimized model with designated training and testing datasets, culminating in the computation of the final SOC and its associated estimation error. To further validate the superiority of the proposed BPNN–BSA model, a comparative study was conducted against alternative approaches, including RBFNN–BSA, GRNN–BSA, and ELM–BSA models. In each case, current, voltage, and temperature served as inputs, and their hyperparameters were similarly tuned using BSA. Evaluation metrics such as RMSE, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and SOC error were calculated for both DST and FUDS cycles across the three temperature settings. For instance, during the DST cycle, the BPNN–BSA model achieved RMSE values of 1.47%, 0.81%, and 0.48% at 0°C, 25 °C, and 45 °C, respectively, while, for the FUDS cycle, RMSE values were 1.74%, 0.91%, and 0.57%. These results indicate significant improvements; at 25 °C, the DST cycle RMSE was reduced by 34%, 62%, and 56% compared to the RBFNN-,BSA, GRNN-,BSA, and ELM-,BSA models, respectively. Similar enhancements were observed for the FUDS cycle, with the proposed method consistently yielding lower error metrics across all temperatures. Beyond RMSE, the proposed model demonstrated notable reductions in MAE and MAPE. For example, the DST cycle MAE reached as low as 0.32% at 45 °C, reflecting improvements of 43%, 68%, and 63% over the alternative methods at 25 °C. Similarly, for the FUDS cycle, the MAE at 0 °C was 0.87%, showing significant decreases compared to higher temperatures and outperforming the competing models by substantial margins. Furthermore, the MAPE for the DST cycle at 25 °C was 7.15% for the proposed model—considerably lower than the corresponding values for the other methods—while the SOC error remained minimal, underscoring the enhanced accuracy and robustness of the BPNN–BSA approach. Overall,

integrating BPNN with BSA not only optimizes the neural network's parameters but also delivers superior performance in terms of accuracy and robustness across varying operating conditions and temperature profiles. Despite the increased computational cost inherent in such optimization, the improved precision in SOC estimation and the model's ability to generalize across different scenarios make the BPNN–BSA approach a highly effective solution for battery management in electric vehicles and energy storage systems. This method with the BSA algorithm is summarized in Figure 13.

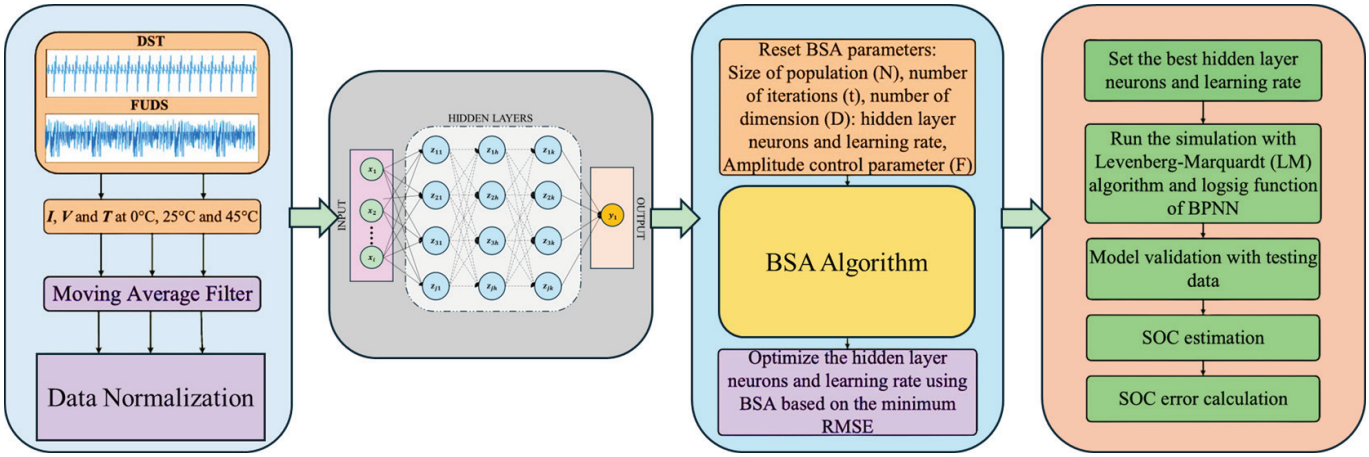


Figure 13. Schematic diagram of SOC estimation with FNN combined with BSA algorithm.

An intriguing recent study presents a highly accurate estimator for the State of Charge (SOC) of lithium-ion batteries by meticulously integrating their underlying physical behavior into the modeling process. The authors introduce a sophisticated hybrid architecture that fuses a physics-based model—as detailed in [63]—with a data-driven deep Feedforward Neural Network (FNN) for battery cells, as described in [64,65]. This innovative framework is implemented within the paradigm of Physics-Informed Neural Networks (PINNs), which uniquely combine traditional neural network architectures with the governing physical laws of the system. Unlike conventional data-driven methods, PINNs inherently respect the electrochemical principles and dynamic processes that govern battery operation, thereby yielding predictions that are not only accurate and reliable but also physically plausible, even in cases where experimental data are limited or contaminated with noise. Central to this methodology is a composite loss function that seamlessly integrates two components. The first is a data fidelity term, which minimizes the discrepancy between the network's predictions and the observed measurements, mirroring the approach taken by conventional neural networks. The second, and more distinctive, component, is the physics fidelity term, which quantifies deviations from established physical laws—such as charge conservation and the kinetics of battery charge and discharge—by directly incorporating the governing differential equations, along with their boundary and initial conditions, into the loss function. This design ensures that the network, which accepts critical input variables such as current, voltage, and temperature, produces SOC predictions that remain consistent with the underlying physics across a broad spectrum of operating conditions. The experimental data used to train the PINN model were gathered under rigorously controlled laboratory conditions. Experiments were conducted at ambient temperatures ranging from −10 °C to 25 °C, thereby capturing the battery's performance across a diverse thermal spectrum. A series of power profiles were generated for an electric vehicle by simulating standard drive cycles—including UDDS, US06, LA92, and HWFET—as well as "Mix" cycles composed of randomized segments of these standard profiles. During these experiments, the battery was housed within a climate chamber where ambient temperature was systematically varied,

and key parameters such as voltage, current, battery temperature, and ampere-hours were continuously recorded. This comprehensive dataset, generated under varying dynamic loads and temperature conditions, provided the necessary depth and diversity to effectively train the PINN model, enabling it to accurately estimate SOC across all tested thermal conditions. To further validate the approach, the study conducted a rigorous comparison between the proposed PINN model, an Adaptive Kalman Filter (AKF), and a conventional FNN. Twelve distinct neural network setups were trained and evaluated using the Li-ion battery dataset. Notably, the battery's behavior at −10 °C posed significant challenges due to substantial voltage deviations under load, heightened nonlinearity, and increased internal resistance. Despite these hurdles, the PINN model demonstrated exceptional performance, with error metrics—such as RMSE, MAE, and maximum error—remaining impressively low. Even at −10 °C, the RMSE was maintained below 2.85% (approximately 0.028310), and the maximum error, indicative of estimation variance, stayed low even as the SOC dropped to 20% (around 0.119825). These results underscore the PINN's adeptness at capturing the intricate, nonlinear dynamics and resistance characteristics under extreme conditions, thereby affirming its reliability and robustness. Furthermore, embedding physical laws within the neural network enhances its interpretability, a significant advantage over conventional "black-box" models. By elucidating the internal mechanisms that govern battery behavior, this approach not only bolsters confidence in the model's predictions but also facilitates the identification and resolution of potential issues under novel or extreme operating scenarios. Although the computational demands are higher—due to the necessity of calculating derivatives and enforcing additional constraints during training—the substantial benefits in terms of prediction accuracy, robustness, and real-time applicability render the PINN-based approach a highly promising solution for advanced battery management systems in electric vehicles. By elegantly merging data-driven techniques with a profound understanding of battery physics, this innovative method represents a significant advancement over traditional SOC estimation techniques, offering a robust, adaptable, and interpretable tool for navigating the complexities inherent in lithium-ion battery behavior. This particular architecture is shown in Figure 14.

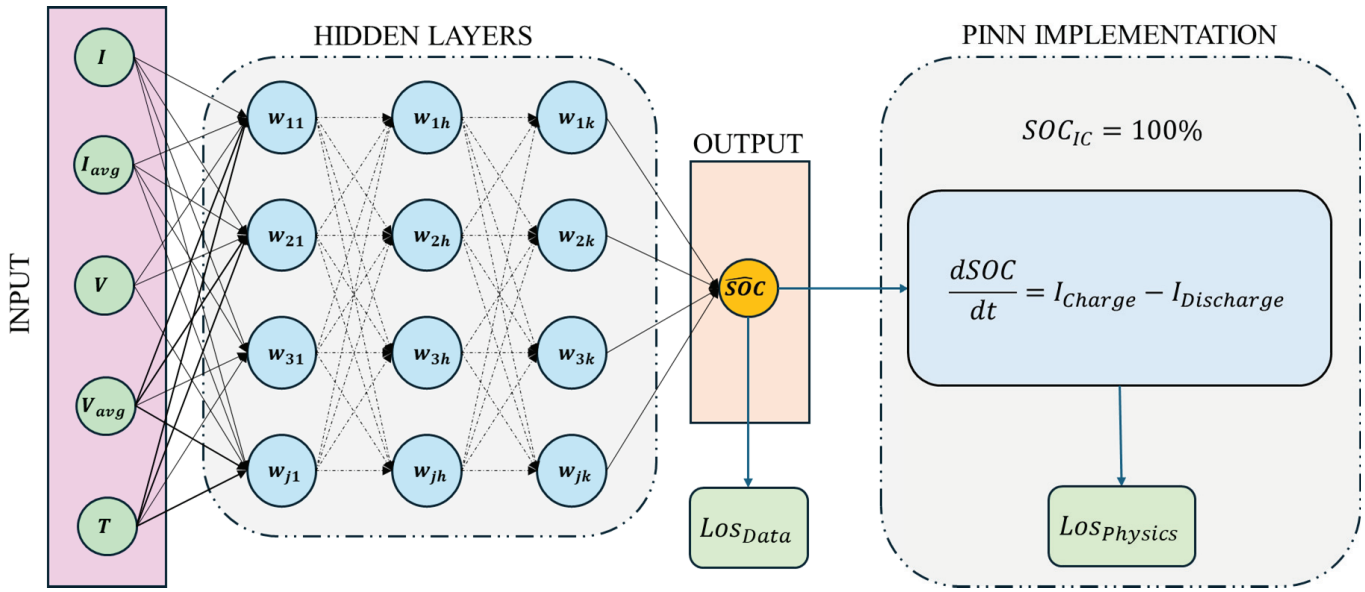


Figure 14. FNNs with PINNs architecture for battery SOC.

Finally, there is another interesting study [66], in which an innovative hybrid method is presented that significantly enhances battery state estimation by integrating the traditional Equivalent Circuit Model (ECM) with Feedforward Neural Networks. The approach

commences with a first-order ECM that characterizes the battery's dynamic behavior using fundamental components such as ohmic resistance, polarization resistance, and polarization capacitance. The initial parameters are determined online via a Recursive Least Squares algorithm with a forgetting factor, ensuring robust calibration against real-world measurements. However, under extreme conditions—such as freezing temperatures or rapid charging and discharging—the ECM alone proves insufficient to capture the battery's pronounced nonlinear behavior. To overcome these limitations, three Feedforward Neural Networks are embedded within the ECM as virtual electronic components, each dedicated to correcting the residual errors in estimating a specific parameter. These networks are trained offline using an iterative strategy that incorporates the battery's state-space equations into a physics-informed loss function. This loss function, derived from the mean squared error between the predicted and measured terminal voltages, directs the backpropagation process to fine-tune the network parameters. The evaluation of the hybrid model is based on a real-world battery dataset, which comprises charging and discharging measurements—such as current, voltage, and cell temperature—collected from a brand-new Panasonic 18650PF cell under diverse operating conditions. To thoroughly assess the model's performance, the authors selected three different operating modes (HPPC, US06, and HWFET) and conducted tests under four ambient temperatures (-20 °C, -10 °C, 0 °C, and 10 °C), resulting in a total of 12 scenarios. To further validate the model's accuracy, the predicted terminal voltage was computed using both the first-order ECM and the proposed hybrid model with the identified parameters. Under the HPPC mode, the traditional ECM demonstrated high accuracy during the initial stages of discharging; however, its estimation error progressively increased due to error accumulation from the recursive FFRLS method and its limited adaptability. In contrast, the hybrid model maintained consistently high accuracy throughout the discharging process, as the neural network modules dynamically compensated for the parameter estimation biases with their superior nonlinear-fitting capabilities. Under the US06 and HWFET operating conditions, the traditional ECM exhibited considerable fluctuations and larger estimation errors, whereas the hybrid model closely tracked the terminal voltage without introducing extraneous noise. This enhanced performance is attributed to the FNN modules' ability to effectively learn the rapid variations in voltage and current characteristic of these conditions. Notably, when discharging at -20 °C—a scenario in which the battery's electrochemical behavior becomes exceedingly complex—the traditional ECM produced the highest errors across all operating modes, while the hybrid model continued to deliver satisfactory accuracy. Following the offline training phase, the refined parameters are deployed for real-time State of Charge (SOC) estimation using an Extended Kalman Filter (EKF). When applied to the dataset, the hybrid model achieved a reduction in SOC estimation errors, ranging from 29% to 64% across various conditions, demonstrating marked improvements in accuracy, adaptability, and robustness compared to conventional ECM approaches. By merging the interpretability of physics-based models with the advanced nonlinear learning capabilities of neural networks, this method offers a precise and dependable solution for battery state estimation in complex and challenging environments. This approach can be summarized as in Figure 15.

Based on the studies analyzed in this section, it is evident that with the integration of advanced optimization techniques and the use of augmented datasets, Feedforward Neural Networks (FNNs) emerge as a highly promising solution for the next generation of battery management systems, particularly in practical and real-world applications [67]. However, in spite of the benefits of what we have seen in this section, FFNs suffer from problems related to the explosion or vanishing of the descending gradient [68], failing always, and especially with complex systems, to handle the correlations that can occur between distant

time instants. To solve these problems mentioned above, the use of Recurrent Neural Networks is often a resort, which we analyze in the next section of the article.

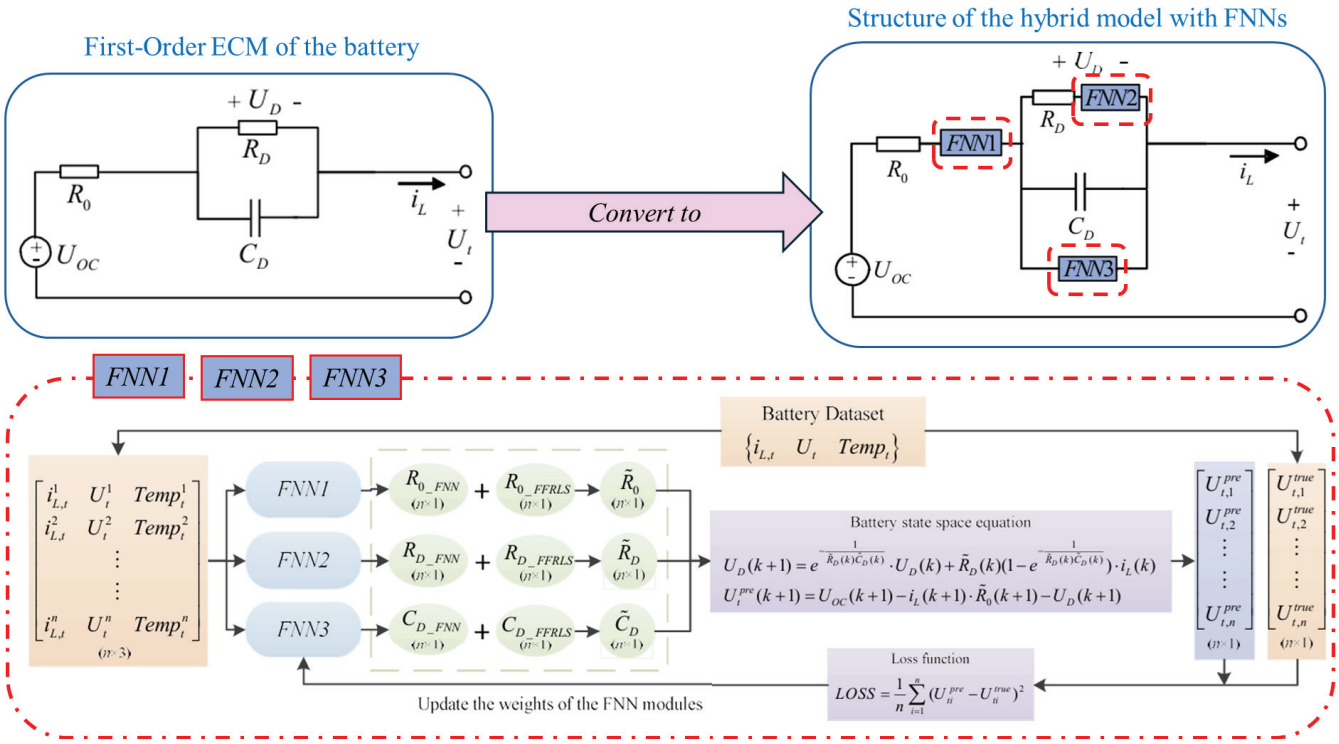


Figure 15. Hybrid ECM approach where we use 3 FNNs for resistance and capacity of battery model.

4. Recurrent Neural Networks (RNNs) for Battery SOC Estimation

In this section, we introduce various types of Recurrent Neural Networks (RNNs) and examine key findings from the recent literature on their application in estimating the State of Charge (SOC) of batteries. RNN-based machine learning models, which retain and leverage past information, show particular promise for such tasks [69,70]. Unlike traditional neural networks, RNNs incorporate temporal dependencies by feeding previous outputs or intermediate states back as inputs [71]. For instance, the SOC from time step $k-1$ can be fed as an input at time k. This architecture is well-suited for capturing short-term sequence dependencies, though it may struggle with the long-term dependencies typically seen in battery systems. However, training RNNs can be challenging due to issues such as vanishing or exploding gradients during backpropagation, which can hinder the learning of long-term dependencies [72]. To address these limitations, several RNN variants have been developed, including Long Short-Term Memory (LSTM) [73], Bidirectional LSTM (BiLSTM) [74], and Gated Recurrent Unit (GRU) [75]. These architectures use specialized gating mechanisms to manage and leverage dependencies over longer sequences, making them suitable for tasks involving extensive sequential data, such as speech recognition and battery SOC estimation [69,76,77].

4.1. LSTM Neural Network

The LSTM network introduces a memory cell c_k to store and manage information over time. It uses three primary gates:

- Forget Gate f_k: Controls how much information from the previous memory cell c_{k-1} should be retained.

$$f_k = \sigma(W_f \cdot [h_{k-1}, x_k] + b_f)$$

- Input Gate i_k: Regulates how much of the new information will be added to the memory cell.

$$i_k = \sigma(W_i \cdot [h_{k-1}, x_k] + b_i)$$

- Output Gate o_k: Determines how much of the information from the memory cell c_k should be used in the output h_k.

$$o_k = \sigma(W_o \cdot [h_{k-1}, x_k] + b_o)$$

These gates work together to update the memory cell and produce the final output:

$$\tilde{c}_k = \tanh(W_c \cdot [h_{k-1}, x_k] + b_c)$$

$$c_k = f_k \cdot c_{k-1} + i_k \cdot \tilde{c}_k$$

$$h_k = o_k \cdot \tanh(c_k)$$

This setup enables the LSTM to selectively retain or discard information, allowing it to capture long-term dependencies [78]. An example of a generic architecture of an LSTM is shown in Figure 16.

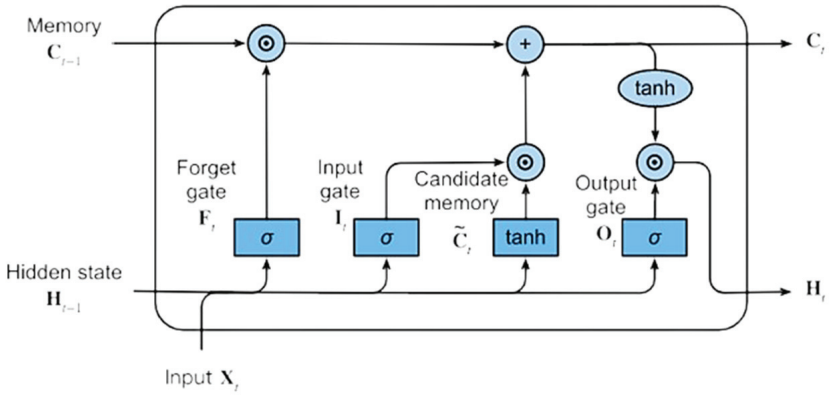


Figure 16. LSTM generic architecture.

In [79], the authors applied a Long Short-Term Memory (LSTM) network to estimate the State of Charge (SOC) using only directly measured battery signals, such as terminal voltage, load current, and ambient temperature. This approach does not rely on additional methods or estimation filters, simplifying the overall implementation. A notable outcome of this work is the LSTM's ability to accurately estimate SOC across varying ambient temperatures. This capability represents a significant advantage over traditional methods that rely on lookup tables (LUTs), which typically require separate LUTs to be created for each temperature range [80]. However, it is important to highlight that the effectiveness of this approach depends on the quality of the dataset used to train the LSTM. Specifically, the dataset must contain sufficient information to encode the temperature variations within the model parameters, which may necessitate a large and comprehensive dataset. For their study, the authors utilized a dataset from a Panasonic 18650PF Li-ion battery cell, acquired at multiple ambient temperatures. This dataset is publicly available for download from [81]. Another interesting solution to improve the use of LSTM neural networks, using its ability to store what happens in the short- to long-term observation time, is the combination of LSTM with other neural networks [82–84].The idea is to use a hybrid CNN–LSTM network to estimate the State of Charge (SOC) of lithium-ion batteries by capturing their complex nonlinear dynamics using voltage, current, and temperature measurements. The method leverages spatial features extracted from the immediate relationships among the

measured variables—including their averages—as well as temporal dependencies derived from historical data. The network begins by processing time-sequenced input data with a one-dimensional convolution layer that transforms raw signals into high-level representations emphasizing internal correlations. These features are then passed to an LSTM network that effectively integrates historical context over both short and long periods. This combined approach enhances performance beyond that of a traditional LSTM alone. To simulate real-world electric vehicle battery loading behaviors, the DST–FUDS–US06 (DFU) profile—an integration of the dynamic stress test (DST), federal urban driving schedule (FUDS), and US06 test profiles—was employed. During discharge, the DFU profile was repeatedly applied until full battery discharge, and tests were conducted under various temperature conditions (0 °C, 10 °C, 20 °C, 30 °C, 40 °C, 50 °C, and room temperature) to build comprehensive testing datasets. More specifically, the network was trained with DST data (8438 samples), FUDS data (8390 samples), and US06 data (7987 samples) at room temperature, while its online SOC estimation performance was evaluated using DFU data (8350 samples) at room temperature. The input vector $x_k = [I_k, , V_k, , T_k, , Iavg_k, , Vavg_k]$ is mapped to the output $y_k = [SOC_k]$. Estimation results at room temperature are presented in Section IV-A, with additional results under varying temperature conditions in Section IV-B. During training, although increasing the number of epochs generally enhances model accuracy, it also extends training time. To determine an optimal number of epochs, the root mean squared error (RMSE) for both training and testing datasets was monitored. The RMSE dropped below 4% after 2000 epochs and stabilized around 2% after 6200 epochs, with the best performance achieved between 8000 and 11,000 epochs. Consequently, 10,000 epochs were selected for training. For performance comparison, the proposed network was benchmarked against a standalone LSTM (i.e., without the convolutional layer) and a CNN comprising three convolutional layers (each with six filters and zero padding to preserve input length). All networks were trained for 10,000 epochs, with training times of 161 min for the proposed network, 231 min for the LSTM, and 102 min for the CNN. On a laptop, the average computation times per time step were 0.098 ms, 0.116 ms, and 0.082 ms for the proposed, LSTM, and CNN networks, respectively. When the battery starts at 100% SOC, the CNN network operating independently of past inputs produces highly fluctuating estimates, whereas both the LSTM and the proposed CNN–LSTM deliver much smoother and more accurate predictions. The proposed network consistently maintains estimation errors within 2%, while the standalone LSTM can sometimes exceed 4%. Although both networks yield satisfactory results, the CNN–LSTM demonstrates slightly better accuracy. Since the initial battery SOC is not always known a priori, robustness against unknown initial conditions is critical. To simulate this, data with lower initial SOC values were generated (for example, by removing data with SOC greater than 80% to simulate an 80% initial SOC). When the SOC starts at 80%, both the LSTM and the proposed networks exhibit a two-stage behavior. Initially, the effect of the unknown state is dominant, and the proposed network takes slightly longer than the LSTM to track the true SOC. Once convergence is achieved, however, the proposed network exhibits lower and more consistent errors, with overall RMSE and mean absolute error (MAE) values of 1.35% and 0.87% compared to 1.43% and 0.95% for the LSTM. When the initial SOC is set to 60%, the proposed network converges to the true SOC faster than the LSTM and subsequently provides significantly more stable and accurate estimates (with an RMSE of 0.92% and MAE of 0.48% versus 2.97% and 2.03% for the LSTM). Although the CNN network is less affected by unknown initial conditions, its inability to model temporal dependencies results in inferior overall performance. Overall, this integrated CNN–LSTM design effectively combines local and temporal feature extraction, adapts to environmental

variations and uncertain initial conditions, and delivers superior performance in estimating the battery's State of Charge. This type of hybrid approach is schematized in Figure 17.

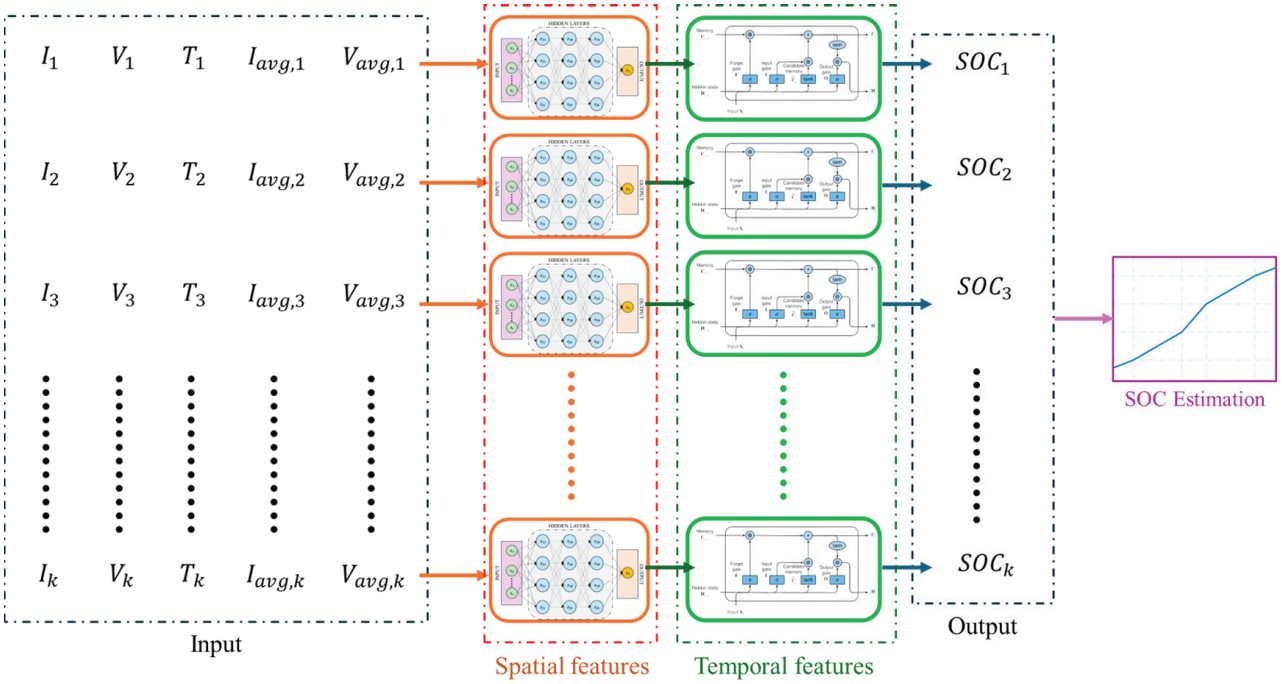


Figure 17. Architecture of combined LSTM with another RNN.

The recent works [85] presents the EI–LSTM–CO, an improved LSTM–RNN model for SOC estimation. By incorporating a sliding window average voltage into the input, the model enhances its ability to capture nonlinear battery behavior. Additionally, a state flow strategy based on ampere-hour integration (AhI) constrains output variations, resulting in smoother and more accurate SOC predictions. This particular LSTM Neural Network can be schematized using the following diagram (Figure 18).

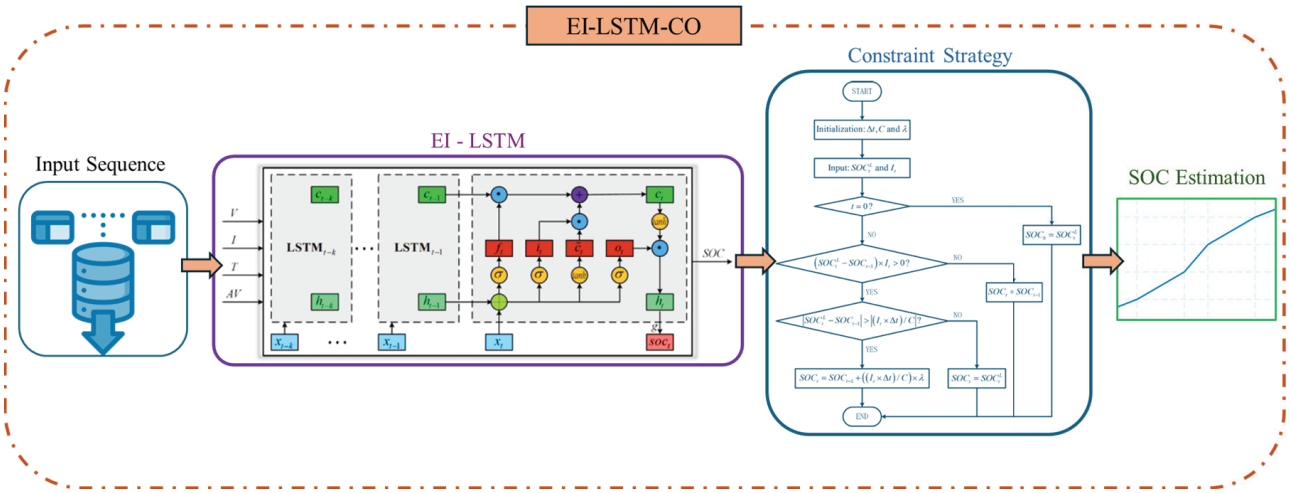


Figure 18. Framework of the SOC estimation using the algorithm EI–LSTM–CO.

An interesting recent work [86] presents a novel approach for battery State of Charge (SOC) estimation—a critical factor for electric vehicle safety and range prediction—using an LSTM neural network optimized via a random search algorithm (RS-LSTM), as depicted in Figure 19. Unlike most SOC studies that rely on manual hyperparameter tuning, which is often random and fails to achieve a globally optimal solution, this method automates

the optimization process. Few studies have attempted automatic hyperparameter tuning; those that have generally lack sufficient validation in real-world scenarios and do not adequately account for the complexities of actual vehicle operation. Real driving conditions introduce numerous measurable battery parameters and including less relevant features can compromise both model accuracy and interpretability. To simulate the diverse conditions encountered by lithium-ion batteries in practical use, the study utilizes a dataset from the University of Maryland (CALCE) based on the INR 18650-20R battery model. The CALCE team conducted a series of dynamic tests under varying temperatures (0 °C, 25 °C, and 45 °C) and initial SOC conditions (80% and 50%) using datasets from the Dynamic Stress Test (DST) and the Federal Urban Driving Schedule (FUDS). To address the challenges associated with manual hyperparameter tuning, the authors first apply a Random Forest algorithm for dimensionality reduction, selecting the most relevant input features (discharge capacity and discharge energy) to reduce noise and prevent overfitting. Building on this foundation, the RS-LSTM method optimizes the LSTM network's hyperparameters—namely, look back, epochs, batch size, and learning rate—within a carefully designed search space tailored to the battery data. This process yields optimal settings of 45 for look back, 177 epochs, a batch size of 64, and a learning rate of 0.0026. The study further compares the feature selection results from the Random Forest algorithm with those derived from the Pearson correlation coefficient to identify the best combination of model inputs from both linear and nonlinear perspectives. Following dimensionality reduction and hyperparameter optimization, the optimized LSTM network is evaluated using standard metrics such as Mean Absolute Error (MAE), root mean squared error (RMSE), and the coefficient of determination (R^2). Experimental validation under various SOC intervals, temperature conditions, and even scenarios with added Gaussian noise confirms the method's superiority, effectiveness, and robustness. Furthermore, to assess the efficiency of the proposed method under complex driving dynamics—including frequent charging, discharging, and pause states—the study incorporates real-world vehicle data collected from three vehicles between 1 and 15 April 2022. The SOC estimation results for both training and validation sets indicate that all vehicles exhibit excellent performance, demonstrating that RS-LSTM effectively captures vehicle-specific SOC change patterns. Although some deviations occur during vehicle pause states—likely due to insufficient training data—the overall performance remains robust. Notably, Vehicle C shows exceptional results (with an R^2 of 0.998, MAE of 0.871, and RMSE of 1.218 in the training set, and similar metrics in the validation set), while Vehicles A and B display only minor increases in error. The RS-LSTM method significantly outperforms alternative approaches such as TCN–LSTM, CNN–BiGRU, CNN, and standard LSTM networks, achieving an MAE of 0.221%, an RMSE of 0.262%, and an R^2 of 0.997. By automating hyperparameter tuning and employing robust feature selection, this integrated approach produces a model that is both highly accurate and stable, overcoming the limitations of manual tuning. The authors recommend that future research further adapt the model to the complex nature of real-world vehicle data and develop more adaptive feature selection methods to dynamically respond to evolving data characteristics. This innovative approach not only represents a significant advancement for battery management systems in electric vehicles but also holds potential for broader applications, such as in portable electronic devices and supercapacitors.

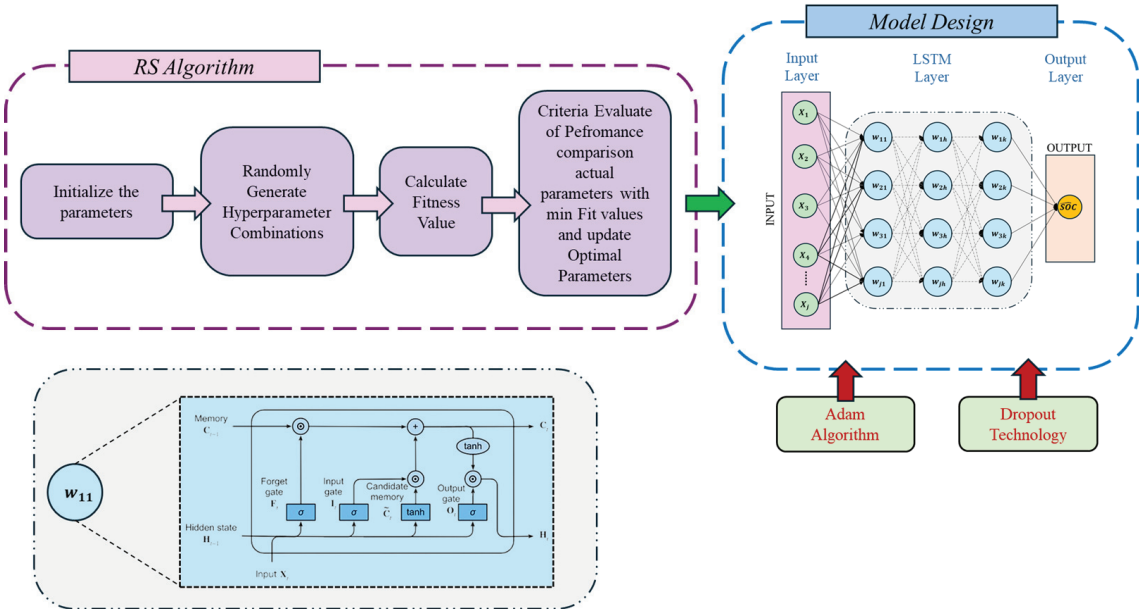


Figure 19. Flowchart for optimization of LSTM neural network by RS random search algorithm.

The use of LSTM networks comes with challenges, such as their unidirectional nature, which limits them to processing data only from past to present, potentially losing valuable future context. They are also prone to overfitting, especially with limited or noisy datasets, and optimizing parameters like look-back length and learning rate can be computationally expensive and time-consuming. Additionally, LSTMs may struggle with capturing complex long-range dependencies and require significant computational resources for processing long sequences. To partially solve these problems, Bidirectional LSTM (BiLSTM) networks address these limitations by processing data in both forward and backward directions, allowing them to capture both past and future contexts [87]. This bidirectional approach enhances their ability to handle noisy or incomplete data, model complex dependencies more effectively, and achieve greater accuracy in tasks like SOC estimation [88]. While BiLSTMs require more computational resources than standard LSTMs, their improved robustness and precision make them a valuable tool for applications requiring detailed sequential data analysis. In the next section of this article, we will discuss this recurrent neural network, highlighting some interesting applications and variants for estimating SOC.

4.2. Bidirectional LSTM Neural Networks

The Bidirectional Long Short-Term Memory (BiLSTM) network is an extension of the traditional LSTM architecture that processes data in both forward and backward directions. By combining the outputs from two LSTMs—one processing the input sequence in its natural order and the other processing it in reverse—BiLSTMs can capture both past (previous time steps) and future (upcoming time steps) context for every point in the sequence [89]. This bidirectional architecture makes BiLSTMs particularly useful for applications where the context from both ends of a sequence significantly impacts predictions, such as text translation, speech recognition, and battery state estimation tasks like State of Charge (SOC) [90] and remaining useful life estimation [91].

- Forward LSTM: Processes the input sequence $x = [x_1, x_2, \ldots, x_T]$ from the first time step $t = 1$ to the last $t = T$, generating a sequence of hidden states.

$$\overrightarrow{h}_t = \text{LSTM}_{\text{forward}}(x_t, \overrightarrow{h}_{t-1})$$

Here, $\overrightarrow{h}_t$ represents the hidden state at time t for the forward LSTM;

- Backward LSTM: Processes the same input sequence in reverse, from $t = T$ back to $t = 1$, generating another sequence of hidden states.

$$\overleftarrow{h}_t = \text{LSTM}_{\text{backward}}(x_t, \overleftarrow{h}_{t+1})$$

$\overleftarrow{h}_t$ is the hidden state at time t for the backward LSTM;

- Combined Output: The output of the BiLSTM at each time step t is a combination of the hidden states from the forward and backward LSTMs. The combination is typically achieved using concatenation, summation, or averaging:

$$h_t = f\left(\overrightarrow{h}_t, \overleftarrow{h}_t\right)$$

where $f(\cdot)$ is a function that combines the two hidden states, such as

  - Concatenation: $h_t = \begin{bmatrix} \overrightarrow{h}_t \\ \overleftarrow{h}_t \end{bmatrix}$
  - Summation: $h_t = \overrightarrow{h}_t + \overleftarrow{h}_t$

Key equations for BiLSTM results including the following:

- Forward LSTM Hidden State:

$$\overrightarrow{h}_t = \tanh\left(W_f x_t + U_f \overrightarrow{h}_{t-1} + b_f\right)$$

where W_f, U_f, and b_f are the weights and biases for the forward LSTM;

- Backward LSTM Hidden State:

$$\overleftarrow{h}_t = \tanh\left(W_b x_t + U_b \overleftarrow{h}_{t+1} + b_b\right)$$

where W_b, U_b, and b_b are the weights and biases for the backward LSTM;

- Output Combination:

$$h_t = f\left(\overrightarrow{h}_t, \overleftarrow{h}_t\right)$$

Depending on the task, the choice of f (concatenation, summation, or averaging) can vary.

This dual-direction approach allows BiLSTM to capture temporal or contextual dependencies from both ends of the sequence, making it ideal for applications such as text translation, where context from both the start and end of a phrase impacts the final interpretation. An example of this architecture is shown in Figure 20.

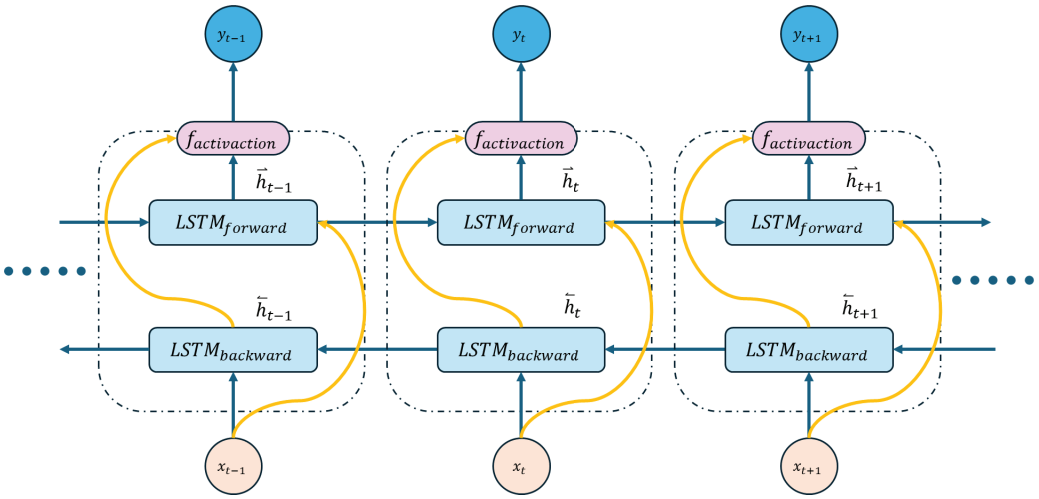


Figure 20. BiLSTM generic architecture.

Many interesting works, such as [92,93], use the concept of BiLSTM and introduce a hybrid deep learning architecture, combining a Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) to enhance the estimation of the State of Charge (SOC) in lithium-ion batteries, critical for electric vehicle efficiency. These models introduce an innovative method for estimating the State of Charge (SOC) in lithium-ion batteries used in electric vehicles by combining spatial and temporal feature analysis within a hybrid deep learning framework. Accurate SOC estimation requires capturing spatial features—which represent the interrelationships among current inputs—and temporal features, which reflect the correlations between the present SOC and historical data. To achieve this, the authors propose a hybrid CNN–BiLSTM network. The one-dimensional Convolutional Neural Network (CNN) extracts advanced spatial features from raw data by adjusting convolution kernel weights and window widths, effectively generating multiple features that serve as inputs for the subsequent recurrent model. The bidirectional long short-term memory (BiLSTM) network then models the temporal dynamics by learning from both past and future dependencies, thereby capturing the evolution of battery behavior over time. A key innovation of this work is the integration of the Group Learning Algorithm (GLA) to automatically optimize the hyperparameters of the CNN–BiLSTM network. By tuning parameters such as the learning rate, the number of LSTM units, and the number of bidirectional layers, GLA enhances the model's predictability and stability while reducing the computational burden and eliminating the need for manual configuration. This strategic approach improves the model's efficiency and adaptability to diverse operating conditions without the risk of overfitting or underfitting. The deep neural network is trained using the mean squared error (MSE) loss function, and its performance is evaluated with metrics including R^2, relative error, normalized RMSE, RMSE, and Mean Absolute Error (MAE). The experiments were conducted on an AMD Ryzen 5 5500U using MATLAB 2022, with data collected from cylindrical INR21700-40T Li-ion batteries provided by the CALCE Research Group. The batteries were discharged under standard industry procedures at temperatures of 25 °C, 0 °C, and 45 °C. Analysis of the data revealed strong positive associations between SOC, voltage, and discharge capacity, as well as the significant influence of urban driving cycles and environmental conditions on battery performance. Evaluations across different temperatures demonstrated that the proposed GLA–CNN–BiLSTM model consistently outperforms alternative approaches (such as CNN, CNN–LSTM, GLA–LSTM, and GLA–CNN) in terms of accuracy and stability. At 0 °C, the model produced significantly lower error metrics than the alternatives, while, at 25 °C and 45 °C, it maintained high prediction accuracy with minimal errors. Further performance assessments, including the Granger causality test, confirmed the model's superior ability to capture the intricate dynamics of battery behavior, achieving remarkably low error rates that surpass state-of-the-art methods. This work presents a robust and efficient methodology for SOC estimation in electric vehicles. By effectively integrating spatial and temporal feature extraction with an innovative hyperparameter tuning strategy via GLA, the hybrid model significantly enhances the accuracy, reliability, and adaptability of battery SOC estimation in real-world applications. .These methods based on GLA–CNN–BiLSTM are summarized in Figure 21.

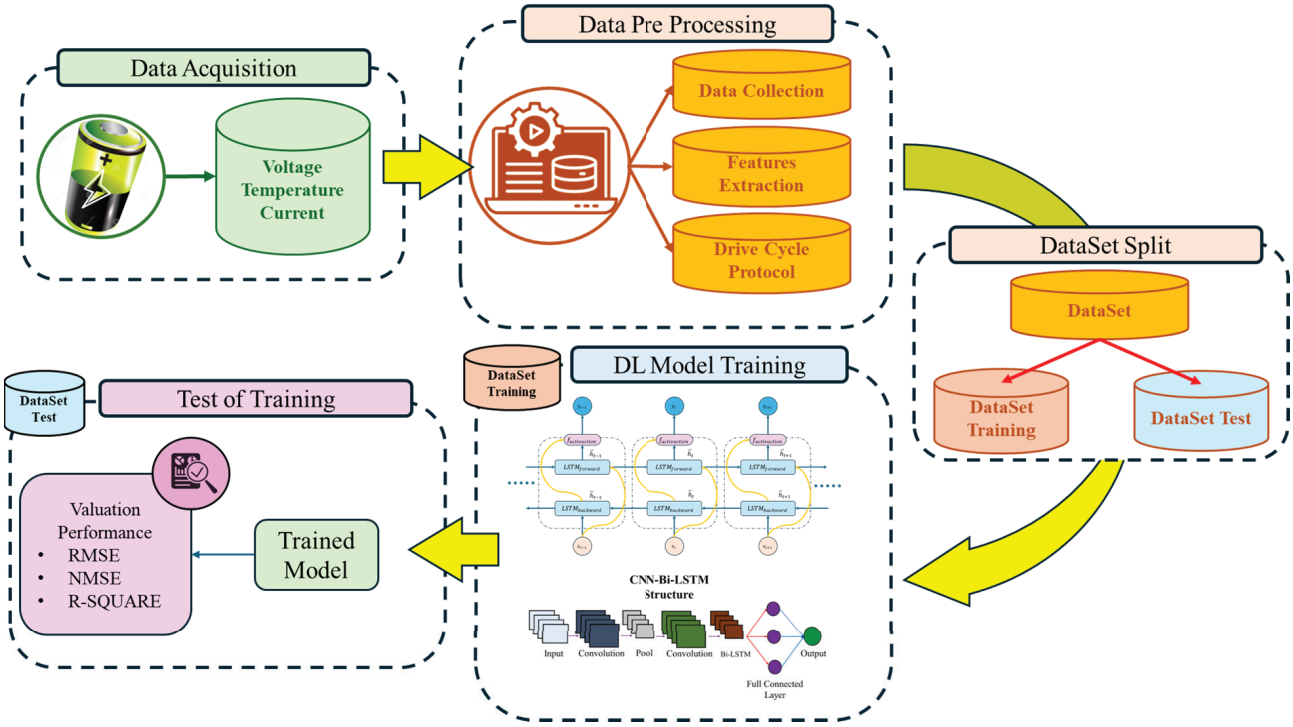


Figure 21. Schematic structure of GLA–CNN–BiLSTM deep learning model.

An interesting enhancement of bidirectional LSTM networks is presented in the work [94], where the authors integrate an attention mechanism to improve performance and introduce multi-dimensional temperature compensation to address variations in temperature. This approach describes an advanced method for estimating the State of Charge (SOC) of lithium-ion batteries using a temperature-compensated Bidirectional LSTM (BiLSTM) network enhanced with an integrated attention mechanism. This approach improves feature extraction by weighting the hidden outputs of multiple LSTM cells, which allows the network to selectively focus on the most relevant inputs in complex time-series battery data. In this method, spatial attention is applied during the encoding phase via a convolutional neural network that learns high-dimensional representations, ensuring that the most critical features are emphasized. During decoding, temporal attention is used to dynamically assign weights across time steps, effectively capturing essential temporal information for accurate SOC prediction. Temperature effects are addressed through a multidimensional compensation strategy. A polynomial model expresses the open-circuit voltage (OCV) as a function of both SOC and temperature, while the reference SOC is computed using ampere-hour integration. This integration of temperature data not only compensates for temperature-induced variations but also improves error convergence during training. Experimental validation was performed using a laboratory test bench that includes a battery testing system, an electrochemical workstation, a host computer, and a thermostat. Lithium-ion cells with a $LiNiCoAlO_2$ cathode, graphite anode, and solid polymer electrolyte (rated at 3.7 V and 2.6 Ah) were tested at temperatures of 0 °C, 10 °C, 25 °C, and 40 °C under two conditions: a dynamic stress test (DST) and an urban dynamometer driving schedule (UDDS). Under DST conditions at 25 °C with a discharge period of 13,500 s sampled every 10 s (yielding 1350 data points), the initial high-quality data allowed for accurate predictions by the LSTM. However, bidirectional errors introduced fluctuations when the dataset was small. The temperature-compensated model with integrated attention (AMBiLSTM TC) effectively mitigated these errors, maintaining stable and lower prediction errors throughout the discharge. Under UDDS conditions, where data are sampled every second (yielding 13,500 data points) and the temperature rises

rapidly within the first 120 s, the AMBiLSTM TC again outperformed standard methods by significantly reducing the root mean squared error (RMSE). Further analysis during training showed that the introduction of attention mechanisms led to rapid error convergence, with an optimal configuration found at 16 hidden units, balancing improved accuracy with only a modest increase in computational time. By integrating both spatial and temporal attention mechanisms with multidimensional temperature compensation and validating the approach through extensive experimental testing, the proposed method markedly improves feature extraction from time-series data, resulting in a robust and accurate SOC estimation under a wide range of operating conditions while maintaining computational efficiency. This technique is summarized in Figure 22.

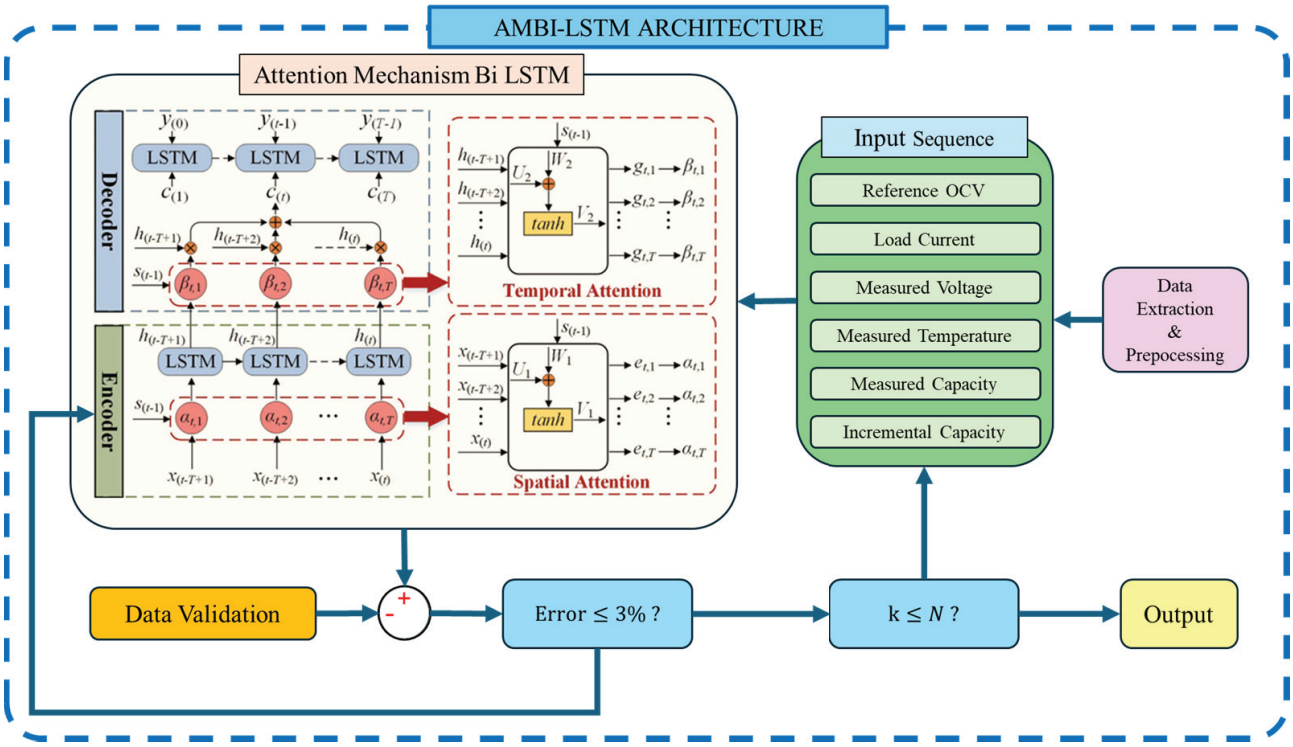


Figure 22. Framework of SOC estimation based on the AMBiLSTM architecture.

Experimental results show that the proposed method achieves maximum errors of less than 1% and improves accuracy by 9.39% and 22.36% under two different current conditions. Furthermore, in the context of battery health prognosis, the approach enhances accuracy by 21.45% compared to standard bidirectional LSTM networks, demonstrating its significant potential for applications in battery management systems. However, BiLSTMs require a sufficient sequence length before outputting predictions, which can introduce computational limits.

4.3. Gated Recurrent Unit Neural Networks

Another effective RNN variant for managing long-term dependencies is the Gated Recurrent Unit (GRU). The GRU employs a simplified gating mechanism compared to the LSTM, using a single gate to control both memory retention and update decisions [95–97]. The Gated Recurrent Unit (GRU) is a simplified variant of the Long Short-Term Memory (LSTM) architecture, designed to address the exploding and vanishing gradient problem in recurrent neural networks (RNNs) while reducing computational complexity [98,99]. An example of this particular architecture is shown in Figure 23.

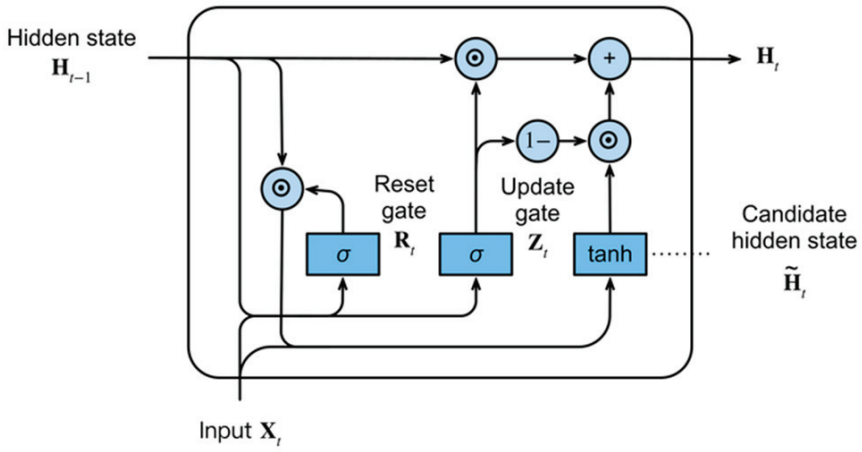


Figure 23. GRU generic architecture.

Unlike the LSTM, which uses separate forget and input gates, the GRU combines these into a single update gate. This design reduces the number of parameters and simplifies training, while still enabling the model to capture long-term dependencies in sequential data [100,101]. The GRU consists of two main gates:

- Reset Gate (r_k): Controls how much of the past information to forget;
- Update Gate (z_k): Determines how much of the previous state to retain and how much to update with new information.

The core equations governing the GRU are as follows:

1. Update Gate:
$$z_k = \sigma(W_z x_k + U_z h_{k-1} + b_z)$$
where x_k is the input vector at time step k, h_{k-1} is the hidden state from the previous time step, W_z and U_z are weight matrices, b_z is the bias term, and σ is the sigmoid activation function;
2. Reset Gate:
$$r_k = \sigma(W_r x_k + U_r h_{k-1} + b_r)$$
Here, W_r, U_r, and b_r are the corresponding weight matrices and bias for the reset gate;
3. Candidate Hidden State: The reset gate determines how much of the past hidden state contributes to the candidate hidden state.
$$\tilde{h}_k = \tanh(W_h x_k + U_h(r_k \odot h_{k-1}) + b_h)$$
where $\odot$ denotes the element-wise multiplication, and tanh is the hyperbolic tangent activation function;
4. Current Hidden State: The update gate determines the final hidden state as a combination of the previous hidden state and the candidate hidden state.
$$h_k = z_k \odot h_{k-1} + (1 - z_k) \odot \tilde{h}_k$$

This streamlined design allows GRUs to achieve performance similar to LSTMs with reduced computational complexity [102]. The GRU's update gate decides which information from the previous and current steps should contribute to the output [103].

An interesting application of a GRU variant is presented in the article [104], where the authors propose an enhanced model for State of Charge (SOC) estimation using a Bidirectional Gated Recurrent Unit (Bi-GRU) network optimized with the Nesterov Accelerated Gradient (NAG) algorithm, as shown in Figure 24.

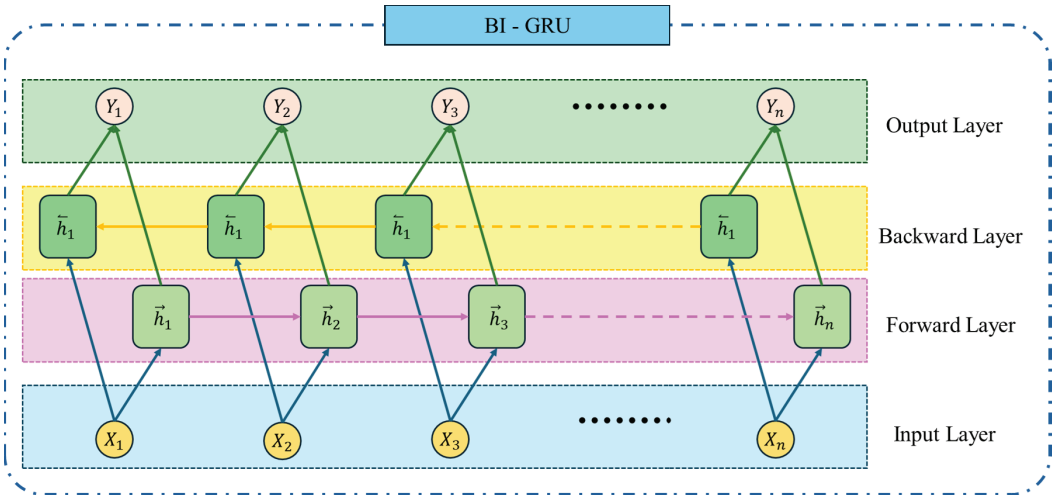


Figure 24. Architecture of Bidirectional GRU (Bi-GRU).

To validate their approach, the authors utilize a well-known lithium-ion battery dataset from the University of Maryland, ensuring a comprehensive evaluation of the model's effectiveness. They achived some interesting results in terms of accuracy in SOC estimation across various ambient temperatures, outperforming existing state-of-the-art data-driven methods and demonstrating its robustness under diverse conditions. In [105], the authors combine Gated Recurrent Units (GRUs) with Recurrent Neural Networks (RNNs) to propose an advanced method for estimating the State of Charge (SOC) in lithium batteries, as shown in Figure 25. The paper presents an innovative approach that integrates a GRU–RNN model with an enhanced momentum gradient optimization algorithm. Traditional techniques—such as ampere-hour integration and open-circuit voltage measurements—often struggle with both accuracy and real-time performance. To address these challenges, the authors develop a model that takes measured voltage and current as inputs to a GRU–RNN, which then outputs an estimated SOC. The GRU, an evolution of conventional recurrent neural networks, effectively manages long-term dependencies while mitigating issues like exploding or vanishing gradients. Additionally, the paper introduces an improved training strategy through a momentum gradient algorithm, a refined version of the classic gradient descent method. The authors first review the conventional gradient descent approach and then present the momentum-enhanced version. This algorithm considers both the current gradient and historical gradient information: when the weight update direction at the current iteration aligns with that of the previous one, the update is reinforced; if not, it is dampened. This approach not only reduces oscillations during weight updates but also accelerates convergence during training. To further enhance generalization and prevent overfitting, random noise is incorporated into the training data. Experimental validation is carried out on a lithium battery test platform that includes a battery tester (NEWARE CT-4008-5V12A-TB), a battery holder, BTcap 21700 (2200 mAh) lithium batteries, and a host computer. The study systematically evaluates various parameters—such as the momentum coefficient, noise variance, number of training epochs, and number of hidden neurons—to optimize performance. The results reveal that a moderate momentum coefficient significantly improves convergence speed and minimizes oscillations, while excessively high values can lead to instability. Moreover, an optimal noise variance of 0.03, combined with 30 hidden neurons, yields the best prediction accuracy, as indicated by improvements in RMSE, MAE, and R^2 metrics. This study demonstrates that the GRU–RNN-based momentum optimization algorithm provides a reliable and efficient method for SOC estimation in battery management systems. The integration of advanced machine learning techniques with tailored optimization strategies represents a significant contribu-

tion to the field and holds promise for application to other dynamic systems, ultimately paving the way for more accurate and efficient battery performance monitoring. This work exemplifies a promising fusion of machine learning and practical battery applications, offering a robust solution to a critical challenge in energy storage technology.

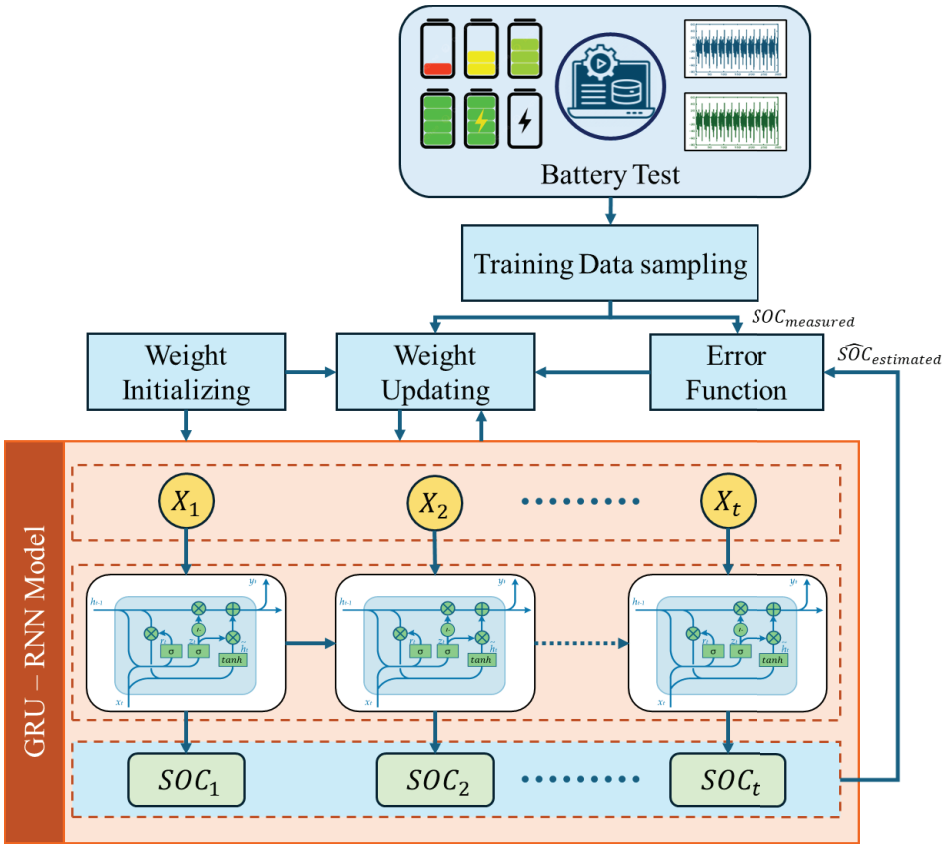


Figure 25. Scheme of GRU–RNN architecture.

The authors of a recent study [106] present a robust and innovative approach for estimating the State of Charge (SOC) in lithium-ion batteries, known as GRU–ASG. This method integrates a Gated Recurrent Unit (GRU) neural network with an adaptive Savitzky–Golay (ASG) filter, enhancing both the accuracy and stability of SOC estimation. The core principles of this approach are illustrated in Figure 26. Accurate SOC estimation is fundamental for ensuring the safety, efficiency, and longevity of energy storage systems. Traditional methods, such as ampere-hour integration and open-circuit voltage measurement, suffer from cumulative errors and rely on precise initial SOC values, making them impractical for many real-world applications. Model-based techniques, like Kalman Filtering, require detailed system parameters, which are often unavailable or difficult to determine. More recently, deep learning approaches, particularly Recurrent Neural Networks (RNNs) such as GRU and LSTM, have demonstrated superior performance in SOC estimation. However, these models struggle with fluctuations in predictions, particularly under dynamic operating conditions. To overcome these challenges, the GRU–ASG method introduces a two-fold solution. First, the "many-to-one" structured GRU network captures long-term dependencies by incorporating multiple past measurements into each SOC estimate, thereby enhancing prediction accuracy. Second, the adaptive Savitzky–Golay filter (ASG) smooths the GRU-generated SOC predictions. Unlike Kalman-based filters, which require a reference SOC, ASG autonomously optimizes its parameters, eliminating the need for predefined initial SOC values. The method operates in two main stages. The GRU network first generates an initial SOC estimate, leveraging historical voltage,

current, and temperature data. The ASG filtering algorithm then refines this estimate by dynamically adjusting the smoothing window length, optimizing the trade-off between reducing fluctuations and preserving accuracy. This adaptive approach eliminates the need for manual parameter tuning, making it particularly well-suited for real-time applications in battery management systems. To validate the effectiveness of the many-to-one GRU architecture, the authors compare its performance against standard RNN and LSTM models. The three models have identical architectures, except for the type of neural network layers. Experimental results indicate that while all three models can capture the general decreasing trend of SOC, their accuracy varies significantly. The GRU model achieves the lowest maximum estimation error at approximately 7%, followed by LSTM at 8% and RNN at 12%. These findings underscore the major reliability of GRU for SOC estimation. In terms of error metrics, GRU achieves a mean squared error (MSE) of 0.17%, which is 59% lower than RNN and 39% lower than LSTM. Similarly, its mean absolute error (MAE) is 3.49%, 33% lower than RNN, and 20% lower than LSTM. These results highlight the clear advantage of the many-to-one GRU structure in improving SOC estimation accuracy. To further enhance its robustness, the ASG filtering algorithm is applied to refine the GRU-generated SOC estimates. The GRU–ASG model is trained using four out of six different operating condition datasets collected from an energy storage plant, with the remaining two datasets used for validation. Experimental results confirm that GRU–ASG consistently outperforms other approaches, including GRU, GRU-MD (Moving Median), and GRU-GA (Gaussian filter), achieving performance comparable to GRU–SG (Savitzky–Golay filtering with an optimally tuned window length). Notably, across all test datasets, MSE remains below 0.15% and MAE does not exceed 3%. Additionally, the GRU–ASG model demonstrates strong generalization capabilities, enabling accurate SOC estimation even under complex and unpredictable operating conditions. When compared to alternative filtering techniques, GRU–ASG shows clear advantages. While all filtering methods enhance the stability of GRU predictions, ASG outperforms both moving median (MD) and Gaussian (GA) filters by dynamically selecting the optimal smoothing window length. Unlike Kalman-based methods, which require an accurate initial SOC, GRU–ASG operates independently, making it highly practical for real-world applications in energy storage. As energy storage systems continue to expand, the increasing availability of real-world battery data will further improve the training and performance of GRU–ASG, strengthening its applicability for large-scale SOC estimation and battery safety management. This scalability highlights its potential for widespread adoption in modern battery management systems. Future research will focus on incorporating additional environmental factors, such as temperature fluctuations and battery capacity degradation, to further enhance robustness under real-world operating conditions. Moreover, transfer learning techniques will be explored to extend the adaptability of GRU–ASG across different battery chemistries and usage scenarios, ensuring even greater versatility in practical applications. The GRU–ASG model represents a significant advancement in SOC estimation, surpassing both traditional and deep learning-based approaches. The many-to-one GRU structure effectively integrates historical data, while the ASG filter enhances stability and adaptability. With its strong generalization capabilities and scalability, GRU–ASG stands as a highly promising an interesting solution for next-generation energy storage systems, where precise SOC estimation is crucial for ensuring safety, efficiency, and longevity.

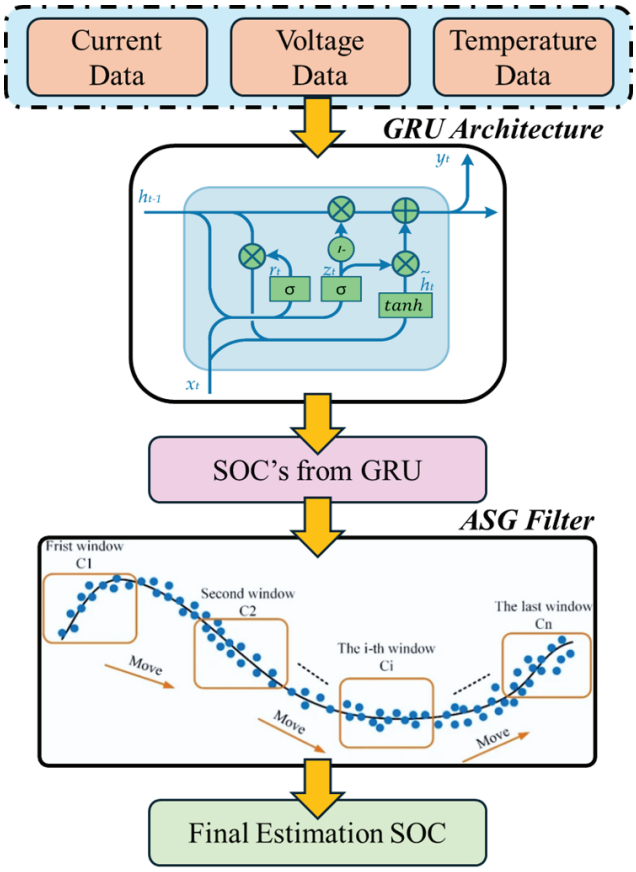


Figure 26. Overall structure of GRU–ASG model.

5. Discussion

The growing demand for machine learning-based methods in estimating the State of Charge (SOC) of EV/HV battery system arises from the limitations of traditional approaches, which struggle with the uncertainty and complexity inherent in battery models. Unlike conventional methods, machine learning techniques exploit large datasets and advanced artificial intelligence algorithms to bypass the need for explicit battery modeling. Instead, they train specialized models to uncover nonlinear relationships between input parameters, such as voltage, current, and temperature, and the SOC. These techniques are recognized for their high accuracy, adaptability, and strong generalization capabilities. This review consolidates recent advancements in machine learning-based SOC estimation, presenting challenges, opportunities, and areas for further research. Additionally, the methods analyzed are summarized in the sequence of Tables 1–4 , which outlines their respective advantages and disadvantages.

Table 1. Comparison of SOC estimation methods.

Article	Innovations	Dataset & EV Tests	Performance	Advantages	Disadvantages	Computational Cost
UKF for SOC Estimation	OCV–SOC-T table with temperature (0–50 °C); UKF for nonlinearity handling; Simplified Rint model.	Lab tests on LiFePO4 (18650) with DST, OCV–SOC-T, FUDS; No direct EV tests.	RMS error < 5% (40 °C), up to 16.4% (0 °C).	Accurate SOC estimation with temperature integration; Real-time capable.	Poor performance at low temperatures; Sensitive to flat OCV–SOC regions.	Low; optimized for BMS real-time use.
Neural-Network-Embedded ECM [66]	ECM with 3 feedforward NNs for residual error correction; Physics-informed loss in training.	Panasonic 18650PF; HPPC, US06, HWFET at 10 °C to −20 °C; Lab-tested under EV conditions.	MSE reduced by 82.5% (US06); RMSE improved 33–64% (US06), 5–29% (HWFET).	High accuracy, especially in extreme conditions; Adaptive to uncertainties.	Initial/final phase errors; Complex offline training, requires retraining.	Offline: 264–450 s; Online: 1.3–2.2 s (real-time suitable).
Physics-Informed NN (PINN) [63]	Deep FNN with physics-constrained loss (data + charge conservation).	McMaster Univ. (LG HG2, 3Ah); UDDS, US06, LA92, HWFET, Mix; −10 °C to 25 °C; Climatic chamber tests.	RMSE < 2.85%; Outperforms NN-only and AKF methods.	High accuracy and robustness with noisy/sparse data; Physically consistent solutions.	Demanding offline training; Complex design.	Offline: demanding; Online: efficient, real-time capable.
NN + UKF for Error Cancellation [45]	Hybrid NN for SOC estimation + UKF for noise/error reduction; No OCV–SOC table needed.	Dynamic tests; NN trained on DST, validated on US06/FUDS; Li-ion (mostly LiFePO4) under EV conditions.	RMS error (US06, 0 °C) reduced from 4.1% (NN only) to 2.4% (NN + UKF).	Higher robustness via UKF filtering; No need for OCV–SOC table.	Risk of NN overfitting; Complex NN–UKF integration.	Offline: structural optimization; Online: real-time with filtered errors.

Table 2. Comparison of SOC estimation methods (part 2).

Article	Innovations	Dataset & EV Tests	Performance	Advantages	Disadvantages	Computational Cost
Hybrid FNNs for SOC [51]	Hybrid FNN model for SOC estimation; NN optimization to improve generalization and reduce errors.	Lab data on Li-ion (LiFePO4) batteries; DST, US06, FUDS at variable temperatures.	RMS error < 3–4% under optimal conditions.	Robust to nonlinearities; No need for OCV–SOC tables.	Complex offline training; Requires careful optimization.	Offline: demanding; Online: real-time capable.
FNN + EKF τ [48]	Hybrid approach: FNN + EKF with parameter τ for filter update; Direct SOC estimation without OCV–SOC table.	Lab data on Li-ion (LiFePO4); DST, US06, FUDS under various temperatures.	Reduced RMS and max errors; Stable SOC estimation.	Accurate and robust; EKF ensures real-time applicability.	EKF less accurate in strong nonlinearities; Requires tuning and sensor calibration.	Offline: demanding; Online: real-time suitable.
BPNN + BSA Model [62]	BPNN with BSA model for direct SOC estimation; Improved data fusion (voltage, current, temperature).	Lab data on Li-ion (LiFePO4); DST, US06, FUDS at various temperatures.	RMS and MAE errors typically < 5%.	Strong nonlinear learning; No OCV–SOC table needed.	Complex training; Requires high-quality data.	Offline: computationally heavy; Online: real-time feasible.
LSTM–RNN for SOC [85]	EI-LSTM-CO model: extended input (avg. voltage via sliding window), output constrained by AhI strategy for stability.	Public dataset (CALCE, LiFePO4); DST for training, US06 and FUDS for validation (7 temperatures, 1s sampling).	RMSE < 1.3%, MAXE < 3.2%; Improved stability vs. standard LSTM–RNN.	Smooth SOC estimation; No need for accurate initial SOC; High accuracy for BMS.	Requires careful tuning (epochs, window size, constraints); More complex than conventional NNs.	Offline: 150 epochs, window size 50; Online: fast, real-time capable.

Table 3. Comparison of SOC estimation methods (part 3).

Article	Innovations	Dataset & EV Tests	Performance	Advantages	Disadvantages	Computational Cost
CNN–LSTM for SOC [82]	Hybrid CNN–LSTM for feature extraction and temporal modeling; Robust to unknown initial SOC and temperature variations.	A123 18650 LiFePO4; DST, FUDS, US06; Climatic chamber tests.	MAE < 1%, RMSE < 2% (unknown initial SOC); RMSE < 2%, MAE < 1.5% (temperature variations).	Stable and accurate SOC estimation; Rapid convergence; Good adaptation to environmental variations.	Complex architecture; Long offline training and parameter tuning.	Offline: 10,000 epochs (161 min on GPU); Online: 0.098 ms per time step (real-time suitable).
RS-LSTM (Random Search LSTM) [86]	LSTM with Random Search optimization; Random Forest for feature selection; Auto-tuned hyperparameters.	CALCE dataset (INR 18650-20R); DST, FUDS, US06; Lab and real vehicle data at different temperatures.	Optimal settings: Look back = 45, Epochs = 177, Batch = 64, LR = 0.0026; MAE = 0.221%, RMSE = 0.262%.	Reduces noise via Random Forest; Auto-optimized hyperparameters; High accuracy and stability.	High offline computation due to hyperparameter search; Complex model tuning.	Offline: intensive training; Online: fast, real-time capable.
GLA–CNN–BiLSTM for SOC [92]	CNN for spatial and BiLSTM for temporal features; GLA for automatic hyperparameter tuning.	Six EV discharge datasets (HWFET, US06, BJDST, DST, FUDS, UDDS); Lab tests on Li-ion.	MAE < 1%, RMSE < 1%, Max error < 2%.	Combines CNN (features) + BiLSTM (time); GLA ensures fast convergence and adaptability.	Intensive offline training; Complex hyperparameter tuning.	Offline: high due to GLA; Online: fast, real-time suitable.

Table 4. Comparison of SOC estimation methods (part 4).

Article	Innovations	Dataset & EV Tests	Performance	Advantages	Disadvantages	Computational Cost
AMBiLSTM with Attention [94]	BiLSTM with spatial/temporal attention; Temperature compensation in OCV and capacity; SOC and SOH co-estimation.	Lab tests on $LiNiCoAlO_2$ (2.6 Ah); 0–40 °C; DST, UDDS; 1 s/10 s sampling.	DST (25 °C): RMSE ↓9.39%; UDDS (25 °C): RMSE ↓22.36%; SOH ↑21.45%.	Improved feature extraction; Robust to temperature variations; Simultaneous SOC/SOH estimation.	Complex model; Requires large dataset and intensive training; Potential short-sequence errors.	Offline: GPU-intensive (RTX 3090); Online: fast, real-time feasible.
GRU–RNN with Momentum [105]	GRU–RNN with momentum gradient algorithm; Noise injection to prevent overfitting.	Lab tests on BTcap 21700 (2.2 Ah); Charge-discharge cycles; Voltage-current measurements.	Sigma = 0.03: RMSE = 0.0092, MAE = 0.0041, R^2 = 0.9990.	Fast convergence; Reduced weight oscillations; High accuracy and generalization.	Requires precise tuning of momentum (beta) and hyperparameters.	Offline: intensive; Online: extremely fast, real-time capable.
GRU–ASG for SOC [106]	GRU with adaptive Savitzky–Golay (ASG) filter; Spearman coefficient for dynamic window selection.	Real data from LiFePO4 (280 Ah, 3.2 V); Energy storage plant; Six discharge datasets.	MSE < 0.15%, MAE < 3%; Outperforms standard GRU and filters.	Robust in varying conditions; Adaptive filter eliminates manual tuning; Effective memory structure.	Complex integration of NN with adaptive filtering; Additional online computation.	Offline: GPU training (GTX 1060, TensorFlow); Online: fast, real-time suitable.

5.1. Challenges in Machine Learning-Based SOC Estimation

From the analysis of the methods proposed in this article, it is possible to define problems that these methods have in common, which may prove to be the challenges to overcome in future work, in order to make these ANN algorithms even more performant. The challenges to be overcome and solved can be summarized in the following key points:

1. Inconsistencies in Model Evaluation: Comparing different models remains challenging due to variability in datasets, hyperparameter settings, optimization algorithms, and computational resources used across studies [107]. These inconsistencies complicate cross-study benchmarking and make it difficult to assess the relative effectiveness of different approaches. Furthermore, the use of complex composite algorithms, which often combine serial or parallel structures, increases model complexity [108]. As the number of modules grows, the model's overall controllability diminishes, making it more vulnerable to disturbances and raising concerns about its robustness in practical applications [109];
2. Dataset Limitations: Large variations in datasets are caused by differences in experimental conditions and the mismatch between training data and real-world electric vehicle operation [110]. This discrepancy impacts even the most sophisticated models, reducing their accuracy and applicability in real-world scenarios [111]. The lack of standardized, comprehensive datasets exacerbates this issue;
3. Computational Complexity: Many advanced SOC estimation algorithms require significant computational power due to their intricate structures and optimization techniques. This increases the demands on hardware resources, potentially hindering the real-time performance of SOC estimation systems [112]. Balancing computational efficiency with accuracy is critical for enabling online SOC estimation;
4. Lack of Open-Source Tools: Few researchers publish their code, and there is often little transparency regarding preprocessing techniques, input parameter choices, and hyperparameter tuning. These gaps hinder reproducibility and prevent effective collaboration within the research community. Without standardized benchmarks, it is difficult to attribute performance improvements to either the model or preprocessing techniques [113];
5. Real-Time Implementation: Implementing advanced machine learning algorithms in real-time systems, especially in resource-constrained vehicles, poses significant technical hurdles [114,115];

6. Generalization to Diverse Conditions: We must adapt SOC estimation models to perform reliably across a wide range of driving conditions and environmental factors [116].

5.2. Areas for Improvement and Key Observations

We use Artificial Neural Network (ANN) methods to estimate the State of Charge (SOC) and these methods have proven effective for their intended tasks. However, our analysis shows that several improvements can boost performance, reliability, and practical applicability. Below, we present key areas for improvement along with emerging challenges that future research must tackle:

1. Standardization of Datasets and Data Quality: Machine learning models often depend on specific datasets that limit generalization. We must develop high-quality, standardized datasets with precise measurements and diverse data types. Researchers should also improve data quality by applying advanced data preprocessing techniques [117], sensor fusion methods, and machine learning-based denoising algorithms to ensure accurate and reliable input data [118];
2. Open-Source Benchmarks: Creating open-source initiatives and establishing standardized benchmarks for data preprocessing, model architectures, and hyperparameter tuning may lead to significant improvements [119]. Adopting uniform techniques to clean, normalize, and prepare data establishes a consistent foundation for experiments. Designing model architectures with comparable features and systematically tuning hyperparameters—by adjusting factors like learning rates and network layers—minimizes biases from arbitrary decisions. This approach promises enhanced reproducibility, fair comparisons among methods, and increased collaboration, ultimately accelerating progress in ANN-based SOC estimation for electric vehicle batteries [120];
3. Balanced Algorithm Complexity: In pursuit of increased accuracy, SOC estimation models must deliver computational efficiency and meet real-world constraints [121]. Evaluating these models from diverse perspectives—including robustness, adaptability, and efficiency—ensures practicality for deployment;
4. Generalization and Transfer Learning: One of the challenges in achieving accurate State of Charge estimation is that vehicles are subjected to a wide range of driving conditions, many of which are not fully represented during the training phase. Ensuring that these models generalize well across such diverse conditions remains a significant concern [122]. Transfer learning is emerging as a promising approach to adapt these models to varying scenarios, with researchers actively exploring how to leverage knowledge from one domain to enhance performance in another;
5. AI Hardware Acceleration: The advent of specialized hardware for AI and machine learning is expected to accelerate the deployment of complex models in real-time applications [123]. In particular, hardware accelerators such as GPUs and TPUs drastically reduce training and inference times, paving the way for the development and integration of increasingly sophisticated state-of-charge estimation models in electric vehicles;
6. Big Data and Cloud Computing: Integrating big data analytics with cloud computing platforms can optimization SOC management. Cloud-based solutions provide scalability and allow remote monitoring and optimization by analyzing data from large fleets, leading to more robust and adaptive management strategies [124];
7. Advanced Machine Learning Techniques: Advanced models—including Convolutional Neural Networks (CNNs) and Transformer-based architectures—promise significant improvements [125]. These techniques excel at feature extraction and capture complex, non-linear relationships in battery data. For example, models using atten-

tion mechanisms (such as AMBiLSTM) and multi-head self-attention in Transformers effectively identify intricate patterns and dynamic operating conditions [126];

8. Hybrid Models and Optimization Strategies: Combining multiple architectures (e.g., GRU–ASG, GLA–CNN–BiLSTM) in hybrid models can improve accuracy and robustness [127]. Researchers must optimize these models to manage increased computational complexity. Likewise, advanced optimization strategies (as seen in RS-LSTM) can boost performance while requiring more computational resources [128].

Addressing these improvements and challenges will enhance the performance, reliability, and applicability of ANN-based SOC estimation methods, driving the development of more efficient and robust battery management systems.

6. Conclusions

This work presents a comprehensive and detailed review of the State of Charge (SOC) estimation methods for battery systems based on machine learning algorithms. By analyzing and comparing various methods, this paper provides a clear understanding of the advantages and limitations of different approaches, offering valuable insights into the evolution and current trends in SOC estimation. A structured methodology for evaluating and benchmarking different models is introduced, enabling a more consistent and fair comparison across diverse studies. This framework addresses key challenges in the field, including dataset variability, computational complexity, and the lack of standardized benchmarks, while offering guidance on how these obstacles can be overcome. Furthermore, the study emphasizes areas for improvement, such as enhancing generalization capabilities, fostering open-source collaboration, and achieving a balance between accuracy and computational efficiency. The potential of machine learning-based SOC estimation lies in its ability to handle the nonlinear dynamics of batteries, its adaptability to various operating conditions, and its robust generalization capabilities.

By addressing challenges such as dataset standardization and promoting the adoption of open-source benchmarks, this work aims to accelerate innovation in SOC estimation methods. Additionally, it advocates for the development of scalable, efficient, and robust machine learning models that align with real-world constraints, fostering advances in battery management systems. Ultimately, this study aims to contribute to the enhancement of EV/HV battery system performance and safety, driving forward the intelligent evolution of SOC estimation techniques. These advancements will play a pivotal role in supporting the sustainable development of electric vehicles and new energy technologies, paving the way for a cleaner, smarter energy future.

Author Contributions: All authors has contributed equally to the the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been partially supported by CN1 Spoke 6; by HARDNESS PE SERICS Spoke 7; and by MIUR project FoReLab.

Data Availability Statement: No new data have been generated.

Conflicts of Interest: Authors declare no conflicts of interest.

References

1. Hindawi, M.; Basheer, Y.; Qaisar, S.M.; Waqar, A. An Overview of Artificial Intelligence Driven Li-Ion Battery State Estimation. In *IoT Enabled-DC Microgrids*; CRC Press: Boca Raton, FL, USA, 2024; pp. 121–157.
2. Dini, P.; Paolini, D.; Saponara, S.; Minossi, M. Leaveraging Digital Twin & Artificial Intelligence in Consumption Forecasting System for Sustainable Luxury Yacht. *IEEE Access* **2024**, 12, 160700–160714.
3. Lu, C.; Li, S.; Lu, Z. Building energy prediction using artificial neural networks: A literature survey. *Energy Build.* **2022**, *262*, 111718. [CrossRef]

4. Moayedi, H.; Mosavi, A. Double-target based neural networks in predicting energy consumption in residential buildings. *Energies* **2021**, *14*, 1331. [CrossRef]
5. Dzedzickis, A.; Subačiūtė-Žemaitienė, J.; Šutinys, E.; Samukaitė-Bubnienė, U.; Bučinskas, V. Advanced applications of industrial robotics: New trends and possibilities. *Appl. Sci.* **2021**, *12*, 135. [CrossRef]
6. Tsapin, D.; Pitelinskiy, K.; Suvorov, S.; Osipov, A.; Pleshakova, E.; Gataullin, S. Machine learning methods for the industrial robotic systems security. *J. Comput. Virol. Hacking Tech.* **2024**, *20*, 397–414. [CrossRef]
7. Manoharan, A.; Begam, K.; Aparow, V.R.; Sooriamoorthy, D. Artificial Neural Networks, Gradient Boosting and Support Vector Machines for electric vehicle battery state estimation: A review. *J. Energy Storage* **2022**, *55*, 105384. [CrossRef]
8. Djaballah, Y.; Negadi, K.; Boudiaf, M. Enhanced lithium–ion battery state of charge estimation in electric vehicles using extended Kalman filter and deep neural network. *Int. J. Dyn. Control* **2024**, *12*, 2864–2871. [CrossRef]
9. Allal, Z.; Noura, H.N.; Salman, O.; Chahine, K. Machine learning solutions for renewable energy systems: Applications, challenges, limitations, and future directions. *J. Environ. Manag.* **2024**, *354*, 120392. [CrossRef]
10. Rinchi, O.; Alsharoa, A.; Shatnawi, I.; Arora, A. The Role of Intelligent Transportation Systems and Artificial Intelligence in Energy Efficiency and Emission Reduction. *arXiv* **2024**, arXiv:2401.14560.
11. Chen, W.; Men, Y.; Fuster, N.; Osorio, C.; Juan, A.A. Artificial intelligence in logistics optimization with sustainable criteria: A review. *Sustainability* **2024**, *16*, 9145. [CrossRef]
12. Qadir, S.A.; Ahmad, F.; Al-Wahedi, A.M.A.; Iqbal, A.; Ali, A. Navigating the complex realities of electric vehicle adoption: A comprehensive study of government strategies, policies, and incentives. *Energy Strategy Rev.* **2024**, *53*, 101379. [CrossRef]
13. Möring-Martínez, G.; Senzeybek, M.; Jochem, P. Clustering the European Union electric vehicle markets: A scenario analysis until 2035. *Transp. Res. Part D Transp. Environ.* **2024**, *135*, 104372. [CrossRef]
14. Raganati, F.; Ammendola, P. CO_2 post-combustion capture: A critical review of current technologies and future directions. *Energy Fuels* **2024**, *38*, 13858–13905. [CrossRef]
15. al Irsyad, M.I.; Firmansyah, A.I.; Harisetyawan, V.T.F.; Supriatna, N.K.; Gunawan, Y.; Jupesta, J.; Yaumidin, U.K.; Purwanto, J.; Inayah, I. Comparative total cost assessments of electric and conventional vehicles in ASEAN: Commercial vehicles and motorcycle conversion. *Energy Sustain. Dev.* **2025**, 101599.. [CrossRef]
16. Hussein, H.M.; Aghmadi, A.; Abdelrahman, M.S.; Rafin, S.S.H.; Mohammed, O. A review of battery state of charge estimation and management systems: Models and future prospective. *Wiley Interdiscip. Rev. Energy Environ.* **2024**, *13*, e507.
17. Nekahi, A.; Madikere Raghunatha Reddy, A.K.; Li, X.; Deng, S.; Zaghib, K. Rechargeable Batteries for the Electrification of Society: Past, Present, and Future. *Electrochem. Energy Rev.* **2025**, *8*, 1–30. [CrossRef]
18. Dini, P.; Saponara, S.; Colicelli, A. Overview on battery charging systems for electric vehicles. *Electronics* **2023**, *12*, 4295. [CrossRef]
19. Ria, A.; Dini, P. A Compact Overview on Li-Ion Batteries Characteristics and Battery Management Systems Integration for Automotive Applications. *Energies* **2024**, *17*, 5992. [CrossRef]
20. Demirci, O.; Taskin, S.; Schaltz, E.; Demirci, B.A. Review of battery state estimation methods for electric vehicles-Part I: SOC estimation. *J. Energy Storage* **2024**, *87*, 111435. [CrossRef]
21. Pillai, P.; Sundaresan, S.; Kumar, P.; Pattipati, K.R.; Balasingam, B. Open-circuit voltage models for battery management systems: A review. *Energies* **2022**, *15*, 6803. [CrossRef]
22. Movassagh, K.; Raihan, A.; Balasingam, B.; Pattipati, K. A critical look at coulomb counting approach for state of charge estimation in batteries. *Energies* **2021**, *14*, 4074. [CrossRef]
23. Lipu, M.H.; Hannan, M.; Hussain, A.; Ayob, A.; Saad, M.H.; Karim, T.F.; How, D.N. Data-driven state of charge estimation of lithium-ion batteries: Algorithms, implementation factors, limitations and future trends. *J. Clean. Prod.* **2020**, *277*, 124110. [CrossRef]
24. Tao, Z.; Zhao, Z.; Wang, C.; Huang, L.; Jie, H.; Li, H.; Hao, Q.; Zhou, Y.; See, K.Y. State of charge estimation of lithium Batteries: Review for equivalent circuit model methods. *Measurement* **2024**, *236*, 115148. [CrossRef]
25. Shrivastava, P.; Soon, T.K.; Idris, M.Y.I.B.; Mekhilef, S. Overview of model-based online state-of-charge estimation using Kalman filter family for lithium-ion batteries. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109233. [CrossRef]
26. Farea, A.; Yli-Harja, O.; Emmert-Streib, F. Understanding physics-informed neural networks: Techniques, applications, trends, and challenges. *AI* **2024**, *5*, 1534–1557. [CrossRef]
27. Reza, M.; Mannan, M.; Mansor, M.; Ker, P.J.; Mahlia, T.I.; Hannan, M. Recent advancement of remaining useful life prediction of lithium-ion battery in electric vehicle applications: A review of modelling mechanisms, network configurations, factors, and outstanding issues. *Energy Rep.* **2024**, *11*, 4824–4848. [CrossRef]
28. Rahman, T.; Alharbi, T. Exploring lithium-Ion battery degradation: A concise review of critical factors, impacts, data-driven degradation estimation techniques, and sustainable directions for energy storage systems. *Batteries* **2024**, *10*, 220. [CrossRef]
29. Marzbani, F.; Osman, A.H.; Hassan, M.S. Electric vehicle energy demand prediction techniques: An in-depth and critical systematic review. *IEEE Access* **2023**, *11*, 96242–96255. [CrossRef]

30. Memon, S.A.; Hamza, A.; Zaidi, S.S.H.; Khan, B.M. Estimating state of charge and state of health of electrified vehicle battery by data driven approach: Machine learning. In Proceedings of the 2022 International Conference on Emerging Technologies in Electronics, Computing and Communication (ICETECC), Jamshoro, Pakistan, 7–9 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–9.
31. Dini, P.; Colicelli, A.; Saponara, S. Review on modeling and soc/soh estimation of batteries for automotive applications. *Batteries* **2024**, *10*, 34. [CrossRef]
32. Ferriol-Galmés, M.; Suárez-Varela, J.; Paillissé, J.; Shi, X.; Xiao, S.; Cheng, X.; Barlet-Ros, P.; Cabellos-Aparicio, A. Building a digital twin for network optimization using graph neural networks. *Comput. Netw.* **2022**, *217*, 109329. [CrossRef]
33. Di Fonso, R.; Teodorescu, R.; Cecati, C.; Bharadwaj, P. A Battery Digital Twin From Laboratory Data Using Wavelet Analysis and Neural Networks. *IEEE Trans. Ind. Inform.* **2024**, *20*, 6889–6899. [CrossRef]
34. Li, W.; Li, Y.; Garg, A.; Gao, L. Enhancing real-time degradation prediction of lithium-ion battery: A digital twin framework with CNN-LSTM-attention model. *Energy* **2024**, *286*, 129681. [CrossRef]
35. Wang, S.; Zhang, J.; Wang, P.; Law, J.; Calinescu, R.; Mihaylova, L. A deep learning-enhanced Digital Twin framework for improving safety and reliability in human–robot collaborative manufacturing. *Robot. Comput.-Integr. Manuf.* **2024**, *85*, 102608. [CrossRef]
36. Iliuţă, M.E.; Moisescu, M.A.; Pop, E.; Ionita, A.D.; Caramihai, S.I.; Mitulescu, T.C. Digital Twin—A Review of the Evolution from Concept to Technology and Its Analytical Perspectives on Applications in Various Fields. *Appl. Sci.* **2024**, *14*, 5454. [CrossRef]
37. Chen, S.; Billings, S.A. Neural networks for nonlinear dynamic system modelling and identification. *Int. J. Control* **1992**, *56*, 319–346. [CrossRef]
38. Arulampalam, G.; Bouzerdoum, A. A generalized feedforward neural network architecture for classification and regression. *Neural Netw.* **2003**, *16*, 561–568. [CrossRef] [PubMed]
39. Wang, Y.; Tian, J.; Sun, Z.; Wang, L.; Xu, R.; Li, M.; Chen, Z. A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renew. Sustain. Energy Rev.* **2020**, *131*, 110015. [CrossRef]
40. Rane, C.; Tyagi, K.; Kline, A.; Chugh, T.; Manry, M. Optimizing performance of feedforward and convolutional neural networks through dynamic activation functions. *Evol. Intell.* **2024**, *17*, 4083–4093. [CrossRef]
41. Szandała, T. Review and comparison of commonly used activation functions for deep neural networks. In *Bio-Inspired Neurocomputing*; Springer: Singapore, 2021; pp. 203–224.
42. Hammad, M. Deep Learning Activation Functions: Fixed-Shape, Parametric, Adaptive, Stochastic, Miscellaneous, Non-Standard, Ensemble. *arXiv* **2024**, arXiv:2407.11090.
43. Narkhede, M.V.; Bartakke, P.P.; Sutaone, M.S. A review on weight initialization strategies for neural networks. *Artif. Intell. Rev.* **2022**, *55*, 291–322. [CrossRef]
44. Dang, X.; Yan, L.; Jiang, H.; Wu, X.; Sun, H. Open-circuit voltage-based state of charge estimation of lithium-ion power battery by combining controlled auto-regressive and moving average modeling with feedforward-feedback compensation method. *Int. J. Electr. Power Energy Syst.* **2017**, *90*, 27–36. [CrossRef]
45. Feng, S.; Li, X.; Zhang, S.; Jian, Z.; Duan, H.; Wang, Z. A review: State estimation based on hybrid models of Kalman filter and neural network. *Syst. Sci. Control Eng.* **2023**, *11*, 2173682. [CrossRef]
46. He, W.; Williard, N.; Chen, C.; Pecht, M. State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter-based error cancellation. *Int. J. Electr. Power Energy Syst.* **2014**, *62*, 783–791. [CrossRef]
47. Murawwat, S.; Gulzar, M.M.; Alzahrani, A.; Hafeez, G.; Khan, F.A.; Abed, A.M. State of charge estimation and error analysis of lithium-ion batteries for electric vehicles using Kalman filter and deep neural network. *J. Energy Storage* **2023**, *72*, 108039.
48. Chen, C.; Xiong, R.; Yang, R.; Shen, W.; Sun, F. State-of-charge estimation of lithium-ion battery using an improved neural network model and extended Kalman filter. *J. Clean. Prod.* **2019**, *234*, 1153–1164. [CrossRef]
49. Srihari, S.; Vasanthi, V. SOC Estimation Using Extended Kalman Filter in Electric Vehicle Batteries. In Proceedings of the 2024 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI), Kannur, India, 21–22 June 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–5.
50. Khan, U.; Kirmani, S.; Rafat, Y.; Rehman, M.U.; Alam, M.S. Improved deep learning based state of charge estimation of lithium ion battery for electrified transportation. *J. Energy Storage* **2024**, *91*, 111877. [CrossRef]
51. Vidal, C.; Haußmann, M.; Barroso, D.; Shamsabadi, P.M.; Biswas, A.; Chemali, E.; Ahmed, R.; Emadi, A. Hybrid energy storage system state-of-charge estimation using artificial neural network for micro-hybrid applications. In Proceedings of the 2018 IEEE Transportation Electrification Conference and Expo (ITEC), Long Beach, CA, USA, 13–15 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1075–1081.
52. Cui, X.; Liu, S.; Ruan, G.; Wang, Y. Data-driven aggregation of thermal dynamics within building virtual power plants. *Appl. Energy* **2024**, *353*, 122126. [CrossRef]
53. Barik, S.; Saravanan, B. Recent developments and challenges in state-of-charge estimation techniques for electric vehicle batteries: A review. *J. Energy Storage* **2024**, *100*, 113623. [CrossRef]

54. Vidal, C.; Gross, O.; Gu, R.; Kollmeyer, P.; Emadi, A. xEV Li-ion battery low-temperature effects. *IEEE Trans. Veh. Technol.* **2019**, *68*, 4560–4572. [CrossRef]
55. Bian, C.; Duan, Z.; Hao, Y.; Yang, S.; Feng, J. Exploring large language model for generic and robust state-of-charge estimation of Li-ion batteries: A mixed prompt learning method. *Energy* **2024**, *12*, 131856. [CrossRef]
56. Rastegarpanah, A.; Asif, M.E.; Stolkin, R. Hybrid Neural Networks for Enhanced Predictions of Remaining Useful Life in Lithium-Ion Batteries. *Batteries* **2024**, *10*, 106. [CrossRef]
57. Badfar, M.; Yildirim, M.; Chinnam, R. State-of-charge estimation across battery chemistries: A novel regression-based method and insights from unsupervised domain adaptation. *J. Power Sources* **2025**, *628*, 235760. [CrossRef]
58. Han, J.; Moraga, C.; Sinne, S. Optimization of feedforward neural networks. *Eng. Appl. Artif. Intell.* **1996**, *9*, 109–119. [CrossRef]
59. Ojha, V.K.; Abraham, A.; Snášel, V. Metaheuristic design of feedforward neural networks: A review of two decades of research. *Eng. Appl. Artif. Intell.* **2017**, *60*, 97–116. [CrossRef]
60. Abdolrasol, M.G.; Hussain, S.S.; Ustun, T.S.; Sarker, M.R.; Hannan, M.A.; Mohamed, R.; Ali, J.A.; Mekhilef, S.; Milad, A. Artificial neural networks based optimization techniques: A review. *Electronics* **2021**, *10*, 2689. [CrossRef]
61. Tong, S.; Lacap, J.H.; Park, J.W. Battery state of charge estimation using a load-classifying neural network. *J. Energy Storage* **2016**, *7*, 236–243. [CrossRef]
62. Hannan, M.A.; Lipu, M.S.H.; Hussain, A.; Saad, M.H.; Ayob, A. Neural network approach for estimating state of charge of lithium-ion battery using backtracking search algorithm. *IEEE Access* **2018**, *6*, 10069–10079. [CrossRef]
63. Vidal, C.; Kollmeyer, P.; Naguib, M.; Malysz, P.; Gross, O.; Emadi, A. Robust xev battery state-of-charge estimator design using a feedforward deep neural network. *SAE Int. J. Adv. Curr. Pract. Mobil.* **2020**, *2*, 2872–2880. [CrossRef]
64. Naguib, M.; Kollmeyer, P.; Vidal, C.; Emadi, A. Accurate surface temperature estimation of lithium-ion batteries using feedforward and recurrent artificial neural networks. In Proceedings of the 2021 IEEE Transportation Electrification Conference & Expo (ITEC), Chicago, IL, USA, 21–25 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 52–57.
65. Baraean, A.; Kassas, M.; Abido, M. Physics-Informed NN for Improving Electric Vehicles Lithium-Ion Battery State-of-Charge Estimation Robustness. In Proceedings of the 2024 IEEE 12th International Conference on Smart Energy Grid Engineering (SEGE), Oshawa, ON, Canada, 18–20 August 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 245–250.
66. Guo, Z.; Li, Y.; Yan, Z.; Chow, M.Y. A Neural-Network-Embedded Equivalent Circuit Model for Lithium-ion Battery State Estimation. *arXiv* **2024**, arXiv:2407.20262.
67. Michailidis, P.; Michailidis, I.; Gkelios, S.; Kosmatopoulos, E. Artificial Neural Network Applications for Energy Management in Buildings: Current Trends and Future Directions. *Energies* **2024**, *17*, 570. [CrossRef]
68. Hu, Y.; Huber, A.; Anumula, J.; Liu, S.C. Overcoming the vanishing gradient problem in plain recurrent networks. *arXiv* **2018**, arXiv:1801.06105.
69. Mienye, I.D.; Swart, T.G.; Obaido, G. Recurrent neural networks: A comprehensive review of architectures, variants, and applications. *Information* **2024**, *15*, 517. [CrossRef]
70. Mousaei, A.; Naderi, Y.; Bayram, I.S. Advancing state of charge management in electric vehicles with machine learning: A technological review. *IEEE Access* **2024**, *12*, 43255–43283. [CrossRef]
71. Graves, A.; Mohamed, A.r.; Hinton, G. Speech recognition with deep recurrent neural networks. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 6645–6649.
72. Pascanu, R. On the difficulty of training recurrent neural networks. *arXiv* **2013**, arXiv:1211.5063.
73. Graves, A.; Graves, A. Long short-term memory. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 37–45.
74. Yao, Q.; Kollmeyer, P.J.; Lu, D.D.C.; Emadi, A. A Comparison Study of Unidirectional and Bidirectional Recurrent Neural Network for Battery State of Charge Estimation. In Proceedings of the 2024 IEEE Transportation Electrification Conference and Expo (ITEC), Chicago, IL, USA, 19–21 June 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6.
75. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.
76. Zhang, S.; Wu, Y.; Che, T.; Lin, Z.; Memisevic, R.; Salakhutdinov, R.R.; Bengio, Y. Architectural complexity measures of recurrent neural networks. *arXiv* **2016**, arXiv:1602.08210. [CrossRef]
77. El Fallah, S.; Kharbach, J.; Vanagas, J.; Vilkelytė, Ž.; Tolvaišienė, S.; Gudžius, S.; Kalvaitis, A.; Lehmam, O.; Masrour, R.; Hammouch, Z.; et al. Advanced state of charge estimation using deep neural network, gated recurrent unit, and long short-term memory models for lithium-Ion batteries under aging and temperature conditions. *Appl. Sci.* **2024**, *14*, 6648. [CrossRef]
78. Duan, J.; Zhang, P.F.; Qiu, R.; Huang, Z. Long short-term enhanced memory for sequential recommendation. *World Wide Web* **2023**, *26*, 561–583. [CrossRef]
79. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Ahmed, R.; Emadi, A. Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries. *IEEE Trans. Ind. Electron.* **2017**, *65*, 6730–6739. [CrossRef]

80. Charkhgard, M.; Farrokhi, M. State-of-charge estimation for lithium-ion batteries using neural networks and EKF. *IEEE Trans. Ind. Electron.* **2010**, *57*, 4178–4187. [CrossRef]
81. Data, M. Panasonic 18650PF Li-ion Battery Data. 2018. Available online: https://data.mendeley.com/datasets/wykht8y7tg/1 (accessed on 1 January 2025).
82. Song, X.; Yang, F.; Wang, D.; Tsui, K.L. Combined CNN-LSTM network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* **2019**, *7*, 88894–88902. [CrossRef]
83. Tian, Y.; Lai, R.; Li, X.; Xiang, L.; Tian, J. A combined method for state-of-charge estimation for lithium-ion batteries using a long short-term memory network and an adaptive cubature Kalman filter. *Appl. Energy* **2020**, *265*, 114789. [CrossRef]
84. Wei, M.; Ye, M.; Li, J.B.; Wang, Q.; Xu, X. State of charge estimation of lithium-ion batteries using LSTM and NARX neural networks. *IEEE Access* **2020**, *8*, 189236–189245. [CrossRef]
85. Chen, J.; Zhang, Y.; Wu, J.; Cheng, W.; Zhu, Q. SOC estimation for lithium-ion battery using the LSTM-RNN with extended input and constrained output. *Energy* **2023**, *262*, 125375. [CrossRef]
86. Chai, X.; Li, S.; Liang, F. A novel battery SOC estimation method based on random search optimized LSTM neural network. *Energy* **2024**, *306*, 132583. [CrossRef]
87. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The performance of LSTM and BiLSTM in forecasting time series. In Proceedings of the 2019 IEEE International conference on big data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3285–3292.
88. Liu, K.; Peng, Q.; Che, Y.; Zheng, Y.; Li, K.; Teodorescu, R.; Widanage, D.; Barai, A. Transfer learning for battery smarter state estimation and ageing prognostics: Recent progress, challenges, and prospects. *Adv. Appl. Energy* **2023**, *9*, 100117. [CrossRef]
89. Alizadegan, H.; Rashidi Malki, B.; Radmehr, A.; Karimi, H.; Ilani, M.A. Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. *Energy Explor. Exploit.* **2024**, *43*, 01445987241269496. [CrossRef]
90. Zhao, D.; Li, H.; Zhou, F.; Zhong, Y.; Zhang, G.; Liu, Z.; Hou, J. Research progress on data-driven methods for battery states estimation of electric buses. *World Electr. Veh. J.* **2023**, *14*, 145. [CrossRef]
91. Begni, A.; Dini, P.; Saponara, S. Design and test of an lstm-based algorithm for li-ion batteries remaining useful life estimation. In Proceedings of the International Conference on Applications in Electronics Pervading Industry, Environment and Society, Genoa, Italy, 26–27 September 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 373–379.
92. Khan, M.K.; Abou Houran, M.; Kauhaniemi, K.; Zafar, M.H.; Mansoor, M.; Rashid, S. Efficient state of charge estimation of lithium-ion batteries in electric vehicles using evolutionary intelligence-assisted GLA–CNN–Bi-LSTM deep learning model. *Heliyon* **2024**, *10*, e35183. [CrossRef]
93. Sherkatghanad, Z.; Ghazanfari, A.; Makarenkov, V. A self-attention-based CNN-Bi-LSTM model for accurate state-of-charge estimation of lithium-ion batteries. *J. Energy Storage* **2024**, *88*, 111524. [CrossRef]
94. Xu, P.; Wang, C.; Ye, J.; Ouyang, T. State-of-charge estimation and health prognosis for lithium-ion batteries based on temperature-compensated Bi-LSTM network and integrated attention mechanism. *IEEE Trans. Ind. Electron.* **2023**, *71*, 5586–5596. [CrossRef]
95. Shiri, F.M.; Perumal, T.; Mustapha, N.; Mohamed, R. A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU. *arXiv* **2023**, arXiv:2305.17473.
96. Su, Y.; Kuo, C.C.J. Recurrent neural networks and their memory behavior: A survey. *APSIPA Trans. Signal Inf. Process.* **2022**, *11*, e26. [CrossRef]
97. Zhou, R.; Dai, X.; Lin, F.; Zhang, J.; Ma, H. TGT: Battery State of Charge Estimation with Robustness to Missing Data. In Proceedings of the 2024 CPSS & IEEE International Symposium on Energy Storage and Conversion (ISESC), Xi'an, China, 8–11 November 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 431–436.
98. Kanai, S.; Fujiwara, Y.; Iwamura, S. Preventing gradient explosions in gated recurrent units. In Proceedings of the NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
99. Salem, F.M.; Salem, F.M. Gated RNN: The Gated Recurrent Unit (GRU) RNN. In *Recurrent Neural Networks: From Simple to Gated Architectures*; Springer: Cham, Switzerland, 2022; pp. 85–100.
100. Rodriguez, S. Gated Recurrent Units-Enhancements and Applications: Studying Enhancements to Gated Recurrent Unit (GRU) Architectures and Their Applications in Sequential Modeling Tasks. *Adv. Deep Learn. Tech.* **2023**, *3*, 16–30.
101. Yang, F.; Li, W.; Li, C.; Miao, Q. State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network. *Energy* **2019**, *175*, 66–75. [CrossRef]
102. Cahuantzi, R.; Chen, X.; Güttel, S. A comparison of LSTM and GRU networks for learning symbolic sequences. In Proceedings of the Science and Information Conference, London, UK, 22–23 June 2023; Springer: Cham, Switzerland, 2023; pp. 771–785.
103. Li, X.; Ma, X.; Xiao, F.; Wang, F.; Zhang, S. Application of gated recurrent unit (GRU) neural network for smart batch production prediction. *Energies* **2020**, *13*, 6121. [CrossRef]

104. Zhang, Z.; Dong, Z.; Gao, M.; Han, Y.; Ji, X.; Wang, M.; He, Y. State-of-Charge Estimation of Lithium-ion Batteries Using the Nesterov Accelerated Gradient Algorithm Based Bi-GRU. In Proceedings of the 2020 8th International Conference on Power Electronics Systems and Applications (PESA), Hong Kong, China, 7–10 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4.
105. Jiao, M.; Wang, D.; Qiu, J. A GRU-RNN based momentum optimized algorithm for SOC estimation. *J. Power Sources* **2020**, *459*, 228051. [CrossRef]
106. Lu, J.; He, Y.; Liang, H.; Li, M.; Shi, Z.; Zhou, K.; Li, Z.; Gong, X.; Yuan, G. State of charge estimation for energy storage lithium-ion batteries based on gated recurrent unit neural network and adaptive Savitzky-Golay filter. *Ionics* **2024**, *30*, 297–310. [CrossRef]
107. Liao, L.; Li, H.; Shang, W.; Ma, L. An empirical study of the impact of hyperparameter tuning and model optimization on the performance properties of deep neural networks. *ACM Trans. Softw. Eng. Methodol. (TOSEM)* **2022**, *31*, 1–40. [CrossRef]
108. Zhan, Z.H.; Shi, L.; Tan, K.C.; Zhang, J. A survey on evolutionary computation for complex continuous optimization. *Artif. Intell. Rev.* **2022**, *55*, 59–110. [CrossRef]
109. Raiaan, M.A.K.; Mukta, M.S.H.; Fatema, K.; Fahad, N.M.; Sakib, S.; Mim, M.M.J.; Ahmad, J.; Ali, M.E.; Azam, S. A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access* **2024**, *12*, 26839–26874. [CrossRef]
110. Pozzato, G.; Allam, A.; Pulvirenti, L.; Negoita, G.A.; Paxton, W.A.; Onori, S. Analysis and key findings from real-world electric vehicle field data. *Joule* **2023**, *7*, 2035–2053. [CrossRef]
111. Fang, W.; Chen, H.; Zhou, F. Fault diagnosis for cell voltage inconsistency of a battery pack in electric vehicles based on real-world driving data. *Comput. Electr. Eng.* **2022**, *102*, 108095. [CrossRef]
112. Wong, R.H.; Sooriamoorthy, D.; Manoharan, A.; Binti Sariff, N.; Hilmi Ismail, Z. Balancing accuracy and efficiency: A homogeneous ensemble approach for lithium-ion battery state of charge estimation in electric vehicles. *Neural Comput. Appl.* **2024**, *36*, 19157–19171. [CrossRef]
113. Sesidhar, D.; Badachi, C.; Green, R.C., II. A review on data-driven SOC estimation with Li-Ion batteries: Implementation methods & future aspirations. *J. Energy Storage* **2023**, *72*, 108420.
114. Nagarale, S.D.; Patil, B. Accelerating AI-Based Battery Management System's SOC and SOH on FPGA. *Appl. Comput. Intell. Soft Comput.* **2023**, *2023*, 2060808. [CrossRef]
115. He, Z.; Shen, X.; Sun, Y.; Zhao, S.; Fan, B.; Pan, C. State-of-health estimation based on real data of electric vehicles concerning user behavior. *J. Energy Storage* **2021**, *41*, 102867. [CrossRef]
116. Hashemi, S.R.; Bahadoran Baghbadorani, A.; Esmaeeli, R.; Mahajan, A.; Farhad, S. Machine learning-based model for lithium-ion batteries in BMS of electric/hybrid electric aircraft. *Int. J. Energy Res.* **2021**, *45*, 5747–5765. [CrossRef]
117. Li, S.; He, H.; Zhao, P.; Cheng, S. Data cleaning and restoring method for vehicle battery big data platform. *Appl. Energy* **2022**, *320*, 119292. [CrossRef]
118. Akbar, K.; Zou, Y.; Awais, Q.; Baig, M.J.A.; Jamil, M. A machine learning-based robust state of health (SOH) prediction model for electric vehicle batteries. *Electronics* **2022**, *11*, 1216. [CrossRef]
119. Zhang, H.; Gui, X.; Zheng, S.; Lu, Z.; Li, Y.; Bian, J. BatteryML: An open-source platform for machine learning on battery degradation. *arXiv* **2023**, arXiv:2310.14714.
120. Zhang, D.; Zhong, C.; Xu, P.; Tian, Y. Deep learning in the state of charge estimation for li-ion batteries of electric vehicles: A review. *Machines* **2022**, *10*, 912. [CrossRef]
121. Rivera-Barrera, J.P.; Muñoz-Galeano, N.; Sarmiento-Maldonado, H.O. SoC estimation for lithium-ion batteries: Review and future challenges. *Electronics* **2017**, *6*, 102. [CrossRef]
122. Liu, F.; Yu, D.; Su, W.; Ma, S.; Bu, F. Adaptive Multitimescale Joint Estimation Method for SOC and Capacity of Series Battery Pack. *IEEE Trans. Transp. Electrif.* **2023**, *10*, 4484–4502. [CrossRef]
123. Timilsina, L.; Hoang, P.H.; Moghassemi, A.; Buraimoh, E.; Chamarthi, P.K.; Ozkan, G.; Papari, B.; Edrington, C.S. A real-time prognostic-based control framework for hybrid electric vehicles. *IEEE Access* **2023**, *11*, 127589–127607. [CrossRef]
124. He, H.; Meng, X.; Wang, Y.; Khajepour, A.; An, X.; Wang, R.; Sun, F. Deep reinforcement learning based energy management strategies for electrified vehicles: Recent advances and perspectives. *Renew. Sustain. Energy Rev.* **2024**, *192*, 114248. [CrossRef]
125. Guirguis, J.; Ahmed, R. Transformer-based deep learning models for state of charge and state of health estimation of li-ion batteries: A survey study. *Energies* **2024**, *17*, 3502. [CrossRef]
126. Salucci, C.B.; Bakdi, A.; Glad, I.K.; Vanem, E.; De Bin, R. A novel semi-supervised learning approach for State of Health monitoring of maritime lithium-ion batteries. *J. Power Sources* **2023**, *556*, 232429. [CrossRef]

127. Zafar, M.H.; Mansoor, M.; Abou Houran, M.; Khan, N.M.; Khan, K.; Moosavi, S.K.R.; Sanfilippo, F. Hybrid deep learning model for efficient state of charge estimation of Li-ion batteries in electric vehicles. *Energy* **2023**, *282*, 128317. [CrossRef]
128. Krzywanski, J.; Sosnowski, M.; Grabowska, K.; Zylka, A.; Lasek, L.; Kijo-Kleczkowska, A. Advanced computational methods for modeling, prediction and optimization—A review. *Materials* **2024**, *17*, 3521. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

The Role of Machine Learning in Enhancing Battery Management for Drone Operations: A Focus on SoH Prediction Using Ensemble Learning Techniques

Büşra Çetinus [1], Saadin Oyucu [2], Ahmet Aksöz [3] and Emre Biçer [1,*]

[1] Battery Research Laboratory, Faculty of Engineering and Natural Sciences, Sivas University of Science and Technology, 58010 Sivas, Türkiye; 200102036@sivas.edu.tr
[2] Department of Computer Engineering, Faculty of Engineering, Adıyaman University, 02040 Adıyaman, Türkiye; saadinoyucu@adiyaman.edu.tr
[3] Mobilers Team, Sivas Cumhuriyet University, 58050 Sivas, Türkiye; aaksoz@cumhuriyet.edu.tr
* Correspondence: emre.bicer@sivas.edu.tr

Abstract: This study considers the significance of drones in various civilian applications, emphasizing battery-operated drones and their advantages and limitations, and highlights the importance of energy consumption, battery capacity, and the state of health of batteries in ensuring efficient drone operation and endurance. It also describes a robust testing methodology used to determine battery SoH accurately, considering discharge rates and using machine learning algorithms for analysis. Machine learning techniques, including classical regression models and Ensemble Learning methods, were developed and calibrated using experimental UAV data to predict SoH accurately. Evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) assess model performance, highlighting the balance between model complexity and generalization. The results demonstrated improved SoH predictions with machine learning models, though complexities may lead to overfitting challenges. The transition from simpler regression models to intricate Ensemble Learning methods is meticulously described, including an assessment of each model's strengths and limitations. Among the Ensemble Learning methods, Bagging, GBR, XGBoost, LightGBM, and stacking were studied. The stacking technique demonstrated promising results: for Flight 92 an RMSE of 0.03% and an MAE of 1.64% were observed, while for Flight 129 the RMSE was 0.66% and the MAE stood at 1.46%.

Keywords: UAV data analysis; machine learning; regression models; Ensemble Learning; Li-ion

1. Introduction

"Drone" is a concept that refers to aircraft of various sizes and features that are controlled by remote control, also known as unmanned aerial vehicles (UAV) which are increasingly preferred in various civilian fields; these areas of use include activities such as live monitoring, expansion of wireless internet access areas, remote sensing, search-and-rescue activities, cargo distribution, security and surveillance, precision agriculture, and control of civil structures. Smart UAV technologies are among the important innovations in UAV technology, with advantages such as reducing risks and providing cost savings, especially in areas such as civil construction management [1].

UAVs' ease of deployment, low maintenance costs, superior mobility, and ability to stay airborne for long periods of time make them a viable option for a range of civilian and military applications. These features make UAVs ideal for these tasks [2]. UAVs primarily rely on either internal combustion engines or batteries for propulsion power sources. In this context, we specifically concentrate on battery-operated drones. Battery-operated drones offer several notable advantages, including zero emissions, simplified maintenance, and

straightforward controllability. However, they also face limitations, particularly concerning flight range and duration [3].

Energy consumption is one of the key factors of UAVs affecting efficient use, endurance, cost, and emissions. Energy consumption is also related to drone design, environment, drone dynamics, and operations [4]. In addition, the battery pack must be capable of providing sufficient electrical power to match the power consumed by aerodynamic drag, considering losses from the propeller, motor, and motor controller, and additionally the avionic system and other electronics [5]. The capacity of a battery is determined based on its application, and this capacity decreases as the battery ages. Additionally, as the battery ages, its impedance increases, which in turn reduces its ability to provide energy effectively.

The state of health (SoH) of a battery reflects its ability to store and supply energy based on its fundamental conditions, considering the energy and power requirements of the application. SoH serves as a crucial indicator for assessing the level of degradation experienced by a battery over time. Monitoring SoH in the operational processes of UAVs is critical for the successful continuation of missions. The condition of Li-ion batteries requires particular attention in order to ensure that UAVs operate uninterruptedly throughout their mission period and do not pose any safety risk in the air. Therefore, battery SoH estimation is a great challenge for operation endurance due to the aging mechanism of the battery. To tackle this situation, various methods have been developed as experimental techniques and adaptive battery models [6]. The determination of SoH in battery packs directly impacts the endurance and operational limits of UAVs. As batteries age, the operational time of drones decreases, affecting their overall performance. Additionally, because State of Charge (SoC) data are linked to SoH, inaccuracies in SoC measurements might mislead the estimation of remaining capacity, potentially leading to premature shutdown or loss of the drone during operation [7]. Andrioaia et al. discussed the necessity of estimating the Remaining Useful Life (RUL) and predicting the capacity of Li-ion batteries to prevent UAVs from losing autonomy, comparing the performance of three machine learning models—Support Vector Machine for Regression (SVMR), Multiple Linear Regression (MLR), and Random Forest (RF)—to estimate the RUL. This method might be implemented in UAVs' Predictive Maintenance (PdM) systems [8]. In addition, Zhang et al. addressed the issue of limited endurance in UAVs for the estimation of the state of charge (SOC) of Li-Po batteries, maximizing energy utilization and enhancing UAV endurance. They used the extended Kalman filter algorithm, based on an equivalent circuit model, to estimate SOC [9]. Accurate prediction of SoH can ensure UAVs battery lifespan and efficiency. Shibl et al. presented a UAV BMS based on machine learning (ML) models predicting SoC and estimating SoH using battery voltage, current, and ambient temperature data. Deep Neural Networks (DNN) and Long Short-Term Memory (LSTM) were used for SoC prediction, while Random Forest (RF) was used for SoH estimation [10].

In various applications, such as steady-level flight for fixed-wing aircraft or hovering conditions for multi-rotor vehicles, a constant power discharge process represented the actual battery loading; however, in the present study, a robust and diverse testing methodology was employed, involving the testing of the battery under various C rates at ten-second intervals to mimic different operational states of the drone [11]. This involved altering the position of an actual drone every 10 s while logging current and voltage data. By analyzing the discharge current associated with position changes, we determined the battery's behavior across different operational scenarios. Subsequently, these data were utilized to calculate SoH information of the battery. A single cell underwent rigorous testing for 285 cycles using a specialized battery test system, during which capacity data were recorded. The effect of discharge rate on available capacity was considered for variable current discharges. Afterwards, having the capacity values in response to discharge current for each cycle, machine learning algorithms were then applied to the data to calculate and assess the SoH information of the battery accurately to determine an efficient source of SoH information (Tables S1 and S2).

Our research group has recently published studies utilizing machine learning for estimating the SoH in electric vehicles (EVs). Specifically, it introduced a methodology for comparative analysis, focusing on classical and deep learning approaches, and assessed improvements made to LSTM and Bi-LSTM methods using evaluation metrics such as MSE, MAE, RMSE, and R-squared. The objective of the study was to drive advancements in EV technology by predicting Li-ion Battery (LIB) performance [12]. Additionally, another study was conducted as an assessment of traditional ML and advanced deep learning models for predictive modeling, encompassing techniques such as Random Forests, XGBoost, Elastic Net, MLP, CNNs, and RNNs, as well as hybrid models such as RNN-LSTM and CNN-LSTM. The findings underscored the potential of deep learning-based hybrid models in improving battery SoH assessments [7]. Accurately managing the capacity of L-ion batteries is crucial for enhancing the cost-effectiveness of large-scale energy storage systems. Therefore, in a study by Oyucu et al., various machine learning models—AdaBoost, gradient boosting, XGBoost, LightGBM, CatBoost, and Ensemble Learning—were used to predict the discharge capacity of LIBs. The study also utilized SHAP (Shapley additive explanations) values within the explainable artificial intelligence (XAI) framework to analyze key features affecting predictions [13].

Apart from the previous studies, in the present study, classical regression models such as Linear Regression (LR), Lasso, ElasticNet, and Ridge Regression (RR) were developed to predict the SoH of UAVs [14,15]. These models were calibrated using comprehensive experimental data obtained from UAV operations to interpret battery aging and operational performance. Following these initial models, Ensemble Learning techniques including Bagging, Boosting, and Stacking were employed to enhance the accuracy of SoH predictions [16]. These methods combined predictions from various regression models to achieve more robust and accurate predictions, thereby playing a critical role in the operational continuity and reliability of UAV batteries.

Two different evaluation metrics were used to measure the accuracy of SoH predictions for UAV batteries using the developed regression and Ensemble Learning models. The first evaluation metric is Root Mean Squared Error (RMSE), indicating how much deviation the model's predictions showed from the actual values by calculating the square root of the mean of squared errors, providing a broad perspective on the magnitude of the model's errors. The second metric was Mean Absolute Error (MAE), which was calculated by taking the mean of the absolute values of errors, presenting the average of the model's errors more directly. By evaluating the model's performance from different perspectives using these two metrics, it was demonstrated how reliable the developed prediction models were.

The results obtained indicated improvements among different models; however, they also demonstrated a decrease in success rates in cases where model complexity increased. This phenomenon was particularly pronounced in models using more complex Ensemble Learning techniques. As model complexity increases, theoretically, the model is expected to learn the data better and generalize; however, in some cases, overfitting issues have arisen, negatively impacting the model's performance on independent datasets. The results have shown that controlling and balancing complexity during the model development process is a significant factor directly affecting success rates. Therefore, while increasing model complexity, the importance of integrating techniques that maintain generalization ability has been emphasized.

This article evaluates the feasibility of machine learning techniques for developing and validating SoH predictions for UAV batteries. It extensively discusses the data collection and processing methods used for estimating battery SoH values. Factors such as temperature and cycle that affect battery performance are examined, and how these factors might be integrated into SoH prediction models is emphasized. The transition from simple regression models to more complex Ensemble Learning methods is described. Each model's implementation and the comparative analysis of their results identifies the advantages and limitations of the models.

Section 3 evaluates the prediction results obtained through different machine learning models using RMSE and MAE metrics to measure model performance. Section 5 focuses on the integration potential of the results into practical scenarios and discusses the limitations encountered in using machine learning methods for UAV battery SoH prediction. The article concludes with suggestions for future research, including the development of new algorithms to improve prediction model accuracy, experiments on more diverse datasets, and enhancing the applicability of SoH prediction models in real-world conditions.

The novel contributions of this study can be listed as below:

- Integration of diverse machine learning techniques: The combination of classical regression models and advanced Ensemble Learning techniques (Bagging, Boosting, and Stacking) is utilized to obtain accurate battery SoH estimation. In particular, the stacking of models is deployed for improving prediction accuracy and robustness by leveraging the strengths of several models.
- Robust testing approach: Within this study, a robust test method is designed for battery SoH estimation on various discharge rates and real operational scenarios. These scenarios mainly simulate the actual usage conditions of drones more accurately than standard tests by changing the drone's operational conditions every 10 s to measure battery behavior in dynamic scenarios. Consequently, this testing process crucially improves the feasibility of the SoH predictions to real-world drone environments.
- Evaluation with real UAV data: For the evaluation process, this study presents a real-world dataset providing calibration and experimental validation of the machine learning models under realistic operational factors, which is critical for practical applications in drone battery management.

2. Materials and Methods

This study aims to estimate the SoH for a UAV by using a literature dataset [13]. Rodrigues and co-workers conducted a study using DJI Matrice 100 Quadcopter (DJI, Shenzhen, China) to work with various parameters such as the battery voltage and current, altitude, wind conditions, wind speed, angular velocity, acceleration, and load within a specific area. In the context of this article, we used some of the datasets.

2.1. Battery Tests and Data Collection

The dataset in this article included not only voltage and current logs but also GPS data, range, altitude, acceleration, aerial time, wind speed, etc. [17]. A total of 209 flights were conducted in the study, and characteristics such as range and altitude were analyzed for these flights. However, the SoH was not shared after each flight. In this article, flights 92 and 129 were considered as having the longest range declared, and current and voltage information was utilized. Based on these data, each current value discharged from the battery was calculated to create a discharge algorithm. Using this algorithm, a test methodology with 285 cycles was conducted on high-energy 18650 batteries. The discharge algorithm, detailed in the supporting information, involved changing the current every ten seconds, repeated approximately 167 times throughout the test. The dataset from these 285 cycles, comprising discharge currents, voltage readings, and capacity values, was then used as input data for estimating the SoH by using machine learning techniques. In feature engineering, firstly data samples are extracted with respect to cycle feature values from 1 to 285. Consequently, a histogram based on the current (A) feature is extracted to analyze current values data distribution. Afterwards, a time-series analysis is conducted between the relative time feature and the capacity (Ah) target feature to analyze how the capacity feature values change over relative time feature values. Ultimately, a correlation matrix is generated between the current and capacity features to better understand the relationship information between capacity feature values and current feature values. Additionally, some 'null' valued samples are separated from the data to provide cleaner data as an input to the models.

2.2. Machine Learning Techniques

This study explores the application of machine learning techniques, specifically classical regression models such as LR, Lasso, ElasticNet, and RR, to predict the SoH of UAV batteries. These models were initially calibrated using comprehensive UAV operational data to interpret battery aging and performance. Subsequently, Ensemble Learning methods such as Bagging, Boosting, and Stacking were employed to enhance SoH prediction accuracy by combining predictions from multiple regression models. Evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were used to measure model accuracy, highlighting the reliability of the developed prediction models [7].

This study found that while model complexity might improve predictions, overly complex models might suffer from decreased success rates, especially with complex Ensemble Learning techniques, due to potential overfitting issues. The finding emphasizes the importance of controlling complexity to maintain generalization ability and improve success rates. The article discusses data collection, processing methods, and the integration of factors such as temperature and cycle into SoH prediction models. It also evaluates model performance using RMSE and MAE metrics and discusses limitations and future research directions, including the development of new algorithms, diverse dataset experiments, and improving the applicability of SoH prediction models in real-world conditions.

Finally, estimation of the battery capacity is the direct measure for determining the SoH state of the battery. Consequently, as capacity decreases, SoH also decreases, which shows the battery aging and degradation. Therefore, SoH estimation and evaluation are intrinsically accomplished by performing the battery capacity estimation and comparing the estimated battery capacity values with ground truth capacity values in the dataset with the machine learning techniques and evaluation metrics mentioned above.

2.2.1. Linear Regression

LR in the context of cycle, current, and capacity for batteries is a statistical modeling technique used to understand and predict how the number of cycles, the magnitude of discharge current during cycles, and the resulting capacity of the battery are related [18]. LR with cycle, current, and capacity for batteries involves modeling the relationship between these variables using a linear equation. The independent variables are cycle and discharge current, representing the number of cycles and the magnitude of discharge current, respectively. The dependent variable is capacity, indicating the amount of charge the battery may hold. By collecting data on cycle currents and corresponding capacities across multiple charge–discharge cycles, LR allowed us to determine the linear relationship between these variables. The LR model provided insights into how changes in cycle and current impacted the capacity of the battery. The resulting regression Equation (1) is given as

$$Capacity = \beta_0 + \beta_1 \times Cycle + \beta_2 \times Current + \epsilon \quad (1)$$

where

- *Capacity* is the battery's capacity;
- *Cycle* the number of cycles;
- *Current* is the discharge current during the cycles;
- β_0 is the intercept, representing the capacity when the cycle current is zero;
- β_1 and β_2 are the coefficients for cycle and current, respectively, indicating how they influence capacity;
- ϵ is the error term.

This equation helps in predicting the capacity of the battery for different cycle currents based on the observed data. LR analysis is valuable for optimizing battery performance, understanding degradation patterns, and designing efficient charging and discharging strategies.

2.2.2. Lasso Regression

Lasso regression is a method that applies L1 norm regularization to independent variables in regression models to approach their coefficients towards zero. This method reduces the complexity of variables in the model and sets the coefficients of insignificant variables to zero, thereby creating simpler and more generalizable models. Lasso stands for "Least Absolute Shrinkage and Selection Operator", referring to its ability to shrink coefficients while performing variable selection [19].

The main difference between LR (Linear Regression) and Lasso lies in Lasso's inclusion of a regularization term during model creation. LR aims to determine coefficients that minimize the relationship between independent and dependent variables, whereas Lasso additionally minimizes the sum of absolute values of coefficients. This additional term makes Lasso more resistant to overfitting and enables it to build models with fewer variables. These characteristics make Lasso more advantageous, especially when the number of variables exceeds the number of observations. Therefore, Lasso regression can be used to select the most effective features from battery data and accurately predict SoH.

2.2.3. ElasticNet Regression

ElasticNet regression performs variable selection and regularization in regression models by using a combination of both L1 and L2 norm regularization. This method combines the variable selection advantages of Lasso with the performance in grouped variables of Ridge regression. ElasticNet provides an effective solution, especially in cases where features exhibit high correlation with each other. The model is optimized with a penalty term that includes both the sum of squares of coefficients (L2) and the sum of absolute values (L1), making the model resistant to overfitting and improving its generalization ability [20].

Unlike Linear and Lasso regression, ElasticNet regression has a more robust modeling structure against the multitude of variables and correlation problems among them, thanks to its regularization terms. LR does not include regularization, so the problem of multicollinearity among variables may negatively impact the model's accuracy. ElasticNet minimizes these issues with regulatory terms, providing more stable and reliable predictions.

2.2.4. Ridge Regression

Ridge Regression (RR), applied to the context of cycle, current, and capacity for batteries, involves using a regularization technique to enhance the stability and accuracy of the LR model [14]. RR is a variant of LR that includes a regularization term to the ordinary least squares (OLS) regression equation. In the case of cycle, current, and capacity analysis for batteries, RR aims to model the relationship between these variables while addressing potential issues such as multicollinearity and overfitting.

The RR Equation (2) with cycle current and capacity as variables is represented as

$$Capacity = \beta_0 + \beta_1 \times Cycle + \beta_2 \times Current + \epsilon + \lambda \sum_{i=1}^{p} {\beta_i}^2 \quad (2)$$

where

- *Capacity* is the number of battery's capacity;
- *Cycle* is the number of cycles;
- *Current* is the discharge current during cycles;
- β_0 is the intercept;
- β_1 and β_2 are the coefficients for cycle and current, respectively;
- ϵ is the error term;
- λ is the regularization parameter that controls the strength of regularization;
- $\sum_{i=1}^{p} {\beta_i}^2$ is the sum of squared coefficients, excluding the intercept.

In our use of the RR method, the regularization term $\lambda \sum_{i=1}^{p} {\beta_i}^2$ was added to the least squares equation. This term penalized large coefficients, preventing overfitting and reducing the impact of multicollinearity in the model. The parameter λ determines the

extent of regularization, with higher values leading to greater shrinkage of coefficients. By incorporating RR into the analysis of cycle, current, and capacity data for batteries, we improved the model's robustness, handled correlated predictors more effectively, and produced more reliable predictions for battery performance.

2.3. Ensemble Modeling in Machine Learning

Ensemble methods in machine learning were designed to integrate predictions from multiple base models to achieve an overall performance improvement. This approach was used to enhance the model's predictive ability for unknown data and to better generalize learning on training data. The underlying motivation for using ensemble methods was to optimize the balance between bias and variance by combining a set of weak prediction models strongly, thus creating more accurate and reliable models. In this context, methods such as the Random Forest algorithm excel by combining multiple Decision Tree models to outperform a single Decision Tree [21].

Ensemble methods may be categorized into three main categories: Bagging, Boosting, and Stacking. Bagging trains many base models independently and creates final predictions by averaging their outputs. Boosting sequentially improves models and tries to correct errors of previous predictions at each step, reducing the model's bias. Stacking combines predictions from different models, using them as input for the final model to generate final predictions. Each method applies different strategies to enhance the model's robustness and performance on data, thus improving the model's reliability and accuracy [22].

3. Results

3.1. Exploratory Data Analysis (EDA) Studies on the Dataset

In this study, datasets from two different flights named Flight 92 and Flight 129 were thoroughly examined. The data from both flights were evaluated separately using exploratory data analysis (EDA) techniques. The analysis process was designed to understand the operational dynamics and potential anomalies specific to each flight.

The examined datasets contain 'CYCLE' numbers ranging from 1 to 285 for each flight, representing different time intervals during the flights. In total, the examined dataset for Flight 92 consists of 46,170 rows and 5 columns, and for Flight 129 it consists of 46,170 rows and 5 columns as well. Figure 1 illustrates the frequency distribution of current values.

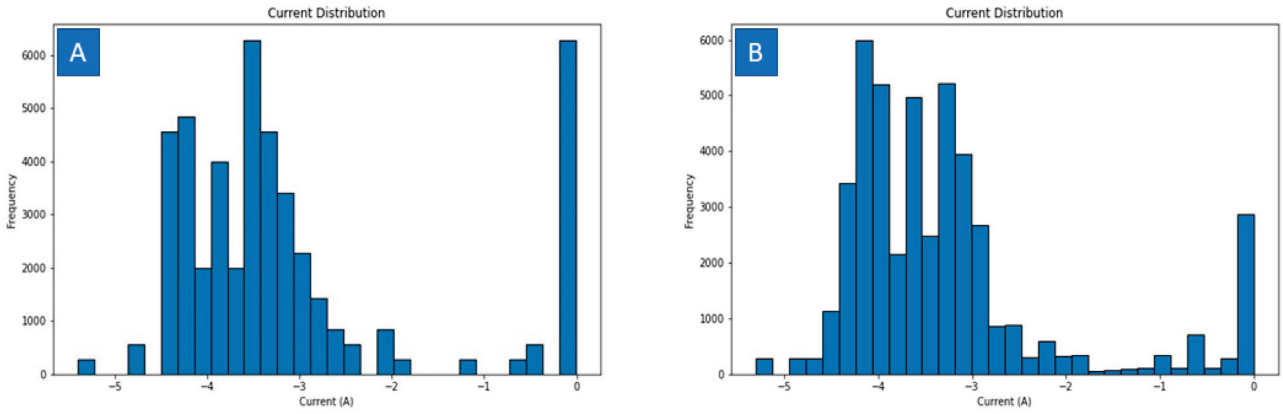


Figure 1. The frequency distribution of current values: (**A**) Flight 92, (**B**) Flight 129.

The current values provided in Figure 1 were classified into different intervals ranging from −5 amps to 0 amps. The shape of the histogram indicates that the majority of current values are concentrated within specific intervals. Particularly, the frequency between −3 and −4 amps shows the highest peak, with approximately 6000 observations within this range for Flight 92, while for Flight 192, the peak frequency occurs near −4 amps. This distribution concludes with a second peak near –1 amp, where approximately 6000 frequency

values were observed. On the other hand, lower frequencies were observed between −1 and −2 amps intervals. The correlation matrix of the data is provided in Figure 2.

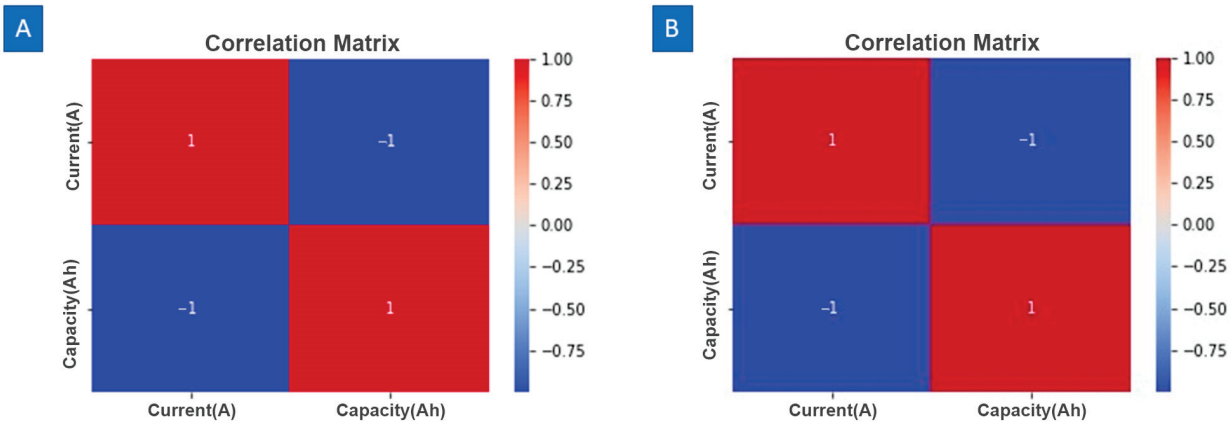


Figure 2. Correlation matrix of (**A**) Flight 92 and (**B**) Flight 129.

The correlation matrix provided in Figure 2 represents the correlation coefficients between current and capacity. These coefficients are shown as –1, indicating a perfect negative correlation between current and capacity. A perfect negative correlation indicates that as one variable increases, the other variable decreases in equal proportion (Equation (3)). This means that an increase in current corresponds to a decrease in capacity, or an increase in capacity corresponds to a decrease in current. This correlation analysis is especially helpful for further understanding the battery capacity degradation effect with respect to the current (A) feature values, accurate SoH modeling, improving predictive maintenance, and improving the accuracy for capacity prediction on models applied in this study.

$$Capacity(Ah) = \beta_0 + \beta_1 * Current(A) \quad (3)$$

According to the correlation matrix, there is a strong negative relationship between "Current (A)" and "Capacity (Ah)" (correlation coefficient of −1). In Equation (3), Capacity (Ah) represents the battery capacity, while Current (A) refers to the current. β_0 is the intercept term, indicating the capacity value when the current is zero. β_1, on the other hand, represents the slope of the relationship between current and capacity, and according to the correlation matrix, this slope is negative.

In this case, the correlation coefficient of −1 implies that the value of β_1 will also be negative. This means that as the current increases, the capacity decreases, resulting in a negative slope for the relationship. Therefore, a clear consistency can be observed between the correlation matrix presented in Figure 2 and Equation (3).

3.2. Dataset Splitting

When examining the data from both flights, it was observed that there were no missing values. The data were divided as shown in Table 1. The training set contains the data used in the learning process of the model. Each observation contains two features (columns). These features represent various dynamics and operational parameters of the flight. Training the model on these data allowed the algorithm to learn from the flight data and generate predictions that might be applied to future data. The "Holdout" set, on the other hand, was a separate data group allocated to evaluate the performance of the model, which was not used during the training process. This set was used to test how the model performed with data it had not seen during training.

Table 1. Dataset splitting values.

Flight 92		**Flight 129**	
Training Set	Holdout Set	Training Set	Holdout Set
30,934.2	15,237.2	30,934.2	15,237.2

3.3. Model Development

The process, starting with the preparation and preprocessing of the dataset, encompassed tasks such as standardization and normalization of the data required for training the models. These steps are of critical importance to enable the model to learn more effectively and make more accurate predictions. Subsequently, prediction models were developed using various regression models and Ensemble Learning techniques. These techniques included LR, Lasso, RR, and ElasticNet regression models. Each model was calibrated using specific error minimization techniques and regularization methods. The hyperparameters of the developed models are provided in Table 2.

Table 2. Model hyperparameters for LR, Lasso, ElasticNet and RR.

Model	**Hyperparameter**	**Value**
LR	fit_intercept	True
	normalize	False
	Positive	False
Lasso	Alpha	0.0005
	fit_intercept	True
	random_state	42
	max_iter	1000
ElasticNet	Alpha	0.0005
	l1_ratio	0.9
	random_state	42
RR	Alpha	2.0
	fit_intercept	True
	Normalize	False

In Table 2, for each type of model, the important hyperparameters and their values are provided. For the LR model, the fit_intercept parameter was set to True, indicating that the model would compute the intercept term. Additionally, the Positive parameter was set to False, allowing coefficients to take negative values. For the Lasso Regression model, the Alpha value was set to 0.0005, indicating the amount of L1 regularization applied in the model. The fit_intercept was also set to True to compute the intercept term of the model. The random_state was set to 42 to determine the randomness of the model, ensuring its reproducibility. Moreover, the max_iter was set to 1000 to allow a sufficient number of iterations during the fitting process of the model. In the RR model, the Alpha value was set to 2.0, indicating the strength of the L2 regularization applied. The hyperparameter settings of each model significantly affected its performance and generalization ability. Therefore, the parameters were carefully selected and tuned.

Development of Ensemble Learning Models

Within the scope of the study, models were developed using regression models by employing Bagging, Boosting, and Stacking methods [21]. In the Bagging process, a Bagging regressor was built on top of a base estimator. The hyperparameters and regression models used in developing Ensemble models are provided in Tables 3 and 4.

Table 3. Model hyperparameters for bagging, boosting and stacking.

Model	Hyperparameter	Value	Predictions
Bagging	n_estimators max_samples bootstrap	10 1.0 True	LR, Lasso, ElasticNet, RR
Boosting	Table 4	Table 4	
Stacking	cv passthrough final_estimator	5 True linear_regression	ElasticNet, RR, Gradient Boosting Regressor

Table 4. Model hyperparameters for GBR, XGBoost and LightGBM.

Boosting Model	Hyperparameter	Value
Gradient Boosting Regressor (GBR)	n_estimators	3000
	learning_rate	0.05
	max_depth	4
	max_features	Sqrt
	min_samples_leaf	15
	min_samples_split	10
	loss	Huber
eXtreme Gradient Boosting Regressor (XGBoost)	colsample_bytree	0.4603
	gamma	0.0468
	learning_rate	0.05
	max_depth	3
	min_child_weight	1.7817
	n_estimators	2200
	reg_alpha	0.4640
	reg_lambda	0.8571
	subsample	0.5213
Light Gradient Boosting Machine Regressor (LightGBM)	num_leaves	5
	learning_rate	0.05
	n_estimators	720
	max_bin	55
	bagging_fraction	0.8
	bagging_freq	5
	feature_fraction	0.2319
	feature_fraction_seed	9
	bagging_seed	9
	min_data_in_leaf	6

In Table 3, the hyperparameters specified for three different Ensemble Learning techniques—Bagging, Boosting, and Stacking—are provided. Key parameters and their values for each technique, along with the summary of prediction models used with these techniques, are outlined. For the Bagging method, the fundamental parameters were specified as n_estimators, max_samples, and bootstrap. The max_samples value was set to 1.0, indicating that the maximum number of samples used for training each base estimator was all samples. The prediction models used in this method were listed as LR, Lasso, ElasticNet, and RR. The Stacking technique combined predictions from different prediction models to produce final predictions through a new meta-model. The cv parameter represented the number of folds used for cross-validation and was set to 5 in this case. The passthrough parameter was set to True, allowing the original features of base estimators to be directly passed through to the final predictor. LR was used as the final estimator, and predictions made with this method utilized ElasticNet, RR, and GBR models. The parameters and values specified for Boosting were provided in Table 5 in this table. Boosting was based on the principle of sequentially improving weak predictors to build a strong model.

Table 5. Comparison of the results of machine learning models.

Model	Flight 92		Flight 129	
	RMSE (%)	**MAE (%)**	**RMSE (%)**	**MAE (%)**
LR	0.04	0.03	0.97	0.20
Lasso	4.56	3.41	4.32	2.98
ElasticNet	4.11	3.07	3.91	2.69
RR	0.04	0.03	0.97	0.20

4. Experimental Results

This section thoroughly examines experimental analyses and the results obtained from evaluating the performance of different regression models. LR, Lasso, ElasticNet, and RR models were utilized within the scope of this study. Each model was trained and tested on appropriate datasets with predefined hyperparameters. RMSE and MAE metrics were employed to measure model accuracy. The predictive performances of the models were comparatively assessed.

4.1. Results of Machine Learning Techniques

During the analysis process, potential risks such as overfitting were considered alongside the predictive capabilities of the models on real data. Specifically, it was observed that the Lasso and ElasticNet regression models provided more stable results by reducing model complexity through regularization terms. Increasing the alpha parameter in RR significantly influenced the model's robustness and generalization ability, improving its adaptation to different datasets. The results of the models developed using machine learning techniques are provided in Table 5.

A performance evaluation conducted on four different regression models (LR, Lasso, ElasticNet, and RR) is presented in Table 5. The results were measured on two different flight datasets named Flight 92 and Flight 129. The LR and RR models exhibited significantly low RMSE and MAE values for both flight datasets. These results indicated high prediction accuracy and good fit of these two models to the data. The findings suggested that the performance of models might vary on datasets from flights with similar characteristics, emphasizing the importance of careful model selection based on flight attributes. This analysis provided important insights into assessing the suitability and generalizability of the model.

4.2. Results of Ensemble Modeling in Machine Learning

The findings obtained from applying ensemble modeling techniques in machine learning and the performance of these approaches were examined. Ensemble modeling aimed to improve overall prediction success by combining multiple prediction models. These techniques worked towards balancing individual model errors and reducing issues such as overfitting to provide more robust and reliable predictions. Within the scope of this study, ensemble methods such as Bagging, Boosting, and Stacking were used, and the effects of these methods on different datasets were analyzed comparatively. The results obtained are provided in Table 6.

Table 6. Result of ensemble modeling in machine learning.

Model	Flight 92		Flight 129	
	RMSE (%)	**MAE (%)**	**RMSE (%)**	**MAE (%)**
Bagging	2.20	1.64	2.22	1.46
GBR	0.14	0.03	0.67	0.10
XGBoost	24.17	19.87	24.42	18.36
LightGBM	1.21	0.23	1.73	0.65
Stacking	0.03	1.64	0.66	1.46

The Bagging method aimed to reduce the model's variance to achieve more stable results, while Boosting focused on correcting errors of individual predictors and thereby reduces the model's bias in this process. The Stacking method, on the other hand, combined predictions obtained from various models through a meta-regressor, ensuring the best possible balance of these predictions. Each method offered different advantages for specific data structures and types of problems.

The experimental results indicated that ensemble models generally provided predictions with higher accuracy compared to individual regression models. Particularly in noisy and complex datasets, the performance of ensemble methods was significantly superior (Figure 3). Moreover, models obtained through the application of these techniques have demonstrated better generalization ability on new and unseen data, enhancing the model's robustness. Figure 3 compares the performances of various machine learning models for the 92nd and 129th flights with their RMSE and MAE values. The GBR and Stacking models had the lowest error rates for both flights and provided the most accurate predictions. In particular, the GBR model gave the most successful results with RMSE 0.14 and MAE 0.03 for the 92nd flight, while it also exhibited very low errors for the 129th flight. On the other hand, the XGBoost model had the highest RMSE and MAE values for both flights and showed the lowest performance compared to the other models. The Bagging and LightGBM models also showed good performance with acceptable error rates.

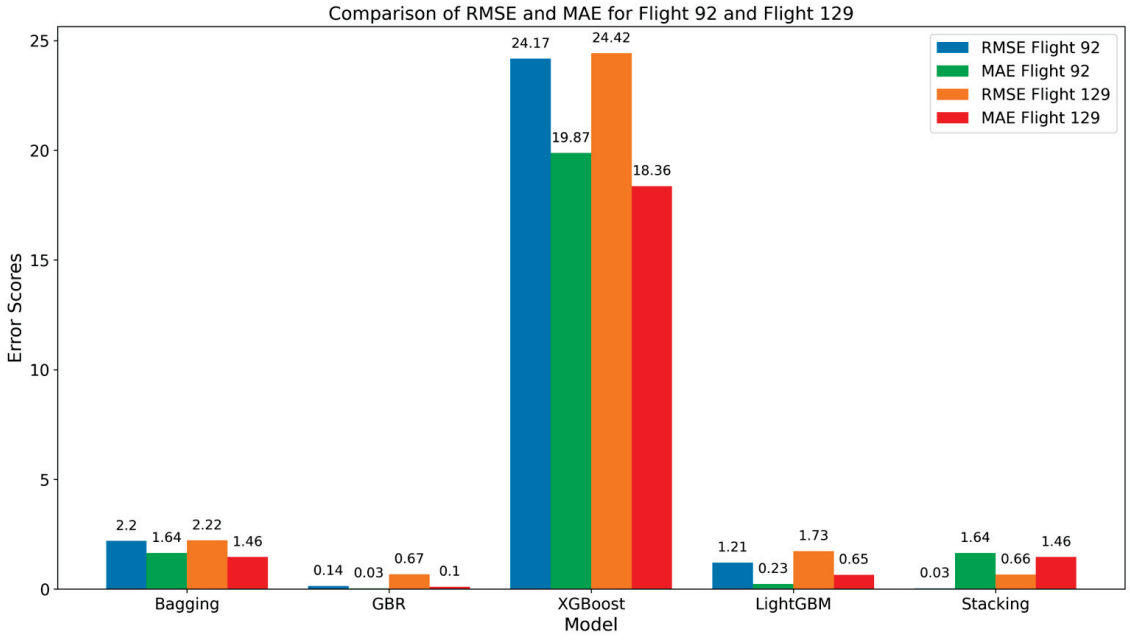


Figure 3. Comparison of RMSE and MAE for Flight 92 and Flight 129.

Consequently, ensemble modeling techniques offered an effective strategy in managing model complexity and maximizing prediction success in machine learning projects. These approaches optimized performance by striking a good balance between various prediction models and leveraging the strengths of each model. The results showed that the choice of hyperparameters and model selection played a crucial role in prediction accuracy. While each regression model offered different advantages for specific data structures and prediction needs, comprehensive testing of these models allowed for the selection of the most suitable model tailored to the application domain. This experimental study provided valuable insights into the development of advanced modeling techniques and the determination of the best model specific to the data.

5. Discussion

This section evaluates the comparative performance of various machine learning techniques applied to battery state estimation, as summarized in Table 7. This analysis highlighted the relative effectiveness of different approaches. Such comparative information with similar studies in the literature provided important insights to advance the field of battery management systems by identifying the most promising methodologies and areas for further research and improvement.

Table 7. Studies in the literature and model performances.

Reference	Method *	Chemistry	Dataset	MAE	RMSE
[23]	RLS	90 Ah $LiFePO_4$	OCV, SOC, and temperature	1.8%	2.3%
[24]	CNN + GRU (FUDS)	1.3 Ah 18650 NMC	Voltage, current, temperature	1.26%	1.54%
[25]	IB-ELM	2 Ah 18650 $LiFePO_4$	Charge, NASA dataset, aging test, voltage, current, temperature	0.010–0.034	0.014–0.039
[26]	LSTM	2.9 Ah 18650 NMC	Voltage, current, temperature	1.39%	1.7%
[27]	GRU	2.3 Ah 26650 $LiFePO_4$	Voltage, temperature	0.49%	0.64%
[28]	DNN	2.3 Ah 26650 $LiFePO_4$	Voltage, current, temperature	-	3.68%
[29]	SVR	2 Ah 18650 $LiFePO_4$	Oxford and NASA dataset	-	3.62% and 2.49%
Proposed Method	Ensemble Modeling	1.5 Ah 18650 NMC	Voltage, current, discharge capacity	0.03%	1.64%

* RLS: recursive least squares, CNN + GRU: convolutional neural network-gated recurrent unit, IB-ELM: improved blinex-extreme learning machine, LSTM: long short-term memory, DNN: deep neural network, SVR: support vector regression.

In Table 7, various machine learning models, ranging from simple regressions to complex deep learning architectures, were evaluated for their accuracy in predicting the SoH status of LIBs. The RLS method applied to the 90 Ah $LiFePO_4$ battery showed reasonable accuracy with MAE and RMSE values of 1.8% and 2.3%, respectively. In particular, the combined CNN+GRU model trained under the Federal Urban Driving Schedule (FUDS) provided improved performance with lower error measures (1.26% MAE and 1.54% RMSE). On the other hand, the performance of the DNN and Support Vector Regression (SVR) models, although generally robust, exhibited higher RMSE values under certain conditions, highlighting potential limitations in their adaptability or tuning.

Our proposed Ensemble Modeling approach, which combines multiple predictive models to leverage the strengths of individual models, achieved an MAE of only 0.03 and an RMSE of 1.64. This approach not only improved the prediction accuracy but also provided robustness to various operating conditions and dataset variabilities, validating the potential of ensemble methods in complex system predictions such as battery states.

The ensemble modeling techniques examined in this study, Bagging, Boosting, and Stacking, were applied to different regression problems, and the performance of each method was comprehensively evaluated. The results shed light on how these techniques might impact the overall success of the model and when each technique should be preferred. It was observed that the Bagging method is particularly effective in datasets with high variance problems. This method increased sample diversity and ensured independence among base predictors, thus minimizing the risk of overfitting. However, Bagging may sometimes be insufficient in reducing bias, which may lead the model to make biased errors in certain cases.

Boosting was used to improve weak predictors and correct sequential errors. While this method balanced low bias and high variance, it might increase complexity in some cases, leading to overfitting. Boosting was suitable for situations designed to manage low bias and variance that needed improvement. Stacking optimized performance by combining the strengths of various prediction models and integrating these predictions through a meta-regressor. This approach maintained generalization ability while increas-

ing model complexity and synergies obtained by combining various prediction models enhanced overall prediction success. The application of Stacking might be highly effective when correct parameters and predictors were chosen, but incorrect implementation might complicate the model's interpretability and practical applicability.

6. Conclusions

The in-depth analysis of ensemble modeling techniques discussed in this study demonstrated how these approaches might be effectively used in various machine learning problems. While Bagging, Boosting, and Stacking methods offered advantages in certain data structures and learning scenarios, the challenges and limitations encountered during their application were also considered. Stacking stands out, especially with a 0.03 and 0.66 RMSE performance.

The Bagging methodology was effective in improving overall performance by reducing the model's variance, particularly in variable datasets where it prevented overfitting, but falls short in low bias problems. Boosting achieved noticeable successes, especially when working with weak learners, by enhancing model accuracy through sequential improvements, but carries risks of overfitting in complex data structures. Stacking created a robust meta-model by integrating predictions from different models and achieved high success rates through synergy from the combination of various prediction models.

In conclusion, the findings of this study shed light on the evolution of ensemble modeling techniques and their future applications. It was expected that with the integration of artificial intelligence algorithms, processes such as automatic adjustment of model parameters and dynamic model selection would be automated. The development of stronger hybrid models through the combination of different ensemble techniques offers the potential to achieve better results in complex data structures. The integration of ensemble modeling approaches with deep learning frameworks requires more attention to their environmental and ethical dimensions, especially regarding data privacy and algorithmic bias issues. In conclusion, ensemble modeling techniques will continue to be a significant part of ongoing developments and innovations in the field of machine learning. These techniques play a critical role in optimizing model performance and producing generalizable solutions on larger datasets, both in academic research and industrial applications.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/batteries10100371/s1. Table S1: Algorithm created for Flight No. 92 in battery test system; Table S2: Algorithm created for Flight No. 129 in battery test system.

Author Contributions: Conceptualization, E.B. and S.O.; methodology, S.O. and E.B.; software, S.O.; validation, B.Ç., S.O., and A.A.; formal analysis, B.Ç.; investigation, B.Ç. and S.O.; resources, A.A.; data curation, S.O.; writing—original draft preparation, E.B. and S.O.; writing—review and editing, E.B. and S.O.; visualization, S.O.; supervision, E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The research was conducted collaboratively by the Battery Research Group at Sivas University of Science and Technology and MOBILERS team at Sivas Cumhuriyet University.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, C.; Zheng, Z.; Xu, T.; Guo, S.; Feng, S.; Yao, W.; Lan, Y. YOLO-based UAV technology: A review of the research and its applications. *Drones* **2023**, *7*, 190. [CrossRef]
2. Hayat, S.; Yanmaz, E.; Muzaffar, R. Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 2624–2661. [CrossRef]
3. Saravanakumar, Y.N.; Sultan, M.T.H.; Shahar, F.S.; Giernacki, W.; Łukaszewicz, A.; Nowakowski, M.; Holovatyy, A.; Stępień, S. Power sources for unmanned aerial vehicles: A state-of-the art. *Appl. Sci.* **2023**, *13*, 11932. [CrossRef]
4. Hong, D.; Lee, S.; Cho, Y.H.; Baek, D.; Kim, J.; Chang, N. Least-energy path planning with building accurate power consumption model of rotary unmanned aerial vehicle. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14803–14817. [CrossRef]
5. Avanzini, G.; de Angelis, E.L.; Giulietti, F. Optimal performance and sizing of a battery-powered aircraft. *Aerosp. Sci. Technol.* **2016**, *59*, 132–144. [CrossRef]
6. Berecibar, M.; Gandiaga, I.; Villarreal, I.; Omar, N.; Van Mierlo, J.; Van Den Bossche, P. Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renew. Sustain. Energy Rev.* **2016**, *56*, 572–587. [CrossRef]
7. Oyucu, S.; Dümen, S.; Duru, İ.; Aksöz, A.; Biçer, E. Discharge capacity estimation for Li-ion batteries: A comparative study. *Symmetry* **2024**, *16*, 436. [CrossRef]
8. Andrioaia, D.A.; Gaitan, V.A.; Culea, G.; Banu, I.V. Predicting the RUL of Li-ion batteries in UAVs using machine learning techniques. *Computers* **2024**, *13*, 64. [CrossRef]
9. Zhang, H.; Zhou, M.; Lan, X. State of charge estimation algorithm for unmanned aerial vehicle power-type lithium battery packs based on the extended kalman filter. *Energies* **2024**, *12*, 3960. [CrossRef]
10. Shibl, M.M.; Ismail, L.S.; Massoud, A.M. A machine learning-based battery management system for state-of-charge prediction and state-of-health estimation for unmanned aerial vehicles. *J. Energy Storage* **2023**, *66*, 107380. [CrossRef]
11. Chopra, I. Small UAS and delivery drones: Challenges and opportunities the 38th Alexander A. Nikolsky honorary lecture. *J. Am. Helicopter Soc.* **2021**, *66*, 1–41. [CrossRef]
12. Hassan, T.J.; Jangula, J.; Ramchadra, A.R.; Sugunaraj, N.; Chandar, B.S.; Rajagopalan, P.; Rahman, F.; Ranganathan, P.; Adams, R. UAS-guided analysis of electric and magnetic field distribution in high-voltage transmission lines (Tx) and multi-stage hybrid machine learning models for battery drain estimation. *IEEE Access* **2024**, *12*, 4911–4939. [CrossRef]
13. Oyucu, S.; Doğan, F.; Aksöz, A.; Biçer, E. Comparative analysis of commonly used machine learning approaches for Li-ion battery performance prediction and management in electric vehicles. *Appl. Sci.* **2024**, *14*, 2306. [CrossRef]
14. Oyucu, S.O.; Ersöz, B.; Sagiroglu, S.; Aksöz, A.; Biçer, E. Optimizing Lithium-ion battery performance: Integrating machine learning and explainable AI for enhanced energy management. *Sustainability* **2024**, *16*, 4755. [CrossRef]
15. Tina, G.M.; Ventura, C.; Ferlito, S.; De Vito, S. A state-of-art-review on machine-learning based methods for PV. *Appl. Sci.* **2021**, *11*, 7550. [CrossRef]
16. Rincy, T.N.; Gupta, R. Ensemble learning techniques and its efficiency in machine learning: A survey. In Proceedings of the 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 28–29 February 2020. [CrossRef]
17. Rodrigues, T.A.; Patrikar, J.; Choudhry, A.; Feldgoise, J.; Arcot, V.; Gahlaut, A.; Lau, S.; Moon, B.; Wagner, B.; Scott Mathews, H.; et al. In-flight positional and energy use data set of a DJI matrice 100 quadcopter for small package delivery. *Sci. Data* **2021**, *8*, 6–13. [CrossRef] [PubMed]
18. Maulud, D.; Abdulazeez, A.M. A Review on linear regression comprehensive in machine learning. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 140–147. [CrossRef]
19. Ranstam, J.; Cook, J.A. LASSO regression. *Br. J. Surg.* **2018**, *105*, 1348. [CrossRef]
20. Chamlal, H.; Benzmane, A.; Ouaderhman, T. Elastic net-based high dimensional data selection for regression. *Expert Syst. Appl.* **2024**, *244*, 122958. [CrossRef]
21. Banfield, R.E.; Hall, L.O.; Bowyer, K.W.; Kegelmeyer, W.P. A comparison of decision tree ensemble creation techniques. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 173–180. [CrossRef]
22. Satoła, A.; Satoła, K. Performance comparison of machine learning models used for predicting subclinical mastitis in dairy cows: Bagging, boosting, stacking and super-learner ensembles versus single machine learning models. *J. Dairy Sci.* **2024**, *107*, 3959–3972. [CrossRef] [PubMed]
23. Duong, V.H.; Bastawrous, H.A.; See, K.W. Accurate approach to the temperature effect on state of charge estimation in the $LiFePO_4$ battery under dynamic load operation. *Appl. Energy* **2017**, *204*, 560–571. [CrossRef]
24. Huang, Z.; Yang, F.; Xu, F.; Song, X.; Tsui, K.L. Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* **2019**, *7*, 93139–93149. [CrossRef]
25. Ma, W.; Cai, P.; Sun, F.; Wang, X.; Gong, J. State of health estimation for Lithium-ion batteries based on extreme learning machine with improved blinex loss. *Int. J. Electrochem. Sci.* **2022**, *17*, 221170. [CrossRef]
26. Ma, L.; Hu, C.; Cheng, F. State of charge and state of energy estimation for lithium-ion batteries based on a long short-term memory neural network. *J. Energy Storage* **2021**, *37*, 102440. [CrossRef]
27. Xiao, B.; Liu, Y.; Xiao, B. Accurate state-of-charge estimation approach for lithium-ion batteries by gated recurrent unit with ensemble optimizer. *IEEE Access* **2019**, *7*, 54192–54202. [CrossRef]

28. How, D.N.T.; Hannan, M.A.; Lipu, M.S.H.; Sahari, K.S.M.; Ker, P.J.; Muttaqi, K.M. State-of-charge estimation of Li-ion battery in electric vehicles: A deep neural network approach. *IEEE Trans. Ind. Appl.* **2020**, *56*, 5565–5574. [CrossRef]
29. Tian, J.; Xiong, R.; Shen, W. State-of-health estimation based on differential temperature for lithium ion batteries. *IEEE Trans. Power Electron.* **2020**, *35*, 10363–10373. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

Battery Scheduling Optimization and Potential Revenue for Residential Storage Price Arbitrage

Nerijus Paulauskas * and Vsevolod Kapustin

Department of Computer Science and Communications Technologies, Vilnius Gediminas Technical University (VILNIUS TECH), 10105 Vilnius, Lithuania; vsevolod.kapustin@vilniustech.lt
* Correspondence: nerijus.paulauskas@vilniustech.lt

Abstract: Residential energy storage systems offer significant potential for price arbitrage by capitalizing on fluctuations in electricity prices throughout the day. This study investigates the potential revenue from optimal battery scheduling for residential storage in different north-western European electricity price zones during 2023. Using Nord Pool day-ahead prices, we applied an optimization model to determine the revenue for two types of batteries: 5 kW/10 kWh and 10 kW/10 kWh. The analysis considered battery capacity, charging and discharging efficiency, and maximum charge/discharge rates. Our results show notable variations in potential revenue across different regions, with the Baltic states demonstrating the highest revenue potential. The findings indicate that while 10 kW batteries can generate higher total revenue, 5 kW batteries are more efficient in terms of revenue per cycle. These regional disparities underscore the need for targeted incentives and policies to enhance the economic viability of residential energy storage. The research results provide valuable insights into optimizing residential battery storage for price arbitrage, offering guidance for consumers, policymakers, and energy providers to maximize economic benefits in various electricity markets.

Keywords: battery scheduling optimization; price arbitrage; day-ahead electricity prices; residential energy storage

1. Introduction

Price arbitrage is based on profiting by charging the battery during periods when electricity prices are low and discharging the battery when prices are high. The greater the price difference between peak and valley prices, the greater the economic benefit will be. Electricity prices fluctuate during the day according to the so-called California "duck curve" [1,2], where the x-axis represents hours and the y-axis represents the quantity of electricity net demand. Fluctuations in electricity prices during the day depend on the dynamics of electricity demand and supply. The integration of renewable energy sources, such as solar and wind power plants, into the grid has a significant impact on the variability of electricity prices. First peak demand is in the morning as consumers wake up and businesses start. From the morning until midday, the net demand decreases as solar production increases. The second demand peak appears in the evening as solar energy production wanes due to the sunset. Energy storage is a notable way to flatten the mismatch level of the overall load that emerges at high solar penetrations.

Based on information reported by Bloomberg in [3], cumulative residential energy storage capacity to be installed globally in 2030 will reach 93 GW/196 GWh. Already, in Germany and Italy, over 70% of new home solar systems have batteries attached. This allows to shift the use of daytime generated solar power to the evening and night time. Customers are concerned with installing residential batteries mainly for three reasons: a desire to increase solar self-consumption, to reduce their bills, and to have backup power. Self-consumption allows to reduce the grid load, especially at peak times, and thereby

reduce the necessity of grid expansion. Changing policies from net metering to net billing and reducing feed-in tariffs are other reasons to incline consumers to install residential batteries and maximize self-consumption.

Energy prices are rising while installation costs of energy storage systems are decreasing. Bloomberg's annual battery price survey finds a 14% drop from 2022 to 2023 [4]. Electricity customers can take advantage of price arbitrage by installing residential batteries without solar power. This type of installation is good for power backup and resilience.

The study presented in [5] proposes two day-ahead battery-behavior-aware operation scheduling strategies to maximize profitability and longevity in residential grid-connected applications with dynamic electricity pricing. The strategy with a low charge/discharge rate extends the battery lifetime but the profit is negative of −3 EUR/kWh/year. Strategies with high and moderate charge/discharge rates resulted in positive profits of 8.3 EUR/kWh/year and 9.2 EUR/kWh/year, despite having shorte-r battery lifetimes estimated at 10.1 years and 13.6 years, respectively.

Four energy arbitrage strategies, including mirror arbitrage, back-to-back arbitrage, and statistical-based operation strategy, are applied to determine the strategy with the best performance for a BESS that would participate in the Colombian market [6]. A financial evaluation shows that the income obtained from battery energy storage systems when only performing energy arbitrage in the Colombian electricity market does not compensate for the investment costs.

The economic viability of hydrogen storage and Li-ion batteries based on price arbitrage in the day-ahead Poland market is presented in [7]. The net present value for Li-ion technology is significantly higher than for hydrogen but the results show that both technologies are unprofitable in the current price fluctuations and capital cost of the systems.

The study presented in [8] proposes a stochastic bidding approach for grid-level battery storage arbitrage profit maximization using day-ahead and real-time market prices. Case studies show that profits using the MILP (Mixed Integer Linear Programming) algorithm are higher than the LP (Linear Programming) algorithm. The profit of the same capacity but lower power batteries is lower as the battery requires a longer time to charge/discharge.

Increasing the penetration level of renewable energy poses challenges to the stability of the power system [9,10]. The integration of the smart grid with renewable energy sources, opportunities, and challenges are discussed in [11]. The advantages and disadvantages of hybrid wind and solar energy integration systems are discussed in [12]. Wind power plants cannot produce electricity when the wind is not strong enough and solar power plants cannot generate at night or on a cloudy day. The opposite case is also possible when, on windy or very sunny days, the amount of energy produced exceeds the net demand. Battery energy storage systems provide a range of benefits and support functions to the power grid, including frequency regulation, peak shaving operating reserves, and others. Batteries can charge and discharge quickly, making them ideal for balancing the grid on demand or production side. Balancing services can be differentiated into frequency containment reserve (FCR), automatic frequency restoration reserve (aFRR), and manual frequency restoration reserve (mFRR) [13]. The study presented in [14] proposes a two-stage optimization technique to schedule batteries owned by commercial consumers through an aggregator, considering energy arbitrage and peak shaving benefits at the consumer level and the frequency regulation benefit at the grid level. Consumers have no possibility to participate directly in the balancing services because the size of residential batteries is too small; therefore, this is organized through an aggregator. The possibility to participate in balancing services gives additional revenues for residential battery owners [15].

The economics for residential storage in Europe are often poor without incentives; therefore, many governments subsidize consumer adoption of batteries. However, customers lack good data on the actual potential revenue [3]. This paper provides a comparative study of the potential revenue for residential battery price arbitrage in the different north-western European power market electricity price zones of 13 countries. The revenue

is obtained by using optimal battery scheduling during day-ahead prices in 2023. We used 5 kW/10 kWh and 10 kW/10 kWh batteries for optimal battery scheduling. The model evaluates the battery capacity, charging, and discharging efficiency, maximum charging, and discharging rate. The obtained results indicate the price zones with the highest and lowest possible revenue and show the revenue and battery cycle differences between 5 kW and 10 kW power batteries.

The remainder of this paper is organized as follows. Section 2 describes the day-ahead prices in 2023 and the battery scheduling optimization model. Section 3 explains the residential battery price arbitrage scenario and analyses the results by applying the optimization model presented in Section 2. Section 4 provides the concluding remarks.

2. Materials and Methods

2.1. Energy Price Data

For the price arbitrage study, we used Nord Pool day-ahead prices in 2023 [16]. Nord Pool is the north-western European power market covering 20 bidding zones in 13 countries. Usually, the day-ahead prices are announced to the market at 12:45 CET or later. Knowing the 24-h day-ahead prices makes it possible to build the optimal schedule for the battery to maximize the revenue during this period. Table 1 summarizes the mean, maximum, minimum, standard deviation, and total sum of day-ahead market prices of 21 zones in 2023.

Table 1. Summary of the Nord Pool day-ahead prices in 2023.

Zone	Country	EUR/MWh				
		Mean	**Max**	**Min**	**SD**	**Sum, Thousand**
AT	Austria	**102.14**	437.47	**−500.00**	44.38	**894.76**
BE	Belgium	97.27	330.36	−120.00	45.87	852.08
DK1	Denmark	86.83	524.27	−440.10	48.82	760.60
DK2		81.25	524.27	−60.04	50.12	711.78
EE	Estonia	90.79	**777.18**	−60.04	55.79	795.31
FI	Finland	56.47	**777.18**	**−500.00**	**56.71**	494.65
FR	France	96.86	276.12	−134.94	45.53	848.46
DE-LU	Germany–Luxembourg	95.18	524.27	**−500.00**	47.58	833.74
LV	Latvia	93.89	**777.18**	−56.55	54.55	822.51
LT	Lithuania	94.44	**777.18**	−56.55	54.87	827.31
NL	Netherlands	95.82	463.77	**−500.00**	49.05	839.36
NO1	Norway	66.95	332.00	−61.84	44.68	586.49
NO2		79.45	261.85	−61.84	36.29	695.95
NO3		38.55	332.00	−10.06	32.79	337.74
NO4		29.95	332.00	−10.06	26.21	262.34
NO5		67.05	261.85	−6.62	43.40	587.32
SE1	Sweden	39.97	332.00	−60.04	34.17	350.15
SE2		39.98	332.00	−60.04	34.16	350.20
SE3		51.70	332.00	−60.04	45.33	452.90
SE4		64.88	332.00	−60.04	50.64	568.34

Bold values highlight the highest values in a column.

The highest electricity mean price of 102.14 EUR/MWh was in Austria. Also, we see that this corresponds to the highest total sum of all hourly prices per year: 894.76 thousand EUR. The price range in the day-ahead market is limited to −500 (minimum) and 4000 (maximum) EUR/MWh. The highest electricity price, 777.18 EUR/MWh, was reached in four zones (EE, FI, LV, and LT) at the same hour on 22 November 2023. It is worth

mentioning that the absolute maximum of prices was reached on 17 August 2022 in the Baltic region (EE, LV, and LT). That day, the price hits 4000 EUR/MWh for one hour, even after the activation of Lithuanian peak load reserves. The absolute price minimum was spotted for two days in 2023. The first time, it was on Sunday 2 July 2023 and included three zones: NL, AT, and DE-LU. High wind and solar generation coincided with low demand. The second time, it was related to an electricity trading error. Kinect Energy sent an incorrect bid for Finland for delivery on 24 November. Starting from 14:00 CET until midnight, the price was at −500 EUR/MWh.

The European Clean Power Pathways Explorer by EMBER provides insights into the energy landscapes of various European countries [17]. For example, Austria relies significantly on renewable energy, particularly hydroelectric power. The availability of water for hydropower can fluctuate, leading to price variability. Belgium's energy mix mainly comprises nuclear and gas-powered generation. The Baltic states rely on imported energy, contributing to higher price variability. Energy prices are affected by their interconnections with neighboring markets and the integration of wind and solar power. Germany is undergoing an energy transition to increase its share of renewable energy, particularly wind and solar. A large portion of energy in Denmark comes from wind turbines. Wind power's intermittent nature leads to price fluctuations. France has a large nuclear fleet that helps to stabilize electricity prices. Periods of maintenance or reduced nuclear output can cause price spikes. The Netherlands is transitioning from natural gas to more renewable energy sources. The integration of wind and solar power contributes to price variability. Finland's energy mix includes nuclear, hydro, and increasing wind power. Norway's electricity prices are influenced by its reliance on hydropower. Sweden's energy mix includes nuclear, hydro, and increasing wind power.

The standard deviation (SD) shows the amount of variation about its mean price. We can expect that a higher standard deviation indicates a higher potential revenue of price arbitrage. But the standard deviation of the whole year's prices is not a good indicator. As prices are announced 24 h ahead, the optimal schedule of the battery can be built based on this period; therefore, the better indicator would be the standard deviation of the day prices. Figure 1 shows the boxplot of the standard deviation of day prices in 2023 ordered from the highest to the lowest. Based on the values presented in Table 1, the highest and the lowest prices were in Finland. The highest standard deviation was in Finland too. However, by analyzing the standard deviation of day prices (Figure 1), the Finland zone is twelfth. The highest standard deviation of day prices is in the Baltic region, positioning Estonia first, Lithuania second, and Latvia third. Based on that, we can make an assumption that revenue also will be the highest in the Baltic region. The exact revenue values of optimal battery scheduling for price arbitrage are presented and discussed in Section 3.

A small deviation in electricity prices refers to a scenario where there is relatively little fluctuation or volatility in electricity costs over time. This stability can offer several advantages, particularly in the context of energy planning, investment, and consumption. Predictable prices help in assessing the true cost of renewables versus conventional energy sources. This facilitates the achievement of grid parity. Grid parity is the point at which the cost of generating electricity from renewable sources becomes equal to or less than the cost of purchasing electricity from the grid. The timing of reaching grid parity is closely related to the price of electricity in a given country. Higher electricity prices make renewable energy more competitive sooner, accelerating the achievement of grid parity. Conversely, lower electricity prices can delay the integration.

The total sum reflects the cumulative cost of electricity over the year, indicating the overall economic impact on consumers and industries.

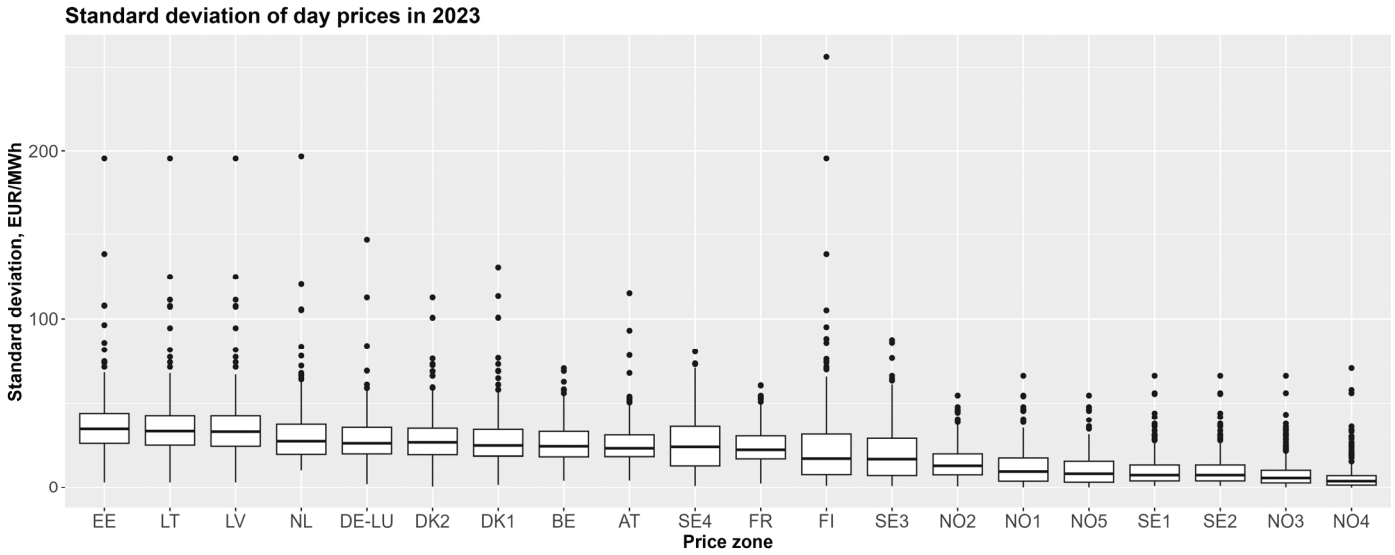


Figure 1. Standard deviation of day prices in 2023 by price zone.

2.2. Battery Scheduling Optimisation Model

The battery scheduling optimization model is based on models presented in [6,18–21]. The main purpose of the implemented model is to determine the optimal battery scheduling for the day-ahead prices. The model accounts for the battery capacity, charging and discharging efficiency, maximum charging and discharging rate, initial charge state, state of the charge, and discharge limit. The time resolution of the model is one hour.

2.2.1. Objective Function

The objective function (Equation (1)) seeks to maximize revenue from the price arbitrage over a day-ahead prices period, as follows:

$$Maximize \sum_{t=1}^{T} C(t)\left(E_{dis}(t) - E_{chg}(t)\right), \quad (1)$$

where t is the time period of one hour, T is the day-ahead optimization period, $C(t)$ is the energy price at time period t, and $E_{dis}(t)$ and $E_{chg}(t)$ correspond to the discharge and charge energy at each time period t.

2.2.2. Constraints

A set of constraints are applied to keep battery operation within its physical limitations during the optimization period.

The state of charge (SOC) of the battery at the time t depends on the previous state of charge and the charging $P_{chg}(t)$ and discharging $P_{dis}(t)$ power, as shown by Equation (2). Δt refers to time between $t-1$ and t.

$$SOC(t) = SOC(t-1) + \left(P_{chg}(t) - P_{dis}(t)\right)\Delta t, \quad (2)$$

The charging and discharging power are also limited by the maximum allowed battery charging and discharging rate (R_{chg_max} and R_{dis_max}) and battery capacity B_{cap}:

$$P_{chg}(t) \leq B_{cap} \cdot R_{chg\_max} \cdot I_{chg}(t) \quad (3)$$

$$P_{dis}(t) \leq B_{cap} \cdot R_{dis\_max} \cdot I_{dis}(t) \quad (4)$$

where $I_{chg}(t)$ and $I_{dis}(t)$ are binary variables used to ensure that the battery can only charge or discharge at each time period t. It is constrained by Equation (5).

$$I_{chg}(t) + I_{dis}(t) \leq 1 \tag{5}$$

The state of charge of the battery at each time period must be within its minimum and maximum limits, as shown in Equation (6),

$$SOC_{min} \leq SOC(t) \leq SOC_{max} \tag{6}$$

The actual energy that is used to charge the battery until a specific *SOC* level or energy, that is removed from the battery and sold to the grid depends on the battery efficiency. The required energy to charge the battery from $SOC(t-1)$ until $SOC(t)$ with efficiency η_{chg} is given by Equation (7).

$$E_{chg}(t) = \frac{\Delta t P_{chg}(t)}{\eta_{chg}} = \frac{\Delta SOC(t)}{\eta_{chg}} \tag{7}$$

The actual energy, which is sold to the grid when the battery state of charge decreases by $\Delta SOC(t)$ with efficiency η_{dis}, is given by Equation (8).

$$E_{dis}(t) = \Delta t P_{dis}(t) \eta_{dis} = \Delta SOC(t) \eta_{dis} \tag{8}$$

3. Results

The potential revenue for residential battery price arbitrage in different north-western European power market electricity price zones is obtained using two the same 10 kWh capacity but different charge–discharge rate batteries. The first battery has 5 kW (0.5 C) charge–discharge power. It takes two hours to fully charge this battery, which means that the daily optimal battery scheduling plan requires at least two hours with the lowest price. The situation is analogous to the discharge battery. Two hours with the highest price are needed to discharge the battery to the minimum level of the specified state of the charge (SOC).

We can make an assumption that the same capacity but 1 C rate battery gives more revenue as the time needed to charge and discharge the battery is shorter. In this case, the advantage is that the battery can be fully charged in an hour, i.e., at the lowest price and fully discharged in an hour at the highest price.

In both cases, we use the following main battery parameters:

- Capacity: 10 kWh;
- Charging/Discharging power: 5 kW (*for Case 1*) and 10 kW (*for Case 2*);
- Charging/Discharging efficiency: 95%;
- SOC upper limit: 90%;
- SOC lower limit: 10%;
- Optimization scenario: max revenue.

3.1. Case Study with a 5 kW/10 kWh Battery

The revenue of price arbitrage with a 5 kW/10 kWh battery in 2023 is shown in Figure 2. The highest revenue is in the Baltic region with the top 1 of 409.78 EUR in the Estonian price zone and the top 2–3 with almost identical revenue in Lithuania and Latvia. The lowest possible revenue is in Norway zone 4. Norway has five price zones where the difference between the lowest revenue zone NO4 and the highest revenue zone NO2 is 2.6 times or 75.44 EUR per year. The similar situation is in Sweden where the SE3 zone gives a revenue of 184 EUR compared to the SE2 zone with a revenue of 86.86 EUR. It can be seen that the price zone revenues highly correlate with the standard deviation of day prices. The Person correlation coefficient is equal to 0.99. But the correlation coefficient between the revenues and standard deviation of the hole year prices is equal to 0.85. This

shows that the standard deviation of day prices is a much better indicator of the possible arbitrage revenue than the standard deviation of the whole year prices.



Figure 2. Price arbitrage revenue of a 5 kW/10 kWh battery in 2023.

The revenue of price arbitrage is a result of battery scheduling during the day. This means that the battery goes through a certain number of charge and discharge cycles during the day. The total number of battery cycles during the whole year is presented in Figure 3. The mean number of cycles per year in all price zones is 523 or 1.4 cycles per day. We can also see that a higher number of cycles does not necessarily mean a higher revenue. Norway, NO2, with a lower number of cycles, gives higher revenue compared with Sweden's SE1 and SE2 zones. The battery cycles of the Netherlands' price zone are lower than in Finland's price zone but the revenue is 68% higher. It depends on the day-ahead price variation. Bigger differences between the hourly prices give higher revenue at the end of the day. It is also worth mentioning that the price difference should compensate for the charging and discharging efficiency of the battery. The optimal battery scheduling model evaluates the battery efficiency and charges or discharges the battery only when the revenue is positive.

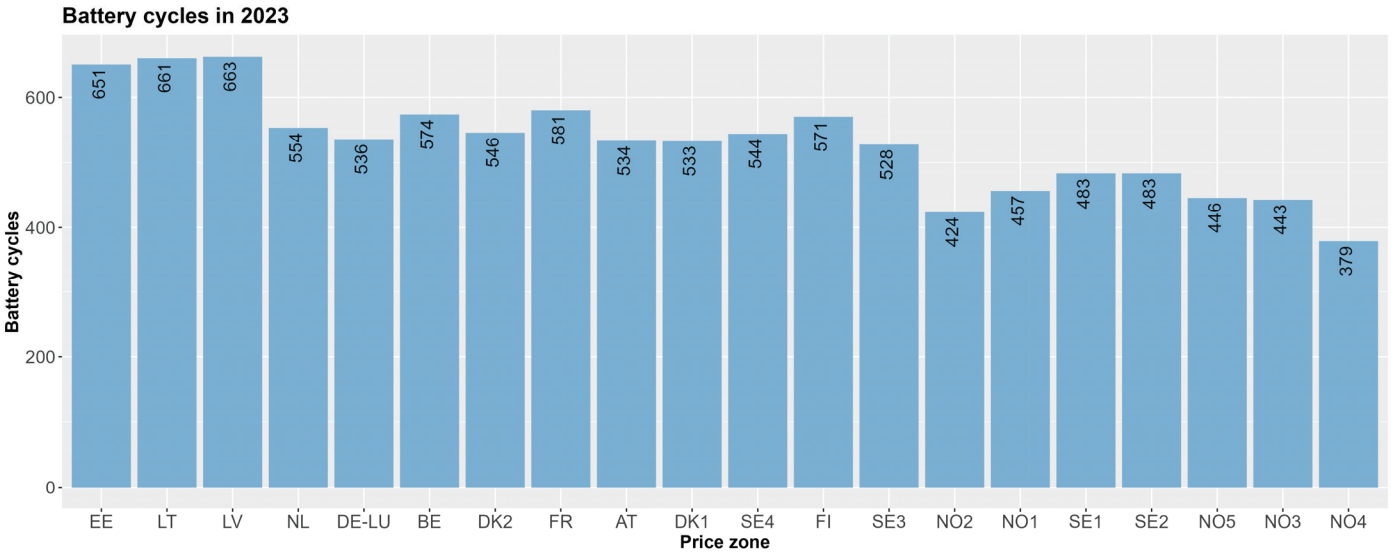


Figure 3. Total battery cycles of a 5 kW/10 kWh battery in 2023.

Revenue per battery cycle is presented in Figure 4. We can see that revenue per one battery cycle in the Netherlands is very close to that in the Baltic states and is equal to 0.601 EUR, which is not the case comparing only the yearly revenue. Therefore, the revenue per one battery cycle is a better criterion to compare the possible revenue of price arbitrage in different price zones. The revenue per battery cycle divided by battery capacity gives the revenue per battery cycle for 1 kWh, which allows to compare revenue of batteries with different capacities. For example, the revenue of one battery cycle gives 1.8 cent/1 kWh in Sweden zone 1 (SE1).

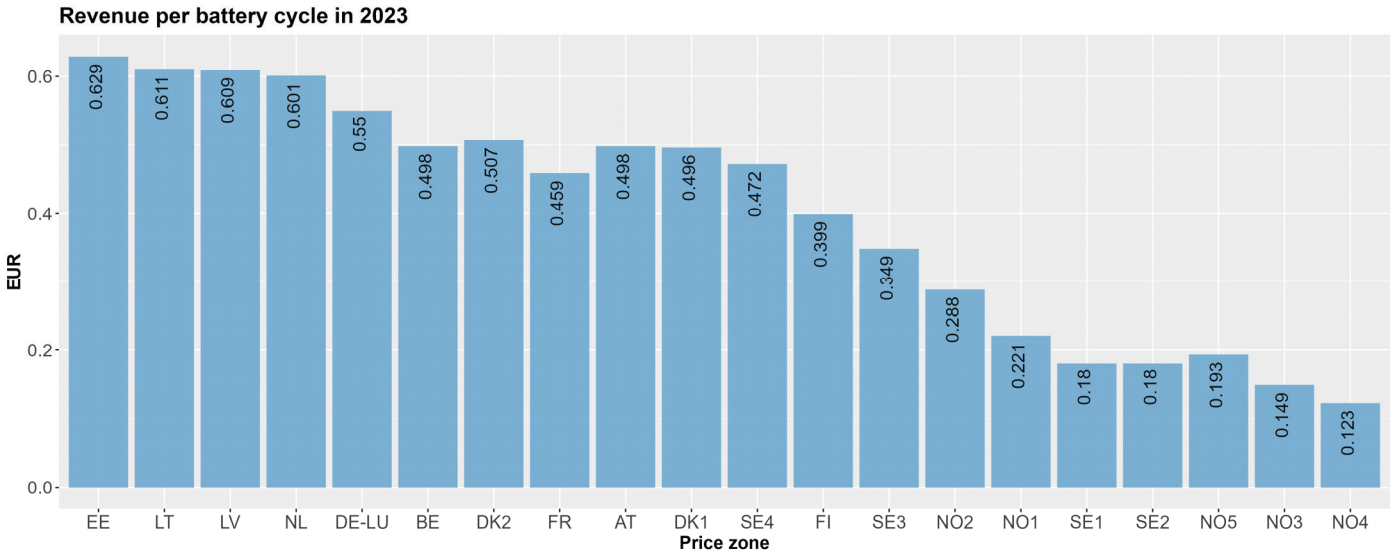


Figure 4. Revenue per battery cycle of a 5 kW/10 kWh battery in 2023.

Making an assumption that revenue per one battery cycle will not change during all of the battery lifetime and assuming that the number of battery cycles is equal to 6000, then the total possible revenue of 5 kW/10 kWh battery in Estonia will be 3774 EUR and only 738 EUR in Norway zone 4 (NO4). The revenue differs more than 5 times. It should also be noted that this revenue does not include taxes on electricity distribution, fixed periodic electricity supplier taxes, power cost, and others, which are different in different price zones.

3.2. Case Study with 10 kW/10 kWh Battery

The 10 kW/10 kWh battery can be fully charged or discharged in one hour. This gives an opportunity to take advantage of cases when there is only one hour during a day with the lowest or highest price. The 5 kW/10 kWh battery needs two hours in a row with the lowest and the highest prices to be charged and discharged to achieve the maximum revenue.

Figure 5 shows the revenue increase of 10 kW power battery compared to a 5 kW battery. The highest revenue increase of 11.6% is in Finland's price zone and the lowest, with 6.14%, is in Germany–Luxembourg. The average increase in all price zones is 7.98%. The biggest monetary increase is in Latvia and is equal to 52.53 EUR and the smallest one is only 3.87 EUR in Norway zone 4 (NO4). The average monetary increase in all price zones is 20.68 EUR.

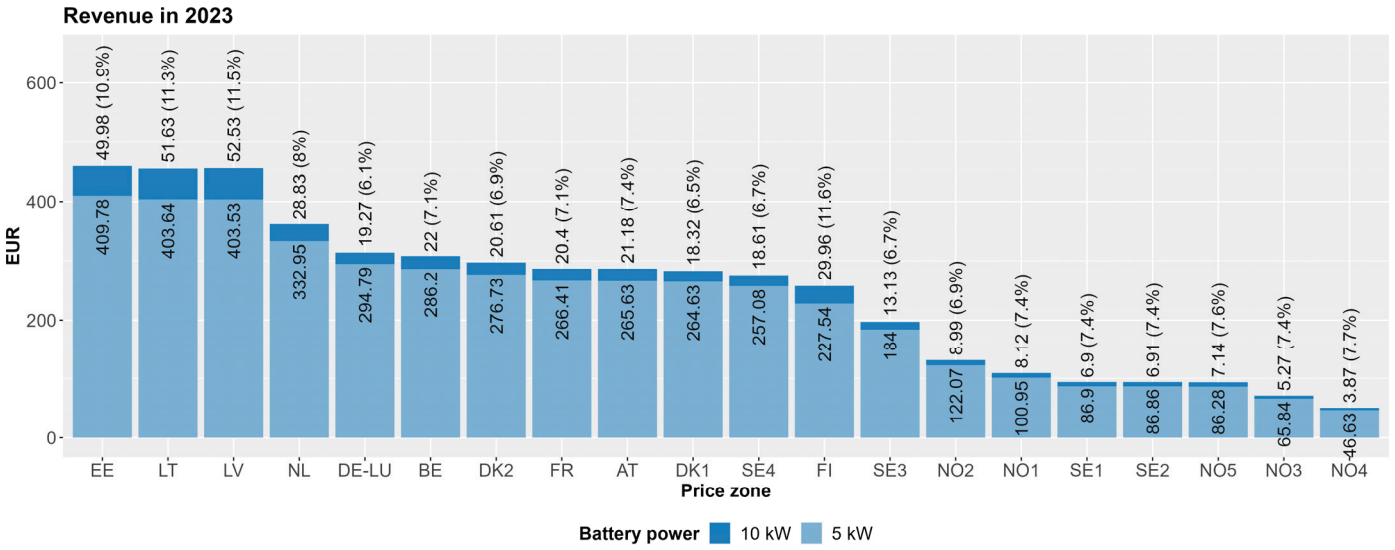


Figure 5. Revenue increase by using a 10 kW power battery compared to a 5 kW power battery.

Figure 6 illustrates the percentage increase in total battery cycles when using a 10 kW power battery instead of a 5 kW power battery. The data show that on average, the 10 kW battery results in a 9.38% increase in battery cycles across various price zones. This increase is greater than the corresponding revenue increase, indicating that although the 10 kW battery is cycled more frequently, the revenue per cycle decreases on average (Figure 7). The highest revenue increase per one battery cycle using 10 kW power battery is in the Netherlands (0.0124 EUR) and in Austria (0.0123 EUR). The highest decrease of −0.0212 EUR is in the Estonia price zone and in Latvia −0.0180 EUR. In Norway zone 2 (NO2), there is no difference between using 10 kW or 5 kW power battery. Prices zones where the increase in cycles does not correspond to a substantial increase in revenue indicate inefficiencies in the use of the higher power battery.

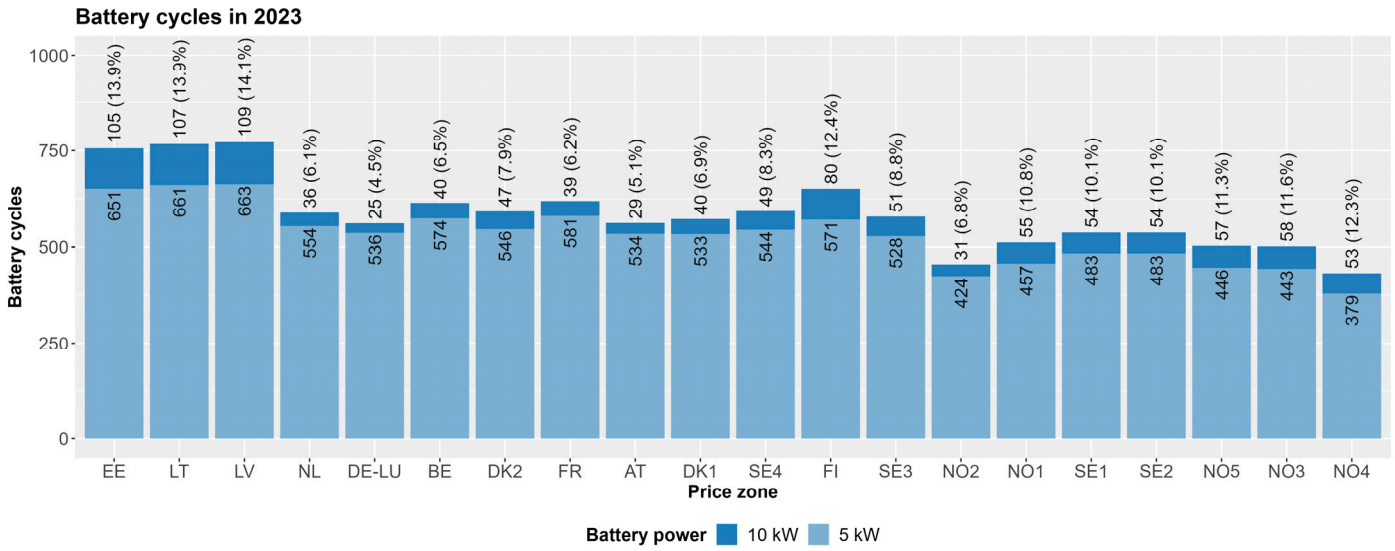


Figure 6. Total battery cycles increase by using a 10 kW power battery compared to a 5 kW power battery.

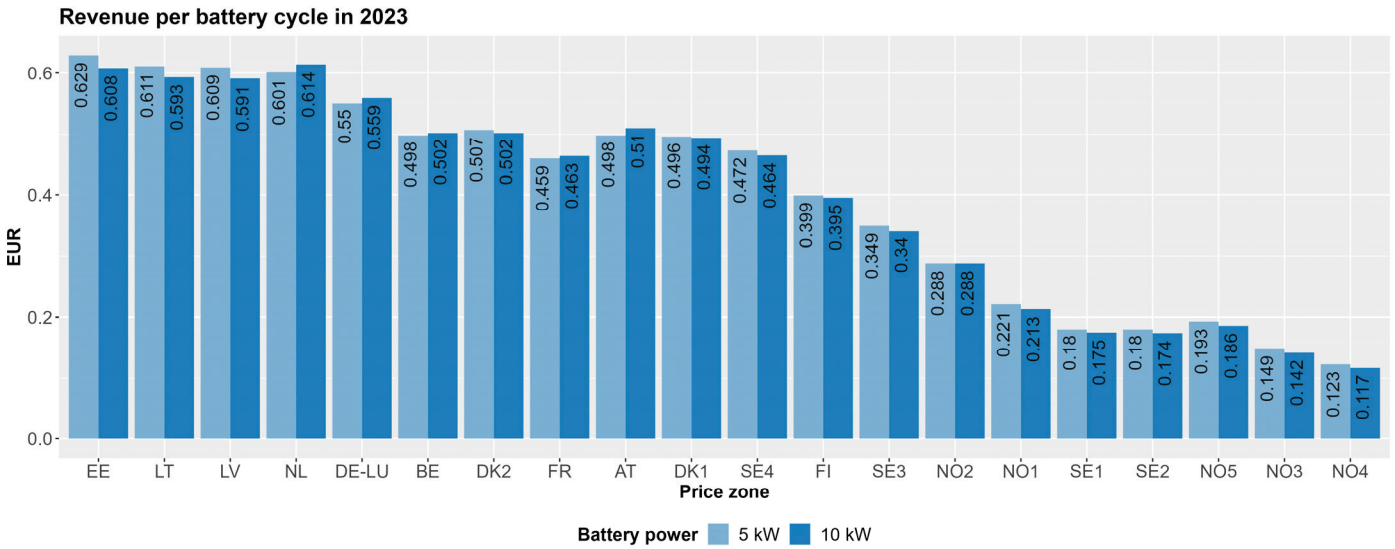


Figure 7. Revenue per 10 kWh battery cycle by using a 10 kW power battery compared to a 5 kW power battery.

The comparative analysis of results reveals that the economic viability of residential energy storage for price arbitrage varies significantly across different north-western European electricity price zones. The 10 kW power batteries generally result in higher total revenue but not necessarily higher revenue per cycle, indicating a trade-off between battery usage and efficiency. Regions with significant price fluctuations, like the Baltic states, offer the best opportunities for maximizing revenue through optimal battery scheduling. Conversely, areas with smaller price variations may see less benefit from higher power batteries, underscoring the need for targeted policies and incentives to enhance the economic appeal of residential energy storage systems.

4. Conclusions

This study provides a comparative analysis of the potential revenue from residential battery price arbitrage across different north-western European power market electricity price zones. By employing optimal battery scheduling during the day-ahead prices of 2023, we evaluated the performance of 5 kW/10 kWh and 10 kW/10 kWh batteries. The results reveal significant differences in potential revenue among various price zones, highlighting both the opportunities and challenges associated with residential energy storage in the region.

The analysis identified specific price zones with the highest and the lowest potential revenues from battery arbitrage, as follows:

- A 5 kW/10 kWh battery gives the highest revenue in the Baltic region with the top 1 of 409.78 EUR in the Estonian price zone and the top 2–3 with almost identical revenue in Lithuania and Latvia;
- The lowest possible revue is in Norway zone 4 (NO4);
- The mean number of cycles per year in all price zones is 1.4 cycles per day;
- The 10 kW battery results on an average 7.98% increase in total revenue across various regions. The average monetary increase in all price zones is 20.68 EUR;
- The findings indicate that while 10 kW batteries can generate higher total revenue, 5 kW batteries are more efficient in terms of revenue per cycle.

This study provides insights into the economic performance of residential battery storage price arbitrage in different north-western European electricity markets. The methodology and results can guide future studies and inform consumers and energy providers on optimizing residential energy storage for maximum economic benefit. As future work, we

plan to analyze battery scheduling strategies when battery storage is coupled with a PV generation system.

Author Contributions: Conceptualization, N.P.; methodology, N.P.; validation, N.P. and V.K.; formal analysis, N.P. and V.K.; investigation, N.P. and V.K.; writing—original draft preparation, N.P. and V.K.; writing—review and editing, N.P. and V.K.; visualization, N.P. and V.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: According Nord Pool's policy the price data cannot be published or redistribute to third parties. We provided an url in the references, number [16], where interested parties can find all necessary information about Nord Pool market data (https://data.nordpoolgroup.com/auction/day-ahead/prices, accessed on 10 July 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

Zone Definitions	**Description**
AT	Austria
BE	Belgium
DK1	Denmark
DK2	Denmark
EE	Estonia
FI	Finland
FR	France
DE-LU	Germany–Luxembourg
LV	Latvia
LT	Lithuania
NL	Netherlands
NO1	Norway
NO2	Norway
NO3	Norway
NO4	Norway
NO5	Norway
SE1	Sweden
SE2	Sweden
SE3	Sweden
SE4	Sweden
Symbols	**Description**
t	Steps of time
Δt	Duration of time period t, 1 h
T	Day-ahead optimization period, 24 h
$C(t)$	Energy price at time period t, EUR/kWh
$E_{chg}(t)$	Charge energy at each time period t, kWh
$E_{dis}(t)$	Discharge energy at each time period t, kWh
$P_{chg}(t)$	Charge power at each time period t, kW
$P_{dis}(t)$	Discharge power at each time period t, kW
R_{chg_max}	Maximum charging rate
R_{dis_max}	Maximum discharging rate
B_{cap}	Battery capacity, kWh
$I_{chg}(t)$	Binary variable indicating that battery is charging
$I_{dis}(t)$	Binary variable indicating that battery is discharging
$SOC(t)$	State of charge of the battery at the time t, kWh
η_{chg}	Battery charge efficiency, %
η_{dis}	Battery discharge efficiency, %

References

1. Sioshansi, F.P. California's 'Duck Curve' Arrives Well Ahead of Schedule. *Electr. J.* **2016**, *29*, 71–72. [CrossRef]
2. Sheha, M.; Powell, K. Using Real-Time Electricity Prices to Leverage Electrical Energy Storage and Flexible Loads in a Smart Grid Environment Utilizing Machine Learning Techniques. *Processes* **2019**, *7*, 870. [CrossRef]
3. BloombergNEF. Scaling the Residential Energy Storage Market. November 2023. Available online: https://assets.bbhub.io/professional/sites/24/Scaling-the-Residential-Energy-Storage-Market.pdf (accessed on 16 April 2024).
4. BloombergNEF. Lithium-Ion Battery Pack Prices Hit Record Low of $139/kWh. November 2023. Available online: https://about.bnef.com/blog/lithium-ion-battery-pack-prices-hit-record-low-of-139-kwh/ (accessed on 16 April 2024).
5. Shabani, M.; Shabani, M.; Wallin, F.; Dahlquist, E.; Yan, J. Smart and optimization-based operation scheduling strategies for maximizing battery profitability and longevity in grid-connected application. *Energy Convers. Manag. X* **2024**, *21*, 100519. [CrossRef]
6. Peñaranda, A.F.; Romero-Quete, D.; Cortés, C.A. Grid-Scale Battery Energy Storage for Arbitrage Purposes: A Colombian Case. *Batteries* **2021**, *7*, 59. [CrossRef]
7. Komorowska, A.; Olczak, P.; Hanc, E.; Kamiński, J. An analysis of the competitiveness of hydrogen storage and Li-ion batteries based on price arbitrage in the day-ahead market. *Int. J. Hydrogen Energy* **2022**, *47*, 28556–28572. [CrossRef]
8. Krishnamurthy, D.; Uckun, C.; Zhou, Z.; Thimmapuram, P.R.; Botterud, A. Energy Storage Arbitrage Under Day-Ahead and Real-Time Price Uncertainty. *IEEE Trans. Power Syst.* **2018**, *33*, 84–93. [CrossRef]
9. Salles, M.B.C.; Huang, J.; Aziz, M.J.; Hogan, W.W. Potential Arbitrage Revenue of Energy Storage Systems in PJM. *Energies* **2017**, *10*, 1100. [CrossRef]
10. Kataray, T.; Nitesh, B.; Yarram, B.; Sinha, S.; Cuce, E.; Shaik, S.; Vigneshwaran, P.; Roy, A. Integration of smart grid with renewable energy sources: Opportunities and challenges—A comprehensive review. *Sustain. Energy Technol. Assess.* **2023**, *58*, 103363. [CrossRef]
11. Soudagar, M.E.M.; Ramesh, S.; Khan, T.M.Y.; Almakayeel, N.; Ramesh, R.; Ghazali, N.N.N.; Cuce, E.; Shelare, S. An overview of the existing and future state of the art advancement of hybrid energy systems based on PV-solar and wind. *Int. J. Low-Carbon Technol.* **2024**, *19*, 207–216. [CrossRef]
12. Lew, D.; Bartlett, D.; Groom, A.; Jorgensen, P.; O'Sullivan, J.; Quint, R.; Rew, B.; Rockwell, B.; Sharma, S. Getting to 100% renewables: Operating experiences with very high penetrations of variable energy resources. *IET Renew. Power Gener.* **2020**, *14*, 3899–3907. [CrossRef]
13. Belmonte, B.B.; Mouratidis, P.; Franke, G.; Rinderknecht, S. Developments in the cost of grid balancing services and the design of the European balancing market. *Energy Rep.* **2023**, *10*, 910–931. [CrossRef]
14. Wang, Z.; Kirschen, D.S. Two-stage optimal scheduling for aggregators of batteries owned by commercial consumers. Transm. Distrib. *IET Gener.* **2019**, *13*, 4880–4887. [CrossRef]
15. Brogan, P.V.; Best, R.; Morrow, J.; Duncan, R.; Kubik, M. Stacking battery energy storage revenues with enhanced service provision. *IET Smart Grid* **2020**, *3*, 520–529. [CrossRef]
16. Nord Pool | Day-Ahead Prices. Available online: https://data.nordpoolgroup.com/auction/day-ahead/prices (accessed on 28 May 2024).
17. European Clean Power Pathways Explorer. Ember. Available online: https://ember-climate.org/data/data-tools/european-clean-power-pathways-explorer/ (accessed on 30 June 2024).
18. Barbour, E.; González, M.C. Projecting battery adoption in the prosumer era. *Appl. Energy* **2018**, *215*, 356–370. [CrossRef]
19. Wankmüller, F.; Thimmapuram, P.R.; Gallagher, K.G.; Botterud, A. Impact of battery degradation on energy arbitrage revenue of grid-level energy storage. *J. Energy Storage* **2017**, *10*, 56–66. [CrossRef]
20. Pena-Bello, A.; Burer, M.; Patel, M.K.; Parra, D. Optimizing PV and grid charging in combined applications to improve the profitability of residential batteries. *J. Energy Storage* **2017**, *13*, 58–72. [CrossRef]
21. Hassan, A.S.; Cipcigan, L.; Jenkins, N. Optimal battery storage operation for PV systems with tariff incentives. *Appl. Energy* **2017**, *203*, 422–441. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

Attention Mechanism-Based Neural Network for Prediction of Battery Cycle Life in the Presence of Missing Data

Yixing Wang and Benben Jiang *

Department of Automation, Tsinghua University, Beijing 100084, China; yx-wang21@mails.tsinghua.edu.cn
* Correspondence: bbjiang@tsinghua.edu.cn

Abstract: As batteries become widespread applications across various domains, the prediction of battery cycle life has attracted increasing attention. However, the intricate internal mechanisms of batteries pose challenges to achieving accurate battery lifetime prediction, and the inherent patterns within temporal data from battery experiments are often elusive. Meanwhile, the commonality of missing data in real-world battery usage further complicates accurate lifetime prediction. To address these issues, this article develops a self-attention-based neural network (NN) to precisely forecast battery cycle life, leveraging an attention mechanism that proficiently manages time-series data without the need for recurrent frameworks and adeptly handles the data-missing scenarios. Furthermore, a two-stage training approach is adopted, where certain network hyperparameters are fine-tuned in a sequential manner to enhance training efficacy. The results show that the proposed self-attention-based NN approach not only achieves superior predictive precision compared with the benchmarks including Elastic Net and CNN-LSTM but also maintains resilience against missing-data scenarios, ensuring reliable battery lifetime predictions. This work highlights the superior performance of utilizing attention mechanism for battery cycle life prognostics.

Keywords: battery cycle life prediction; self-attention mechanism; neural network; missing data

1. Introduction

Recent years have seen a significant focus on battery degradation and its impact on the operational lifespan of energy storage systems. As the demand for reliable and efficient energy storage solutions grows, accurate prediction of battery cycle life has become paramount. Batteries, being crucial for a myriad of applications from portable electronics [1,2] to electric vehicles [3–5], energy storages [6,7], and renewable energy systems [8–10], suffer from degradation influenced by a complex interplay of physical, chemical, and environmental factors. These complexities pose a substantial challenge for accurately modeling and predicting battery performance over time.

Researchers have been dedicated to extending battery lifespan and enhancing system performance by exploring various methodologies for predicting battery cycle life. Traditional methods typically rely on empirical models that may oversimplify and thus fail to contain the complex dynamics of battery behavior. More sophisticated approaches such as model-based approaches can capture more intricate details but often at the cost of computational intensity and the need for battery's physical and chemical property information due to the complex electrochemical nature of the modeling [11]. The development of battery models has been critical in this endeavor, with a pioneer work on the electrochemical model (ECM) by Newman and Tiedemann [12]. The ECM, evolving through time, provides a comprehensive description of battery properties due to its modelling capabilities [13]. This enables researchers to accurately estimate the state of health (SOH) and predict the remaining useful life (RUL) of batteries [14]. Additionally, the ECM has been enhanced with thermal models to account for phenomena such as solid electrolyte interface (SEI) layer growth—a key indicator of battery degradation [15]—furthering the precision of lifetime prediction methods [16]. The pseudo-two-dimensional (P2D) model [17], a prominent

ECM variant, applies porous electrode theory principles, focusing on the order reduction and simplification of kinetic equations [18], and also serves as a competitive model for SOH estimations [19]. This model has been utilized by researchers such as Ashwin et al. [20] to simulate battery aging processes, incorporating temperature effects and SEI layer thickness and thereby modeling battery capacity fade under cyclic loading conditions. Innovations have continued with variants such as the single-particle model (SPM) [21], the simplified P2D model [17], the improved SPM model [22], and the reduced-order model (ROM) [23–26]. For instance, Parhizi et al. [27] investigated the SEI layer development using the SPM model to understand capacity degradation under various operational conditions, aiming to predict capacity fade.

Concurrently, machine learning methods have surfaced as potent alternatives, harnessing data-driven insights to capture hidden patterns within battery performance data [28]. These techniques range from regression-based algorithms and ensemble methods to recurrent neural networks (RNNs), which are particularly adept at capturing temporal dependencies. Severson et al. [29] leveraged the elastic net (EN) regression to predict battery cycle life based on early cycle charging and discharging characteristics, using features derived from voltage, capacity, temperature, and internal resistance as machine learning inputs. Zhang et al. [30] also achieved battery lifetime predictions based on feature construction, where the physical features were namely ECM parameters. In addition, Hong et al.'s approach [31] to predict a battery's RUL using voltage curve data from only a few early cycles included the use of dilated convolutional neural networks (CNNs) to handle the time-sequential voltage data and implicit sampling to gauge the neural network's output uncertainty. Thelen et al. [32] also focused on RUL prediction while using various data-driven methods, including EN or random forest, to enhance the predictive performance of an empirical capacity fade model implemented by a Gaussian process (GP) model. Nuhic et al. [33] utilized the support vector machine (SVM) for the diagnosis and prognostics of battery SOH and RUL, taking into account the influence of environmental and load conditions. Considering the usage of fully discharged voltage and internal resistance, Tseng et al. [34] built a regression model for SOH estimation and RUL prediction that incorporated particle swarm optimization for optimal parameter determination. Mansouri et al. [35] and Guo et al. [36] expanded the suite of regression techniques, implementing algorithms like LASSO, multi-layer perceptron (MLP), support vector regression (SVR), and gradient boosted trees (GBT) for RUL prediction, with some even incorporating Bayesian methods for managing uncertainty. Neural networks have also been pivotal, with studies by Khumprom et al. [37] and Ren et al. [38] employing deep neural network (DNN) structures including autoencoders for feature extraction, while Zhang et al. [39] utilized RNNs to capture capacity degradation patterns. Meanwhile, the model-based and data-driven methods can also be combined together to achieve SOH and RUL prediction, where these types of methods are commonly addressed as grey box modeling [40]. For example, Thelen et al. [32] proposed a framework where various machine learning methods, including EN, random forest (RF), GP, and neural networks (NNs), were employed to correct and enhance the predictive performance of an empirical model in the context of RUL prediction. Liao et al. [41] raised another method fusion framework where a particle filter model was assisted by two data-driven methods to measure the internal state of the model and predict future measurements, eventually predicting the RUL of batteries.

Model-based approaches are recognized for their detailed simulation of battery behaviors, providing explicit models of internal processes that accurately represent battery dynamics. On the other hand, data-driven methods have garnered acclaim for their data processing capabilities, yielding impressive predictive accuracy even in the absence of intricate domain-specific knowledge. Nevertheless, the sequential nature of data hidden in battery experiments poses a significant challenge; the above-mentioned methodologies have yet to surmount the complexity and non-linearity embedded within the data sequences. Notably, RNNs have been employed to handle temporal sequences in many

researches [39], although this advantage often entails high computational demand or a compromise in prediction precision.

In the domain of natural language processing (NLP), a groundbreaking concept known as the self-attention mechanism has been introduced with noteworthy success. Initiated by A. Vaswani et al. [42] to address translational intricacies, the attention mechanism shows a great capability for dealing with serial data without recurrent structure, reducing the complexity of networks and calculation burden. The self-attention mechanism outperforms conventional RNNs with two merits: its capability for parallel computation of temporal sequences, markedly accelerating operational throughput, and its intrinsic proficiency in maintaining focus across long-time series, overcoming the fundamental limitation encountered in traditional RNN architectures.

Meanwhile, the issue of missing data is a prevalent concern in real-world sequential data, and its occurrence can introduce unpredictable disruptions to model training and evaluative processes. To tackle the challenge of data loss, various methods have been proposed in the literature. For example, Severson et al. [43] introduced a method employing expectation maximization and sparse discriminant analysis for high-dimensional datasets, training classifiers across varying patterns of missing data. Complementary to this, PCA-based techniques have also shown promise in addressing data insufficiencies [44]. Jeong et al. [45] tackled fairness in model training amid absent values, devising a fair mixed-integer programming (MIP) forest algorithm based on decision trees with the missing incorporated in attribute (MIA) technique [46], while Chen et al. [47] integrated a multi-head self-attention mechanism within a variational auto-encoder (VAE) network for the industrial soft-sensing problem with missing data, and the best method could satisfy the requirements of high precision in complex industrial processes. Additionally, Wu et al. [48] applied an attention-based AimNet structure for data imputation and data cleaning issues, and Yu et al. [49] proposed a novel network structure consisting of convolution modules and an attention module, along with a specially designed loss function, successfully improving the performance for the consecutively missing seismic data reconstruction problem.

In the context of battery life prediction, where the intricacies of temporal correlations and the varying importance of attributes are decisive, the ability of the attention mechanism to spotlight pertinent information holds substantial promise. However, there are still few studies focusing on the application of the self-attention mechanism on battery life prediction [50]. By effectively capturing the interplay between attributes and their temporal progression, the attention mechanism presents an opportunity to elevate prediction accuracy. Building on the progress made in attention-based models and the growing recognition of their potential in capturing temporal dependencies, this study introduces a novel application: the integration of the attention mechanism into neural networks for battery life prediction, where we combined the calculation of attention value and typical fully connected networks together to enhance battery life forecasting. To enhance the predictive performance of the attention network during training, a two-stage training approach was employed, where some hyperparameters of the network need to be tuned between the training stages. Moreover, this study delves into the attention mechanism's innate capacity to handle missing data, examining its efficacy under different data-missing possibilities across three distinct data-loss patterns: random missing, cycle missing, and time-step missing. The results demonstrate the attention mechanism's exceptional predictive performance under varied data-missing scenarios and various degree of data missingness.

The remaining sections of this paper are outlined as follows. Section 2 elaborates on the principles of the self-attention mechanism and its deployment in contexts of data-loss scenarios. Section 3 details the used dataset and the categorization of data missingness, and then Section 4 presents the predictive results for battery cycle life under different data conditions. Finally, Section 5 provides the conclusion of the work.

2. Methods

2.1. The Self-Attention Mechanism

Denote $X = \left[\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}, \ldots, \boldsymbol{x}_L^{\mathrm{T}}\right]^{\mathrm{T}}$ as the input sequence, and $\boldsymbol{x}_i^{\mathrm{T}} = [x_{i1}, x_{i2}, \ldots, x_{im}]$, where L is the sequence length, and $m = d_{\text{model}}$ is the dimension of the input sequence. The calculation of attention mainly involves three aspects, i.e., query Q, key K, and value V, derived by $Q = XW^Q$, $K = XW^K$, $V = XW^V$, where W^Q, W^K, W^V are the weight matrices with the dimension of $d_{\text{model}} \times d_Q, d_{\text{model}} \times d_K$, and $d_{\text{model}} \times d_V$, respectively. As the requirement of the calculation, $d_Q = d_K$ should hold, and $d_K = d_V$ is a common setup in practical applications similar to our work.

Given the Q, K, V, the calculation of self-attention can be obtained by (1). As the $\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_K}}\right)$ part is a square matrix, the SoftMax operator is applied on each of its rows.

$$\text{Attention}(Q, K, V) = \text{SoftMax}\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_K}}\right)V \tag{1}$$

Inspired by the idea of attention, multi-head self-attention is more popular and more capable in practical applications, where different heads can focus on different parts of the input sequence, covering important portions as much as possible. The calculation of multi-head self-attention mainly consists of a concatenation of attention heads and a linear transformation, as shown in (2):

$$\begin{aligned} \text{MultiHead}(Q, K, V) &= [\text{head}_1, \text{head}_2, \ldots, \text{head}_h]W^O \\ \text{where head}_i &= \text{Attention}\left(XW_i^Q, XW_i^K, XW_i^V\right) \end{aligned} \tag{2}$$

In addition, to emphasize location information, positional encoding is adopted before calculating the attention value. As a common choice in practice, we use sine and cosine functions for the positional encoding in the work, as shown in (3):

$$\begin{aligned} PE_{(pos,\,2i)} &= \sin\left(pos/10,000^{2i/d_{\text{model}}}\right) \\ PE_{(pos,\,2i+1)} &= \cos\left(pos/10,000^{2i/d_{\text{model}}}\right) \end{aligned} \tag{3}$$

where pos represents the position, and i indicates the dimension of input.

After the multi-head self-attention has been calculated, a fully connected network is employed to reduce dimension and obtain the target output. The overall structure of the entire attention-based network is described in Figure 1.

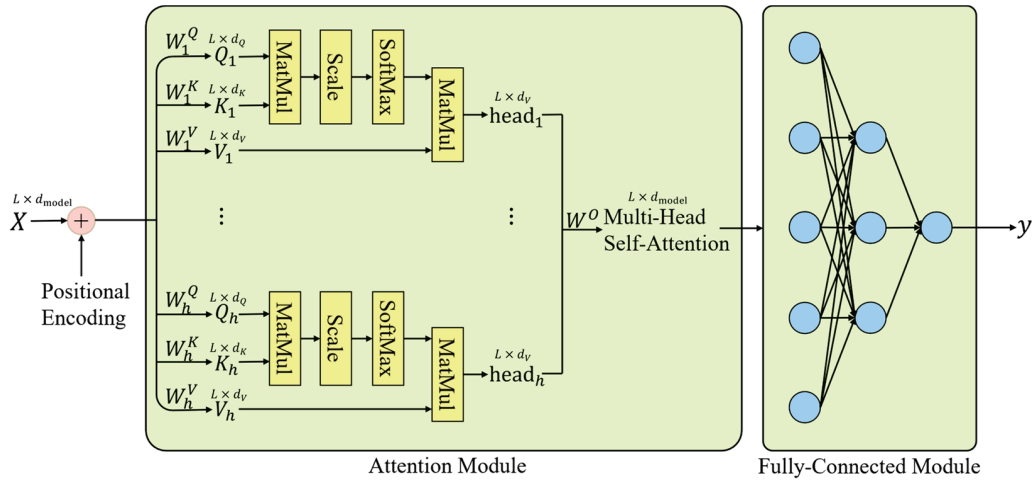


Figure 1. The structure of the attention-based network. The arrows indicate the data flow and the blue circles represent the general neurons in the network.

2.2. Handling Missing Data with the Self-Attention Mechanism

Within the domain of sequential data analysis, the self-attention mechanism has emerged as a cornerstone for capturing contextual relationships. Its innovative deployment in the face of missing data further underscores its versatility. In real-world datasets, where incomplete data are commonplace, the self-attention mechanism showcases its inherent adaptability to these challenges. Under the data-missing condition, the affected positions in the data input X will be set to zero, yet their positional encoding preserves the location's informational context. Particularly in the context of self-attention, where queries, keys, and values are derived from the same input sequence, the mechanism's intrinsic similarity computation and attention weight allocation process exhibit a degree of adaptability that can accommodate partial data availability.

In the case of missing data within self-attention, the absence of values at specific positions—manifested as random zeros within the input X—directly affects their influence on the similarity calculations between queries and keys. Denote the data input with missing values as X'; then, the query, key, and value can be articulated as $Q' = (X' + PE)W^Q$, $K' = (X' + PE)W^K$, and $V' = (X' + PE)W^V$, respectively. While missing values do not disrupt the attention values' computation, the resultant attention scores inherently reflect the data's absence and maintain positional context, as depicted in Figure 2. Consequently, the mechanism intuitively assigns greater attention weights to the present data, ensuring that the calculated scores naturally diminish the impact of any missing elements, thereby implicitly moderating the undue emphasis on missing values. This intrinsic modulation of attention allows the model to concentrate on the segments of data that are available, resonating with the fundamental principle of the self-attention mechanism to prioritize segments enriched with information.

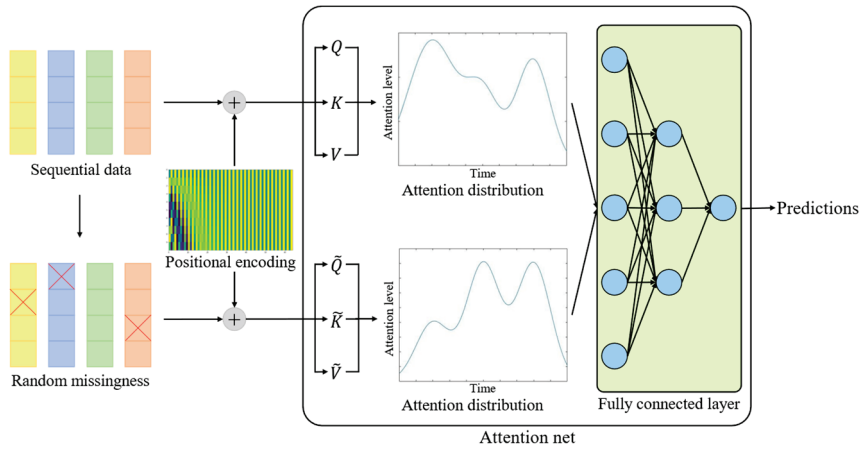


Figure 2. The self-attention mechanism for handling data missingness. When sequences with missing data (the pattern with red crosses) are fed into the network along with positional encodings, the attention-based network retains its original computational structure but adapts its self-attention values computation in response to the missing data, consequently altering attention distribution across different positions, thereby autonomously handling missing data scenarios.

Leveraging the attention mechanism to handle time-series data suffering from random losses presents several advantages. Its inherent adaptability to missing data scenarios is particularly notable, as it dynamically recalibrates significance to the intact portions of data, adeptly bridging the voids left by absent values. This flexibility ensures the model's predictive performance remains robust, even when confronted with incomplete sequences. Additionally, the attention mechanism's ability to selectively emphasize informative segments while minimizing the impact of missing or noisy data facilitates the extraction of pivotal patterns and dependencies, ultimately enhancing prediction accuracy.

2.3. Training Process

Overfitting poses a significant challenge in network training. To avoid the overfitting issue, our approach incorporates a two-stage training methodology, as illustrated in Figure 3. This strategy involves distinct settings for the regularization coefficient and batch size in each stage. Initially, the training commences with a smaller regularization coefficient and a reduced batch size for a duration of 50 epochs. The network parameters are then preserved at the point where the training error reaches its minimum during this initial phase. Before the second stage, the division of the dataset into training and validation sets is re-randomized. The subsequent stage of training proceeds from the previously saved network parameters, now employing an increased regularization coefficient and a larger batch size. This two-stage technique for training enables the network to learn with enhanced precision while effectively preventing overfitting.

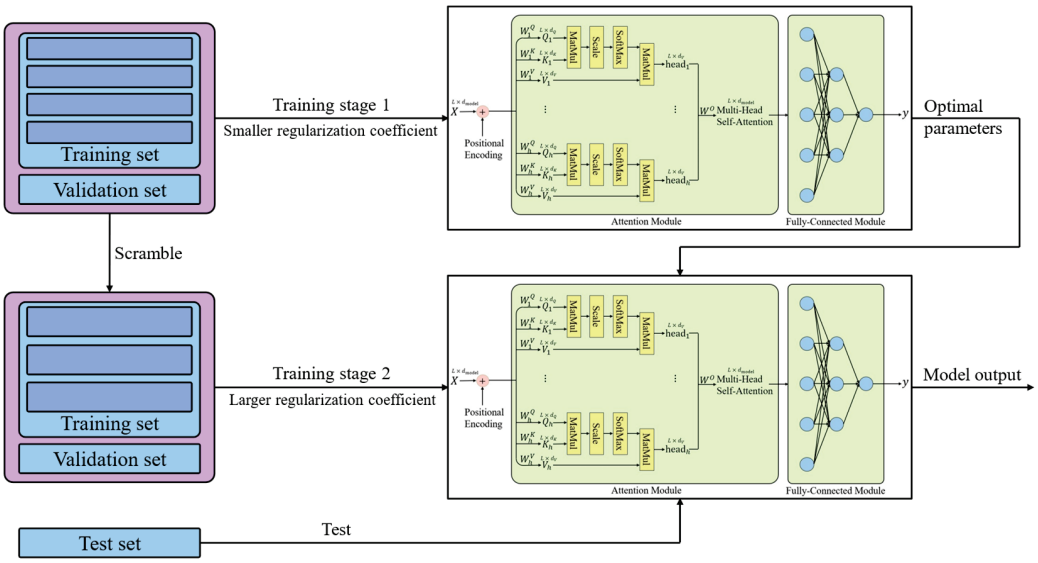


Figure 3. Flowchart of the two-stage training process.

3. Description of Dataset and Data-Missing Patterns

3.1. Dataset and Input Data Construction

This work utilizes the dataset referenced in [29], which comprises 124 commercial high-power LFP/graphite battery cells, each with a nominal capacity of 1.1 Ah and a nominal voltage of 3.3 V. These cells underwent cycling tests under varying constant current–constant voltage (CC-CV) charging protocols with several different charging rates while maintaining a constant environment temperature of 30 °C within experimental chambers. The CC stage of the charging process featured C rates ranging from 2 C to 6 C across different SOC intervals. Upon reaching 80% SOC, a CV stage ensued until the completion of the charging process. For discharging, all cells underwent another CC-CV discharge process at a rate of 4 C down to 2.0 V, with a current cutoff of C/50. Under such cycling tests, the lifespan of all batteries ranged from 148 to 2233 cycles.

Standard practice in battery cycling experiments involves the recording of several parameters in each cycle, including time, current, voltage, charging capacity, discharging capacity, and temperature, as depicted in Figure 4a–e. Recognizing the significance of the interplay among capacity, temperature, and voltage in battery analysis, we further computed the scaled discharging capacity and discharging temperature through fitting and interpolation against the voltage curve, as shown in Figure 4f,g. In addition, to obtain deeper insights from the cycling data, we derived the gradient of discharging capacity in relation to voltage (Figure 4h), i.e., the cycle-wise incremental capacity of the cells. In total, nine attributes of battery cycling data are employed in our work.

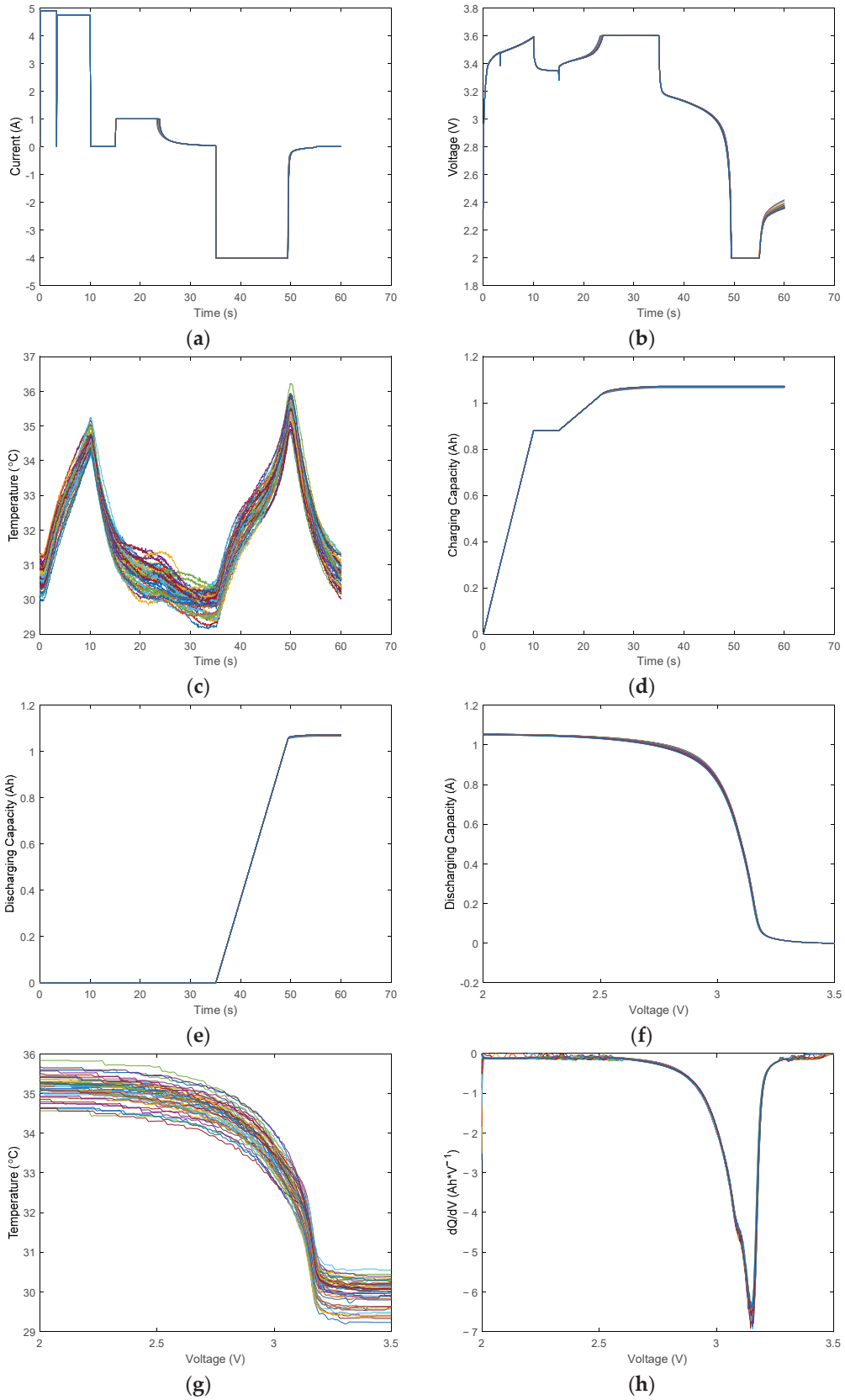


Figure 4. The contents of input data. (**a**) Current. (**b**) Voltage. (**c**) Temperature. (**d**) Charging capacity. (**e**) Discharging Capacity. (**f**) Discharging capacity variation against voltage. (**g**) Temperature variation against voltage. (**h**) Incremental capacity. The different line colors indicate different cycles of the displayed battery cell.

To optimize the utilization of sequential data from the cycling experiments, we adopted a unique structure for input data, diverging from conventional intuitive input formats, as illustrated in Figure 5. Our input construction (as depicted in Figure 5a) involves vertically concatenating the nine attributes to form an extended array for each cycle, while the horizontal dimension represents the different cycles. The principal benefit of this structured approach is the strategic reconstruction of the matrix's dimensions, coupled with a reduction in its length. This effectively mitigates the computational demands, particularly in the subsequent phases of attention calculation, ensuring a more efficient processing workflow.

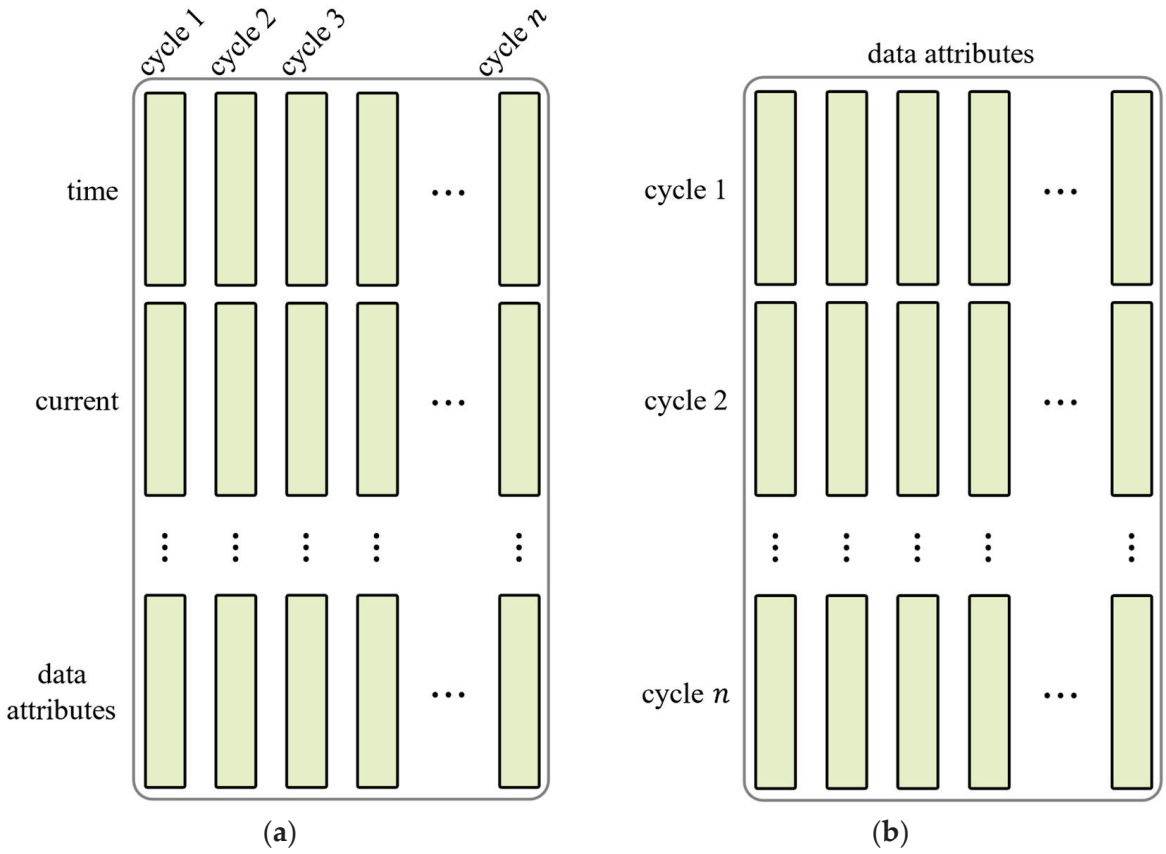
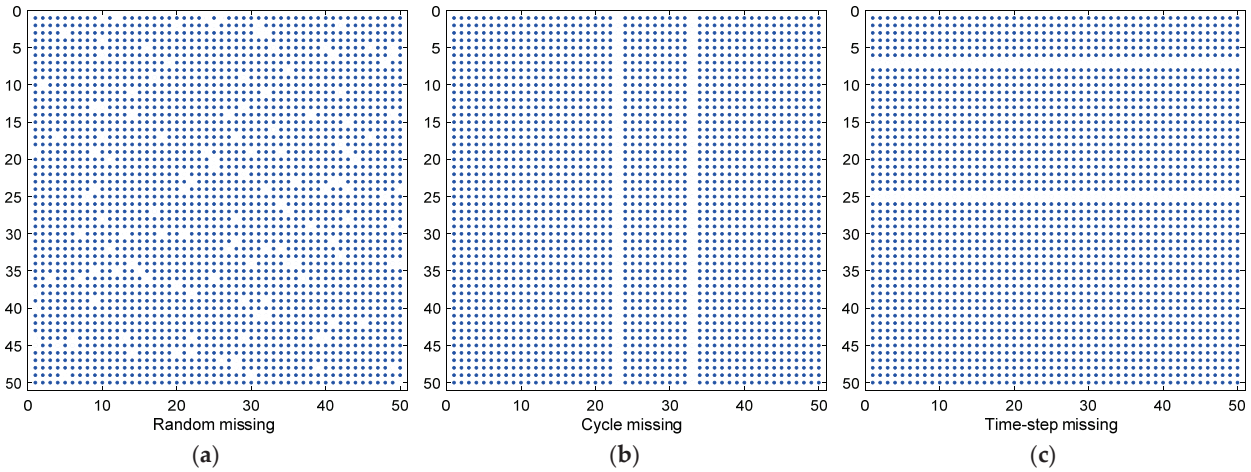


Figure 5. The structure of input data. (**a**) The input structure used in this work. (**b**) The intuitive input structure.

3.2. Addition of Missing Data

This work primarily investigates scenarios involving data loss [43]. We consider three different patterns of missing data: random missing, cycle missing, and time-step missing, which is illustrated in Figure 6. Random missing drops data points in a completely random manner, while cycle missing and time-step missing drop data points by entire columns or rows, reflecting real-world instances of data loss in specific cycles or at particular time steps.

In addition, it is important to note that while the attention mechanism inherently possesses the ability to adjust to missing data, this capability is inversely related to the extent of the data gaps. To maintain the balance during training and ensure the effectiveness of the mechanism, we set the dropout probability at a default of 0.01, with an upper limit of 0.1 under any circumstances. This approach enables us to systematically evaluate the robustness of the attention mechanism across different levels of data missing while preventing excessive data loss that could compromise the model's performance.



(**a**) (**b**) (**c**)

Figure 6. Examples of the various types of data missing [43]: (**a**) random missing, (**b**) cycle missing, and (**c**) time-step missing.

4. Results

To measure the predictive performance of the attention-based network approach, the metrics of average percentage error (APE) and root mean square error (RMSE) are used in this work, which is defined as follows:

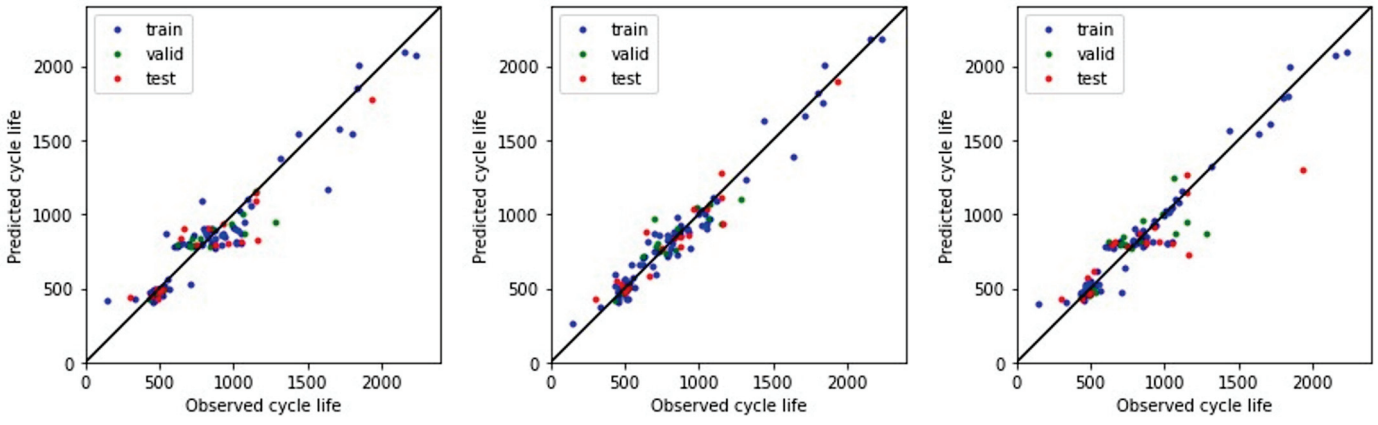
$$\text{APE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \tag{4}$$

and

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5}$$

where y_i is the observed cycle life, and $\hat{y}_i$ is the predicted cycle life.

The prediction results of the proposed method, in scenarios without data missing, are displayed in Figure 7, wherein the conducted battery lifetime prediction using the initial 20, 50, and 100 cycles, respectively, is shown. As can be seen from the results in Figure 7, the cycle life predictions in all three conditions are relatively accurate (i.e., close to the ideal prediction line, represented by the black solid line in the figures), with the test APE of 13.02%, 10.59%, and 15.16%, respectively. In addition, compared to the results presented in [29], it is evident that our approach effectively addresses the challenge of outliers, showcasing enhanced predictive accuracy and reliability.



(**a**) (**b**) (**c**)

Figure 7. The prediction result of the proposed attention-based network approach by using different cycles of battery data without data missing: (**a**) 20 cycles, (**b**) 50 cycles, and (**c**) 100 cycles.

Additionally, we conducted an analysis to examine the impact of altering the number of cycles utilized for lifetime prediction, with the findings showcased in Figure 8. As demonstrated in Figure 8, utilizing the first 50 cycles for prediction obtains the optimal performance, where the average test APE of ten different dataset splits reaches 9.34%. In contrast, employing either fewer or more cycles results in increased APEs. The reason for this phenomenon is that a lower number of cycles may lead to underutilization of the network's capacity to process available information, and the lack of information may also deteriorate the performance. Conversely, an excessively high cycle number can introduce redundant data, potentially impeding the network's training efficiency and overall predictive performance. This balance highlights the importance of selecting appropriate information to achieve optimal prediction accuracy when solving battery lifetime prediction problems.

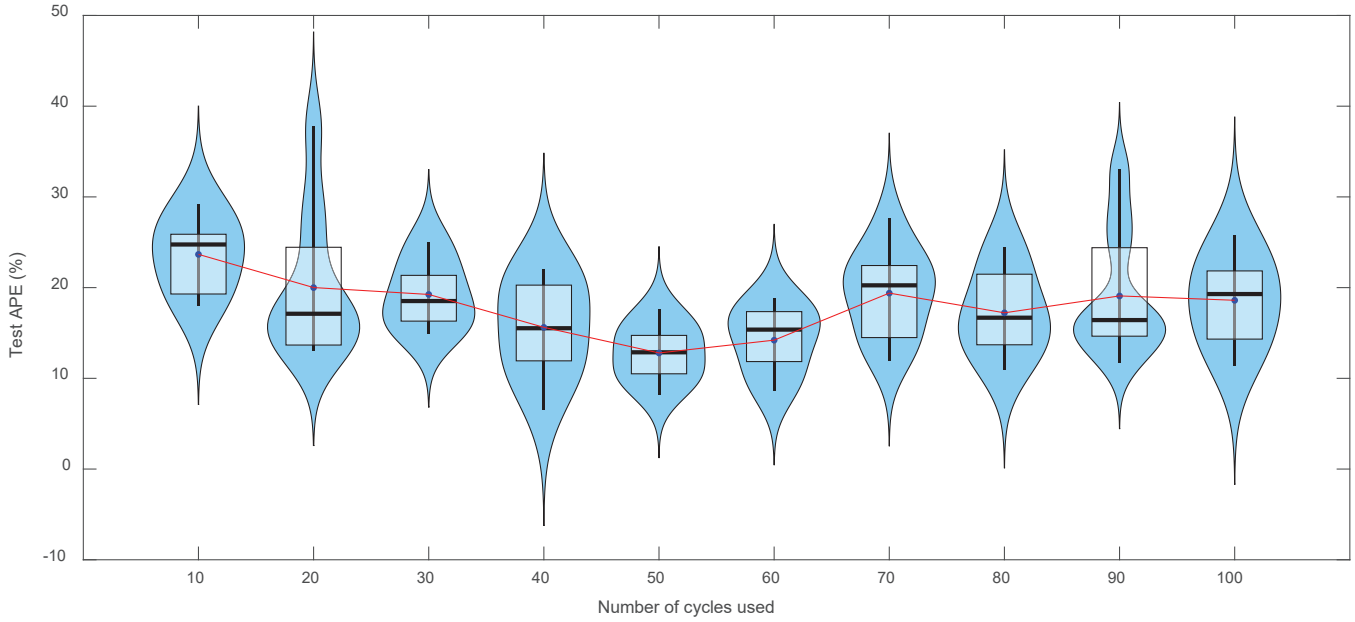


Figure 8. The violin plot for the influence of using different number of cycles. Under each number, 10 different splits of train/valid/test datasets were conducted. In the figure, the bold lines represent median APE values, the boxes represent quartiles of the APE values, and blue dots represent mean APE values with the red line connecting them, and the blue areas represent the probability distribution of the APE values.

Therefore, we opted to use 50 cycles of battery data to examine how the attention network performs under different scenarios of missing data, including random missing, cycle missing, and time-step missing. The results of this analysis are detailed in Table 1. Notably, the attention network exhibited exceptional stability in scenarios characterized by cycle missing. In such cases where data points were consecutively absent in a cyclic pattern, the attention mechanism demonstrated its resilience, effectively preserving its capability to discern relationships within the data sequence, notwithstanding these gaps. The result highlights the mechanism's ability to adapt to predictable patterns of data gaps.

Table 1. Model metrics for different data-missing types.

	APE (%)			**RMSE (Cycles)**		
	Train	**Validation**	**Test**	**Train**	**Validation**	**Test**
Random missing	9.4	10.8	11.3	82	119	104
Cycle missing	8.8	9.3	10.3	77	102	100
Time-step missing	9.8	9.8	11.1	80	108	99

On the other hand, in scenarios characterized by random missing data and time-step missing data, the attention network's performance closely approached its full potential. Despite the unpredictability and challenges associated with the two types of data absence, the network effectively leveraged the available information, maintaining a considerable level of predictive accuracy.

These findings shed light on the adaptability and robustness of the attention network across various scenarios of data absence. Its consistent performance in cycle-missing scenarios and substantial resilience in situations of random missing data and time-step missing data underscore its capability to handle patterns of missing data effectively. The overall capability of the attention mechanism in lifetime prediction demonstrates its versatility and potential in diverse data conditions.

To benchmark the performance of our method under conditions with data missing, we also conducted validations against several other methods from existing literature, specifically focusing on the cycle-missing scenario. The results of these comparative tests are presented in Table 2. As the method described in [29] solely utilizes the 2nd and 100th cycles of battery data for feature construction, the replication of its result did not incorporate scenarios of data missing. Instead, the basic performance of this method was used as a benchmark for comparison. The findings indicate that, under identical conditions of missing data, our proposed attention-based method demonstrates superior adaptability and enhanced predictive performance compared to the alternatives. This further validates the efficacy of the proposed attention-based network approach in predicting battery lifetime, particularly in situations with missing data.

Table 2. Model metrics for different prediction methods with data missing.

	APE (%)			**RMSE (Cycles)**		
	Train	**Validation**	**Test**	**Train**	**Validation**	**Test**
Attention net	8.8	9.3	10.3	77	102	100
Elastic Net [29]	11.2	-	14.4	112	-	165
CNN	17.8	14.1	18.7	189	135	239
CNN-LSTM [51]	12.7	12.3	15.8	98	96	140

Lastly, to further explore the adaptability and capabilities of the attention-based network in the scenarios of extreme data missing, we assessed the predictive performance of our proposed method under varying data-missing degrees. The results of these tests are illustrated in Figure 9. In the experimental settings, all the cases are randomly repeated five times, and the circle in the figure represents the mean value of test APEs, while the shaded areas represent the APE values within one standard deviation. The results demonstrate that the attention-based network consistently exhibits robust predictive performance in scenarios of cycle missing and random missing, even as the degree of data missingness escalates to 10%. However, it encounters slight limitations in the context of time-step missing, preserving its effectiveness up to a maximum of 5% degree of data missingness. Overall, the superior ability of the attention-based network for battery lifetime prediction with missing data underscores its significance in real-world settings where data loss can be a prevailing issue.

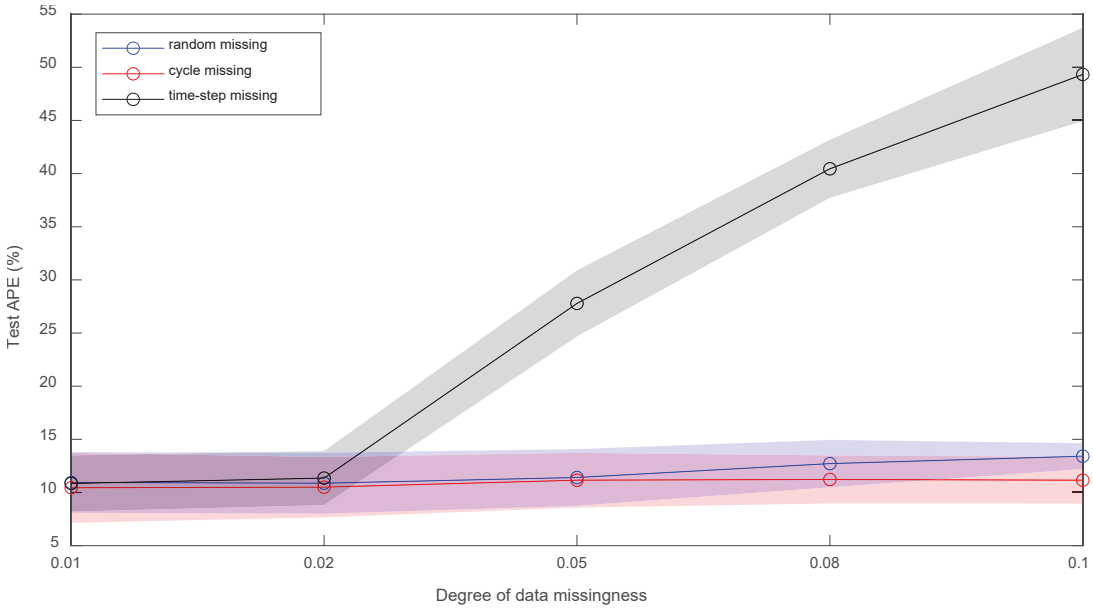


Figure 9. The impact of data-missing degree on battery lifetime prediction in different data-missing scenarios. The shaded areas represent the values within one standard deviation across the results of five different train/valid/test splits.

5. Conclusions

In this article, we proposed an innovative neural network that incorporates the self-attention mechanism, specifically designed to tackle the challenges of extracting meaningful information from sequential data in battery cycling experiments for accurate battery cycle life prediction. The results from our study indicate that our proposed self-attention mechanism-based network approach performs higher accuracy compared to other benchmarks in this field under the conditions without data missingness, where the average test APE reaches 9.34%, improving the performance by approximately 35.1% relative to the original article of the dataset [29]. Furthermore, our proposed method exhibits robust adaptability across three distinct data-missing scenarios, i.e., random missing, cycle missing, and time-step missing, achieving test APEs of 11.3%, 10.9%, and 11.1%, respectively, and maintaining good performance under all data-missing scenarios. This adaptability plays a pivotal role in preserving the accuracy of the predictions under various data-loss conditions. Moreover, even in the face of these data-missing scenarios, our proposed method consistently outperforms the benchmarking methods within the field, with an average test APE of 10.3%, signifying its superior predictive power. Additionally, the attention-based network maintains good predictive performance even as the extent of data missing intensifies. Notably, the test APEs under cycle-missing and random-missing scenarios remained under 13.5% even as the data-missing degree escalated to 10%, showcasing its resilience in increasingly challenging circumstances.

When considering a broader range of applications, our proposed self-attention-based neural network, as a supervised machine learning approach, has the potential to remain applicable under different conditions, such as varying depths of discharge (DOD). DOD is a critical factor that directly influences battery performance and lifetime, with shallower discharges generally leading to longer battery life compared to deeper discharges. The primary requirement for adapting our method to diverse conditions is the collection of relevant data under the respective conditions to serve as the training set. With appropriate training, our method can predict battery lifetime or other performance metrics under those specific conditions. The successful implementation of the attention mechanism in this context not only highlights its potential in battery cycle life prediction but also underscores its applicability in critical areas such as battery management systems that rely on accurate cycle life prediction.

Author Contributions: Conceptualization, Y.W. and B.J.; methodology, Y.W.; software, Y.W.; formal analysis, Y.W.; resources, B.J.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, B.J.; visualization, Y.W.; supervision, B.J.; funding acquisition, B.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (2022YFE0197600), the National Natural Science Foundation of China under grant number 62273197, and the Beijing Natural Science Foundation under grant number L233027.

Data Availability Statement: The authors have not generated any new experimental data in this research, and the data utilized in the research is available at https://data.matr.io/1 (accessed on 28 April 2024), under the "data-driven prediction of battery cycle life before capacity degradation" project.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liang, Y.; Zhao, C.Z.; Yuan, H.; Chen, Y.; Zhang, W.; Huang, J.Q.; Yu, D.; Liu, Y.; Titirici, M.M.; Chueh, Y.L.; et al. A Review of Rechargeable Batteries for Portable Electronic Devices. *InfoMat* **2019**, *1*, 6–32. [CrossRef]
2. Olabi, A.G.; Abbas, Q.; Shinde, P.A.; Abdelkareem, M.A. Rechargeable Batteries: Technological Advancement, Challenges, Current and Emerging Applications. *Energy* **2023**, *266*, 126408. [CrossRef]
3. König, A.; Nicoletti, L.; Schröder, D.; Wolff, S.; Waclaw, A.; Lienkamp, M. An Overview of Parameter and Cost for Battery Electric Vehicles. *World Electr. Veh. J.* **2021**, *12*, 21. [CrossRef]
4. Lebrouhi, B.E.; Khattari, Y.; Lamrani, B.; Maaroufi, M.; Zeraouli, Y.; Kousksou, T. Key Challenges for a Large-Scale Development of Battery Electric Vehicles: A Comprehensive Review. *J. Energy Storage* **2021**, *44*, 103273. [CrossRef]
5. Tran, M.K.; Bhatti, A.; Vrolyk, R.; Wong, D.; Panchal, S.; Fowler, M.; Fraser, R. A Review of Range Extenders in Battery Electric Vehicles: Current Progress and Future Perspectives. *World Electr. Veh. J.* **2021**, *12*, 54. [CrossRef]
6. Cheng, X.; Liu, H.; Yuan, H.; Peng, H.; Tang, C.; Huang, J.; Zhang, Q. A Perspective on Sustainable Energy Materials for Lithium Batteries. *SusMat* **2021**, *1*, 38–50. [CrossRef]
7. Hannan, M.A.; Wali, S.B.; Ker, P.J.; Rahman, M.S.A.; Mansor, M.; Ramachandaramurthy, V.K.; Muttaqi, K.M.; Mahlia, T.M.I.; Dong, Z.Y. Battery Energy-Storage System: A Review of Technologies, Optimization Objectives, Constraints, Approaches, and Outstanding Issues. *J. Energy Storage* **2021**, *42*, 103023. [CrossRef]
8. Weiss, M.; Ruess, R.; Kasnatscheew, J.; Levartovsky, Y.; Levy, N.R.; Minnmann, P.; Stolz, L.; Waldmann, T.; Wohlfahrt-Mehrens, M.; Aurbach, D.; et al. Fast Charging of Lithium-Ion Batteries: A Review of Materials Aspects. *Adv. Energy Mater.* **2021**, *11*, 2101126. [CrossRef]
9. Kebede, A.A.; Kalogiannis, T.; Van Mierlo, J.; Berecibar, M. A Comprehensive Review of Stationary Energy Storage Devices for Large Scale Renewable Energy Sources Grid Integration. *Renew. Sustain. Energy Rev.* **2022**, *159*, 112213. [CrossRef]
10. Chen, T.; Jin, Y.; Lv, H.; Yang, A.; Liu, M.; Chen, B.; Xie, Y.; Chen, Q. Applications of Lithium-Ion Batteries in Grid-Scale Energy Storage Systems. *Trans. Tianjin Univ.* **2020**, *26*, 208–217. [CrossRef]
11. Park, S.; Pozzi, A.; Whitmeyer, M.; Perez, H.; Kandel, A.; Kim, G.; Choi, Y.; Joe, W.T.; Raimondo, D.M.; Moura, S. A Deep Reinforcement Learning Framework for Fast Charging of Li-Ion Batteries. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 2770–2784. [CrossRef]
12. Newman, J.; Tiedemann, W. Porous-electrode Theory with Battery Applications. *AIChE J.* **1975**, *21*, 25–41. [CrossRef]
13. Elmahallawy, M.; Elfouly, T.; Alouani, A.; Massoud, A.M. A Comprehensive Review of Lithium-Ion Batteries Modeling, and State of Health and Remaining Useful Lifetime Prediction. *IEEE Access* **2022**, *10*, 119040–119070. [CrossRef]
14. Khodadadi Sadabadi, K.; Jin, X.; Rizzoni, G. Prediction of Remaining Useful Life for a Composite Electrode Lithium Ion Battery Cell Using an Electrochemical Model to Estimate the State of Health. *J. Power Sources* **2021**, *481*, 228861. [CrossRef]
15. Lam, F.; Allam, A.; Joe, W.T.; Choi, Y.; Onori, S. Offline Multiobjective Optimization for Fast Charging and Reduced Degradation in Lithium Ion Battery Cells. In Proceedings of the 2021 American Control Conference (ACC), New Orleans, LA, USA, 25–28 May 2021; pp. 4441–4446. [CrossRef]
16. O'Kane, S.E.J.; Ai, W.; Madabattula, G.; Alonso-Alvarez, D.; Timms, R.; Sulzer, V.; Edge, J.S.; Wu, B.; Offer, G.J.; Marinescu, M. Lithium-Ion Battery Degradation: How to Model It. *Phys. Chem. Chem. Phys.* **2022**, *24*, 7909–7922. [CrossRef] [PubMed]
17. Jokar, A.; Rajabloo, B.; Désilets, M.; Lacroix, M. Review of Simplified Pseudo-Two-Dimensional Models of Lithium-Ion Batteries. *J. Power Sources* **2016**, *327*, 44–55. [CrossRef]
18. Kong, X.; Plett, G.L.; Scott Trimboli, M.; Zhang, Z.; Qiao, D.; Zhao, T.; Zheng, Y. Pseudo-Two-Dimensional Model and Impedance Diagnosis of Micro Internal Short Circuit in Lithium-Ion Cells. *J. Energy Storage* **2020**, *27*, 101085. [CrossRef]
19. Liu, B.; Tang, X.; Gao, F. Joint Estimation of Battery State-of-Charge and State-of-Health Based on a Simplified Pseudo-Two-Dimensional Model. *Electrochim. Acta* **2020**, *344*, 136098. [CrossRef]
20. Ashwin, T.R.; Chung, Y.M.; Wang, J. Capacity Fade Modelling of Lithium-Ion Battery under Cyclic Loading Conditions. *J. Power Sources* **2016**, *328*, 586–598. [CrossRef]
21. Han, X.; Ouyang, M.; Lu, L.; Li, J. Simplification of Physics-Based Electrochemical Model for Lithium Ion Battery on Electric Vehicle. Part I: Diffusion Simplification and Single Particle Model. *J. Power Sources* **2015**, *278*, 802–813. [CrossRef]

22. Han, X.; Ouyang, M.; Lu, L.; Li, J. Simplification of Physics-Based Electrochemical Model for Lithium Ion Battery on Electric Vehicle. Part II: Pseudo-Two-Dimensional Model Simplification and State of Charge Estimation. *J. Power Sources* **2015**, *278*, 814–825. [CrossRef]
23. Lee, J.L.; Chemistruck, A.; Plett, G.L. Discrete-Time Realization of Transcendental Impedance Models, with Application to Modeling Spherical Solid Diffusion. *J. Power Sources* **2012**, *206*, 367–377. [CrossRef]
24. Stetzel, K.D.; Aldrich, L.L.; Trimboli, M.S.; Plett, G.L. Electrochemical State and Internal Variables Estimation Using a Reduced-Order Physics-Based Model of a Lithium-Ion Cell and an Extended Kalman Filter. *J. Power Sources* **2015**, *278*, 490–505. [CrossRef]
25. Deng, Z.; Hu, X.; Lin, X.; Xu, L.; Li, J.; Guo, W. A Reduced-Order Electrochemical Model for All-Solid-State Batteries. *IEEE Trans. Transp. Electrif.* **2021**, *7*, 464–473. [CrossRef]
26. Lai, X.; He, L.; Wang, S.; Zhou, L.; Zhang, Y.; Sun, T.; Zheng, Y. Co-Estimation of State of Charge and State of Power for Lithium-Ion Batteries Based on Fractional Variable-Order Model. *J. Clean. Prod.* **2020**, *255*, 120203. [CrossRef]
27. Parhizi, M.; Pathak, M.; Ostanek, J.K.; Jain, A. An Iterative Analytical Model for Aging Analysis of Li-Ion Cells. *J. Power Sources* **2022**, *517*, 230667. [CrossRef]
28. Ng, M.F.; Zhao, J.; Yan, Q.; Conduit, G.J.; Seh, Z.W. Predicting the State of Charge and Health of Batteries Using Data-Driven Machine Learning. *Nat. Mach. Intell.* **2020**, *2*, 161–170. [CrossRef]
29. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D.; et al. Data-Driven Prediction of Battery Cycle Life before Capacity Degradation. *Nat. Energy* **2019**, *4*, 383–391. [CrossRef]
30. Zhang, Y.; Feng, X.; Zhao, M.; Xiong, R. In-Situ Battery Life Prognostics amid Mixed Operation Conditions Using Physics-Driven Machine Learning. *J. Power Sources* **2023**, *577*, 233246. [CrossRef]
31. Hong, J.; Lee, D.; Jeong, E.R.; Yi, Y. Towards the Swift Prediction of the Remaining Useful Life of Lithium-Ion Batteries with End-to-End Deep Learning. *Appl. Energy* **2020**, *278*, 115646. [CrossRef]
32. Thelen, A.; Li, M.; Hu, C.; Bekyarova, E.; Kalinin, S.; Sanghadasa, M. Augmented Model-Based Framework for Battery Remaining Useful Life Prediction. *Appl. Energy* **2022**, *324*, 119624. [CrossRef]
33. Nuhic, A.; Terzimehic, T.; Soczka-Guth, T.; Buchholz, M.; Dietmayer, K. Health Diagnosis and Remaining Useful Life Prognostics of Lithium-Ion Batteries Using Data-Driven Methods. *J. Power Sources* **2013**, *239*, 680–688. [CrossRef]
34. Tseng, K.H.; Liang, J.W.; Chang, W.; Huang, S.C. Regression Models Using Fully Discharged Voltage and Internal Resistance for State of Health Estimation of Lithium-Ion Batteries. *Energies* **2015**, *8*, 2889–2907. [CrossRef]
35. Mansouri, S.S.; Karvelis, P.; Georgoulas, G.; Nikolakopoulos, G. Remaining Useful Battery Life Prediction for UAVs Based on Machine Learning. *IFAC-PapersOnLine* **2017**, *50*, 4727–4732. [CrossRef]
36. Guo, J.; Li, Z.; Pecht, M. A Bayesian Approach for Li-Ion Battery Capacity Fade Modeling and Cycles to Failure Prognostics. *J. Power Sources* **2015**, *281*, 173–184. [CrossRef]
37. Khumprom, P.; Yodo, N. A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries Based on a Deep Learning Algorithm. *Energies* **2019**, *12*, 660. [CrossRef]
38. Ren, L.; Zhao, L.; Hong, S.; Zhao, S.; Wang, H.; Zhang, L. Remaining Useful Life Prediction for Lithium-Ion Battery: A Deep Learning Approach. *IEEE Access* **2018**, *6*, 50587–50598. [CrossRef]
39. Zhang, Y.; Xiong, R.; He, H.; Liu, Z. A LSTM-RNN Method for the Lithuim-Ion Battery Remaining Useful Life Prediction. In Proceedings of the 2017 Prognostics and System Health Management Conference, PHM-Harbin 2017—Proceedings, Harbin, China, 9–12 July 2017; pp. 1–4.
40. Guo, W.; Sun, Z.; Vilsen, S.B.; Meng, J.; Stroe, D.I. Review of "Grey Box" Lifetime Modeling for Lithium-Ion Battery: Combining Physics and Data-Driven Methods. *J. Energy Storage* **2022**, *56*, 105992. [CrossRef]
41. Liao, L.; Köttig, F. A Hybrid Framework Combining Data-Driven and Model-Based Methods for System Remaining Useful Life Prediction. *Appl. Soft Comput. J.* **2016**, *44*, 191–199. [CrossRef]
42. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008. [CrossRef]
43. Severson, K.A.; Monian, B.; Love, J.C.; Braatz, R.D. A Method for Learning a Sparse Classifier in the Presence of Missing Data for High-Dimensional Biological Datasets. *Bioinformatics* **2017**, *33*, 2897–2905. [CrossRef] [PubMed]
44. Severson, K.A.; Molaro, M.C.; Braatz, R.D. Principal Component Analysis of Process Datasets with Missing Values. *Processes* **2017**, *5*, 38. [CrossRef]
45. Jeong, H.; Wang, H.; Calmon, F.P. Fairness without Imputation: A Decision Tree Approach for Fair Prediction with Missing Values. In Proceedings of the 36th AAAI Conference on Artificial Intelligence, Virtual, 22 February–1 March 2022.
46. Twala, B.E.T.H.; Jones, M.C.; Hand, D.J. Good Methods for Coping with Missing Data in Decision Trees. *Pattern Recognit. Lett.* **2008**, *29*, 950–956. [CrossRef]
47. Chen, L.; Xu, Y.; Zhu, Q.X.; He, Y.L. Adaptive Multi-Head Self-Attention Based Supervised VAE for Industrial Soft Sensing with Missing Data. *IEEE Trans. Autom. Sci. Eng.* **2023**, 1–12. [CrossRef]
48. Wu, R.; Zhang, A.; Ilyas, I.F.; Rekatsinas, T. Attention-Based Learning for Missing Data Imputation in HoloClean. In Proceedings of the 3rd MLSys Conference, Austin, TX, USA, 15 March 2020.
49. Yu, J.; Wu, B. Attention and Hybrid Loss Guided Deep Learning for Consecutively Missing Seismic Data Reconstruction. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5902108. [CrossRef]

50. Ansari, S.; Ayob, A.; Hossain Lipu, M.S.; Hussain, A.; Saad, M.H.M. Remaining Useful Life Prediction for Lithium-Ion Battery Storage System: A Comprehensive Review of Methods, Key Factors, Issues and Future Outlook. *Energy Rep.* **2022**, *8*, 12153–12185. [CrossRef]
51. Ma, G.; Zhang, Y.; Cheng, C.; Zhou, B.; Hu, P.; Yuan, Y. Remaining Useful Life Prediction of Lithium-Ion Batteries Based on False Nearest Neighbors and a Hybrid Neural Network. *Appl. Energy* **2019**, *253*, 113626. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Review

Status and Prospects of Research on Lithium-Ion Battery Parameter Identification

Jianlin Li [1], Yuchen Peng [1,*], Qian Wang [1,*] and Haitao Liu [2]

[1] National User-Side Energy Storage Innovation Research and Development Center, North China University of Technology, Beijing 100144, China; dkyljl@163.com
[2] Jiangsu Collaborative Innovation Center for Smart Distribution Network, Nanjing Institute of Technology, Nanjing 211167, China; 13851424346@163.com
* Correspondence: 2023312080101@mail.ncut.edu.cn (Y.P.); wangqian@mail.ncut.edu.cn (Q.W.); Tel.: +86-199-3688-9338 (Y.P.)

Abstract: Lithium-ion batteries are widely used in electric vehicles and renewable energy storage systems due to their superior performance in most aspects. Battery parameter identification, as one of the core technologies to achieve an efficient battery management system (BMS), is the key to predicting and managing the performance of Li-ion batteries. However, due to the complex chemical reactions and thermodynamic processes inside lithium-ion batteries, coupled with the influence of the external environment, accurate identification of lithium-ion battery parameters has become an urgent problem to be solved. In addition, data-driven parameter identification can enable battery models to better understand battery behavior, which is one of the focuses of future research. For this reason, this paper comprehensively reviews the application of data-driven parameter identification methods in different scenarios. Firstly, the research briefly explains the working principle of lithium-ion batteries and the key parameters affecting their performance. Secondly, this paper deeply discusses data-driven methods for parameter identification, which are widely used nowadays, and provides improvement ideas to address the shortcomings of traditional methods. Finally, the paper discusses the challenges faced by parameter identification technology for lithium-ion batteries and envisages future prospects.

Keywords: battery management system; data-driven method; lithium-ion battery; parameter identification

1. Introduction

With the increasing installed capacity of new energy and the increasing popularity of electronic equipment, the rational use of large-scale integrated energy storage battery cells has become a current trend, and lithium-ion batteries are widely used because of their stable performance, long life, low pollution, and fast charging speed [1,2]. According to a white paper, the total shipment of lithium-ion batteries in the world in 2023 will be 1202.6 GW·h, a year-on-year increase of 25.6%. Global shipments of lithium-ion batteries are expected to reach 1926.0 GW·h by 2025 [3], so research on lithium-ion batteries is becoming more significant as the use of lithium-ion batteries increases worldwide.

At present, lithium-ion batteries are one of the main energy sources for electrochemical energy storage power stations and new energy vehicles, but due to the complex operating conditions of lithium-ion batteries, it is necessary to establish an effective battery management system (BMS) to monitor the batteries continuously. As an integrated electronic system, in addition to effectively monitoring the health and safety of the battery and ensuring its optimal performance, its key responsibilities also include calculating the main parameters such as voltage, current, temperature, state of charge (SOC), and state of health (SOH) of the battery [4,5], as well as protecting the battery from overcharge, overdischarge, abnormal temperature, and short circuits and implementing functions such as data logging

and fault diagnosis, etc. In addition, advanced BMS may also include thermal management, energy optimization, and research on battery charging and discharging strategies [6–8]. While battery model parameter identification plays a crucial role in realizing efficient battery management systems, traditional battery parameter identification methods often rely on complex empirical models or electrochemical models (EM), which require a large amount of experimental data and computational time. The empirical model is often called Equivalent Circuit Model (ECM) [9,10], and the basic idea is to use electrical characteristics of circuit components to describe the characteristics of electrochemical systems; in other words, it can describe batteries with voltage sources, resistors, and capacitors based on the circuit. ECM has fewer parameters and a simpler structure than other models, so it is the most widely used battery model in BMS. In another aspect, electrochemical models are mainly divided into pseudo two-dimensional (P2D) models, single particle (SP) models, and enhanced single particle (ESP) models [11,12]. In recent years, some researchers have also reviewed the methods of battery parameter identification. Barcellona et al. [13] and Madani et al. [14] have comprehensively summarized different models and parameter identification techniques for lithium-ion batteries.

However, with the rise of artificial intelligence technology, especially the rapid development of deep learning, data-driven parameter identification methods provide new research ideas for battery management. And most of the existing research focuses on specific types of data-driven identification methods, lacking a comprehensive comparison and in-depth analysis. In particular, how to effectively improve the accuracy of traditional physical models and rationally use flexible data-driven models to adapt complex and diverse battery application scenarios have become urgent problems to be solved. In addition, with the development of battery requirements in the direction of higher energy density, longer life, and lower cost, more new battery types and application scenarios have emerged; therefore, the parameter identification methods are also facing new challenges and requirements. In view of this, this paper aims to systematically analyze and compare the existing data-driven battery model parameter identification methods while providing optimization ideas and improvement directions for the problems existing in different identification methods. In addition, considering the latest battery technology and application requirements, this paper also involves emerging parameter identification methods that are in line with the current development trend of science and technology, which provides a reference for the subsequent selection of appropriate parameter identification methods in different scenarios and also helps to promote the development of battery technology and BMS. Finally, this paper discusses the current challenges of parameter identification of lithium-ion batteries and predicts possible research directions in the future.

The remaining sections are organized as follows. Section 2 introduces the basic principles and key parameters of the battery. Section 3 explains the parameter identification method based on least squares and its derivative algorithms and proposes modification ideas. Section 4 presents the existing data-driven parameter identification method and summarizes the analysis. The challenges and perspectives are provided in Section 5. The conclusions are provided in Section 6.

2. Structural Characteristics of Lithium-Ion Batteries

2.1. Internal Mechanism of Lithium-Ion Battery

A lithium-ion battery is a type of secondary battery that typically contains two kinds of compounds capable of embedding and detaching Li^+, serving as the anode and cathode of the battery, using the currently commercialized lithium-ion battery with graphite as the anode and lithium cobaltate as the cathode as an example. The electrolyte adopts ethylene carbonate (EC) and dimethyl carbonate (DMC), and the charging process is as follows: During the charging process, the potential of the power supply forces Li^+ to move from lithium cobaltate to the cathode, passing through the electrolyte and diaphragm embedded in graphite. The lithium-ion moves through the electrolyte to the negative electrode, passing through the electrolyte and separator to be embedded in graphite. Simultaneously,

electrons flow from the positive electrode to the negative electrode through an external circuit to replenish the charge balance formed by the lithium-ions in the negative electrode. This process is accompanied by the oxidation of Co^{3+} in the positive electrode material. The discharge process is reversed; Li^+ is released through an internal electrochemical drive and moves through the electrolyte and becomes embedded in the negative electrode. Electrons reach the positive electrode from the external circuit and trigger the reduction of high-valent cobalt, in which the battery releases electrical energy through the external circuit to power the device [15,16]. The basic principle is shown in Figure 1. The electrochemical reaction equation of the lithium-ion battery during cycling is as follows:

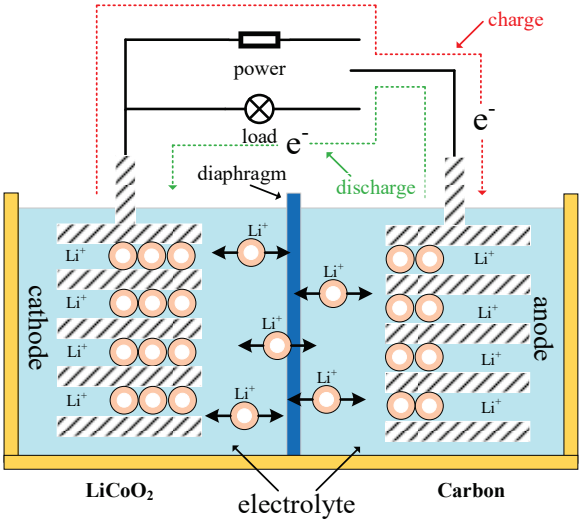


Figure 1. Working principles diagram of a rechargeable lithium-ion battery.

Anodic equation:

$$LiCoO_2 \underset{\text{discharge}}{\overset{\text{charge}}{\rightleftarrows}} Li_{1-x}CoO_2 + xLi^+ + xe^- \quad (1)$$

Cathode equation:

$$6C + xLi^+ + xe^- \underset{\text{discharge}}{\overset{\text{charge}}{\rightleftarrows}} Li_xC_6 \quad (2)$$

Overall reaction equation:

$$LiCoO_2 + 6C \underset{\text{discharge}}{\overset{\text{charge}}{\rightleftarrows}} Li_{1-x}CoO_2 + Li_xC_6 \quad (3)$$

2.2. Basic Parameters of a Lithium-Ion Battery

In order to understand and study the performance of lithium-ion batteries, it is necessary to start from the internal parameters of lithium-ion batteries, and the basic parameters of lithium-ion batteries are as follows:

1. Open-circuit voltage

Open circuit voltage (OCV) refers to the potential difference between the positive and negative poles of the battery when the battery is not connected to any load or power supply [17,18], which is one of the most significant indicators of the electrochemical state of the lithium-ion batteries, and it also plays an important role in evaluating the electrode materials [19] and aging state [20]. Besides, there is a nonlinear relationship between the OCV and SOC of the battery in a steady state, while different models of batteries have their

specific SOC-OCV curves, and accurate OCV values can provide a great foundation for the state estimation and study of optimization strategies of batteries [21].

2. Capacity

Battery capacity is one of the important indicators to measure the performance of the battery; it represents the amount of electricity released by the battery under certain conditions such as discharge rate, temperature, termination voltage, etc., which is generally divided into rated capacity and actual capacity [22]. The rated capacity refers to the capacity of the battery to work continuously for a long time under certain working conditions under the condition of an ambient temperature of 20 ± 5 °C, while the actual capacity refers to the actual amount of power that the battery can give under certain current density, termination voltage, and other conditions. The actual capacitance will be affected by various factors such as temperature, humidity, charge, and discharge rate during the actual use of the battery [23,24].

3. Internal resistance

The internal resistance of a lithium-ion battery is the resistance that hinders the passage of current during the conductive process of the battery. It is mainly divided into ohmic resistance and polarization resistance [25], and polarization resistance can be further classified into electrochemical polarization internal resistance and concentration polarization internal resistance. The ohmic internal resistance is affected by the electrode material, electrolyte, etc., resulting in continuous changes during the detection process, and the polarization internal resistance will change to varying degrees with the process of electrochemical reaction inside the battery, which has the dynamic characteristics of time and current [26]. The main measurement methods of battery internal resistance include the DC internal resistance method, AC impedance spectroscopy method, pulse discharge method, etc.

4. State of charge

The state of charge of a battery refers to the percentage of the battery's remaining charge, which is usually used to express the charge level of the battery. There are many ways to describe SOC; one that is widely used as in [27] is:

$$SOC(t_1) = SOC(t_0) - \frac{1}{C_r}\int_{t_0}^{t_1} I(t)dt \tag{4}$$

where the t_0 and t_1 is the start and end time, respectively. It is generally accepted that the SOC = 1 when the battery is fully charged using a standard charging method. Cr is the rated capacity of the battery. In the actual use process, the lithium-ion battery SOC will be affected by many factors such as charge and discharge current, ambient temperature, and self-discharge, which make it difficult to measure directly [28]. Accurate estimation of SOC value can ensure the stable and safe operation of related equipment; therefore, the estimation of SOC is also one of the important research directions at domestic and overseas.

5. Battery self-discharge rate

The self-discharge rate of a battery is also known as the charge retention capacity, which refers to the ability to store power under the open-circuit voltage condition [29]. It is seriously affected by temperature, and usually the lower the ambient temperature, the lower the self-discharge rate. Compared with other rechargeable batteries, lithium-ion batteries have a lower self-discharge rate, generally 2~5% at room temperature [30].

3. Parameter Identification Based on Least Squares and Its Derivative Algorithms

3.1. Least Squares

Least squares(LS) parameter identification is usually used for offline parameter identification of batteries, which is usually carried out under non-operating conditions, and its process is similar to the optimization method. The purpose is to match the measurement

curve with the target curve and find the optimal solution of the state parameters by solving the minimum value of the objective function [31,32]. The flowchart of the offline parameter identification is shown as an example in Figure 2.

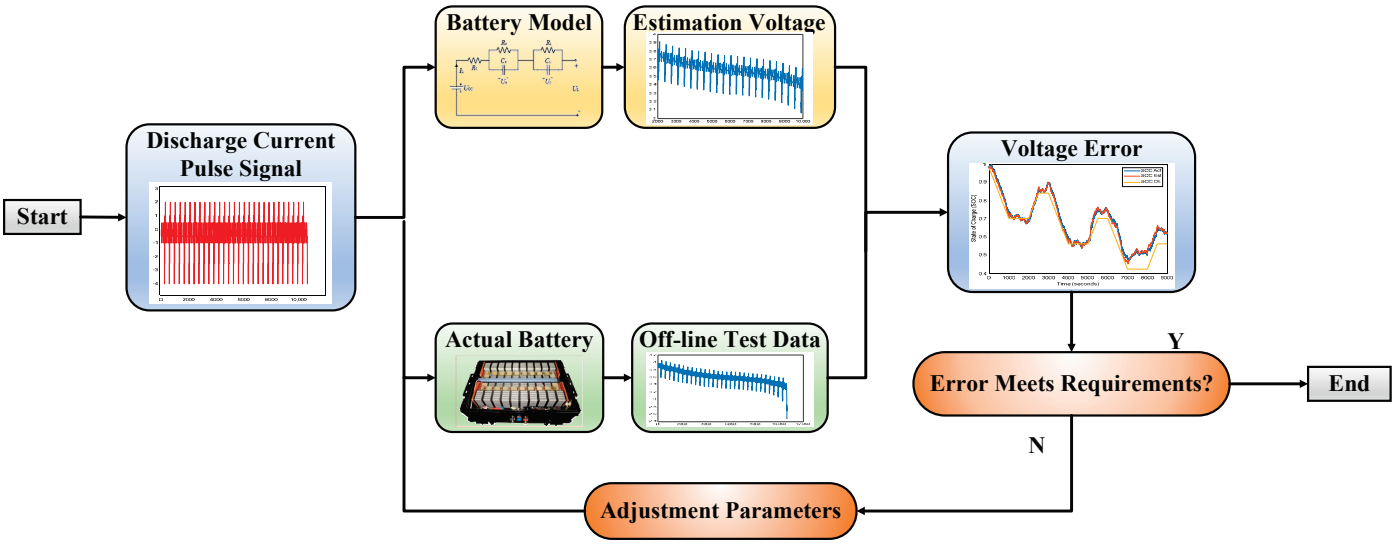


Figure 2. Flowchart of offline parameter identification.

Before the least squares method can be used to identify parameters, the battery needs to undergo a series of tests, such as charge-discharge cycles, pulse tests, etc. Its essence is to estimate the parameters by minimizing the sum of squares of the errors between the model prediction output and the experimental data [33], which is simple, time-consuming, and easy to identify [34], and the relevant algorithm process can be summarized as follows.

The mathematical model of the battery system is assumed to be:

$$A\left(z^{-1}\right)Y(i) = B\left(z^{-1}\right)X(i) + e(i) \tag{5}$$

where the *Y(i)* and the *X(i)* is the output and input of the system, and the *e(i)* is the systematic measurement error. The difference equation of the system is expressed as:

$$Y(i) = \sum_{k=1}^{n} a_i Y(i-k) + \sum_{k=1}^{n} b_i X(i-k) + e(i) \tag{6}$$

Writing Equation (5) in least squares form can be expressed as:

$$Y(i) = \varphi^T(i)\theta + e(i) \tag{7}$$

where the model parameters vector to be estimated and the input data vector can be expressed as follows:

$$\begin{cases} \theta = [a_1\ a_2 \cdots a_n\ b_1\ b_2 \cdots b_n]^T \\ \varphi(i) = [Y(i-1)Y(i-2)\cdots Y(i-n)X(i-1)X(i-2)\cdots X(i-n)]^T \end{cases} \tag{8}$$

After m observations, let the standard function J be:

$$J = \sum_{i=1}^{m} \left(Y(i) - \varphi(i)^T\theta\right)^2 = \left(Y - \varphi^T\theta\right)^T\left(Y - \varphi^T\theta\right) \tag{9}$$

where $\varphi(i)^T\theta$ is the value predicted by the model at the ith data point by the parameter vector θ. By minimizing the standard function J, when

$$\frac{\partial J}{\partial \theta} = \frac{\partial}{\partial \theta}[(Y - \varphi^T\theta)^T(Y - \varphi^T\theta)] = 0 \tag{10}$$

The best parameter vector estimate that can be found is:

$$\hat{\theta}_{LS} = (\varphi^T\varphi)^{-1}\varphi^T Y \tag{11}$$

Then the least squares form for $Y(i)$ can be expressed as:

$$Y(i) = \varphi^T(i)\theta \tag{12}$$

The core of the LS method to identify battery parameters aims to find a set of parameters that allow the mathematical model to best fit the behavior of the actual battery, and its advantage lies in the ability to analyze and optimize the parameters in detail without affecting the performance of the battery; therefore it is usually used in battery performance evaluation or the early stage of battery design [35]. However, the disadvantages of the LS method are also very obvious in that its fixed mode cannot meet the more complex working conditions; the identification of the parameter values cannot accurately reflect the real-time changes of the battery, so it is only applicable to some fixed scenarios [36].

3.2. Recursive Least Squares

Recursive least square (RLS) is one of the most common estimation methods currently in use, which is mainly suitable for the online parameter identification of batteries. The RLS method not only reduces a large number of preliminary experiments but also largely solves the time-varying problem of battery parameters [37]. The basic principle of online parameter identification is to use a unified input u(t), input to the system to be identified to generate the output y(t), and then superimpose with the measurement noise v(t) to form the observation output Z(t), compare with the model output generated in the system model to obtain the measurement error $\widetilde{Z}(t)$, and then import it into the identification algorithm to correct the error through the generated parameter estimation vector. The schematic is shown as an example in Figure 3.

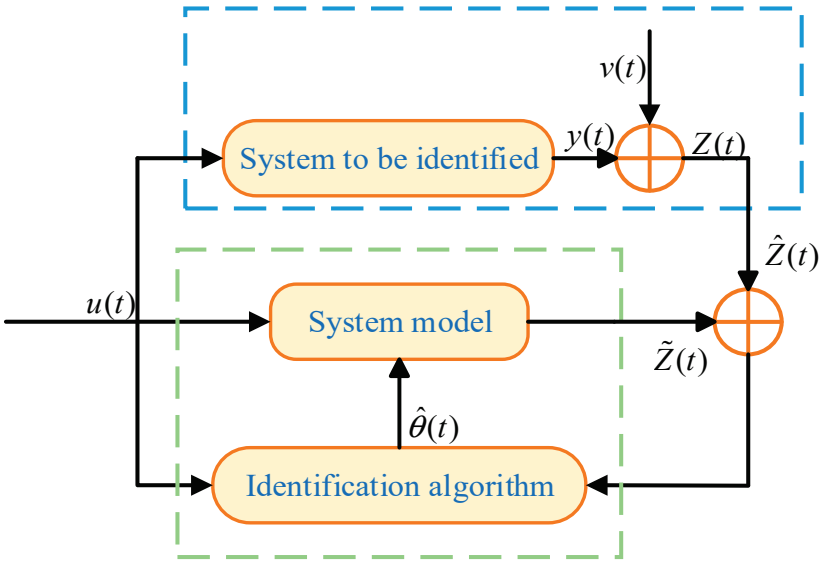


Figure 3. Basic principle of online identification.

The principle of the RLS method is to add a time-varying link on the basis of the LS method, which can update the parameter estimates in real time according to each new data

point, reducing the calculation of a large number of datasets. The nonlinear form of the RLS method is provided in Ref. [38], and the process evolves from Equation (11) to the recursive form is shown as:

$$\begin{cases} \hat{\theta}(i+1) = \hat{\theta}(i) + K(i+1) \times \left[Y(i+1) - \varphi^T(i+1)\hat{\theta}(i) \right] \\ e(i+1) = Y(i+1) - \varphi^T(i+1)\theta(i) \\ K(i+1) = \frac{P(i)\varphi(i+1)}{1+\varphi^T(i+1)P(i)\varphi(i+1)} \\ P(i+1) = P(i) - K(i+1)\varphi^T(i+1)P(i) \end{cases} \tag{13}$$

where the $\hat{\theta}$ is the parameter to be recognized, e is the estimated error of the system, K is the system gain, and P is the covariance matrix. According to the formula, the RLS method will calculate the result of the current LS according to the current new input data combined with the data calculated at the previous moment. However, in actual calculations, there are many areas that need to be improved in the process of parameter identification using the RLS method. Listed below are three aspects where the RLS method needs to be improved.

1. Data saturation

In the recursive process, as the number of cycles increases, the old data will continue to increase, and when accumulated to a certain extent, it will inevitably affect the recognition of new data, which ultimately leads to the termination of the parameter identification process. In order to reasonably allocate the proportion of old and new data and reduce the influence of old data on new data, researchers added a forgetting factor on the basis of RLS to match the parameter identification in different environments. The improved algorithm is called Forgetting Factor Recursive Least Squares (FFRLS). The recursive principle of FFRLS is based on Equation (13) and is shown as:

$$\begin{cases} \hat{\theta}(i+1) = \hat{\theta}(i) + K(i+1) \times \left[Y(i+1) - \varphi^T(i+1)\hat{\theta}(i) \right] \\ e(i+1) = Y(i+1) - \varphi^T(i+1)\theta(i) \\ K(i+1) = \frac{P(i)\varphi(i+1)}{1+\varphi^T(i+1)P(i)\varphi(i+1)} \\ P(i+1) = \frac{1}{\lambda}[P(i) - K(i+1)\varphi^T(i+1)P(i)] \end{cases} \tag{14}$$

where the λ is the forgetting factor, usually between 0.95 and 1, and different λ corresponds to different forgetting speeds; when $\lambda = 1$, it means no forgetting, and when $\lambda = 0$, it means all forgetting. Many scholars have tried to apply the FFRLS algorithm to different scenarios and found that the identification results are significantly improved compared with the traditional RLS algorithm. Xia et al. [39] used the FFRLS algorithm to continuously update the parameters in the battery model and then used the Kalman filter to estimate the SOC of the battery and verified the algorithm under different ambient temperatures and driving conditions, proving that the RLS algorithm can show good adaptability and accuracy in different scenarios. In addition, the selection of the forgetting factor in the FFRLS algorithm is also very important; different forgetting factors will affect the time-varying and accuracy of parameter identification. Based on this, Shi et al. [40] adjusted the forgetting factor by using the mean square value of the difference between the battery OCV and the terminal voltage in the sliding window mode. Lao et al. [41] proposed a variable forgetting factor least squares method Recursive Least Squares (AFFRLS), the principle of which is to adjust the size according to the error, and its relationship to the error can be expressed as:

$$\begin{cases} \lambda(k) = \lambda_{\min} + (1 - \lambda_{\min})^{\alpha(k)} \\ \alpha(k) = 2^{\rho e^2(k)} \end{cases} \tag{15}$$

where the $e(k)$ is the measurement error, $\alpha(k)$ is the intermediate variable, ρ is the undetermined coefficient, and its value range is within $[10^4, 5 \times 10^4]$. The most suitable λ-value is determined by adjusting the values of $e^2(k)$, but the convergence speed of this method is affected by the value of ρ, so the requirements for the adjustment of the coefficients are high. At the same time, considering the instability of the parameter results and the inaccuracy

of the initial state, Fan et al. [42] proposed an Adaptive Forgetting Factor Recursive Least Squares (AFFRLS) method based on Equation (15), which has both accuracy and stability, and it can make $e(k)$ quickly converge to the minimum value when λ is larger values, and when $e(k)$ is within the allowable error range, λ can rapidly approach the maximum. Its principle can be expressed as:

$$\begin{cases} \lambda(k) = \lambda_{\max} - (\lambda_{\max} - \lambda_{\min}) \cdot \arctan(\mu(k)) \cdot \frac{\pi}{2} \\ \mu(k) = round(\left| \frac{e(k)}{e_0} \right|^n) \end{cases} \tag{16}$$

where the $\lambda_{\max}$ and $\lambda_{\min}$ are the expected maximum and minimum values of λ, e_0 is the benchmark error. When $e(k)$ is smaller than e_0, the ratio by n exponentiation is scaled down, so that λ can quickly close to the $\lambda_{\max}$; when $e(k)$ is greater than e_0, the ratio expands and λ can more quickly converge to the $\lambda_{\min}$. Where the n is generally taken as 2 or 4, and the rounding function $round(\ \ )$ is used to reduce the intermediate of the transition value.

2. Unequal supply and demand

When using the LS method for parameter identification, the most obvious problem is the "inequality of supply and demand" caused by the incentive and identification parameters. In short, the excitation provided for parameter identification is less than the number of parameters that need to be identified. Taking the second-order RC equivalent circuit as an example, we need to identify five parameter results by only two inputs of voltage and current as the model excitation; in the higher order equivalent circuit, this problem will be more obvious, and in the actual process, we also need to consider the influence of various environmental factors, which results in a more complicated process. Therefore, it is very important to select the appropriate excitation input in order to fully reflect the internal characteristics of the lithium-ion battery. Considering that the parameter discrimination ability of lithium-ion batteries is poor when the input and output data are not suitable, Song et al. [43] found that the data selected for battery state estimation needs to have certain adaptability, while analyzing the accuracy of single-parameter and multi-parameter identification schemes, and verified that the recognition accuracy can be further improved when the current excitation meets certain criteria.

3. Different time scales

The dynamic characteristics of the battery are distributed in a wide frequency range. It contains both fast and slow response processes under the action of excitation, so the second-order RC circuit model will set two different time constants to correspond to different response links [44]. However, this lead to the result that the process of parameter identification cannot be ensured to take the corresponding links of each time constant into account at the same time, and when identifying one of the links, the other link may cause information loss or data saturation, which will greatly affect the accuracy of parameter identification [45,46]. In order to solve the above problems, relevant scholars propose a multi-time scale identification method. Taking ECM as an example, the basic principle is to apply different time constants to different time scales for identification, and its improvement idea is shown in Figure 4.

At present, many scholars use this method to improve the accuracy of ECM. Yang et al. [47] used a combination of the fixed memory recursive least squares method and fading extended Kalman filter method to obtain the parameters of the fast and slow dynamic links of the ECM, respectively, and the experimental results showed that the method had high consistency under different working conditions. In addition, many parameters of lithium-ion batteries change over time, so they can be improved along this line. Based on this, Shi et al. [48] established a multiple-time scales characteristics model for on-board lithium-ion batteries and distinguished between fast and slow changes in the model parameters of the resistance-capacitance links. Furthermore, Zou et al. [39] proposed a multi-time scale estimator that can adjust parameters on different time scales and verified the effectiveness of the method in verifying SOC and SOH. In general, multi-time-scale

identification allows the model to react sudden changes on fast time scales while adapting to sluggish changes on slow scales. This allows the model to have better responsiveness and flexibility while maintaining high accuracy.

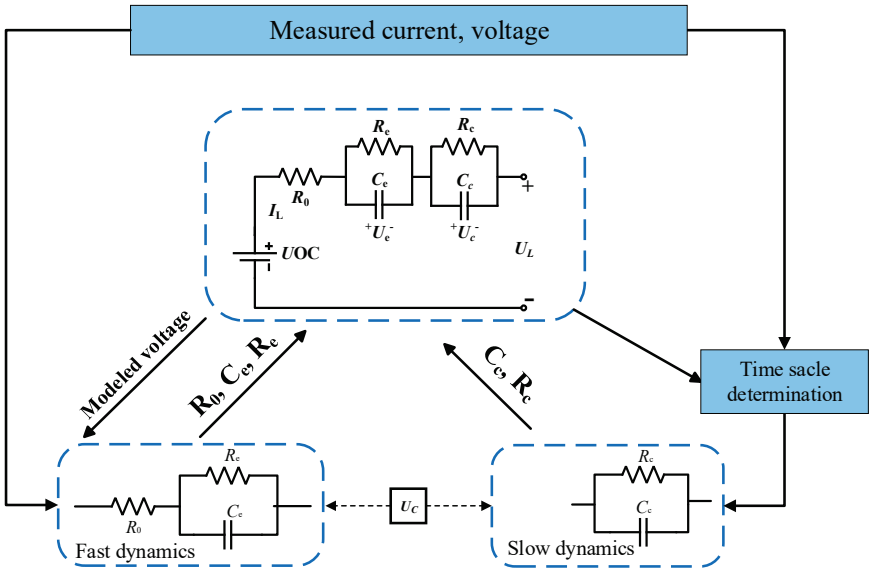


Figure 4. Improvement idea based on ECM with multiple time scales.

4. Data-Driven Parameter Identification

4.1. Heuristic Algorithms

Heuristic algorithms (HAs) are a kind of solution method constructed on the basis of specific empirical rules, which are often used to solve global optimal problems [49], but the traditional HA needs to evaluate the performance of the system by deriving all nonlinear constraint functions when solving practical problems [50], which is a very complex process. To solve this problem, relevant scholars have proposed meta-heuristic algorithms (MHA). This algorithm has better global search ability and higher robustness and has been effectively applied to constraint optimization problems in various fields [51–53], so it is also suitable for parameter identification of lithium-ion batteries, which can find the parameters closest to the actual data of the battery through iteration. Taking the most typical Genetic Algorithms (GA) [54] and Particle Swarm Optimization (PSO) in MHA as an example [55,56], when identifying battery parameters, the basic principle is to setting the parameters to be identified as the initial population, setting the error between the output of the battery model and the actual measured data as the fitness function, by evaluating the individuals separately through the fitness function and iterating continuously until the algorithm stop condition is satisfied, and then finding the optimal solution and output the optimal model parameters [57]. The process of GA and PSO is shown as an example in Figure 5.

Similar methods also include Tabu Search (TS) [58], Simulated Annealing (SA) [59], and Ant Colony Optimization (ACO) [60]. However, when using these algorithms, researchers have found that the convergence speed is relatively slow when used alone, and it is possible for GA and PSO to converge in the local optimal solution before reaching their global optimal solution. To face this problem, Srinivas et al. [61] applied adaptive probability to the intersection and mutation links of genetic algorithms, which effectively optimized the population diversity and convergence ability of genetic algorithms, but the method was limited to some links of GA and did not analyze from the whole part. In order to overcome this challenge, San et al. [62] applied adaptive links to all links of GA and proposed an Adaptive Genetic Algorithm (AGA), which provides an improved way to solve the identification problem of nonlinear systems. Moreover, considering that individual algorithms have obvious advantages and disadvantages when used, in order to maximize

the optimization ability of the algorithm, Garg H [50] proposed a PSO-GA hybrid algorithm to deal with nonlinear problems. The basic idea is to optimize the vector by using the PSO while using the genetic operator of GA to modify the decision vector, and it is considered as an effective algorithm that is suitable for various optimization problems. In addition, MHA is also suitable for model-specific optimization. Ren et al. [63] used PSO to optimize the key parameters of the long short-term memory (LSTM) network to match the data features of the lithium-ion battery with the network topology and experimentally verified that the method can effectively reduce the error. Recently, Nouri et al. [64] established an intelligent charge and discharge management system for electric vehicles and adopted a new method combining ANN and PSO to achieve data collection under different conditions, which greatly improved the robustness of the system.

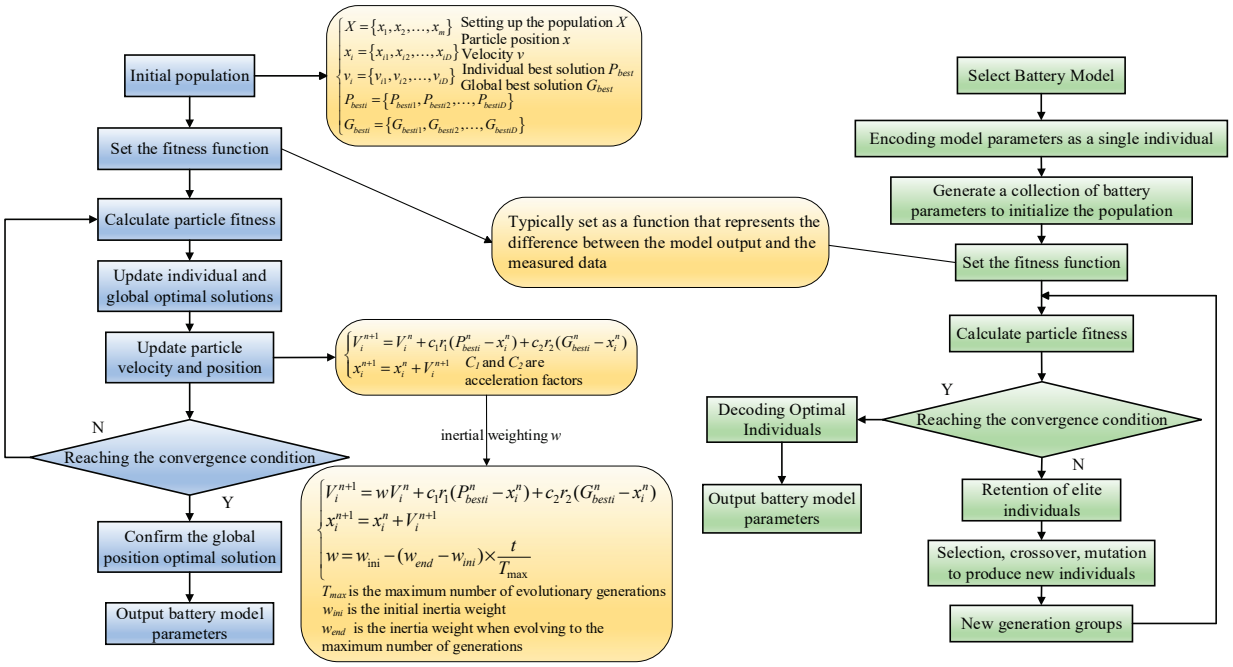


Figure 5. Flowchart of GA and PSO identification parameters.

4.2. Kalman Filter and Its Derivative Algorithms

Kalman filtering (KF) is an algorithm capable of estimating the state of a dynamic system from a series of noisy measurements. It was first proposed by E. Kalman in his paper on solving the linear filtering problem, which was published in 1960 [65]. The process mainly includes defining the equation of state and the equation of observation, initializing the state, predicting, updating, and iterating. The schematic diagram of KF is shown as an example in Figure 6.

EKF Most of the traditional KF algorithms are only applicable to state vectors in linear systems and cannot deal with the nonlinear equations in battery models. In order to solve this problem, researchers added a real-time linear Taylor approximation process to the estimation of the previous state of the system equation in the steps of state estimation and prediction and proposed an extended Kalman filtering (EKF) algorithm [66]. Based on this, Plett of the University of Colorado (Boulder, CO, USA), firstly applied it to the SOC estimation and parameter identification of lithium-ion batteries [67,68] and obtained good results. However, considering that the monotonic EKF algorithm cannot well solve the uncertainty and system noise of complex lithium-ion battery models, to address this challenge, He et al. [69] proposed an adaptive extended Kalman filter (AEKF) and verified that this algorithm can greatly improve the dependence of traditional filtering algorithms on the battery model. At the same time, in order to reduce the error caused by incorrect prior covariance in the EKF algorithm, Jaehyun et al. [70] also adaptively modified the

covariance by using the AEKF method, which improves the reliability and accuracy of the measurement process. Recently, with the gradual development of cloud platforms, Wang et al. [71] proposed a noise matrix self-adjustment-extended Kalman filter (NMSA-EKF) algorithm based on the powerful computing power and huge storage capacity of the cloud platform to estimate the SOC of cloud-based discharging fragments and verified its accuracy.

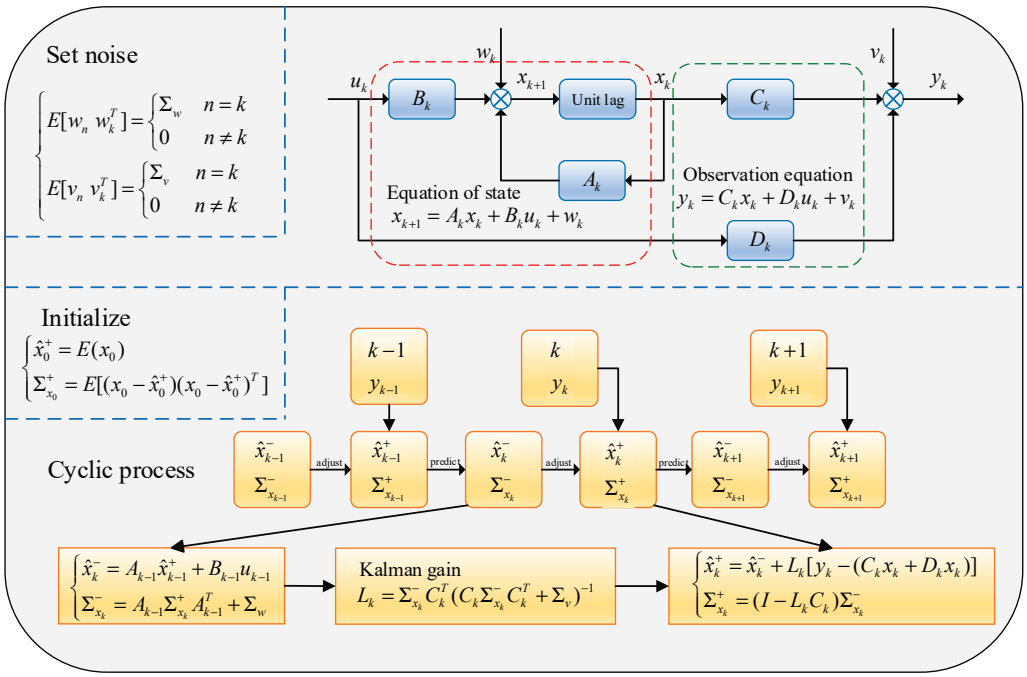


Figure 6. Schematic diagram of Kalman filtering.

UKF The traditional EKF only uses the first- or second-order terms of the Taylor series to approximate the nonlinear model. In order to solve the problem that the EKF produces a large error when facing a highly nonlinear system, Tian et al. [72] proposed unscented Kalman filtering (UKF) to estimate the SOC of the battery and verified that the UKF can capture the distribution characteristics of the nonlinear system through a series of carefully selected sigma points, which improves the accuracy of prediction. At the same time, He et al. [73] compared the EKF and UKF methods experimentally and further verified that the UKF algorithm has high accuracy and convergence in SOC estimation. Based on this, Meng et al. [74] proposed an adaptive unscented Kalman filter (AUKF) to adaptively adjust the process measurement noise and noise covariance, reducing the computational burden when updating the covariance matrix. Of course, the UKF is not without its weaknesses; for example, in the case of high measurement noise, the UKF can lose stability and cause divergence problems. In addition, due to the uncertainty of the battery electrochemical model, the UKF may not be able to produce good robustness [75].

PF In 1993, Gordon et al. [76] proposed a SIS-based bootstrap nonlinear filtering method, which laid the foundation for the particle filter (PF) algorithm. PF is also a nonlinear state estimator based on the Bayesian model, but the algorithm eliminates the constraint that the random quantity must meet the Gaussian distribution when solving the nonlinear filtering problem, so it is also suitable for parameter identification of lithium-ion battery models. Schwunk et al. [77] used PF to estimate lithium-iron phosphate batteries. In addition, considering that the classical PF technique may have a lack of samples or a decrease in prediction accuracy when the weight of the simulated sample is small or the diversity of sampled particles is reduced, Li et al. [78] proposed a mutated particle filter (MPF) algorithm for predicting the Remaining Useful Life (RUL) of the battery, which utilizes mutated particles to search for the extended region of a priori Probability Density

Function (PDF) thus exploring the posterior PDF more comprehensively. Other than that, Lai [79] proposed the combination of PF and EKF, which enables the omission of calculating the Jacobi matrix of battery model and reduces the complexity of the identification process. At present, the research on the PF algorithm is relatively scarce, and it is still a major challenge for battery parameter identification in practical applications. Recently, Wang et al. [80] used PSO to solve the degradation and diversity reduction of the PF algorithm and then proposed a double-scale dual particle filter (D-PF) method to adjust parameters in real time during battery aging and temperature changes, which has certain practical value.

4.3. New Algorithms Based on Machine Learning

With the rapid development of machine learning and artificial intelligence technologies, it has become a research hotspot to apply these advanced methods to the field of battery parameter identification and further optimize model performance and improve the identification accuracy by improving algorithms and calculation strategies. These new methods can not only process a large amount of complex data but also learn the deep characteristics of battery behavior, which is of great significance for the development of battery management systems in the future.

ANN The neural network (NN), also known as the artificial neural network (ANN), is a type of machine learning algorithm that achieves learning from experience by simulating the transmission of signals between neurons or interconnected nodes in the hierarchical structure of the human brain [81], and its schematic diagram is shown as an example in Figure 7. It was first proposed by neurophysiologist Warren McCulloch and mathematician Walter Pitts in 1943 [82]. In 1986, David Rumelhart et al. used the backpropagation (BP) algorithm to train multi-layer neural networks, which greatly improved the learning ability of neural networks [83]. Its ability to describe nonlinear relationships between inputs and outputs has been recognized in various fields. Related derived neural networks also include the recurrent neural network (RNN) [84,85], the convolutional neural network (CNN) [86,87], and long short-term memory (LSTM) [88,89], etc. Therefore, researchers have gradually discovered the potential of ANN in the field of batteries, which can achieve accurate estimation of various battery states without considering the internal electrochemical state of the battery due to the feature extraction and fitting ability of neural networks [90,91]. Cui et al. [92] provided an introduction to the use of different types of neural networks for estimating the SOC of lithium-ion batteries and summarized the principles, advantages, disadvantages, and current status. Based on this, Yang et al. [93] used the parameters obtained from the Hybrid Pulse Power Characterization (HPPC) test to train a three-layer BP neural network to verify its accuracy in SOH estimation. However, with the increasing requirements for battery prediction accuracy, more and more researchers are improving on classic neural network algorithms. In order to eliminate the influence of battery degradation on the estimation accuracy of the original training model, Kang et al. [94] proposed a new radial basis function neural network (RBFNN) and compared it with the traditional neural network through experiments. The results show that the model can effectively improve the accuracy of system SOC estimation and has better robustness for different battery aging cycles, temperatures, and load curves. In addition, in order to solve the dependence of traditional neural network systems on Kalman filtering, Chemali et al. [95] proposed a deep feedforward neural network (DFNN), which can directly map observable signals such as voltage and current to the battery SOC and can learn to estimate the SOC at different ambient temperatures through learning algorithms such as gradient descent. Recently, considering the increasing importance of studying the thermal behavior of batteries such as temperature change and heating rate, Pang et al. [96] integrated the positive and negative electrode surface concentrations in the thermodynamic single event model (SPMT) as physical information into bidirectional long short-term memory networks (BiLSTM) and constructed a new physics-based information neural network (PINN) while verifying its effectiveness under different driving conditions

and ambient temperatures, which opens up a new research direction for neural networks in the field of battery state estimation.

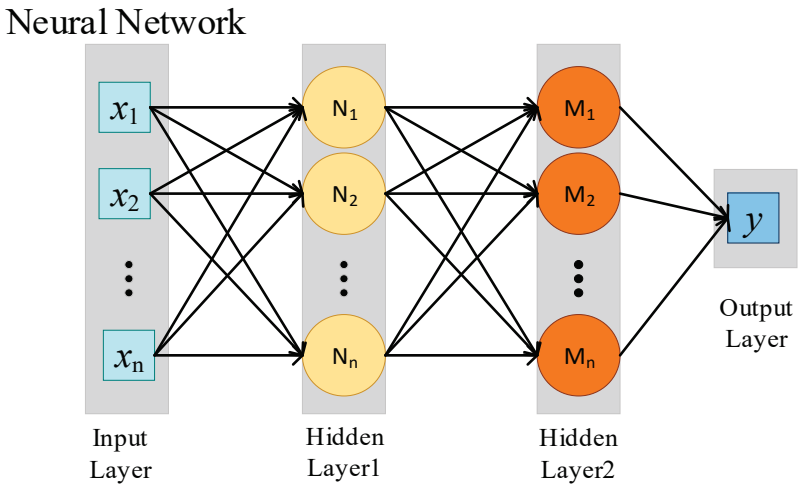


Figure 7. Schematic diagram of ANN.

SVM Support Vector Machines (SVM) is also one of the most popular machine learning algorithms, and its core idea is to find a hyperplane in the feature space based on the principle of statistical structured risk minimization to maximize the separation of different classes of data points. The schematic diagram of SVM is shown as an example in Figure 8. When dealing with nonlinear problems, SVM can map data to high-dimensional spaces through kernel functions and find separating planes in high-dimensional spaces [97]. Compared with the ANN, SVW has advantages in solving the problem of high-dimensional data model construction under the limited samples conditions, and because of its good generalization ability, it can perform better on small sample datasets [98]. Besides, it can effectively deal with nonlinear problems with the help of suitable kernel functions, so there are more scholars who have gradually have tried to use SVM for the parameter identification of lithium-ion batteries and found that SVM has excellent performance on battery life prediction. Based on this, Patil et al. [99] used SVM to establish a classification and regression model for the RUL of batteries and used the Support Vector Regression (SVR) to predict the accurate RUL of the battery. The results showed that the method could improve the prediction accuracy and calculation speed. In addition, in order to address the problem of long SVR training time, Hu et al. [100] used a two-step search to improve the training efficiency of SVR, avoiding the behavior of blindly searching for parameters over a wide range and improving the accuracy and robustness of battery SOC estimation when using SVR under more complex working conditions. Recently, considering the problem of data anomalies in the training of SVM models, Xiong et al. [101] proposed a weighted least squares SVM-based method for early prediction method for the life of lithium-ion batteries, which improved the prediction results through the error square term and weight coefficient, and verified the effectiveness of the method through experiments. It provides a theoretical basis for the battery system faults hierarchical management strategy.

ANN and SVM are essentially data-driven algorithms, and similar algorithms include Gaussian regression [102,103], logistic regression [104,105], gradient enhancement algorithms [106,107], and other machine learning algorithms. These algorithms show very good adaptability in the face of nonlinearity and relative complexity and can find the optimal model and parameters in a large amount of data. The use of appropriate learning algorithms in the BMS is often effective in improving the performance of the battery.

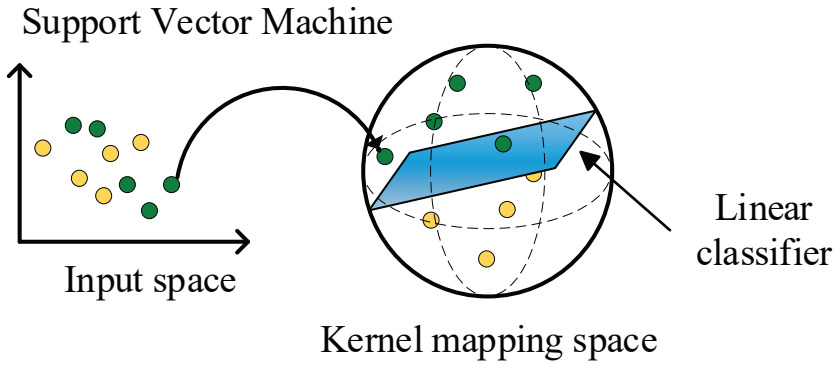


Figure 8. Schematic diagram of Support Vector Machines.

4.4. Comparison of Parameter Identification Methods

In an in-depth discussion of lithium-ion battery management and optimization strategies, it is essential to fully understand and accurately identify the key parameters of the battery. These parameters not only directly affect the performance, lifetime, and safety of the battery but also serve as the basis for optimizing the design of the battery management system and implementing efficient energy utilization strategies. Since different parameter identification methods have their own unique advantages and application scenarios, choosing the most appropriate method has become a prerequisite for achieving efficient battery management. To this end, Table 1 summarizes the current parameter identification methods for lithium-ion batteries, analyzes their disadvantages, and discusses their improvement directions, which provides a reference for the development and optimization of battery technology and subsequent research on parameter identification.

In addition to the improvement directions of each method discussed above, the application scenarios of different parameter identification methods are also different. The current application scenarios that are widely used can be summarized as small-medium capacity battery systems mainly for portable electronic devices and electric vehicles and large-scale battery systems used as energy storage devices in power systems. In one aspect, regarding small-medium capacity systems, linear regression methods based on RLS and algorithms based on KF can exert effective results, while these methods have excellent performance in real-time monitoring and basic performance prediction and can effectively estimate the state variables of the batteries in dynamic systems. Besides, it is also possible to simulate the charging and discharging behavior of small-capacity batteries by establishing smaller neural networks, which can also achieve good results. Moreover, SVM is suitable for the case of small feature spaces, so it can effectively monitor the health condition of small-scale batteries. In another aspect, large-scale neural networks such as LSTM have the ability to deal with high-dimensional and high-nonlinear problems and thus can effectively deal with richer and more complex data sets in large-scale energy storage battery systems. In addition, heuristic algorithms such as GA and PSO are able to show better accuracy in optimizing battery management strategies and life prediction in large-scale battery systems. In actual use, the most appropriate parameter identification methods should be selected according to the different battery sizes and application scenarios. If necessary, different methods also can be used in combination to complement each other in order to maximize the effect of each method to achieve a more efficient battery management system.

Table 1. Summary comparison of parameter identification methods.

Algorithms	Problems	Improving Directions	Literature
LS	Cannot deal with non-linear conditions	Combine with online models	[31–36,108]
RLS	Data saturation	Add the forgetting factor	[39–42]
	Unequal supply and demand	Choose the right incentive	[43]
	Different time scales	Segment at different time scales	[44–48]
GA	Slower convergence	Combine with local search algorithms	[109]
	Difficult to find the global optimal solution when facing high-dimensional problems	Combine with specific heuristic algorithms	[50,110]
	Higher requirements for parameterization	Use adaptive parameter tuning strategies	[61,62]
PSO	May fall into local optimization in solving complex problems	combine local search algorithms	[111]
	Convergence speed may drop when approaching the global optimal solution	Introduce a convergence factor	[112]
	Higher requirements for parameterization	Use adaptive parameter tuning strategies	[111]
EKF	Jacobi matrices increase computational complexity	Combine new algorithms to simplify EKF calculations	[79,113]
	Algorithm performance is more dependent on the initial state estimate	Optimize initial state estimation by preprocessing data or using a priori estimation	[114]
	The accuracy is affected by noise	Add an adaptive mechanism to adjust the noise	[69,70]
UKF	The performance depends heavily on the selection of the Sigma point	Add an adaptive mechanism to adjust the distribution of sigma points based on estimated performance	[74]
PF	The reduction in diversity creates sample impoverishment	Use resampling techniques to increase the diversity of particles	[77]
	Reduction of effective particles will affect the efficiency of the filter.	Add an adaptive mechanism adjust the number of particles	[115]
ANN	Complicated process, long time needed, large amount of calculation	Use efficient neural network architectures	[84–87]
	Difficult to adjust due to too many parameters	Apply automated machine—learning frameworks or adding hyperparametric optimization techniques	[116–119]
SVM	Higher requirements for kernel function setup	Add optimization algorithms to select parameters	[120,121]
	Kernel function cannot deal with non-linear conditions	Visualization of support vectors and decision boundaries	[122]

5. Challenges and Perspectives

As one of the core industries in the current new energy field, lithium-ion batteries have gradually revealed their key influence in many fields. With the rapid improvement and innovation of science and technology, the demand for performance stability, safety, and economy of lithium-ion battery continues to increase, which increases the reliance on battery management systems. As the core components of a BMS, battery modeling and parameter identification are important means to ensure its accuracy and reliability. Although current researchers have achieved many important results in this field, accurate and efficient studying of lithium-ion batteries still faces serious challenges. Therefore, this section focuses on three aspects to look forward to the research trend of battery modeling and parameter identification, hoping to provide more in-depth theoretical support for battery performance optimization, service life extension, and the guarantee of safety.

5.1. Efficient and Accurate Parameter Identification

Currently, there are relatively mature parameter identification methods for simple ECM models such as least squares algorithms, filtering algorithms, and heuristic algorithms such as genetic algorithms and particle swarm algorithms. However, when it comes to higher-order or complex electrochemical models, we still need to continuously improve the efficiency and accuracy of the model parameter identification. One approach is to apply artificial intelligence techniques to parameter identification, such as the deep learning models DNN, CNN, etc., which can learn from the battery operation data and reduce the computation time while maintaining the identification accuracy. In addition, with the development of chip technology, using emerging technologies such as cloud computing and edge computing for parameter identification can take advantage of more powerful computing resources and more efficient data processing capabilities while realizing the rapid deployment of practical applications and is also gradually becoming a popular area of current research.

5.2. Integration of Multiple Identification Methods

In practical applications, different identification methods are suitable for specific scenarios and requirements due to their inherent advantages and limitations. For example, although physical model-based approaches can provide in-depth insights into the interior of the battery, they have higher computational complexity and may not be suitable for situations that require fast response. In contrast, data-driven approaches are more suitable for real-time monitoring and control applications due to their lower computational requirements and rapid adaptability. By integrating multiple algorithms, such as combining the accuracy of the physical models with the fast responsiveness of data-driven models, the respective deficiencies can be complemented to a certain extent to achieve better identification results. This multi-method fusion strategy can not only improve the accuracy of the model but also enhance the adaptability of the model to different operating conditions. Therefore, future research will pay more attention to algorithms innovation and the development of multi-algorithm fusion technology while strengthening the analysis of the requirements of different application scenarios to guide the selection and optimization of the battery parameter identification methods and further enhance the overall performance and reliability of the battery management system.

5.3. Consideration of Multiple Categories of Influences

In the process of actual parameter identification, how to comprehensively consider the impact of temperature changes, battery aging, and other factors on the performance of lithium-ion batteries is also one of the challenges faced. The following are three aspects of the current research trend:

(1) The integration of multi-physical field models

In the multi-scale model, the combination of microscopic electrochemical models and macroscopic thermal and mechanical models can be achieved through the formation of multi-scale, multi-physical field comprehensive models in order to comprehensively reflect the behavior of the battery under different working conditions; under multiple environmental conditions, a model identification framework that can accurately perform parameter identification under various conditions such as different temperatures, humidity, pressure, etc., can be developed to improve the applicability and robustness of the model.

(2) Consideration of the battery aging phenomenon

The study of the battery aging model and the integration of battery aging mechanisms into the parameter identification model can improve the reliability of the parameter identification results and provide the basis for the health management and timely maintenance of lithium-ion batteries.

(3) Consideration of multiple aspects

Combining the research results of many disciplines, such as chemistry, material science, electronic engineering, computer science, etc., can systematically solve the problems in the process of lithium-ion battery parameter identification and improve the comprehensive performance and application scope of identification. For example, among the selection of cathode materials for lithium-ion batteries, Ni-rich cathodes materials (NCM) are becoming increasingly dominant in the field of lithium-ion batteries, while future research will continue to be directed towards high-nickel batteries that can significantly increase energy density [123]. In addition, lithium-sulfur is becoming a popular battery system due to its high capacity and excellent cycle efficiency [124]. Besides, comprehensively considering the voltage, capacity, and temperature sensitivity of the battery will further enable higher performance, cost efficiency, and sustainability.

6. Conclusions

Lithium-ion batteries are a widely used energy source in electric vehicles and renewable energy storage systems, and the parameter identification of these batteries is essential for achieving an efficient BMS. With the continuous development of technology, data-driven parameter identification methods have attracted more attention in recent years. However, most of the current studies are only for specific data-driven methods; this type of method has not been comprehensively reviewed. To this end, this paper mainly summarizes the applications of data-driven parameter identification methods in different scenarios, and advantages and limitations of each method are also discussed while exploring the challenges and future research directions of these methods.

Upon analyzing and reviewing the related literature, it has been found that in the face of higher-order or more complex electrochemical models, effectively improving the accuracy and stability of model parameter identification is still an important research direction. In addition, although the RLS method is more mature, there is still much room for improving the adaptive ability, reducing the computational complexity, and improving the robustness of the algorithm. For the filtering algorithm, how to flexibly deal with the strong nonlinearity of the electrochemical model and adjust the system noise efficiently is still an urgent problem to be solved. Furthermore, with the rapid development of machine learning and artificial intelligence technology, the reasonable use of deep learning models can effectively reduce the computation time and maintain the accuracy of recognition.

Combining the future perspectives, it is realized that although existing research has made remarkable achievements in battery parameter identification, lithium-ion battery parameter identification still faces many challenges and opportunities compared with the complex and changing application scenarios and continuously increasing performance requirements. This paper provides research scholars with a reference for battery parameter identification in the hope of realizing the development of more advanced and efficient battery management systems.

Author Contributions: Conceptualization, J.L., Y.P., Q.W. and H.L.; methodology, J.L. and Y.P.; validation, J.L., Y.P. and Q.W.; investigation, Y.P.; resources, J.L. and H.L.; data curation, J.L., Y.P. and Q.W.; writing—original draft preparation, Y.P.; writing—review and editing, J.L., Y.P. and Q.W.; visualization, J.L.; supervision, J.L. and Q.W.; project administration, J.L. and H.L.; funding acquisition, J.L. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China [NO. 52277211]; North China University of Technology Scientific Research Start-up Fund Project: Organized Scientific Research 2-National Energy User-side Energy Storage Innovation R&D Center. [NO: 110051360024XN149-09].

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts.

References

1. Etacheri, V.; Marom, R.; Elazari, R.; Salitra, G.; Aurbach, D. Challenges in the development of advanced Li-ion batteries: A review. *Energy Environ. Sci.* **2011**, *4*, 3243–3262. [CrossRef]
2. Nitta, N.; Wu, F.; Lee, J.T.; Yushin, G. Li-ion battery materials: Present and future. *Mater. Today* **2015**, *18*, 252–264. [CrossRef]
3. White Paper on China's New Energy Storage Industry 2024 [R/OL]. Available online: https://www.esacn.com.cn/20240310/dd7cf7271bfc45759089a733fcb6bbd1/c.html (accessed on 10 March 2024).
4. Shen, M.; Gao, Q. A review on battery management system from the modeling efforts to its multiapplication and integration. *Int. J. Energy Res.* **2019**, *43*, 5042–5075. [CrossRef]
5. Lelie, M.; Braun, T.; Knips, M.; Nordmann, H.; Ringbeck, F.; Zappen, H.; Sauer, D.U. Battery Management System Hardware Concepts: An Overview. *Appl. Sci.* **2018**, *8*, 534. [CrossRef]
6. Habib, A.K.M.A.; Hasan, M.K.; Issa, G.F.; Singh, D.; Islam, S.; Ghazal, T.M. Lithium-Ion Battery Management System for Ele ctric Vehicles: Constraints, Challenges, and Recommendations. *Batteries* **2023**, *9*, 152. [CrossRef]
7. Thakur, A.K.; Sathyamurthy, R.; Velraj, R.; Saidur, R.; Pandey, A.K.; Ma, Z.; Singh, P.; Hazra, S.K.; Sharshir, S.W.; Prabakaran, R.; et al. A state-of-the art review on advancing battery thermal management systems for fast-charging. *J. Appl. Therm. Eng.* **2023**, *226*, 120303. [CrossRef]
8. Hasan, M.K.; Mahmud, M.; Habib, A.K.M.; Motakabber, S.M.A.; Islam, S. Review of electric vehicle energy storage and management system: Standards, issues, and challenges. *J. Energy Storage* **2021**, *41*, 102940. [CrossRef]
9. Hu, X.; Li, S.; Peng, H. A comparative study of equivalent circuit models for Li-ion batteries. *J. Power Sources* **2012**, *198*, 359–367. [CrossRef]
10. He, H.; Xiong, R.; Fan, J. Evaluation of lithium-ion battery equivalent circuit models for state of charge estimation by an experimental approach. *Energies* **2011**, *4*, 582–598. [CrossRef]
11. Wang, Y.; Li, J.; Zhang, J.; Pecht, M. Lithium-iron-phosphate battery electrochemical modelling under a wide range of ambient temperatures. *J. Electroanal. Chem.* **2021**, *882*, 115041. [CrossRef]
12. Kanamura, K.; Yamada, Y.; Annaka, K.; Nakata, N.; Munakata, H. Electrochemical Evaluation of Active Materials for Lithium Ion Batteries by One (Single) Particle Measurement. *Electrochemistry* **2016**, *84*, 759–765. [CrossRef]
13. Barcellona, S.; Piegari, L. Lithium ion battery models and parameter identification techniques. *Energies* **2017**, *10*, 2007. [CrossRef]
14. Madani, S.S.; Schaltz, E.; Kær, S.K. A review of different electric equivalent circuit models and parameter identification methods of lithium-ion batteries. *ECS Trans.* **2018**, *87*, 23. [CrossRef]
15. Ingemar, K.; Victorien, K. Modeling battery cells under discharge using kinetic and stochastic battery models. *Appl. Math. Model.* **2016**, *40*, 7901–7915.
16. Asef, P.; Milan, M.; Lapthorn, A.; Padmanaban, S. Future trends and aging analysis of battery energy storage systems for electric vehicles. *Sustainability* **2021**, *13*, 13779. [CrossRef]
17. Elumalai, N.K.; Uddin, A. Open circuit voltage of organic solar cells: An in-depth review. *Energy Environ. Sci.* **2016**, *9*, 391–410. [CrossRef]
18. Pattipati, B.; Balasingam, B.; Avvari, G.V.; Pattipati, K.R.; Bar-Shalom, Y. Open circuit voltage characterization of lithium-ion batteries. *J. Power Sources* **2014**, *269*, 317–333. [CrossRef]
19. Tian, J.; Xiong, R.; Shen, W.; Sun, F. Electrode ageing estimation and open circuit voltage reconstruction for lithium ion batteries. *Energy Storage Mater.* **2021**, *37*, 283–295. [CrossRef]
20. Lavigne, L.; Sabatier, J.; Francisco, J.M.; Guillemard, F.; Noury, A. Lithium-ion Open Circuit Voltage (OCV) curve modelling and its ageing adjustment. *J. Power Sources* **2016**, *324*, 694–703. [CrossRef]
21. Zhang, C.; Jiang, J.; Zhang, L.; Liu, S.; Wang, L.; Loh, P.C. A generalized SOC-OCV model for lithium-ion batteries and the SOC estimation for LNMCO battery. *Energies* **2016**, *9*, 900. [CrossRef]
22. Rong, P.; Pedram, M. An analytical model for predicting the remaining battery capacity of lithium-ion batteries. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2006**, *14*, 441–451. [CrossRef]
23. Ma, S.; Jiang, M.; Tao, P.; Song, C.; Wu, J.; Wang, J.; Deng, T.; Shang, W. Temperature effect and thermal impact in lithium-ion batteries: A review. *Prog. Nat. Sci. Mater. Int.* **2018**, *28*, 653–666. [CrossRef]
24. Djian, D.; Alloin, F.; Martinet, S.; Lignier, H.; Sanchez, J.Y. Lithium-ion batteries with high charge rate capacity: Influence of the porous separator. *J. Power Sources* **2007**, *172*, 416–421. [CrossRef]
25. Chiang, Y.H.; Sean, W.Y.; Ke, J.C. Online estimation of internal resistance and open-circuit voltage of lithium-ion batteries in electric vehicles. *J. Power Sources* **2011**, *196*, 3921–3932. [CrossRef]
26. Schweiger, H.-G.; Obeidi, O.; Komesker, O.; Raschke, A.; Schiemann, M.; Zehner, C.; Gehnen, M.; Keller, M.; Birke, P. Comparison of several methods for determining the internal resistance of lithium ion cells. *Sensors* **2010**, *10*, 5604–5625. [CrossRef]
27. Hannan, M.A.; Lipu, M.S.H.; Hussain, A.; Mohamed, A. A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations. *Renew. Sustain. Energy Rev.* **2017**, *78*, 834–854. [CrossRef]
28. Rivera-Barrera, J.P.; Muñoz-Galeano, N.; Sarmiento-Maldonado, H.O. SoC estimation for lithium-ion batteries: Review and future challenges. *Electronics* **2017**, *6*, 102. [CrossRef]
29. Sloop, S.E.; Kerr, J.B.; Kinoshita, K. The role of Li-ion battery electrolyte reactivity in performance decline and self-discharge. *J. Power Sources* **2003**, *119*, 330–337. [CrossRef]

30. Niu, J.; Conway, B.E.; Pell, W.G. Comparative studies of self-discharge by potential decay and float-current measurements at C double-layer capacitor and battery electrodes. *J. Power Sources* **2004**, *135*, 332–343. [CrossRef]
31. Ahmed, R.; Rahimifard, S.; Habibi, S. Offline parameter identification and soc estimation for new and aged electric vehicles batteries. In Proceedings of the 2019 IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, USA, 19–21 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
32. Miniguano, H.; Barrado, A.; Lázaro, A.; Zumel, P.; Fernández, C. General parameter identification procedure and comparative study of Li-Ion battery models. *IEEE Trans. Veh. Technol.* **2019**, *69*, 235–245. [CrossRef]
33. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [CrossRef]
34. Lawson, C.L.; Hanson, R.J. *Solving Least Squares Problems*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1995.
35. Unterrieder, C.; Zhang, C.; Lunglmayr, M.; Priewasser, R.; Marsili, S.; Huemer, M. Battery state-of-charge estimation using approximate least squares. *J. Power Sources* **2015**, *278*, 274–286. [CrossRef]
36. Markovsky, I.; Van Huffel, S. Overview of total least-squares methods. *Signal Process.* **2007**, *87*, 2283–2302. [CrossRef]
37. Duong, V.H.; Bastawrous, H.A.; Lim, K.C.; See, K.W.; Zhang, P.; Dou, S.X. Online state of charge and model parameters estimation of the LiFePO4 battery in electric vehicles using multiple adaptive forgetting factors recursive least-squares. *J. Power Sources* **2015**, *296*, 215–224. [CrossRef]
38. Engel, Y.; Mannor, S.; Meir, R. The kernel recursive least-squares algorithm. *IEEE Trans. Signal Process.* **2004**, *52*, 2275–2285. [CrossRef]
39. Xia, B.; Lao, Z.; Zhang, R.; Tian, Y.; Chen, G.; Sun, Z.; Wang, W.; Sun, W.; Lai, Y.; Wang, M.; et al. Online Parameter Identification and State of Charge Estimation of Lithium-Ion Batteries Based on Forgetting Factor Recursive Least Squares and Nonlinear Kalman Filter. *Energies* **2018**, *11*, 3. [CrossRef]
40. Shi, J.; Guo, H.; Chen, D. Parameter identification method for lithium-ion batteries based on recursive least square with sliding window difference forgetting factor. *J. Energy Storage* **2021**, *44*, 103485. [CrossRef]
41. Lao, Z.; Xia, B.; Wang, W.; Sun, W.; Lai, Y.; Wang, M. A novel method for lithium-ion battery online parameter identification based on variable forgetting factor recursive least squares. *Energies* **2018**, *11*, 1358. [CrossRef]
42. Fan, X.; Feng, H.; Zhang, X. Research on optimization of least squares algorithm and its application in parameter identification of lithium-ion batteries. *Trans. China Electrotech. Soc.* **2024**, *39*, 1577–1588.
43. Song, Z.; Hofmann, H.; Lin, X.; Han, X.; Hou, J. Parameter identification of lithium-ion battery pack for different applications based on Cramer-Rao bound analysis and experimental study. *Appl. Energy* **2018**, *231*, 1307–1318. [CrossRef]
44. Zou, C.; Manzie, C.; Nešić, D.; Kallapur, A.G. Multi-time-scale observer design for state-of-charge and state-of-health of a lithium-ion battery. *J. Power Sources* **2016**, *335*, 121–130. [CrossRef]
45. Hua, Y.; Cordoba-Arenas, A.; Warner, N.; Rizzoni, G. A multi time-scale state-of-charge and state-of-health estimation framework using nonlinear predictive filter for lithium-ion battery pack with passive balance control. *J. Power Sources* **2015**, *280*, 293–312. [CrossRef]
46. Ye, M.; Guo, H.; Xiong, R.; Yu, Q. A double-scale and adaptive particle filter-based online parameter and state of charge estimation method for lithium-ion batteries. *Energy* **2018**, *144*, 789–799.
47. Yang, Z.; Wang, X. An improved parameter identification method considering multi-timescale characteristics of lithium-ion batteries. *J. Energy Storage* **2023**, *59*, 106462.
48. Shi, H.; Wang, S.; Wang, L.; Xu, W.; Fernandez, C.; Dablu, B.E.; Zhang, Y. On-line adaptive asynchronous parameter identification of lumped electrical characteristic model for vehicle lithium-ion battery considering multi-time scale effects. *J. Power Sources* **2022**, *517*, 230725. [CrossRef]
49. Rashedi, E.; Nezamabadi-Pour, H.; Saryazdi, S. GSA: A gravitational search algorithm. *Inf. Sci.* **2009**, *179*, 2232–2248. [CrossRef]
50. Garg, H. A hybrid PSO-GA algorithm for constrained optimization problems. *Appl. Math. Comput.* **2016**, *274*, 292–305.
51. He, Q.; Wang, L. An effective co-evolutionary particle swarm optimization for constrained engineering design problems. *Eng. Appl. Artif. Intell.* **2007**, *20*, 89–99.
52. Dimopoulos, G.G. Mixed-variable engineering optimization based on evolutionary and social metaphors. *Comput. Methods Appl. Mech. Eng.* **2007**, *196*, 803–817. [CrossRef]
53. Lee, K.S.; Geem, Z.W. A new meta-heuristic algorithm for continuous engineering optimization: Harmony search theory and practice. *Comput. Methods Appl. Mech. Eng.* **2005**, *194*, 3902–3933. [CrossRef]
54. Goldberg, D.E. *Genetic Algorithm in Search, Optimization and Machine Learning*; Addison-Wesley Publishing Company, Inc.: Reading, MA, USA, 1989.
55. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; IEEE: Piscataway, NJ, USA, 1995; Volume 4, pp. 1942–1948.
56. Eberhart, R.; Kennedy, J. A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995*; IEEE: Piscataway, NJ, USA, 1995; pp. 39–43.
57. Moldovan, N.; Picos, R.; Garcia-Moreno, E. Parameter extraction of a solar cell compact model usign genetic algorithms. In Proceedings of the 2009 Spanish Conference on Electron Devices, Santiago de Compostela, Spain, 11–13 February 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 379–382.

58. Abido, M.A. Optimal power flow using tabu search algorithm. *Electr. Power Compon. Syst.* **2002**, *30*, 469–483. [CrossRef]
59. van Laarhoven, P.J.M.; Aarts, E.H.L. Simulated annealing. In *Simulated Annealing: Theory and Applications*; Mathematics and Its Applications; Springer: Dordrecht, The Netherlands, 1987; Volume 37. [CrossRef]
60. Dorigo, M.; Blum, C. Ant colony optimization theory: A survey. *Theor. Comput. Sci.* **2005**, *344*, 243–278. [CrossRef]
61. Srinivas, M.; Patnaik, L.M. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Trans. Syst. Man Cybern.* **1994**, *24*, 656–667. [CrossRef]
62. Ren, Z.; San, Z. Study on the improvement of adaptive genetic algorithm and its application in system identification. *J. Syst. Simul.* **2006**, *18*, 41–43+66.
63. Ren, X.; Liu, S.; Yu, X.; Dong, X. A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM. *Energy* **2021**, *234*, 121236. [CrossRef]
64. Nouri, A.; Lachheb, A.; El Amraoui, L. Optimizing efficiency of Vehicle-to-Grid system with intelligent management and ANN-PSO algorithm for battery electric vehicles. *Electr. Power Syst. Res.* **2024**, *226*, 109936. [CrossRef]
65. Kalman, R.E. A new approach to linear filtering and prediction problems. *J. Basic Eng. Mar.* **1960**, *82*, 35–45. [CrossRef]
66. Chui, C.K.; Chen, G. *Kalman Filtering with Real-Time Applications*, 4th ed.; Tsinghua University Press: Beijing, China, 2013; pp. 93–94.
67. Plett, G.L. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 1. Background. *J. Power Sources* **2004**, *134*, 252–261. [CrossRef]
68. Plett, G.L. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2. Moding and identification. *J. Power Sources* **2004**, *134*, 262–276. [CrossRef]
69. He, H.; Xiong, R.; Zhang, X.; Sun, F.; Fan, J. State-of-charge estimation of the lithium-ion battery using an adaptive extended Kalman filter based on an improved Thevenin model. *IEEE Trans. Veh. Technol.* **2011**, *60*, 1461–1469.
70. Han, J.; Kim, D.; Sunwoo, M. State-of-charge estimation of lead-acid batteries using an adaptive extended Kalman filter. *J. Power Sources* **2009**, *188*, 606–612. [CrossRef]
71. Wang, L.; Gao, K.; Han, J.; Zhao, X.; Liu, L.; Pan, C.; Li, G.; Wang, Y. Battery pack SOC estimation by Noise Matrix Self Adjustment-Extended Kalman Filter algorithm based on cloud data. *J. Energy Storage* **2024**, *84*, 110706. [CrossRef]
72. Tian, Y.; Xia, B.; Sun, W.; Xu, Z.; Zheng, W. A modified model based state of charge estimation of power lithium-ion batteries using unscented Kalman filter. *J. Power Sources* **2014**, *270*, 619–626. [CrossRef]
73. He, H.; Qin, H.; Sun, X.; Shui, Y. Comparison study on the battery SoC estimation with EKF and UKF algorithms. *Energies* **2013**, *6*, 5088–5100. [CrossRef]
74. Meng, J.; Luo, G.; Gao, F. Lithium polymer battery state-of-charge estimation based on adaptive unscented Kalman filter and support vector machine. *IEEE Trans. Power Electron.* **2015**, *31*, 2226–2238. [CrossRef]
75. Wang, Y.; Tian, J.; Sun, Z.; Wang, L.; Xu, R.; Li, M.; Chen, Z. A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renew. Sustain. Energy Rev.* **2020**, *131*, 110015. [CrossRef]
76. Gordon, N.J.; Salmond, D.J.; Smith AF, M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE proceedings F (radar and signal processing). *IET Digit. Libr.* **1993**, *140*, 107–113.
77. Schwunk, S.; Armbruster, N.; Straub, S.; Kehl, J.; Vetter, M. Particle filter for state of charge and state of health estimation for lithium–iron phosphate batteries. *J. Power Sources* **2013**, *239*, 705–710. [CrossRef]
78. Li, D.Z.; Wang, W.; Ismail, F. A mutated particle filter technique for system state estimation and battery life prediction. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 2034–2043. [CrossRef]
79. Lai, X.; Huang, Y.; Han, X.; Gu, H.; Zheng, Y. A novel method for state of energy estimation of lithium-ion batteries using particle filter and extended Kalman filter. *J. Energy Storage* **2021**, *43*, 103269. [CrossRef]
80. Wang, J.; Meng, J.; Peng, Q.; Liu, T.; Peng, J. An electrochemical-thermal coupling model for lithium-ion battery state-of-charge estimation with improve dual particle filter framework. *J. Energy Storage* **2024**, *87*, 111473. [CrossRef]
81. Anderson, J.A. *An Introduction to Neural Networks*; MIT press: Cambridge, MA, USA, 1995.
82. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]
83. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
84. Liu, J.; Saxena, A.; Goebel, K.; Saha, B.; Wang, W. An adaptive recurrent neural network for remaining useful life prediction of lithium-ion batteries. In Proceedings of the Annual Conference of the PHM Society, Portland, OR, USA, 10–14 October 2010; p. 2.
85. Yang, F.; Li, W.; Li, C.; Miao, Q. State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network. *Energy* **2019**, *175*, 66–75. [CrossRef]
86. Ren, L.; Dong, J.; Wang, X.; Meng, Z.; Zhao, L.; Deen, M.J. A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life. *IEEE Trans. Ind. Inform.* **2020**, *17*, 3478–3487. [CrossRef]
87. Song, X.; Yang, F.; Wang, D.; Tsui, K.L. Combined CNN-LSTM network for state-of-charge estimation of lithium-ion batteries. *IEEE Access* **2019**, *7*, 88894–88902. [CrossRef]
88. Yang, F.; Zhang, S.; Li, W.; Miao, Q. State-of-charge estimation of lithium-ion batteries using LSTM and UKF. *Energy* **2020**, *201*, 117664. [CrossRef]

89. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Ahmed, R.; Emadi, A. Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries. *IEEE Trans. Ind. Electron.* **2017**, *65*, 6730–6739. [CrossRef]
90. Piller, S.; Perrin, M.; Jossen, A. Methods for state-of-charge determination and their applications. *J. Power Sources* **2001**, *96*, 113–120. [CrossRef]
91. Weigert, T.; Tian, Q.; Lian, K. State-of-charge prediction of batteries and battery–supercapacitor hybrids using artificial neural networks. *J. Power Sources* **2011**, *196*, 4061–4066. [CrossRef]
92. Cui, Z.; Wang, L.; Li, Q.; Wang, K. A comprehensive review on the state of charge estimation for lithium-ion battery based on neural network. *Int. J. Energy Res.* **2022**, *46*, 5423–5440. [CrossRef]
93. Yang, D.; Wang, Y.; Pan, R.; Chen, R.; Chen, Z. A neural network based state-of-health estimation of lithium-ion battery in electric vehicles. *Energy Procedia* **2017**, *105*, 2059–2064. [CrossRef]
94. Kang, L.W.; Zhao, X.; Ma, J. A new neural network model for the state-of-charge estimation in the battery degradation process. *Appl. Energy* **2014**, *121*, 20–27. [CrossRef]
95. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Emadi, A. State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach. *J. Power Sources* **2018**, *400*, 242–255. [CrossRef]
96. Pang, H.; Wu, L.; Liu, J.; Liu, X.; Liu, K. Physics-informed neural network approach for heat generation rate estimation of lithium-ion battery under various driving conditions. *J. Energy Chem.* **2023**, *78*, 1–12. [CrossRef]
97. Joachims, T. Making Large-Scale SVM Learning Practical. Technical Report, 1998. Available online: https://hdl.handle.net/1041 9/77178 (accessed on 15 April 2024).
98. Chen, W.H.; Hsu, S.H.; Shen, H.P. Application of SVM and ANN for intrusion detection. *Comput. Oper. Res.* **2005**, *32*, 2617–2634. [CrossRef]
99. Patil, M.A.; Tagade, P.; Hariharan, K.S.; Kolake, S.M.; Song, T.; Yeo, T.; Doo, S. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. *Appl. Energy* **2015**, *159*, 285–297. [CrossRef]
100. Hu, J.N.; Hu, J.J.; Lin, H.B.; Li, X.P.; Jiang, C.L.; Qiu, X.H.; Li, W.S. State-of-charge estimation for battery management system using optimized support vector machine for regression. *J. Power Sources* **2014**, *269*, 682–693. [CrossRef]
101. Xiong, W.; Xu, G.; Li, Y.; Zhang, F.; Ye, P.; Li, B. Early prediction of lithium-ion battery cycle life based on voltage-capacity discharge curves. *J. Energy Storage* **2023**, *62*, 106790. [CrossRef]
102. Williams, C.; Rasmussen, C. Gaussian processes for regression. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1995; p. 8.
103. Schulz, E.; Speekenbrink, M.; Krause, A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *J. Math. Psychol.* **2018**, *85*, 1–16. [CrossRef]
104. Menard, S. *Applied Logistic Regression Analysis*; Sage: Newcastle upon Tyne, UK, 2002.
105. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
106. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
107. Yang, F.; Wang, D.; Xu, F.; Huang, Z.; Tsui, K.L. Lifespan prediction of lithium-ion batteries based on various extracted features and gradient boosting regression tree model. *J. Power Sources* **2020**, *476*, 228654. [CrossRef]
108. She, C.; Li, Y.; Zou, C.; Wik, T.; Wang, Z.; Sun, F. Offline and online blended machine learning for lithium-ion battery health state estimation. *IEEE Trans. Transp. Electrif.* **2021**, *8*, 1604–1618. [CrossRef]
109. Chen, L.; Wang, Z.; Lü, Z.; Li, J.; Ji, B.; Wei, H.; Pan, H. A novel state-of-charge estimation method of lithium-ion batteries combining the grey model and genetic algorithms. *IEEE Trans. Power Electron.* **2017**, *33*, 8797–8807. [CrossRef]
110. Zhang, X.; Wang, Y.; Liu, C.; Chen, Z. A novel approach of battery pack state of health estimation using artificial intelligence optimization algorithm. *J. Power Sources* **2018**, *376*, 191–199. [CrossRef]
111. Zhou, S.; Liu, X.; Hua, Y.; Zhou, X.; Yang, S. Adaptive model parameter identification for lithium-ion batteries based on improved coupling hybrid adaptive particle swarm optimization-simulated annealing method. *J. Power Sources* **2021**, *482*, 228951. [CrossRef]
112. Rahman, M.A.; Anwar, S.; Izadian, A. Electrochemical model parameter identification of a lithium-ion battery using particle swarm optimization method. *J. Power Sources* **2016**, *307*, 86–97. [CrossRef]
113. Guo, R.; Shen, W. A model fusion method for online state of charge and state of power co-estimation of lithium-ion batteries in electric vehicles. *IEEE Trans. Veh. Technol.* **2022**, *71*, 11515–11525. [CrossRef]
114. Nejad, S.; Gladwin, D.T. Online battery state of power prediction using PRBS and extended Kalman filter. *IEEE Trans. Ind. Electron.* **2019**, *67*, 3747–3755. [CrossRef]
115. Ye, M.; Guo, H.; Cao, B. A model-based adaptive state of charge estimator for a lithium-ion battery using an improved adaptive particle filter. *Appl. Energy* **2017**, *190*, 740–748. [CrossRef]
116. Roman, D.; Saxena, S.; Robu, V.; Pecht, M.; Flynn, D. Machine learning pipeline for battery state-of-health estimation. *Nat. Mach. Intell.* **2021**, *3*, 447–456. [CrossRef]
117. Ng, M.F.; Zhao, J.; Yan, Q.; Conduit, G.J.; Seh, Z.W. Predicting the state of charge and health of batteries using data-driven machine learning. *Nat. Mach. Intell.* **2020**, *2*, 161–170. [CrossRef]
118. Vidal, C.; Malysz, P.; Kollmeyer, P.; Emadi, A. Machine learning applied to electrified vehicle battery state of charge and state of health estimation: State-of-the-art. *IEEE Access* **2020**, *8*, 52796–52814. [CrossRef]
119. Kong, D.; Wang, S.; Ping, P. State-of-health estimation and remaining useful life for lithium-ion battery based on deep learning with Bayesian hyperparameter optimization. *Int. J. Energy Res.* **2022**, *46*, 6081–6098. [CrossRef]

120. Antón, J.C.Á.; Nieto, P.J.G.; de Cos Juez, F.J.; Lasheras, F.S.; Vega, M.G.; Gutiérrez, M.R. Battery state-of-charge estimator using the SVM technique. *Appl. Math. Model.* **2013**, *37*, 6244–6253. [CrossRef]
121. Junping, W.; Quanshi, C.; Binggang, C. Support vector machine based battery model for electric vehicles. *Energy Convers. Manag.* **2006**, *47*, 858–864. [CrossRef]
122. Klass, V.; Behm, M.; Lindbergh, G. Capturing lithium-ion battery dynamics with support vector machine-based battery model. *J. Power Sources* **2015**, *298*, 92–101. [CrossRef]
123. Zhou, F.; Wang, R.; He, S.; Liu, X.; Liu, S.; Shao, H.; Liu, X.P.; Xiao, Z.; Liu, J. Defect-Rich Hierarchical Porous Mn-Doped CoP Hollow Microspheres Accelerate Polysulfide Conversion. *Adv. Funct. Mater.* **2023**, *33*, 2211124. [CrossRef]
124. Liu, X.; Wang, R.; Liu, S.; Pu, J.; Xie, H.; Wu, M.; Liu, D.; Li, Y.; Liu, J. Organic Eutectic Salts-Assisted Direct Lithium Regeneration for Extremely Low State of Health Ni-Rich Cathodes. *Adv. Energy Mater.* **2023**, *13*, 2302987. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

Hybrid Neural Networks for Enhanced Predictions of Remaining Useful Life in Lithium-Ion Batteries

Alireza Rastegarpanah [1,2,*,†], Mohammed Eesa Asif [1,†] and Rustam Stolkin [1,2]

[1] Department of Metallurgy & Materials Science, University of Birmingham, Birmingham B15 2TT, UK; mxa1586@student.bham.ac.uk (M.E.A.); r.stolkin@bham.ac.uk (R.S.)
[2] The Faraday Institution, Quad One, Harwell Science and Innovation Campus, Didcot OX11 0RA, UK
* Correspondence: a.rastegarpanah@bham.ac.uk
† These authors contributed equally to this work.

Abstract: With the proliferation of electric vehicles (EVs) and the consequential increase in EV battery circulation, the need for accurate assessments of battery health and remaining useful life (RUL) is paramount, driven by environmentally friendly and sustainable goals. This study addresses this pressing concern by employing data-driven methods, specifically harnessing deep learning techniques to enhance RUL estimation for lithium-ion batteries (LIB). Leveraging the Toyota Research Institute Dataset, consisting of 124 lithium-ion batteries cycled to failure and encompassing key metrics such as capacity, temperature, resistance, and discharge time, our analysis substantially improves RUL prediction accuracy. Notably, the convolutional long short-term memory deep neural network (CLDNN) model and the transformer LSTM (temporal transformer) model have emerged as standout remaining useful life (RUL) predictors. The CLDNN model, in particular, achieved a remarkable mean absolute error (MAE) of 84.012 and a mean absolute percentage error (MAPE) of 25.676. Similarly, the temporal transformer model exhibited a notable performance, with an MAE of 85.134 and a MAPE of 28.7932. These impressive results were achieved by applying Bayesian hyperparameter optimization, further enhancing the accuracy of predictive methods. These models were bench-marked against existing approaches, demonstrating superior results with an improvement in MAPE ranging from 4.01% to 7.12%.

Keywords: deep learning; lithium-ion batteries; battery management systems; EV battery recycling; battery degradation

1. Introduction

The prediction of remaining useful life (RUL) for lithium-ion batteries is a critical task for ensuring the safe and optimal operation of battery packs, especially in applications like electric vehicles (EVs) [1]. This literature delves into two primary approaches for RUL prediction: physics-based models (PBMs) and data-driven models (DDMs). PBMs leverage the fundamental principles of electrochemistry and battery physics to simulate battery behaviour over time [2]. These models consider factors such as ion diffusion, electrode reactions [3], discharge capacity [4,5], cycles, capacity fade [6], and thermal effects [7]. While PBMs provide valuable insights into degradation mechanisms, they are computationally intensive, require detailed knowledge of electrochemical processes [2,8], and may struggle to capture real-world complexities. DDMs, on the other hand, use machine learning algorithms to learn patterns and relationships directly from available data [9,10]. They have gained prominence due to their ability to capture complex and nonlinear relationships that exist in the data, making them more adaptable and flexible than PBMs.

DDMs in machine learning are broadly categorized into two distinct groups: statistical machine learning and deep learning, each offering unique strengths and applications in various domains. Statistical machine learning, often referred to as shallow learning, typically

encompasses models that are less complex and computationally intensive. This category includes a range of methods such as support vector machines (SVMs) [10], Gaussian process regression (GPR) [11–13], random forest [14,15], and Bayesian approaches [16–18]. These techniques are particularly effective for smaller datasets where prior knowledge about the generative process of the data is available. However, they often encounter challenges when addressing complex scenarios, such as accurately capturing intricate battery characteristics and long-term dependencies within datasets [2,19].

In contrast, deep learning methods, which utilize neural networks with multiple hidden layers, are adept at handling large, complex datasets, often where there is limited a priori knowledge about the underlying processes or the most suitable features. This category includes advanced techniques like recurrent neural networks (RNNs) [20,21], convolutional neural networks (CNNs) [22,23], and various hybrid models [24,25]. Deep learning approaches are renowned for their ability to discern and learn from intricate patterns present in raw data, making them particularly effective in the analysis of multivariate time-series information. This capability makes them especially valuable in fields where understanding and predicting complex behaviours over time is crucial.

Recent advancements in sequence-to-sequence deep learning in the domain of natural language processing further contribute to the discourse on multi-horizon time-series forecasting (MTSF). Sutskever et al. introduced a powerful end-to-end approach employing multilayered long short-term memory (LSTM) networks for sequence learning, showcasing impressive results in translation tasks [26]. Similarly, Yang et al. explored incorporating a cross-entity attention mechanism in MTSF in [27]. This method minimizes assumptions on sequence structures, particularly in tasks where large labelled training sets are available. The use of time-series forecasting has also been adopted in RUL prediction, as shown by [28], which introduces a Capsule Neural Network–LSTM, a hybrid model for precise RUL prediction in mechanical systems, enhancing sensitivity to spatial features in time-series sensor data.

However, in the context of RUL prediction, improvements in robustness, generalizability, and addressing challenges like variable sampling rates and incomplete data remain areas of focus [29]. Hence, despite the progress in RUL prediction, current research has several knowledge gaps and limitations. These include (1) the need for more robust and accurate models, for enhanced generalizability, (2) the comprehensive exploration of various deep learning architectures, (3) the utilisation of complete datasets with varying sampling rates, and the consideration of factors like battery ageing and non-stationary signals.

Rationale for Advanced Model Architectures over Regression

The progression in methodologies for predicting the remaining useful life (RUL) of lithium-ion batteries underscores a pivotal shift toward enhancing the operational efficiency and reliability of battery-powered systems. Building upon the groundwork laid by Severson et al. (2019) [30], who utilized regression-based methods for early-cycle data prediction, our study uses the same dataset for advanced machine learning exploration. Diverging significantly, we employ sophisticated hybrid deep learning architectures, such as convolutional long short-term memory deep neural networks (CLDNN) and temporal transformers, which are adept at intricately modelling the spatial and temporal dynamics of battery degradation. This strategic departure from linear regression to deep learning is instrumental in capturing the multifaceted and nonlinear aspects of battery ageing, thereby elevating the precision and dependability of our predictions within a dynamic 100-cycle window. The rationale behind exploring temporal deep learning approaches based on neural networks over vanilla regression lies in their superior ability to process and learn from sequential data, providing a nuanced understanding of degradation patterns that enhance predictive accuracy and reliability, and broadening the scope for practical applications in battery life-cycle management.

Further, we refine the dataset through a rigorous feature engineering and selection process, including techniques such as linear interpolation and principal component analysis

(PCA), which enhance model training and provide a broader application range. Coupled with precise data preprocessing and sampling techniques, such as stratified random sampling and outlier removal, our methodology assures robust and dependable predictions across varied battery use cases. The incorporation of Bayesian optimization for hyperparameter tuning is a testament to our dedication to achieving optimal model performance, tailored specifically to the dataset's nuances (an overview can be seen in Figure 1). It encompasses crucial parameters such as discharge capacity, temperature, internal resistance, and discharge time.

To effectively address the heterogeneity within our dataset, we adopted a hybrid approach known as CLDNN proposed by Sainath et al. [31], which stands for convolutional, long short-term memory, and dense neural networks. CLDNN harnesses the collective power of these neural network architectures, providing a solution to the multifaceted nature of LIB RUL prediction. In addition, we repurposed the hybrid model called the temporal transformer (TT) proposed by Chadha et al. [32] to enhance prediction accuracy and robustness in LIB RUL. The TT model combines the strengths of transformer self-attention layers and LSTM architectures, presenting a unique approach to sequential modelling that effectively addresses the challenge of capturing long-term dependencies [33]. While the temporal transformer shares a name with the temporal fusion transformer (TFT) introduced by Lim et al. [34], it is important to note their architectural differences. The TFT is designed for handling multi-modal data, allowing it to incorporate various relevant features for forecasting tasks. In contrast, our dataset did not require such multi-modal capabilities, leading to divergent architectural choices in our models.

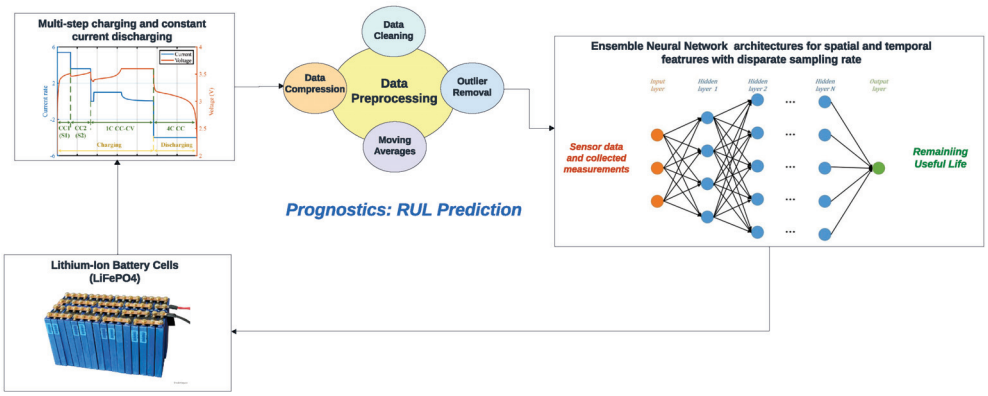


Figure 1. An overview of the proposed advanced data refinement and model optimization.

The remainder of this article is structured as follows. In Section 2, we delve into the related works within the field of lithium-ion battery RUL prediction, focusing on deep-learning approaches. Section 3 provides a comprehensive overview of our proposed methodology, followed by a detailed account of the experimental procedure. Section 4 is dedicated to discussions regarding the outcomes of our experiments. Lastly, Section 5 closes this article with our conclusion and future work.

2. Related Works

Physics-based models (PBMs) provide valuable insights into the underlying mechanisms of lithium-ion batteries but often face challenges when predicting their remaining useful life (RUL) [2]. These models perform optimally within narrow domains, requiring a detailed understanding of battery electrochemical processes, which can be time-consuming and computationally intensive [3,8,35]. The simplified assumptions often made in these models limit their ability to generalize, affecting RUL prediction accuracy. Furthermore, parametrisation in PBMs is challenging, due to manufacturing variations and the difficulty of obtaining precise model parameters [36,37]. These models also exhibit limited adaptability to dynamic environments and high computational complexity, which hampers their real-time applicability. Calibration and validation procedures for these models demand

extensive data and resources [38]. Consequently, data-driven methods like machine learning have gained prominence for their ability to learn complex nonlinear relationships from data, overcoming uncertainties, and providing a promising alternative for RUL prediction in LIB [39–42].

2.1. Statistical Machine Learning

As previously noted, statistical machine learning methods, including support vector machines (SVMs) [10], support vector regression with Kalman filters [43], Gaussian process regression (GPR) [11–13], random forest [14,15], and Bayesian approaches [16–18], have been employed for predicting battery RUL. Tong et al. [44] introduced a deep-learning-based algorithm called adaptive dropout long short-term memory (ADLSTM) and Monte Carlo (MC) simulations, achieving high precision with only 25% of degradation data. Similarly, Kwon et al. [45] integrated a multi-linear regression approach with a recurrent neural network to model LIB degradation.

2.2. Temporal Models

Temporal models, like recurrent neural networks (RNNs) and more advanced variants such as LSTM cells and gated recurrent units (GRUs), are employed to capture the sequential nature of battery data. Lipu et al. [20] introduced a nonlinear auto-regressive with exogenous input (NARX)-based neural network (NARXNN) for state of charge (SOC) estimation, utilizing a lighting search algorithm (LSA) to optimise hyperparameters. Signal decomposition techniques like discrete wavelet transform (DWT), empirical mode decomposition (EMD), and variational mode decomposition (VMD) are combined with the NARX model in [46] to predict capacity degradation trajectories. In [47], Zheng et al. employed a deep LSTM network followed by multiple fully connected (FC) layers, whereas Wang et al. in [48] introduced a bidirectional LSTM architecture with additional FC layers. Chemali et al. [21] proposed an RNN architecture with LSTM for RUL estimation, achieving low mean absolute error (MAE) values and demonstrating generalization across different conditions.

2.3. Convolutional Models

Shen et al. [23] presented a deep convolutional neural network (DCNN) that achieved higher accuracy than traditional methods, such as filter-based models [49–52]. However, limitations include the fixed-size input matrix and a limited understanding of the method's interpretability. Following a similar approach, Li et al. [53] proposed another DCNN, employing a time window approach for sample preparation and feature extraction. In another work, [54], Babu et al. proposed a novel deep CNN-based regression method for RUL estimation. Similarly, Shen et al. in [14] proposed a DCNN with transfer learning (DCNN–TL) model, which was then integrated with ensemble learning to form DCNN–ETL. The effectiveness of the DCNN–ETL model was then benchmarked against five other data-driven methods including random forest regression, Gaussian process regression, and DCNN. In another study, a temporal convolutional network (TCN) model was used along with causal and dilated convolutions to capture local capacity degradation, but its limitations include its applicability to other battery types and real-time implementation challenges [24].

2.4. Hybrid Models

Ren et al. proposed an auto-CNN–LSTM model for RUL estimation, incorporating an auto-encoder to enhance feature vectors through dimensionality reduction and feature learning, achieving accurate predictions [25]. Identified knowledge gaps include the need for enhanced generalizability, further exploration of various temporal networks, and improved utilization of datasets with varying feature sampling rates. Additionally, it is suggested that future research should address the ageing effects, the impact of non-stationary signals, the incorporation of relevant features, and the capture of spatial information.

Li et al. introduced a directed acyclic graph network that combined the LSTM with the CNN, offering a novel approach to RUL prediction [55]. Tan et al. focused on a multi-variate time-series approach using a lightweight CNN with attention mechanisms to enhance prediction accuracy [56]. Hybrid models, such as the one combining a GRU and a CNN for state of health (SOH) estimations, presented by Fan et al., and the attention-assisted temporal convolutional memory-augmented network (ATCMN) for RUL prediction from limited data, proposed by Fei et al., demonstrate promise but require further exploration and rigorous testing on more diverse datasets [57,58]. In previous work, the effectiveness of CNN–LSTM and CNN–LSTM–Differential-Neural-Computer models was demonstrated, showcasing not only superior predictive accuracy but also efficiency in learning temporal dependencies with fewer epochs [59,60].

In the following section, we aim to address critical knowledge gaps identified in the literature regarding the prediction of the RUL of LIBs. These gaps encompass issues of robustness, accuracy, and generalizability in DDMs for battery management systems (BMSs). To overcome these limitations, we developed end-to-end hybrid models capable of utilizing complete datasets [61], including features with varying sampling rates. By doing so, we intend to enhance the performance metrics for RUL estimation in LIBs and move closer to enabling real-time implementation of these models within BMS for practical applications. Additionally, we explored various deep learning approaches, such as convolutional neural networks (CNNs) [62], long short-term memory cells (LSTMs) [33], the Transformer network [63], autoencoder [64], neural Turing machines [65], differentiable neural computers [66], and hybrid models, while considering the impact of ageing, non-stationary signals, relevant feature incorporation, spatial information capture, and variable ambient temperature conditions.

3. Methodology

This section presents the framework (Figure 2) of the proposed LIB RUL prediction method. An important part of this research was the optimization of model hyperparameters using Bayesian optimization techniques aimed at maximizing the models' predictive accuracy. The primary objective was to develop a robust predictive model for estimating the RULs of LIBs, accounting for intricate temporal battery degradation dynamics and feature variations in sensor measurements across different battery batches.

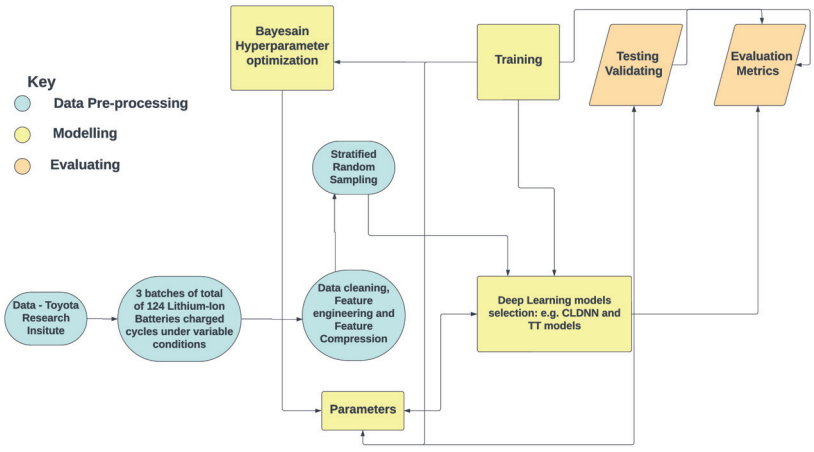


Figure 2. The study pipeline begins with a dataset comprising 124 commercial LIBs that have undergone extensive cycling until failure. Next, we implement feature selection to extract crucial features that enhance the accuracy of RUL predictions. Following this, a series of data preprocessing techniques are employed to clean and compress the dataset. Subsequently, stratified random sampling is utilized to create a representative sample. Finally, we develop a variety of hybrid deep network architectures, apply hyperparameter optimization, and evaluate these models using a testing set to determine the best-performing model.

3.1. Dataset Description

The dataset employed in this study consists of 124 commercial lithium iron phosphate (LFP)/graphite lithium-ion batteries, each with a nominal capacity of 1.1 Ah and a nominal voltage of 3.3 V. These batteries were subjected to cycling tests until failure under conditions aimed at fast charging within a temperature-controlled environment set at 30 °C. Our dataset is derived from extensive data collection efforts, capturing critical parameters such as voltage, current, temperature, and internal resistance, which were continuously measured throughout the cycling process. The term 'continuously measured' refers to the collection of data at frequent, regular intervals throughout each battery's charging and discharging cycles. This high-resolution data capture allows for an in-depth analysis of battery performance over time, enabling the precise modelling of degradation patterns. The charging protocol involved an initial phase utilizing a current of C1 until reaching a state-of-charge (SOC) of S1, followed by a transition to a current of C2 until achieving a SOC of S2, consistently maintained at 80% for all cells. Subsequently, a constant current–constant voltage (CC–CV) charging technique was applied for the transition from 80% to 100% SOC at a rate of 1 C, culminating at a cut-off voltage of 3.6 V [67]. The batteries' lifetimes, defined by the cycle count at which the capacity declined to 80% of its initial value, varied from 150 to 2300 cycles, showcasing a diverse range of degradation behaviors.

3.2. Stratified Random Sampling

Stratified random sampling was employed to create a representative dataset for the model's training and testing. This approach ensured that cells from different battery batches, characterised by varying quality control protocols, were proportionally included in the dataset. By preventing models from overfitting to specific batch attributes, this method improved model robustness and generalisation across different LIB batches.

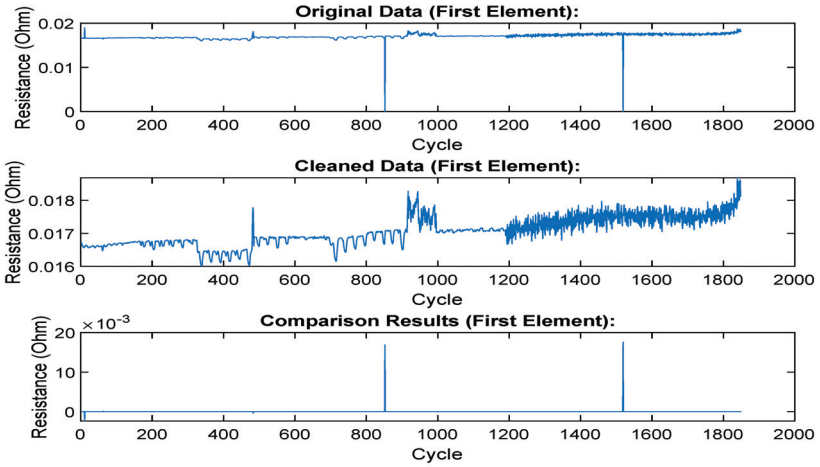
3.3. Feature Selection

The dataset consists of three groups of approximately 48 lithium-ion battery cells each, to identify features critical for predicting the remaining useful life (RUL) of the batteries. Our analysis focuses on key parameters, which include linearly interpolated discharge capacity (Qdlin) and linearly interpolated temperature (Tdlin). These features are standardized through linear interpolation to maintain uniform analysis conditions across all cells, ensuring accurate predictions by aligning data on a consistent timeline. The interpolation process applied maintains a consistent sampling rate for all cells, a critical factor for the accurate prediction of RUL and ensuring the reliability of time-series forecasting models. The selection also encompasses internal resistance (IR) and discharge characteristics, which are vital for evaluating battery health and degradation. This approach ensures the reliability of our forecasting models by providing a detailed view of battery performance over time. These features provide descriptive, summary, and cycle data, offering a holistic view of the batteries' operational conditions, performance metrics, and cycle-by-cycle behaviour. This supports a nuanced analysis, leveraging Qdlin and Tdlin alongside other selected parameters to assess battery health and predict its lifespan effectively. The output metric, remaining cycles, is calculated to estimate each battery's potential duration until its capacity falls to 80% of its original value, serving as a key indicator of its useful life.

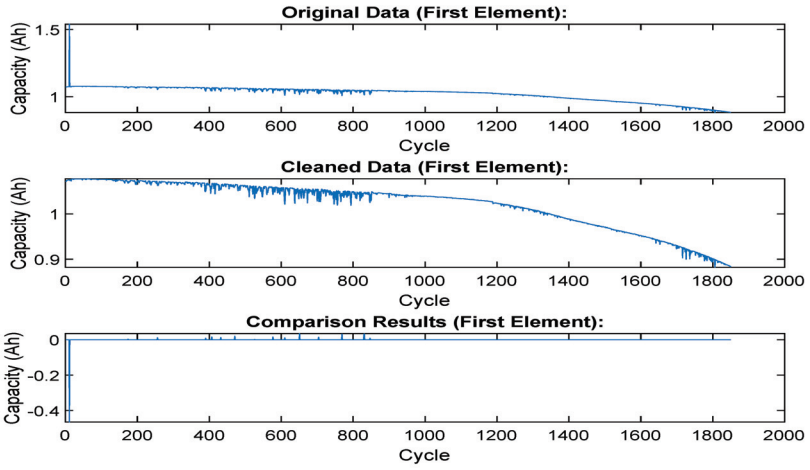
3.4. Data Preprocessing

To prepare the data for deep neural networks, we implemented a comprehensive preprocessing pipeline. Outliers in internal resistance, discharge time, and discharge quantity were removed using the fill outliers function with the cubic spline method and a moving average window of 100 (Figure 3a, Figure 3b and Figure 3c, respectively). These figures show the features before and after preprocessing (note: these plots depict the first cell in the first batch). Smoothing techniques, such as moving average filters with a window size of 15, were applied to discharge data. To capture temporal dynamics in Qdlin and

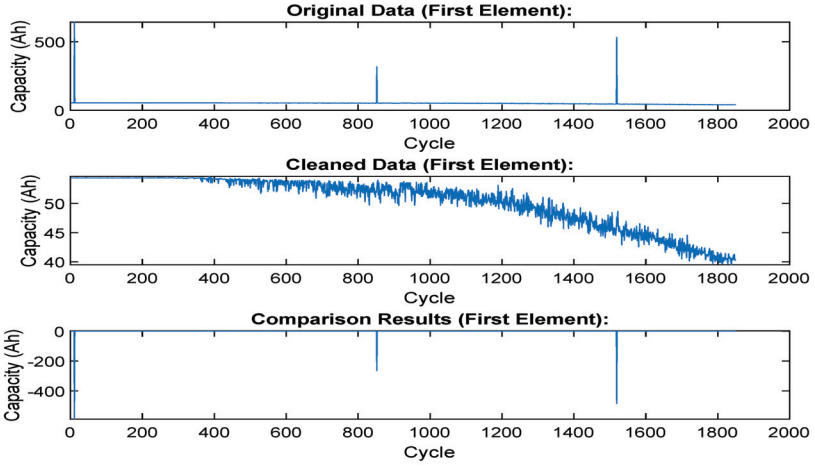
Tdlin, the sampling rate was standardized to 1000 entries per cycle. Finally, principal component analysis (PCA) was employed to reduce the dimensionality of Qdlin and Tdlin while preserving their salient features.



(**a**)



(**b**)



(**c**)

Figure 3. Comparison of data features before and after preprocessing. (**a**) Internal resistance, (**b**) discharge time, (**c**) discharge Quantity.

3.5. RUL Estimation

Traditional supervised learning approaches rely on labelled training data, where each data point is associated with a known target value. However, in the context of prognostics, this assumption often does not hold. The remaining useful life (RUL) of individual components is typically not known a priori, making it challenging to train predictive models using standard regression techniques. To overcome this hurdle, researchers have explored alternative approaches, such as employing physics-based models or utilizing machine learning algorithms to estimate the RUL of training data. While assigning a constant RUL value to all training points may seem like a straightforward solution, this can lead to inaccurate representations of the actual degradation patterns and hinder the model's ability to generalize effectively. A more sophisticated approach involves estimating the RUL based on a suitable model, as demonstrated by the use of a deep convolution neural network [23]. This approach offers a more realistic representation of the degradation process and can potentially enhance the model's predictive performance.

3.6. Metrics

A variety of metrics were used in this study. The success of the model was calculated based on the deviation (e_i) of the predicted number (model prediction $\hat{y}_i : RUL_{predicted}$) of cycles from the actual number (ground truth $y_i : RUL_{truth}$) of cycles remaining after every 100 cycle count, as shown in Equation (1). The choice of loss function significantly impacts the outcome of RUL prediction. Along with the mean absolute error (*MAE*), shown in Equation (3), and the mean absolute percentage error (*MAPE*), shown in Equation (4), which simply average the absolute errors and calculate the percentage, respectively, we also tracked the root mean squared error (*RMSE*), shown in Equation (2), which squares the errors before averaging, placing greater emphasis on larger deviations. This sensitivity to larger errors makes *RMSE* a more suitable metric for prognostics, where accurately predicting RUL is crucial, and substantial errors can lead to poor performance. Additionally, the rationale behind tracking these specific metrics is that they allow for meaningful comparisons against state-of-the-art approaches. Success in our study is quantitatively measured by the accuracy of our RUL prediction models. The *MAE* provides a direct measure of prediction accuracy in the same units as the dataset (number of cycles), *MAPE* offers a percentage-based error that is independent of the scale, and *RMSE* gives a sense of the error distribution and penalizes larger errors more severely. Our success criterion was to minimize these error metrics, thereby improving the predictive accuracy of our models.

$$e_i = \hat{y}_i - y_i \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |e_i| \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|y_i - \hat{y}_i|}{|y_i|}\right) \times 100 \tag{4}$$

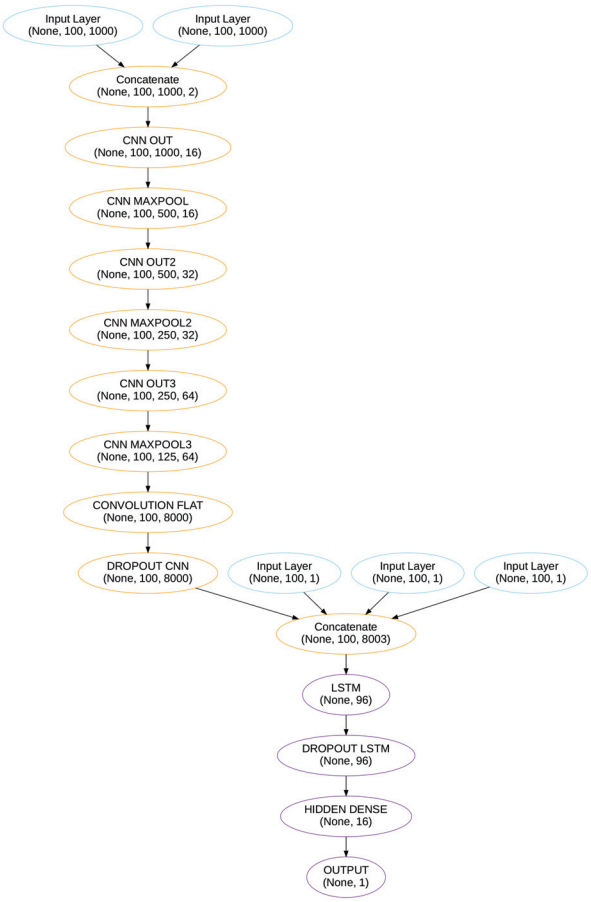
3.7. Proposed Architectures

It is worth noting that our experimentation phase encompassed the exploration of multiple temporal hybrid models; however, we only detail the most effective models. The most successful architectures in our study were the convolutional long short-term memory deep neural network (CLDNN), originally proposed by Sainath et al. [31] for natural language processing and specifically used for large vocabulary continuous speech recognition (LVCSR) tasks [31]. In their work, CLDNN outperformed the Gaussian mixture

model (GMM) and hidden Markov model (HMM) systems [68]. We repurposed the CLDNN architecture for the specific task of RUL estimation.

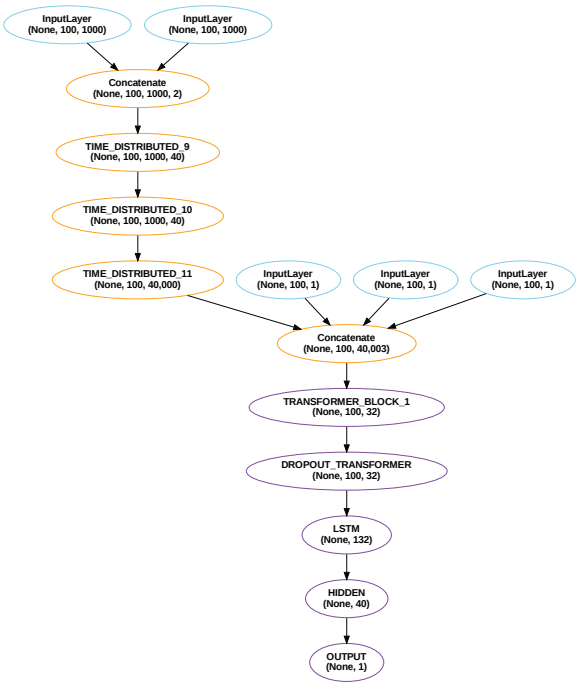
Another noteworthy model in our investigation was the temporal transformer (TT), initially introduced by Chadha et al. [32] for RUL estimation using the commercial modular aero-propulsion system simulation (C-MAPSS) dataset. The TT model demonstrated effectiveness in predicting the remaining useful life of aircraft engines. Ma et al. [69] presented a similar use case where their model utilized multi-head attention to capture global features from various representation sub-spaces. Although originally designed for predicting the remaining useful life of aerospace engines, we adapted these models for estimating the remaining useful life of lithium-ion batteries (LIB). This adaptation involved specific architectural deviations and adjustments to hyperparameters.

Our approach follows Occam's razor principle, which emphasizes simplicity and efficiency in model selection without sacrificing the predictive accuracy needed for effective RUL estimation [70]. This method allows for scalable solutions in battery degradation prediction by using the advantages of neural network architectures while avoiding unnecessary complexity. Regarding dataset size concerns, our hybrid deep learning models aim to balance complexity with dataset limitations. We enhanced data utilization through preprocessing strategies and incorporated dropout, regularization, and early stopping in our training processes to prevent overfitting, ensuring consistent performance (please refer to Figure 4a,b).



(**a**)

Figure 4. *Cont.*



(**b**)

Figure 4. Comparison of the CLDNN and temporal transformer architecture. (**a**) The CLDNN model. (**b**) The temporal transformer model.

3.7.1. The Convolutional Long Short-term Memory Deep Neural Network (CLDNN)

Adapting a convolutional long short-term memory deep neural network (CLDNN), initially developed for natural language processing (NLP) tasks with tokenized sentences, to predict remaining useful life (RUL) in lithium-ion batteries (LIB) necessitates several architectural and algorithmic modifications, as follows:

- Input Representation: In NLP tasks, CLDNN takes tokenized sentences as input. For RUL prediction, the input representation needs to be tailored to the characteristics of battery data. Time-series data from sensors measuring various parameters (voltage, current, temperature, etc.) were used as input. The input data were reshaped into a format suitable for time-series analysis.
- Sequence Length and Padding: LIB data have variable lengths of sequences as the cycle count for each battery differs, unlike fixed-length sentences in NLP. Padding or trimming sequences to a uniform length was not necessary. The network architecture was able to handle variable-length input sequences.
- Temporal Features: LIB data are inherently temporal, reflecting the degradation of the battery over time. The CLDNN architecture incorporated mechanisms to capture temporal dependencies effectively. Long short-term memory (LSTM) layers allowed us to model temporal patterns.
- Feature Extraction: The features relevant to RUL prediction in LIB differ from those important for NLP tasks. Modifications to the convolutional layers had to be made to extract features that are indicative of the battery's health and degradation.
- Hyperparameter Tuning: The hyperparameters, such as learning rate (l_r), filter size (n_f), kernel parameters (k_s), activation functions (a), and dropout rates (d_r) needed adjustment for the new task. Bayesian hyperparameter tuning was used to optimize the model for RUL prediction.
- Fine-tuning: The model was fine-tuned with different optimizer choices to minimize loss functions.

The final optimized CLDNN architecture excels at predicting the remaining useful life (RUL) of lithium-ion batteries. It comprises a total of 1,518,665 trainable parameters and combines convolutional neural networks (CNN), long short-term memory (LSTM), and dense neural networks (DNN). Each component of the network serves a specific function, as follows:

1. **CNN:** The CNN layers are used for feature extraction. They capture intricate spatial patterns within the input data, which, in this case, are time-series data from sensors measuring various parameters like voltage, current, and temperature.
2. **LSTM:** The LSTM layers are used to model temporal patterns in the data. They capture the temporal dependencies in the data, reflecting the degradation of the battery over time.
3. **DNN:** The dense neural networks are used for prediction. They contribute to the model's regularization and refined prediction capabilities.

The CLDNN architecture for our task is shown in Algorithm 1 and Figure 4a. This model demonstrates a strong efficiency in handling the heterogeneity of the dataset.

Algorithm 1 CLDNN

procedure DEFINESPATIALLAYER($Q_{dlin}, T_{dlin}, IR, DT, QD, hp$)
Input: Q_{dlin}, T_{dlin}: Input data tensors, IR, DT, QD: Additional inputs, hp: Hyperparameters
Concatenate: (Q_{dlin} and T_{dlin})
Apply Convolutional Layers and Dropout:
$h_1 \leftarrow \text{Conv1D}([Q_{dlin}, T_{dlin}], n_f, k_s, s, a, \text{padding} = \texttt{same})$
$h_2 \leftarrow \text{MaxPooling1D}(h_1)$
$h_3 \leftarrow \text{Conv1D}(h_2, 2n_f, k_s, s, a, \text{padding} = same)$
$h_4 \leftarrow \text{MaxPooling1D}(h_3)$
$h_5 \leftarrow \text{Conv1D}(h_4, 4n_f, k_s, s, a, \text{padding} = same)$
$h_6 \leftarrow \text{MaxPooling1D}(h_5)$
$h_7 \leftarrow \text{Flatten}(h_6)$
$h_8 \leftarrow \text{Dropout}(h_7, d_r)$
Features Concatenation:
$h_9 \leftarrow \text{Concatenate}(h_8, IR, DT, QD)$
Return h_9
end procedure
procedure DEFINELSTMLAYER(h_9, n_u, a_u)
Input: h_9, n_u, a_u
$h_{10} \leftarrow \text{LSTM}(h_9, n_u, a_u)$
Return h_{10}
end procedure
procedure DEFINEDENSELAYERS(h_{10}, n_d, a_d)
$h_{11} \leftarrow \text{Dense}(h_{10}, n_d, a_d)$
$RUL \leftarrow \text{Dense}(h_{11}, 1, relu\_cut)$
Return: RUL
end procedure

3.7.2. Temporal Transformer

Adapting the temporal transformer (TT) model, which was originally designed for the estimation of the remaining useful life (RUL) of aircraft engines, for the estimation of RUL in lithium-ion batteries (LIB), involves several architectural differences and adjustments. The following are some modifications that might be considered.

- Input Representation: Adjustments to the input representation to accommodate the characteristics of LIB data were required. The original model took input sequences related to engine parameters. LIB data consists of time-series measurements of capacity, temperature, resistance, and discharge time.

- Attention Mechanisms: Multi-head attention mechanisms, which were used in the original model by Ma et al. [69], needed adjustments for LIB data. Attention mechanisms were tailored to focus on features relevant to battery degradation patterns about the linearly interpolated feature.
- Model Size and Complexity: The overall size and complexity of the LIB dataset required an increase in the size and complexity of the TT model. This involved adding layers, adjusting attention mechanisms, and increasing the model depth, which led to 3,936,281 trainable parameters.
- Hyperparameter Tuning: Fine-tuning hyperparameters using Bayesian optimization specific to LIB data was required. This included learning rates, the number of attention heads, embedding dimensions, layer sizes, and dropout rates.

The temporal transformer (TT) model we used is a hybrid network that combines the transformer and long short-term memory (LSTM). Each component of the network serves a specific function, as follows:

1. **Transformer:** The transformer's multi-head self-attention mechanism allows the model to decipher complex temporal dependencies within the dataset. By parallel processing different parts of the input sequence, it extracts a rich contextual understanding of each data point.
2. **LSTM:** The LSTM units capture long-term relationships between features, enhancing the model's predictive capabilities.
3. **Feed-forward neural networks (FFNs):** The architecture employs FFNs for further refinement, facilitating the modelling of non-linear data relationships.

The TT algorithm is outlined in the provided Algorithm 2 and Figure 4b (note: $x \in \mathbb{R}^{B\times T\times D}$: input data with dimensions $B \times T \times D$, where B is the batch size, T is the time dimension, and D is the feature dimension).

The SelfAttention function in Algorithm 2 computes weighted representations of input data by considering inter-dependencies across multiple dimensions, employing a scaled dot-product attention mechanism. The TransformerBlock further refines these representations through layer normalization and feed-forward networks, enhancing their expressiveness while retaining sequential relationships. This modular and hierarchical structure allows the LSTM–transformer to capture patterns in sequential data, making it versatile for offering a robust solution for accurate RUL predictions in lithium-ion batteries.

3.8. Hyperparameter Optimization

The hyperparameter tuning for the proposed LIB RUL prediction models was achieved using Bayesian Optimization. Hyperparameter tuning is a critical step in the development of machine learning models, involving the search for optimal configurations to enhance predictive accuracy [71]. In this study, Bayesian optimization was chosen due to its effectiveness in handling non-linear and complex search spaces. Unlike traditional grid search or cross-validation methods [72], Bayesian optimization uses probabilistic models to predict the performance of different hyperparameter configurations, guiding the search toward promising regions [73]. This is particularly beneficial in high-dimensional spaces, where an exhaustive search becomes computationally expensive.

The specific implementation of this tuning process utilized the Keras Tuner library [74], a choice motivated by its compatibility and ease of integration with deep learning models (which can be seen in Algorithm 3). The process began with the definition of a hypermodel class, referred to as "`MyHyperModel`", derived from Keras Tuner's "`HyperModel`" class. This class encapsulated not only the architecture of the LIB RUL prediction models, including CLDNN or TT but also the hyperparameters' search space. Furthermore, a specialized Bayesian optimization tuner class, named "`MyBayesianOptimizationTuner`", was created by extending the "`BayesianOptimization`" class from Keras Tuner. This custom tuner was configured to minimize the validation mean absolute error (MAE), thereby optimizing the models' performance.

Algorithm 2 Temporal Transformer

procedure DEFINETIMEDISTRIBUTEDLAYER(Q_{dlin}, T_{dlin}, hp)
Input: $Q_{dlin}, T_{dlin} \in \mathbb{R}^{B\times T\times D}$ and Hyperparameters hp ▷ $Qdlin, Tdlin$: Input data tensors, hp: Hyperparameters
Concatenate: (Q_{dlin} and T_{dlin}) ▷ Merge input features
Apply TimeDistributed Dense Layer and Dropout: ▷ Feature transformation and regularization
$h \leftarrow$ TimeDistributedDense($[Q_{dlin}, T_{dlin}], hp$)
Flatten h ▷ Prepare for Transformer input
Return $[Q_{dlin}, T_{dlin}]$ ▷ Linearly interpolated features compressed
end procedure
procedure DEFINESTRUCTURING(hp, IR, DT, QD)
Extract Hyperparameters: embed_dim, num_heads, ff_dim, lstm_units, dense_units ▷ Model configuration parameters
Define Inputs: IR, DT, QD ▷ Additional model inputs
Features Concatenation: $[Q_{dlin}, T_{dlin}, IR, DT, QD]$ ▷ Combine all input features
Apply TimeDistributed Dense Layer and Dropout: ▷ Further feature transformation
    $h \leftarrow$ TimeDistributedDense($[Q_{dlin}, T_{dlin}, IR, DT, QD], hp$)
Flatten h ▷ Prepare for Transformer block
Return $[Q_{dlin}, T_{dlin}, IR, DT, QD]$ ▷ We now proceed to apply the attention mechanism across each data point
end procedure
procedure DEFINEMULTIHEADSELFATTENTION(x, D_h) ▷ Uses three sets of weight matrices W_q, W_k, W_v to transform the input data into query (Q), key (K), and value (V)
Input: $x \in \mathbb{R}^{B\times T\times D}$, D_h ▷ Obtained from DefineStructuring()
Parameters: $W_q, W_k, W_v \in \mathbb{R}^{D\times D_h}, W_o \in \mathbb{R}^{D_h\times D}$ ▷ D_h is number of attention heads, W is weight matrices
$Q, K, V \leftarrow xW_q, xW_k, xW_v$
$H \leftarrow$ Attention(Q, K, V) ▷ Computing attention scores Q and K representations, obtained by linear transformations using W_q and W_k.
$h \leftarrow HW_o$
Return h
end procedure
procedure DEFINETRANSFORMERENCODERBLOCK(x, D_h, F)
Input: $x \in \mathbb{R}^{B\times T\times D}$, D_h, F
Parameters: $W_q, W_k, W_v \in \mathbb{R}^{D\times D_h}, W_o \in \mathbb{R}^{D_h\times D}, W_1, W_2 \in \mathbb{R}^{D\times F}, b_1, b_2 \in \mathbb{R}^F$ ▷ F is the feedforward dimension
$h \leftarrow$ MultiHeadSelfAttention(x, D_h)
$h \leftarrow$ LayerNorm$(h + x)$ ▷ Apply layer normalization to $h + x$
$u \leftarrow W_1 h + b_1$ ▷ Perform a linear transformation on h and add bias b_1
$u \leftarrow$ ReLU(u) ▷ Apply the ReLU activation function for non-linearity
$v \leftarrow W_2 u + b_2$
$z \leftarrow$ LayerNorm$(v + h)$ ▷ Perform another linear transformation
    Return $z \in \mathbb{R}^{B\times T\times D}$
end procedure
Dropout Layer
LSTM Layers: ▷ LSTM for sequential data processing
$h_1 \leftarrow$ LSTM$(z, lstm_units)$
Dense Layers: ▷ Dense layer for feature extraction
$h_2 \leftarrow$ Dense$(h_1, dense_units, activation)$ ▷ Output layer for remaining useful life prediction
$RUL \leftarrow$ Dense$(h_2, 1, relu_cut)$
Return: $RUL \in \mathbb{R}^{B\times 1}$ ▷ Final model output

Algorithm 3 Hyperparameter Optimization with Keras Tuner

procedure DEFINEHYPERMODELCLASS
Class MyHyperModel(HyperModel):
Inputs: ▷ Define hyperparameters and model architecture
Define build: ▷ Build and compile models with varying hyperparameters (e.g., CLDNN, TRANSFORMER–LSTM)
Return model
end procedure
procedure DEFINECUSTOMBAYESIANOPTIMIZATIONTUNER
Class MyBayesianOptimizationTuner(BayesianOptimization):
Define initialization: ▷ Define custom Bayesian Optimization tuner
Define on_error: ▷ Handle errors during the optimization for increased modularity and robustness
end procedure
procedure HYPERMODELINSTANCEANDTUNER
Instantiate Hypermodel and Tuner:
hypermodel ← MyHyperModel(...)
tuner ← MyBayesianOptimizationTuner(hypermodel, objective=val_mae) ▷ Optimize for the lowest validation MAE
tuner.search(max_trials=100, epochs=10, dataset_train, batch_size=512)
end procedure
procedure BESTHYPERPARAMETERSANDMODEL ▷ Extract optimized hyperparameters and associated weights
best_hp ← tuner.oracle.get_best_trials(1)
best_model.set_weights(tuner.get_best_trial().get_weights())
end procedure

The hyperparameter search was conducted over a substantial span of 100 trials, encompassing 10 epochs for each trial on the training dataset. This comprehensive approach ensured a thorough exploration of the hyperparameter space, thereby identifying configurations that yield the most accurate predictions. Finally, the performance of the optimized models was rigorously evaluated on the validation dataset, ensuring the reliability and effectiveness of the hyperparameter tuning process in enhancing the predictive capabilities of the LIB RUL prediction models.

After completing the search, the optimal hyperparameter set was retrieved, and a new model was built using these parameters. This process ensured that the final CLDNN and TT model was fine-tuned for optimal performance, leveraging the power of Bayesian optimization to navigate the hyperparameter space efficiently.

4. Results and Discussion

This section presents and compares our findings and results to other algorithms, leveraging the validation data. During model development, it became evident that the most effective models necessitated two cardinal attributes. Firstly, they were required to be capable of managing sparse data by proficiently extracting significant features. Secondly, they were expected to possess the capability to learn both temporal and spatial relationships between the features, and the remaining cycles of each lithium-ion battery.

After these findings, a comparative analysis was undertaken among an assortment of hybrid models, which embraced these two critical characteristics. The performance of these models is visually presented in Figure 5a–c where key performance indicators such as loss, mean absolute error (MAE), mean absolute percentage error (MAPE) and mean squared error (MSE) are prominently depicted. It is noteworthy that each figure includes a comparison of all the hybrid neural network architectures developed, encompassing the baseline CLDNN, the optimized CLDNN, the transformer–LSTM, and its optimized counterpart.

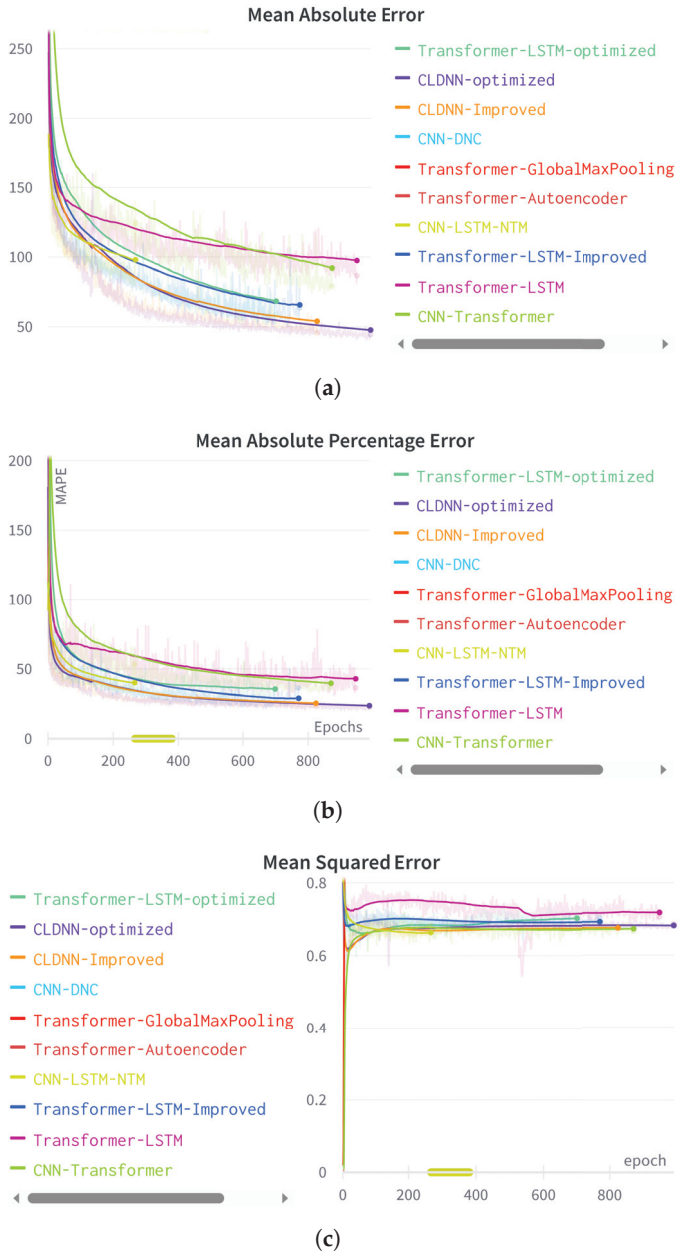


Figure 5. Comparison of error metrics between the top-performing models and other tested models, as detailed in Table 1. (**a**) MAE—mean absolute error for the best performing models. (**b**) MAPE—mean absolute percentage error for the best performing models. (**c**) MSE—mean squared error for the best performing models.

4.1. Comparing All the Tested Temporal Models

When placed in contrast with other models that embody the desired characteristics, such as convolution neural network–differential neural computer (CNN–DCN), CNN–LSTM–neural Turing machine (CNN–LSTM–NTM), CNN–transformers, and transformer–autoencoder, it becomes evident that these models exhibit higher error metrics, thereby underscoring their relative inefficacy in accurately predicting the RUL for lithium-ion batteries, as can be seen in Table 1.

Table 1. Results of six different ensemble NN models.

Model	MSE	MAE	MAPE	RMSE %
CLDNN	0.6754	84.012	25.676	0.8218
CNN–DCN	1.402	95.6365	46.408	1.1841
CNN–LSTM–NTM	1.333	284.887	-	1.1546
Transformer–LSTM	0.7136	85.134	28.7932	0.8444
CNN–transformers	0.6783	92.127	36.981	0.8236
Transformer–autoencoder	1.524	288.951	-	1.2345

4.2. Best Performing Models

In the course of assessing these results, two models, namely the convolutional, LSTM, densely connected (CLDNN) and the transformer–LSTM (temporal transformer), emerged as the most proficient in predicting the RUL, as outlined in Table 1 and Figure 5. The CLDNN model exhibited an MAE of 84.012, a MAPE of 25.676, and a MSE of 0.6754. Conversely, the temporal transformer model recorded a MAE of 85.134, a MAPE of 28.7932, and a MSE of 0.7136. To view the model's training duration, please refer to Figure 6.

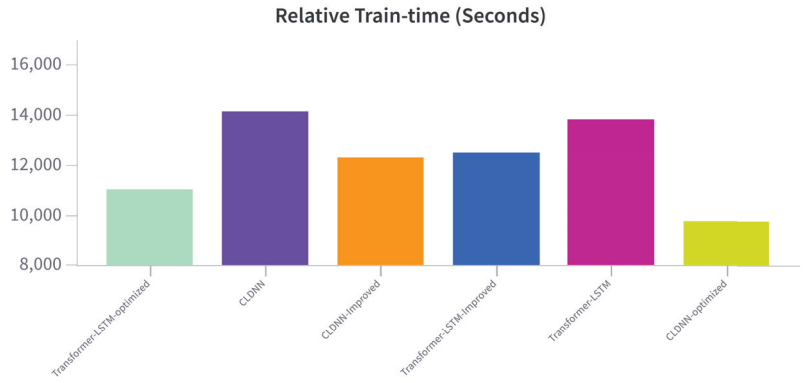


Figure 6. Training times of the baseline and optimized models in seconds.

4.3. Observations

Several key observations were drawn from the hyperparameter setups (Tables 2 and 3) of the best-performing hybrid models.

- The optimized temporal transformer had fewer embedding dimensions and a lower number of attention heads compared to the original model. This meant that the original model was over-parametrised.
- Both transformer–LSTM and CLDNN models have 'optimized' versions with distinguishable hyperparameter configurations. For instance, the dense layer in the optimized models contains an increased number of units, with the optimized transformer–LSTM model having 64 units, compared to its original 40. Additionally, the optimized configurations have a reduced dropout rate and learning rate.
- With a learning rate of `0.001`, the original transformer–LSTM model is ten times more robust than its optimized counterpart, which has a learning rate of `0.0001`, preventing gradient explosion and overshooting the minimum in the optimized model.

Table 2. Hyperparameter comparison for CLDNN.

Parameter	Original	Optimized
Convolution filter	56	44
Convolution kernel	27	12
Dense layer Activation	tanh	tanh
Dense layer units	40	64
LSTM layer activation	tanh	tanh
LSTM layer units	132	108
Dropout rate–CNN	0.45	0.3
Dropout rate–LSTM	0.4	0.3
Output activation	relu	relu_cut
Learning rate	0.001	0.0001

Table 3. Hyperparameter comparison for transformer–LSTM.

Hyperparameter	Original	Optimized
Embedding Dimension (embed_dim)	64	32
Number of Attention Heads (num_heads)	8	2
Hidden layer size (ff_dim)	32	32
Dropout after attention	0.3	0.2
Dense layer activation	tanh	tanh
Dense layer units	40	64
LSTM layer activation	relu	tanh
LSTM layer units	132	108
Dropout rate–LSTM	0.4	0.3
Output activation	relu_cut	relu
Learning rate	0.001	0.0001

4.4. Comparing Convolutional–Long Short-Term Memory–Deep Neural Network (CLDNN) and Temporal Transformer (TT) Models to Existing Approaches

In comparing our results to other approaches in the literature, several standout models, namely autoencoder–DNN, auto–CNN–LSTM, and ATCMN, exhibit relevance and importance to our research. These models, as presented in Table 4, share a hybrid deep learning architecture similar to ours, combining different network architectures to enhance prediction. Notably, they distinguish themselves by predicting the remaining useful life (RUL) of batteries in terms of remaining cycle counts, aligning with our research objectives and providing a more actionable metric for battery health management compared to other models that predominantly focus on binary classification tasks (predicting whether the battery has remaining cycles or has surpassed the end of life threshold). While these models bear significance due to their hybrid architecture and cycle count prediction approach, our CLDNN and transformer–LSTM models demonstrate superior performance in RUL prediction for lithium-ion batteries (LIBs), especially when compared to other temporal models like LSTM–RNN, Deep-CNN, and Temporal-CNN [75]. Table 4 also highlights the average inference time during the validation of each of the approaches. The average inference time of the CLDNN and TT is also reasonable, making them efficient for real-time applications.

The ATCMN model, as described by Fei et al. [58], utilizes discharging time, voltage, and capacity for RUL estimation. In contrast, our models demonstrate superior performance in predicting remaining cycles for LIBs within a 100-cycle moving window, employing a CC–CV (constant current–constant voltage) charge policy. The CC–CV charging approach involves initially charging the battery at a constant current until a set voltage threshold is reached, followed by maintaining this voltage while the current gradually decreases as the battery becomes fully charged. This method is widely used due to its effectiveness in extending the battery's lifespan and optimizing its performance. Typical parameters of the CC–CV charge policy include the constant charge current, the voltage

threshold at which the transition to constant voltage occurs, and the termination current at which the charge cycle is considered complete. These parameters are not dynamic but are predetermined based on the battery's characteristics and the desired charging performance. The Auto–CNN–LSTM model by Ren et al. [25] operates on a distinct dataset with less variation and limited diversity in input parameters, which the authors acknowledge. This disparity in datasets, compared to our more extensive and diverse dataset, may lead to variations in outcomes between the two models. Our comprehensive dataset provides a robust foundation for modelling and prediction, ultimately enhancing the reliability and applicability of our findings, especially when considering real-time variables such as voltage, capacity, battery body temperature, and internal resistance within the 100-cycle moving window.

Table 4. Caparison of CLDNN and TT against other deep learning algorithms.

Model	MAE	RMSE	MAPE	Avg. Inference Time
Auto–CNN–LSTM [25]	-	5.03	-	-
ATCMN [58]	84	-	32.8%	18 ms
Deep CNN [23]	-	1.986	-	12.3 ms
Deep CNN transfer-learning [14]	-	1.361	-	133.9 s
CLDNN (ours)	84.012	0.8218	25.676 %	15.5 ms
Transformer–LSTM (ours)	85.134	0.8444	28.7923%	16.7 ms

5. Conclusions and Future Work

This study addresses the critical challenge of accurately estimating the RUL of LIBs within the context of electric vehicles. By leveraging deep learning techniques and utilizing a rich dataset from the Toyota Research Institute, we have developed and evaluated two hybrid models: CLDNN and TT. Our contributions encompass the creation of a pre-processed high-quality dataset through stratified random sampling, by implementation of a comprehensive data preprocessing pipeline, and the development of two hybrid models. This pipeline ensures feature consistency and captures temporal dynamics, thereby laying the foundation for precise RUL predictions.

Both the CLDNN and TT models exhibited commendable performance, surpassing existing approaches with mean absolute errors (MAEs) of 84.012 and 85.134, respectively. Furthermore, they demonstrated improvements in mean absolute percentage error (MAPE), which ranged from 4.01% to 7.12%. These models prove to be well-suited for LIB RUL prediction, making substantial contributions to battery recycling and sustainability within the electric vehicle industry. Additionally, these models demonstrated superior inference time, highlighting their efficiency and applicability in real-world scenarios where rapid decision-making is crucial. This enhancement in inference speed, combined with their predictive accuracy, positions the CLDNN and TT models as highly effective tools for advancing the reliability and efficiency of battery systems in the electric vehicle sector.

While the current achievements in battery health estimation models are commendable, there are several avenues for improvement. Expanding the validation of these models to a wider array of real-world datasets is essential to strengthening confidence in their practical applicability. Investigating advanced techniques such as complex data augmentation, and alternative ensemble methods, and exploring the potential of liquid neural networks (LNNs) could lead to improvements in model performance, offering more adaptable and robust solutions for battery health estimation. Additionally, the application of graph neural

networks (GNNs) in explaining the lithium-ion battery's (LIB) remaining useful life (RUL) has been suggested as a means to reduce the number of parameters while potentially outperforming traditional physics-based models. Incorporating diverse battery chemistries into future research is also critical. The current dataset is limited to a specific type of battery (lithium ferrous phosphate/Gr, 1.1 Ah, 3.3 V), which may not fully represent the performance across different battery types. By including a variety of battery chemistries, such as LFP batteries with a capacity of 3 Ah or nickel cobalt manganese (NCM) batteries with a capacity of 3 Ah and a nominal voltage of 3.8 V, models can be trained to generalize better across different battery systems. The benefit of including diverse battery chemistries lies in the ability to create more universal and robust predictive models that can adapt to various battery behaviours and degradation patterns, ultimately leading to improved safety, reliability, and efficiency in battery usage. Additionally, a more exhaustive comparison with existing RUL prediction methods, encompassing both traditional statistical models and contemporary machine learning approaches, could be conducted.

Despite these achievements, several areas for future improvement have been identified. Real-world implementation and validation on a broader dataset are crucial for bolstering confidence in the models' applicability. Exploring complex augmentation methods, alternative ensemble solutions, and liquid neural networks (LNNs) [76–78] could further refine model performance and introduce more efficient, adaptable, and robust approaches to battery health estimation. There is also potential for the use of explaining the LIB RUL using graph neural networks, which have also been shown to significantly reduce parameter count and perform better than their traditional physics-based model counterparts [79,80]. Future research may leverage LNNs or GNNs, known for their dynamic adaptability, to potentially enhance RUL prediction for LIBs. With reduced computational intensity, these networks may offer superior generalization and efficiency for large-scale applications like electric vehicle battery management systems.

Reducing parameter counts to enhance model efficiency, measuring processing times in online scenarios, and investigating the alignment between hyperparameter optimization and a comparison between physics-based models are promising avenues for future research. In conclusion, while this study represents a significant step forward in battery health estimation, ongoing research should focus on diversifying data sources, simplifying model complexities, and exploring emerging technologies, such as LNNs, to further advance this field.

Author Contributions: Conceptualization, M.E.A.; methodology, M.E.A.; software, M.E.A.; formal analysis, M.E.A.; investigation, A.R.; resources, A.R.; data curation, M.E.A.; writing—original draft preparation, M.E.A.; writing—review and editing, A.R. and R.S.; supervision, A.R. and R.S.; project administration, A.R. and R.S.; funding acquisition, A.R. and R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Faraday Institution grant number "FIRG057" and UK Research and Innovation grant number "101104241".

Data Availability Statement: The dataset used in this study, was published by the Toyota Research Institute, is accessible at https://data.matr.io/ and was retrieved on 10 June 2023.

Acknowledgments: This work was supported in part by the project called "Research and Development of a Highly Automated and Safe Streamlined Process for Increase Lithium-ion Battery Repurposing and Recycling" (REBELION) under Grant 101104241 and in part by the UK Research and Innovation (UKRI) project "Reuse and Recycling of Lithium-Ion Batteries" (RELIB) under RELiB2 Grant FIRG005 and RELIB3 Grant FIRG057.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
AE	Autoencoder
BMS	Battery management system
CLDNN	CNN–LSTM–deep neural networks
CNN	Convolution neural network
DDA	Data-driven approaches
DL	Deep learning
EOL	End of life
EM	Electrochemical model
GPR	Gaussian process regression
IR	Internal resistance
LSTM	Long short-term memory
LIB	Lithium-ion batteries
MAE	Mean absolute error
MAPE	Mean absolute percentage error
QD	Quantity of discharge
Qdlin	Linearly interpolated discharge capacity
DDM	Data-driven model
RUL	remaining useful life
RMSE	Root mean square error
RNN	Recurrent neural network
SOC	State of charge
SOH	State of health
SVM	Support vector machine
Tdlin	Linearly interpolated temperature
TT	Temporal transformer

References

1. Plett, G.L. *Battery Management Systems. Volume I, Battery Modeling*; Artech House: Boston, MA, USA, 2015.
2. Yang, B.; Wang, J.; Cao, P.; Zhu, T.; Shu, H.; Chen, J.; Zhang, J.; Zhu, J. Classification, summarization and perspectives on state-of-charge estimation of lithium-ion batteries used in electric vehicles: A critical comprehensive survey. *J. Energy Storage* **2021**, *39*, 102572. [CrossRef]
3. Dickinson, E.J.; Wain, A.J. The Butler-Volmer equation in electrochemical theory: Origins, value, and practical application. *J. Electroanal. Chem.* **2020**, *872*, 114145. [CrossRef]
4. Tian, Y.; Chen, C.; Xia, B.; Sun, W.; Xu, Z.; Zheng, W. An Adaptive Gain Nonlinear Observer for State of Charge Estimation of Lithium-Ion Batteries in Electric Vehicles. *Energies* **2014**, *7*, 5995–6012. [CrossRef]
5. Ouyang, M.; Liu, G.; Lu, L.; Li, J.; Han, X. Enhancing the estimation accuracy in low state-of-charge area: A novel onboard battery model through surface state of charge determination. *J. Power Sources* **2014**, *270*, 221–237. [CrossRef]
6. Wang, S.; Zhang, J.; Gharbi, O.; Vivier, V.; Gao, M.; Orazem, M.E. Electrochemical impedance spectroscopy. *Nat. Rev. Methods Prim.* **2021**, *1*, 41. [CrossRef]
7. Ramadesigan, V.; Boovaragavan, V.; Pirkle, J.C.; Subramanian, V.R. Efficient reformulation of solid-phase diffusion in physics-based lithium-ion battery models. *J. Electrochem. Soc.* **2010**, *157*, A854. [CrossRef]
8. Han, X.; Ouyang, M.; Lu, L.; Li, J. Simplification of physics-based electrochemical model for lithium ion battery on electric vehicle. Part II: Pseudo-two-dimensional model simplification and state of charge estimation. *J. Power Sources* **2015**, *278*, 814–825. [CrossRef]
9. Li, Y.; Liu, K.; Foley, A.M.; Zülke, A.; Berecibar, M.; Nanini-Maury, E.; Van Mierlo, J.; Hoster, H.E. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109254. [CrossRef]
10. Anton, J.C.A.; Nieto, P.J.G.; Viejo, C.B.; Vilán, J.A.V. Support vector machines used to estimate the battery state of charge. *IEEE Trans. Power Electron.* **2013**, *28*, 5919–5926. [CrossRef]
11. Deng, Z.; Hu, X.; Lin, X.; Che, Y.; Xu, L.; Guo, W. Data-driven state of charge estimation for lithium-ion battery packs based on Gaussian process regression. *Energy* **2020**, *205*, 118000. [CrossRef]
12. Si, X.S.; Wang, W.; Hu, C.H.; Zhou, D.H. Remaining useful life estimation—A review on the statistical data driven approaches. *Eur. J. Oper. Res.* **2011**, *213*, 1–14. [CrossRef]
13. Liu, J.; Chen, Z. Remaining useful life prediction of lithium-ion batteries based on health indicator and Gaussian process regression model. *IEEE Access* **2019**, *7*, 39474–39484. [CrossRef]

14. Shen, S.; Sadoughi, M.; Li, M.; Wang, Z.; Hu, C. Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. *Appl. Energy* **2020**, *260*, 114296. [CrossRef]
15. Alfarizi, M.G.; Tajiani, B.; Vatn, J.; Yin, S. Optimized random forest model for remaining useful life prediction of experimental bearings. *IEEE Trans. Ind. Inform.* **2022**, *19*, 7771. [CrossRef]
16. Hu, X.; Jiang, J.; Cao, D.; Egardt, B. Battery Health Prognosis for Electric Vehicles Using Sample Entropy and Sparse Bayesian Predictive Modeling. *IEEE Trans. Ind. Electron.* **2016**, *63*, 2645–2656. [CrossRef]
17. Kraus, M.; Feuerriegel, S. Forecasting remaining useful life: Interpretable deep learning approach via variational Bayesian inferences. *Decis. Support Syst.* **2019**, *125*, 113100. [CrossRef]
18. Mosallam, A.; Medjaher, K.; Zerhouni, N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *J. Intell. Manuf.* **2016**, *27*, 1037–1048. [CrossRef]
19. Guan, Q.; Wei, X. The Statistical Data-driven Remaining Useful Life Prediction—A Review on the Wiener Process-based Method. In Proceedings of the 2023 Prognostics and Health Management Conference (PHM), Paris, France, 31 May–2 June 2023; pp. 64–68.
20. Lipu, M.S.H.; Hannan, M.A.; Hussain, A.; Saad, M.H.M.; Ayob, A.; Blaabjerg, F. State of Charge Estimation for Lithium-Ion Battery Using Recurrent NARX Neural Network Model Based Lighting Search Algorithm. *IEEE Access* **2018**, *6*, 28150–28161. [CrossRef]
21. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Ahmed, R.; Emadi, A. Long Short-Term Memory Networks for Accurate State-of-Charge Estimation of Li-ion Batteries. *IEEE Trans. Ind. Electron.* **2018**, *65*, 6730–6739. [CrossRef]
22. How, D.N.; Hannan, M.A.; Lipu, M.S.H.; Sahari, K.S.; Ker, P.J.; Muttaqi, K.M. State-of-charge estimation of li-ion battery in electric vehicles: A deep neural network approach. *IEEE Trans. Ind. Appl.* **2020**, *56*, 5565–5574. [CrossRef]
23. Shen, S.; Sadoughi, M.; Chen, X.; Hong, M.; Hu, C. A deep learning method for online capacity estimation of lithium-ion batteries. *J. Energy Storage* **2019**, *25*, 100817. [CrossRef]
24. Zhou, D.; Li, Z.; Zhu, J.; Zhang, H.; Hou, L. State of health monitoring and remaining useful life prediction of lithium-ion batteries based on temporal convolutional network. *IEEE Access* **2020**, *8*, 53307–53320. [CrossRef]
25. Ren, L.; Dong, J.; Wang, X.; Meng, Z.; Zhao, L.; Deen, M.J. A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life. *IEEE Trans. Ind. Inform.* **2020**, *17*, 3478–3487. [CrossRef]
26. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. *Neural Inf. Process. Syst.* **2014** , *27*. [CrossRef]
27. Yang, S.; Eisenach, C.; Madeka, D. MQRetNN: Multi-Horizon Time Series Forecasting with Retrieval Augmentation. *arXiv* **2022**, arXiv:2207.10517. [CrossRef].
28. Zhao, C.; Huang, X.; Li, Y.; Li, S. A Novel Cap-LSTM Model for Remaining Useful Life Prediction. *IEEE Sens. J.* **2021**, *21*, 23498–23509. [CrossRef]
29. Xiong, R.; Li, L.; Tian, J. Towards a smarter battery management system: A critical review on battery state of health monitoring methods. *J. Power Sources* **2018**, *405*, 18–29. [CrossRef]
30. Severson, K.A.; Attia, P.M.; Jin, N.; Perkins, N.; Jiang, B.; Yang, Z.; Chen, M.H.; Aykol, M.; Herring, P.K.; Fraggedakis, D.; et al. Data-driven prediction of battery cycle life before capacity degradation. *Nat. Energy* **2019**, *4*, 383–391. [CrossRef]
31. Sainath, T.N.; Vinyals, O.; Senior, A.; Sak, H. Convolutional, long short-term memory, fully connected deep neural networks. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, 19–24 April 2015; pp. 4580–4584.
32. Chadha, G.S.; Shah, S.R.B.; Schwung, A.; Ding, S.X. Shared temporal attention transformer for remaining useful lifetime estimation. *IEEE Access* **2022**, *10*, 74244–74258. [CrossRef]
33. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
34. Lim, B.; Arık, S.Ö.; Loeff, N.; Pfister, T. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764. [CrossRef]
35. Lyu, P.; Liu, X.; Qu, J.; Zhao, J.; Huo, Y.; Qu, Z.; Rao, Z. Recent advances of thermal safety of lithium ion battery for energy storage. *Energy Storage Mater.* **2020**, *31*, 195–220. [CrossRef]
36. Tremblay, O.; Dessaint, L.A.; Dekkiche, A. A Generic Battery Model for the Dynamic Simulation of Hybrid Electric Vehicles. In Proceedings of the 2007 IEEE Vehicle Power and Propulsion Conference, Arlington, TX, USA, 9–12 September 2007; pp. 284–289.
37. Zhao, X.; Wang, Y.; Sahinoglu, Z.; Wada, T.; Hara, S.; Callafon, R. State-of-charge estimation for batteries: A multi-model approach. In Proceedings of the American Control Conference, Portland, OR, USA, 4–6 June 2014; pp. 2779–2785. [CrossRef]
38. Song, W.; Wu, D.; Shen, W.; Boulet, B. A Remaining Useful Life Prediction Method for Lithium-ion Battery Based on Temporal Transformer Network. *Procedia Comput. Sci.* **2023**, *217*, 1830. [CrossRef]
39. Yang, Y. A machine-learning prediction method of lithium-ion battery life based on charge process for different applications. *Appl. Energy* **2021**, *292*, 116897. [CrossRef]
40. Yang, D.; Wang, Y.; Pan, R.; Chen, R.; Chen, Z. A neural network based state-of-health estimation of lithium-ion battery in electric vehicles. *Energy Procedia* **2017**, *105*, 2059–2064. [CrossRef]
41. Lotfi, N.; Li, J.; Landers, R.G.; Park, J. Li-ion Battery State of Health Estimation based on an improved Single Particle model. In Proceedings of the 2017 American Control Conference (ACC), Seattle, WA, USA, 24–26 May 2017; pp. 86–91. [CrossRef]
42. El-Dalahmeh, M.; Al-Greer, M.; El-Dalahmeh, M.; Bashir, I. Physics-based model informed smooth particle filter for remaining useful life prediction of lithium-ion battery. *Measurement* **2023**, *214*, 112838. [CrossRef]

43. Xue, Z.; Zhang, Y.; Cheng, C.; Ma, G. Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression. *Neurocomputing* **2020**, *376*, 95–102. [CrossRef]
44. Tong, Z.; Miao, J.; Tong, S.; Lu, Y. Early prediction of remaining useful life for Lithium-ion batteries based on a hybrid machine learning method. *J. Clean. Prod.* **2021**, *317*, 128265. [CrossRef]
45. Kwon, S.J.; Han, D.; Choi, J.H.; Lim, J.H.; Lee, S.E.; Kim, J. Remaining-useful-life prediction via multiple linear regression and recurrent neural network reflecting degradation information of 20 Ah $LiNi_xMnyCo_{1-x-y}O_2$ pouch cell. *J. Electroanal. Chem.* **2020**, *858*, 113729. [CrossRef]
46. Bashir, I.; Al-Greer, M.; et al. Lithium-ion Batteries Capacity Degradation Trajectory Prediction Based on Decomposition Techniques and NARX Algorithm. In Proceedings of the 2022 57th International Universities Power Engineering Conference (UPEC), Istanbul, Turkey, 30 August–2 September 2022; pp. 1–6.
47. Zheng, S.; Ristovski, K.; Farahat, A.; Gupta, C. Long Short-Term Memory Network for Remaining Useful Life estimation. In Proceedings of the 2017 IEEE International Conference on Prognostics and Health Management (ICPHM), Dallas, TX, USA, 19–21 June 2017; pp. 88–95. [CrossRef]
48. Wang, J.; Wen, G.; Yang, S.; Liu, Y. Remaining Useful Life Estimation in Prognostics Using Deep Bidirectional LSTM Neural Network. In Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing), Chongqing, China, 26–28 October 2018; pp. 1037–1042. [CrossRef]
49. Spagnol, P.; Rossi, S.; Savaresi, S.M. Kalman filter SoC estimation for Li-ion batteries. In Proceedings of the 2011 IEEE International Conference on Control Applications (CCA), Denver, CO, USA, 28–30 September 2011; pp. 587–592.
50. Hu, C.; Youn, B.D.; Chung, J. A multiscale framework with extended Kalman filter for lithium-ion battery SOC and capacity estimation. *Appl. Energy* **2012**, *92*, 694–704. [CrossRef]
51. Dai, H.; Wei, X.; Sun, Z.; Wang, J.; Gu, W. Online cell SOC estimation of Li-ion battery packs using a dual time-scale Kalman filtering for EV applications. *Appl. Energy* **2012**, *95*, 227–237. [CrossRef]
52. Lambert, B. *A Student's Guide to Bayesian Statistics*; SAGE: New York, NY, USA, 2018.
53. Li, X.; Ding, Q.; Sun, J.Q. Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliab. Eng. Syst. Saf.* **2018**, *172*, 1–11. [CrossRef]
54. Sateesh Babu, G.; Zhao, P.; Li, X.L. Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life. In *Database Systems for Advanced Applications*; Navathe, S.B., Wu, W., Shekhar, S., Du, X., Wang, X.S., Xiong, H., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 214–228.
55. Li, J.; Li, X.; He, D. A Directed Acyclic Graph Network Combined With CNN and LSTM for Remaining Useful Life Prediction. *IEEE Access* **2019**, *7*, 75464–75475. [CrossRef]
56. Tan, W.M.; Teo, T.H. Remaining Useful Life Prediction Using Temporal Convolution with Attention. *AI* **2021**, *2*, 48–70. [CrossRef]
57. Fan, Y.; Xiao, F.; Li, C.; Yang, G.; Tang, X. A novel deep learning framework for state of health estimation of lithium-ion battery. *J. Energy Storage* **2020**, *32*, 101741. [CrossRef]
58. Fei, Z.; Zhang, Z.; Yang, F.; Tsui, K.L. A deep attention-assisted and memory-augmented temporal convolutional network based model for rapid lithium-ion battery remaining useful life predictions with limited data. *J. Energy Storage* **2023**, *62*, 106903. [CrossRef]
59. Rastegarpanah, A.; Contreras, C.A.; Stolkin, R. Hyperparameter-optimized CNN and CNN-LSTM for Predicting the Remaining Useful Life of Lithium-Ion Batteries. In Proceedings of the 2023 Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 21–23 November 2023; pp. 110–115. [CrossRef]
60. Rastegarpanah, A.; Contreras, C.A.; Stolkin, R. Harnessing CNN-DNC and CNN-LSTM-DNC Architectures for Enhanced Lithium-Ion Remaining Useful Life Prediction. In Proceedings of the 2023 Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 21–23 November 2023; pp. 116–121. [CrossRef]
61. Li, P.; Zhang, Z.; Grosu, R.; Deng, Z.; Hou, J.; Rong, Y.; Wu, R. An end-to-end neural network framework for state-of-health estimation and remaining useful life prediction of electric vehicle lithium batteries. *Renew. Sustain. Energy Rev.* **2022**, *156*, 111843. [CrossRef]
62. LeCun, Y.; Bengio, Y.; et al. Convolutional networks for images, speech, and time series. In *The Handbook of Brain Theory and Neural Networks*; MIT Press: Cambridge, MA, USA, 1995; Volume 3361, p. 1995.
63. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2017; Voume 30.
64. Zhai, J.; Zhang, S.; Chen, J.; He, Q. Autoencoder and its various variants. In Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7–10 October 2018; pp. 415–419.
65. Collier, M.; Beel, J. Implementing neural turing machines. In Proceedings of the Artificial Neural Networks and Machine Learning—ICANN 2018: 27th International Conference on Artificial Neural Networks, 2018. Available online: https://www.nature.com/articles/nature20101 (accessed on 10 January 2023).
66. Graves, A.; Wayne, G.; Reynolds, M.; Harley, T.; Danihelka, I.; Grabska-Barwińska, A.; Colmenarejo, S.G.; Grefenstette, E.; Ramalho, T.; Agapiou, J.; et al. Hybrid computing using a neural network with dynamic external memory. *Nature* **2016**, *538*, 471–476. [CrossRef]
67. Xu, L.; Deng, Z.; Xie, Y.; Lin, X.; Hu, X. A Novel Hybrid Physics-Based and Data-Driven Approach for Degradation Trajectory Prediction in Li-Ion Batteries. *IEEE Trans. Transp. Electrif.* **2023**, *9*, 2628–2644. [CrossRef]

68. Hinton, G.; Deng, L.; Yu, D.; Dahl, G.E.; Mohamed, A.R.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.N.; et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Process. Mag.* **2012**, *29*, 82–97. [CrossRef]
69. Ma, Q.; Zhang, M.; Xu, Y.; Song, J.; Zhang, T. Remaining Useful Life Estimation for Turbofan Engine with Transformer-based Deep Architecture. In Proceedings of the 2021 26th International Conference on Automation and Computing (ICAC), Portsmouth, UK, 2–4 September 2021; pp. 1–6. [CrossRef]
70. Blumer, A.; Ehrenfeucht, A.; Haussler, D.; Warmuth, M.K. Occam's razor. *Inf. Process. Lett.* **1987**, *24*, 377–380. [CrossRef]
71. Probst, P.; Boulesteix, A.L.; Bischl, B. Tunability: Importance of hyperparameters of machine learning algorithms. *J. Mach. Learn. Res.* **2019**, *20*, 1934–1965.
72. Malakouti, S.M.; Menhaj, M.B.; Suratgar, A.A. The usage of 10-fold cross-validation and grid search to enhance ML methods performance in solar farm power generation prediction. *Clean. Eng. Technol.* **2023**, *15*, 100664. [CrossRef]
73. Wu, J.; Chen, X.Y.; Zhang, H.; Xiong, L.D.; Lei, H.; Deng, S.H. Hyperparameter optimization for machine learning models based on Bayesian optimization. *J. Electron. Sci. Technol.* **2019**, *17*, 26–40.
74. Gulli, A.; Pal, S. *Deep Learning with Keras*; Packt Publishing Ltd.: Birmingham, UK, 2017.
75. Wang, S.; Jin, S.; Bai, D.; Fan, Y.; Shi, H.; Fernandez, C. A critical review of improved deep learning methods for the remaining useful life prediction of lithium-ion batteries. *Energy Rep.* **2021**, *7*, 5562–5574. [CrossRef]
76. Hasani, R.; Lechner, M.; Amini, A.; Rus, D.; Grosu, R. Liquid time-constant networks. In Proceedings of the AAAI Conference on Artificial Intelligence, 2021; Volume 35, pp. 7657–7666. Available online: https://ojs.aaai.org/index.php/AAAI/article/view/16936 (accessed on 10 January 2023).
77. Hasani, R. Interpretable Recurrent Neural Networks in Continuous-Time Control Environments. Ph.D. Thesis, Technische Universität Wien, Vienna, Austria, 2020.
78. Hashemi, S.R.; Bahadoran Baghbadorani, A.; Esmaeeli, R.; Mahajan, A.; Farhad, S. Machine learning-based model for lithium-ion batteries in BMS of electric/hybrid electric aircraft. *Int. J. Energy Res.* **2021**, *45*, 5747–5765. [CrossRef]
79. Wang, Z.; Yang, F.; Xu, Q.; Wang, Y.; Yan, H.; Xie, M. Capacity estimation of lithium-ion batteries based on data aggregation and feature fusion via graph neural network. *Appl. Energy* **2023**, *336*, 120808. [CrossRef]
80. Lam, R.; Sanchez-Gonzalez, A.; Willson, M.; Wirnsberger, P.; Fortunato, M.; Alet, F.; Ravuri, S.; Ewalds, T.; Eaton-Rosen, Z.; Hu, W.; et al. Learning skillful medium-range global weather forecasting. *Science* **2023**, 382, 1416–1421. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

Enhancing Lithium-Ion Battery Manufacturing Efficiency: A Comparative Analysis Using DEA Malmquist and Epsilon-Based Measures

Chia-Nan Wang [1], Fu-Chiang Yang [1,*], Nhut T. M. Vo [1,2,*] and Van Thanh Tien Nguyen [1,3,*]

[1] Department of Industrial Engineering and Management, National Kaohsiung University of Science and Technology, Kaohsiung 80778, Taiwan; cn.wang@nkust.edu.tw
[2] Thu Dau Mot University, Binh Duong 75000, Vietnam
[3] Industrial University of Ho Chi Minh City, Nguyen Van Bao Street, Ward 4, Go Vap District, Ho Chi Minh City 70000, Vietnam
* Correspondence: fuchiang@nkust.edu.tw (F.-C.Y.); vonhut@tdmu.edu.vn (N.T.M.V.); thanhtienck@naver.com (V.T.T.N.)

Abstract: Innovative carbon reduction and sustainability solutions are needed to combat climate change. One promising approach towards cleaner air involves the utilization of lithium-ion batteries (LIB) and electric power vehicles, showcasing their potential as innovative tools for cleaner air. However, we must focus on the entire battery life cycle, starting with production. By prioritizing the efficiency and sustainability of lithium-ion battery manufacturing, we can take an essential step toward mitigating climate change and creating a healthier planet for future generations. A comprehensive case study of the leading LIB manufacturers demonstrates the usefulness of the suggested hybrid methodology. Initially, we utilized the Malmquist model to evaluate these firms' total efficiency while dissecting their development into technical and technological efficiency change components. We employed the Epsilon-Based Measure (EBM) model to determine each organization's efficiency and inefficiency scores. The findings show that the EBM approach successfully bridged the gap in the LIB industry landscape. Combined with the Malmquist model, the resulting framework offers a powerful and equitable evaluation paradigm that is easily applicable to any domain. Furthermore, it accurately identifies the top-performing organizations in specific aspects across the research period of 2018–2021. The EBM model demonstrates that most organizations have attained their top level, except for A10, which has superior technology adoption but poor management. A1, A2, A4, A6, A8, A9, and A10 were unable to meet their targets because of the COVID-19 pandemic, despite productivity improvements. A12 leads the three highest-scoring enterprises in efficiency and total productivity changes, while A3 and A5 should focus on innovative production techniques and improved management. The managerial implications provide vital direction for green energy practitioners, enhancing their operational effectiveness. Concurrently, consumers can identify the best LIB manufacturers, allowing them to invest in long-term green energy solutions confidently.

Keywords: EBM; lithium-ion batteries; DEA Malmquist; manufacturing efficiency; green energy; Epsilon-Based Measure (EBM); a DEA application procedure

1. Introduction

1.1. Overview of the Lithium-Ion Batteries Industry

Lithium-ion batteries have emerged as a dominant technology for portable electronics, electric vehicles, and renewable energy storage due to their high energy density, long life cycle, and environmentally friendly characteristics [1]. As the demand for lithium-ion batteries continues to grow, it becomes imperative to assess the efficiency of lithium-ion battery manufacturers to optimize their performance and ensure sustainable production practices. The global lithium-ion battery industry has experienced remarkable growth, with

a market value of USD 42.30 billion in 2020 [2–5]. It is projected to reach approximately USD 160.21 billion by 2026, growing at a CAGR of around 26.04% [6,7]. This growth can be attributed to several key factors driving the expansion of lithium-ion battery technology.

Firstly, increasing environmental concerns and the need to mitigate carbon emissions from conventional automobiles have spurred the adoption of electric vehicles worldwide [8]. Strict emission standards imposed by the highest authorities in many countries have accelerated the shift towards electric cars, which has fueled the demand for lithium-ion batteries as a preferred choice for powering electric vehicles [7].

Furthermore, the decreasing lithium-ion battery prices and rising investment by market players in research and development to launch batteries with an advanced capacity have contributed to the market growth. Introducing new market participants and model variations for electric vehicles has intensified competition and led to innovations to reduce production costs, further propelling the demand for lithium-ion batteries [9].

Governments of several developing countries are also promoting the adoption of electric vehicles through assistance and incentives for production, consumption, and the development of public charging infrastructure [4]. This encouragement has created a favorable environment for the expanding demand for lithium-ion batteries. Besides that, lithium-ion batteries' small size, excellent energy efficiency, and low price make them an attractive choice for various applications, including manufacturing, automobile, electronic devices, healthcare gadgets, telecommunication buildings, and other sectors [3,10,11]. The expanding applications of lithium-ion batteries in diverse industries such as military, aviation, smart grid, and passenger cars are expected to boost the market growth further [4,12,13]. The global LIB industry is segmented based on category, structure, employment, market competition, and geographic distribution. A flourishing industry propels the strong demand for lithium-ion battery technology in the thriving automotive sector, an amplified allocation of electric vehicles, and a growing presence of market players in this domain [7].

The global LIB industry is witnessing robust growth driven by increasing environmental concerns, declining prices, investments in research and development, government incentives, and expanding applications [14]. The projected growth in the automotive and traction segment and the overall market presents significant opportunities for manufacturers, investors, and other stakeholders in the lithium-ion battery industry.

1.2. Research Gap and Research Motive

Despite the increasing demand and widespread use of lithium-ion batteries in various applications, there is still a research gap in evaluating the efficiency of lithium-ion battery manufacturers. The current research mainly focuses on assessing the performance of lithium-ion batteries in terms of energy storage capacity, durability, and safety features. However, limited research addresses the efficiency of manufacturers producing these types of batteries. An efficiency evaluation is crucial for manufacturers because it provides detailed information about their operational performance and identifies areas that need improvement. Traditional efficiency evaluation methods, such as Data Envelopment Analysis (DEA) and its variations, have been widely used in various industries to measure efficiency. However, there is still a lack of research applying DEA and other advanced methods to evaluate the efficiency of lithium-ion battery manufacturers.

The primary motive of this study is to bridge the research gap by assessing the efficiency of lithium-ion battery manufacturers using a DEA approach, especially the Malmquist and the Epsilon-Based Measure (EBM) model. The DEA Malmquist model is a widely used method for evaluating the efficiency of a Decision-Making Unit (DMU) over time. The EBM model is a relatively new approach incorporating undesirable outputs in the efficiency assessment process.

By employing these advanced methods, this research aims to provide a comprehensive and accurate assessment of the efficiency of lithium-ion battery manufacturers. This assess-

ment can help identify best practices, benchmarking targets, and areas for improvement in the manufacturing processes of lithium-ion batteries.

Furthermore, the research motive extends to academic contributions by adding to the existing literature on efficiency assessment methods for lithium-ion battery manufacturers. This research can contribute to operations management, industrial engineering, and sustainable energy research by applying advanced efficiency assessment techniques to a specific context, i.e., lithium-ion battery manufacturing. The findings of this study can serve as a reference for future research and provide insights for researchers interested in efficiency assessment methods in the context of battery manufacturing and other related industries.

This study is to fill the research gap by assessing the efficiency of lithium-ion battery manufacturers using advanced methods such as the DEA Malmquist and EBM models. We will compare the results obtained from this approach to provide a comprehensive and compelling assessment of the efficiency of LIB manufacturers. This comparison will highlight the similarities and differences between the two models in evaluating the efficiency of LIB manufacturers, further enhancing the rigor and credibility of our research. The findings of this research can have practical implications for manufacturers and policymakers in the battery industry, as well as academic contributions to the literature on efficiency assessment methods in the context of battery manufacturing.

The findings are presented in the study using a systematic framework. Section 2 thoroughly examines DEA models and their specific applications in the literature. Section 3 provides an overview of the research method and dives into the theoretical features of the Malmquist and EBM models. Section 4 presents a case study concentrating on the LIB industry as an example of the suggested methodology's efficacy and relevance in solving performance assessment issues in the marine industry. The report summarizes the most relevant findings, highlights contributions, recognizes potential limits, and suggests future research prospects in Section 5.

2. Study Process and Related Works

2.1. Study Process

This paper presents an innovative and incorporated approach for assessing the efficiency of the top twelve lithium-ion battery companies from 2018 to 2021. Our proposed model combines Data Envelopment Analysis (DEA) Malmquist and Efficiency-Based Measure (EBM) techniques, offering a comprehensive and sophisticated framework for evaluating efficiency in this context [15,16]. The research process for assessing the efficiency of lithium-ion battery manufacturers using the DEA Malmquist and EBM model can be outlined in three main phases, as demonstrated in Figure 1.

Phase 1: Problem Analysis and Objective Definition

We expect to find and examine the challenges connected to analyzing the efficiency of lithium-ion producers throughout this phase of the research. We established a precise research objective with specified targets and measurable outcomes.

Phase 2: Data Collection and Analysis

This phase involves selecting the appropriate inputs and outputs for the DEA Malmquist model based on the study objectives and available data. Total assets, liabilities, and SG&A expenses are determined for entry. Based on the study objectives and open data, revenue and gross income are chosen as the outputs for the DEA Malmquist model at this step. A Pearson correlation test assesses data homogeneity and isotonicity, which ensures the study's validity. As part of this total productivity evaluation, the DEA Malmquist model measures lithium-ion producers' "technical efficiency change (catch-up)" and "technological investment (frontier shift)" [17–20]. Before proceeding to the next stage, a diversity affinity test is run to double-check the diversity and affinity coefficient indices [16].

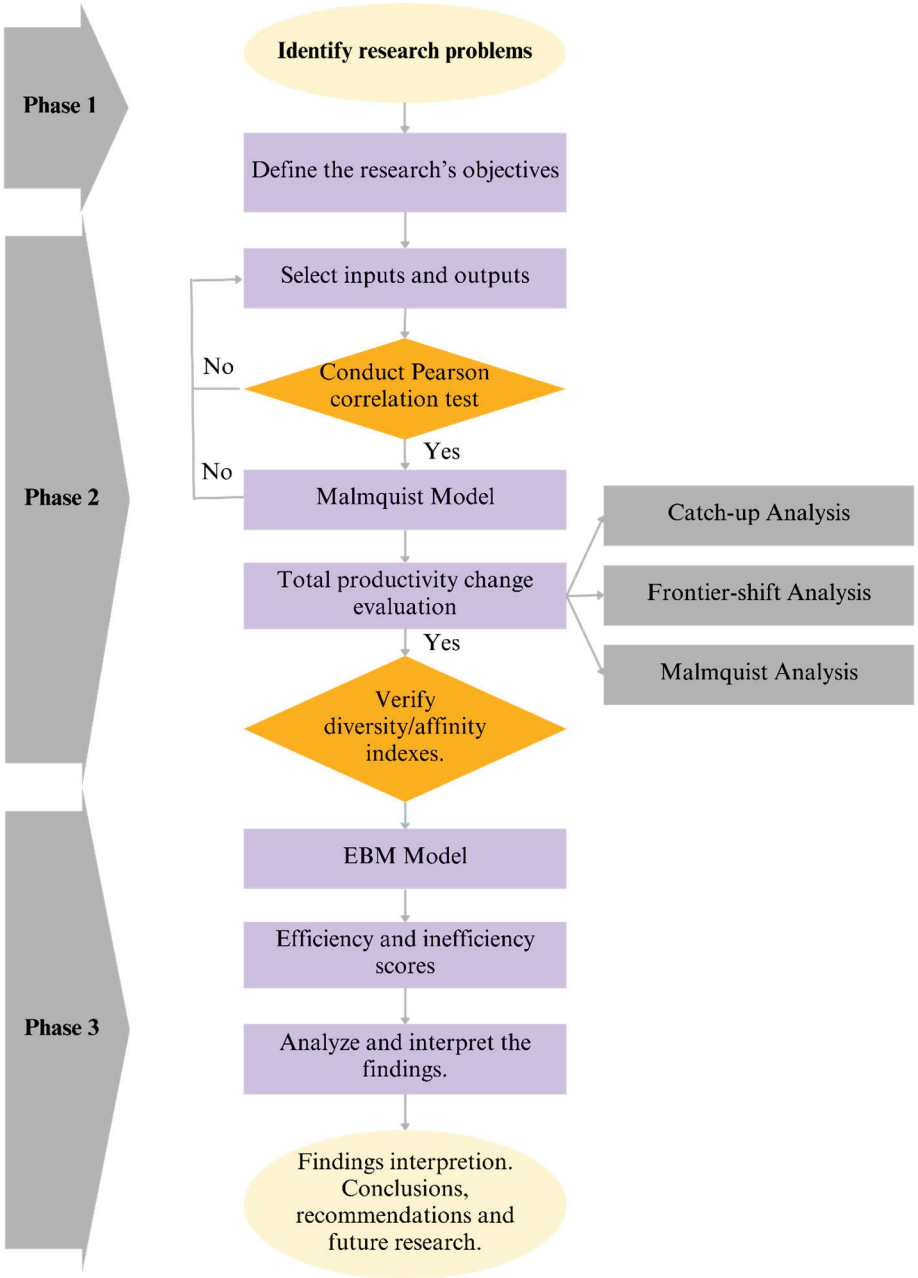


Figure 1. The research framework for a lithium-ion industry assessment.

Phase 3: EBM Model and Results Analysis

In this phase, we utilized the Epsilon-Based Measure (EBM) model to determine scores for the efficiency and inefficiency of the Decision-Making Units (DMUs), also known as lithium-ion battery producers, based on the DEA Malmquist model outputs. These scores rank DMUs based on their ability to manufacture lithium-ion batteries. The efficiency and inefficiency scores are evaluated, and the findings are interpreted considering the research objectives. The implications of the results are discussed, as well as the research's limitations and potential paths for additional exploration. The study draws several conclusions and makes recommendations to various stakeholders, including practitioners and policymakers in the lithium-ion manufacturing industry.

Combining the DEA Malmquist and EBM models, this three-phase research process offers a comprehensive approach for evaluating lithium-ion manufacturers' efficiency, considering total productivity change and efficiency scores while ensuring data integrity through a correlation analysis and diversity affinity testing.

The integrated DEA Malmquist and EBM models proposed in this manuscript offer a cutting-edge approach for assessing efficiency in the lithium-ion battery industry, accounting for technical efficiency change and technological investment. Using a Pearson correlation analysis and diversity and affinity coefficient index verification further strengthens our findings' robustness. The results of this study are expected to provide valuable insights for stakeholders, including policymakers and researchers, in the lithium-ion battery industry and contribute to the existing literature on efficiency assessments in the field of battery technology.

2.2. Related Works

Some noteworthy studies implemented the efficiency evaluation problems for businesses and manufacturers by combining the Malmquist and EBM models in diverse approaches. Mykhalovskiy et al., 2004 give thoughts, reviews, and evaluations to assist in developing more comprehensive social scientific research of EBM [21]. Li et al., 2020 utilized a modified meta-two-stage EBM Malmquist approach to investigate regional disparities in thirty-one Chinese cities' economies, energies, environments, health, and media during 2014–2016 [22]. A quantitative method was used by Rusli et al. to evaluate the logistics sector in Malaysia before and after the COVID-19 pandemic by comparing the sector's effectiveness and efficacy using the EBM and Malmquist index approaches. [23]. Data envelopment analysis is used to analyze the energy efficiency of China's coastal regions in terms of air emissions between 2000 and 2012 (Qin et al., 2017) [24]. Carbon dioxide, sulfur dioxide, and nitrogen oxide emissions are all negative consequences of energy consumption.

Under the new regulation, the process by which banks develop appropriate internal judgments for absorbing international strategic capital becomes crucial for managing banks. Constant productivity improvements will lead to long-term growth; thus, the goals of this study are as follows: (1) to figure out the connection between international strategic ventures and improvements in the output of China's banks and to validate the effectiveness of implementing overseas strategic investing; (2) to determine the best overseas ownership percentage; and (3) to illustrate the impact of overseas strategic expenditures on China's financial institution efficiency, i.e., the way it transmits between institutions [25].

Cheng et al., 2019 discover that the Malmquist trend of total factor productivity indicators corresponds with the findings from the best practice gap change (BPC) and pure technological catch-up indexes (PTCU), indicating that the BPC and PTCU indexes' innovation effects are the primary factors responsible for productivity improvement [26]. Lu et al. 2020 use a three-stage DEA model and a period neural network framework to assess and forecast total factor productivity in Chinese petroleum enterprises [27]. From 2009 to 2018, the panel data from 50 publicly traded Chinese petroleum companies were used. A three-stage Data Envelopment Analysis (DEA) model was used to exclude environmental and random effects. As a result, the radial basis function neural network prediction model was employed to forecast the total factor productivity of publicly traded petroleum businesses over the following two years.

A two-phase DEA methodology based on EBM and Malmquist is used to investigate the effectiveness of maritime transportation in European countries. The results identify the most prosperous nations across multiple economic sectors from 2016 to 2019 and demonstrate that the research gap in applying the EBM method to marine transport has been successfully filled [28].

3. Resources and Procedures

3.1. The Malmquist Model Theory

It is necessary to precisely check the positive correlation between the variables utilized for analysis to guarantee the validity and reliability of the Malmquist model. The Pearson correlation test is the answer to this problem. Statistical analysis is a standard tool in the academic world. The Pearson correlation coefficient, denoted by the symbol (H_{pq}), is a

standardized measure of the linear association between two variables. It is computed using Equation (1) and takes on values between +1 and −1, with larger values indicating a stronger linear relationship. A correlation coefficient around +1 indicates an almost ideal linear relationship between the variables, lending credence to the study. Thus, confirming the validity and accuracy of the study data is crucial by carefully applying this Pearson correlation test at the outset of employing the DEA Malmquist model.

All variables utilized in the analysis must have a negative correlation before the Malmquist model can be applied successfully. A correlation test should be used to guarantee this criterion is met. When the coefficient of correlation between two variables is high, they have a strong relationship. However, the Pearson correlation coefficient value is directly proportional to the link between the two parameters. When the value of the correlation is lower, the link is weaker. The correlation coefficient has a fixed value of −1 to +1, which is almost perfect if it falls within a range of ±1.

$$H_{pq} = \frac{\sum_{i=1}^{n}(p_i - \overline{p})(q_i - \overline{q})}{\sqrt{\sum_{i=1}^{n}(p_i - \overline{p})^2 \sum_{i=1}^{n}(q_i - \overline{q})^2}} \quad (1)$$

The fundamental goal of the MPI is to investigate alterations in the performance of the production of many DMUs during a period when assessed by the outcome of relative efficiency change (catch-up) and "technological change (frontier)" [29]. Catch-up efficiency refers to a DMU's extraordinary reaction to a difference in effectiveness. The phrase "frontier shift" relates to how DMUs withstand technological innovation from one period to another.

The two periods in the DEA analysis are known as (m_i, n_i) for the first and (m_i^2, n_i^2) for the second with a particular DMU_i. To determine the efficiency score $DMU_i\ (m_i,\ n_i)^{\ t_1}$, the frontier's effectiveness is t_2: $d^{t_2}((m_i,\ n_i)^{t_1}(t_1 = 1, 2$ and $t_2 = 1, 2)$

With a certain DMU_i, the analysis' two-time frames have been introduced as (m_i, n_i) for the initial duration and (m_i^2, n_i^2) for the following duration. Frontier effectiveness is t_2: $d^{t_2}(m_i,\ n_i)^{t_1}(t_1 = 1, 2$ and $t_2 = 1, 2)$ to evaluate the score of efficiency $DMU_i\ (m_i,\ n_i)^{\ t_1}$. The relative efficiency shift in Equations (2)–(4) will be calculated as follows using the formulas for the frontier-shift index (FSI), Malmquist productivity index (*MPI*), and catch-up index (*CA*).

$$CA = \frac{d^2\ ((m_i,\ n_i))^2}{d^1\ ((\ m_i, n_i\ ))^1} \quad (2)$$

$$FR = \left[\frac{d^1((m_i, n_i))^1}{d^2((m_i, n_i))^1} \times \frac{d^1((m_i, n_i))^2}{d^2((m_i, n_i))^2}\right]^{\frac{1}{2}} \quad (3)$$

$$MPI = \left[\frac{d^1\left((m_i, n_i)^2\right)}{d^1\left((m_i, n_i)^1\right)} \times \frac{d^2\left((m_i, n_i)^2\right)}{d^2\left((m_i, n_i)^1\right)}\right]^{\frac{1}{2}} \quad (4)$$

Detailed explanations of all the variables:

Catch-up Index (CI): Fare et al. divided the entire change in productivity into those that are related to a shift in the efficiency border that separates the times of t and the t + 1 period, with the cause being due to the efficiency of the unit "catch-up The catch-up ratio represents the change in efficiency over the cross-section of the unit operating as we go from t to 1 + t [30,31]. The phrase "boundary shift" refers to the difference in the efficiency boundary between period t and time 1 regarding the amount of input needed to keep a particular output level while operating optimally.

Frontier-shift Index (FSI): FSI indicates the new research and development (R&D) innovations and skills; for example, breakthrough creation of procedures and systems or technological change. It is essential to know how far one is from the R&D technical frontier

and how rapidly one may reach it via upgrading machinery and procedures. A formula calculates the "boundary shift" in R&D advancement.

Malmquist Productivity Index (MPI): The index tracks productivity growth over time, breaking it down into efficiency and technological advancement increases using a DEA-like nonparametric technique, separating productivity into technological advancement and efficiency catch-up calls for using current data and temporal adjustments in the research period. The MPI is a distance function from the equations' measurements at times t and t + 1. It is calculated as the result of DEA-derived catch-up (recovery) and frontier-shift (innovation) elements using the nonparametric approach.

Using the abovementioned methodologies, we can identify whether a DMU's overall efficiency factor rises or falls. There is a chance that efficiency will either increase or decrease because of catch-up or frontier efficiency. As calculated above, the DMU's total effectiveness of factors reflects comparative and technologically innovative performance improvements or decreases. CA, FSI, and MPI figures may vary by more than, less than, or equal to one, showing if a DMU is moving forward, backward, or remaining constant between the two periods.

3.2. The EBM Theory

Since it involves determining efficiency, DEA can handle several input and output parameters. The Charnes–Cooper–Rhodes (CCR) model finds an optimal proportional change in input and output quantities while ignoring the emergence of excess inputs or deficient outputs in a DMU (known as a radical approach because it only considers proportional changes in inputs and outputs) [32–34]. Although it does not focus on the proportion of output and input changes (non-radial technique), the Slack-Based Measure (SBM) must deal with slacks directly. SBM simulations (non-radial slack variable efficiency), are based on the efficiency of slack variables but do not use the radial estimate assumption, and aim to optimize output and input inefficiencies by picking locations farthest from the border [16,35–37]. They lack data on the ratio used to compute the efficiency front projection during the technique. The final findings are rarely as accurate as the estimates because there is so much possibility for improvement. To solve this issue, Tone offered three Epsilon-Based Measurement (EBM) models with "radial and non-radial components" in 2010 [16]. Among the designs were "input-oriented, output-oriented, and non-oriented" and "non-oriented" [16,25,37]. The models considered both non-radial and radial features. When input-oriented EBM (EBM I-C) for $DMU_o = (x_o, x_o)$ is used, the standard "unguided EBM calculation model" can be stated in Equation (5) as follows.

$$\begin{gathered}\delta^* = \min_{\theta,\lambda,\, s^-} \theta - \varepsilon_x \sum_{i=1}^{m} \frac{\omega_i^- s_i^-}{x_{io}} \\ \text{Subject to} \\ \sum_{j=1}^{n} x_{ij}\lambda_j = \theta x_{io} - s_{i,}^-,\ i = 1, \ldots, m \\ \sum_{j=1}^{n} y_{ij}\lambda_j \geq y_{ro,r} = 1, \ldots, s \\ \lambda_j \geq 0, j = 1, 2, \ldots, n s_{i,}^- \geq\ 0, i = 1, 2, \ldots, n\end{gathered} \tag{5}$$

$s_{i,}^-$ and ω_i^- reflect the weight and slack present in the λ_j input, respectively, ε_x is a variable that uses input scattering to reflect peripheral characteristics, and "o" indicates that the DMU has been checked.

$s_{i,}^-$ and ω_i^- describe how much slack and weight are present in the input, respectively, ε_x is a variable demonstrating the radial characteristics and affects the amount of scattering present in the inputs, and "o" indicates that the DMU is being evaluated [25,38,39].

The variable ε_x demonstrates the radial properties and impacts the amount of scattering present in the inputs, whereas $s_{i,}^-$ and ω_i^- define the slack and weight of the inputs, respectively. An evaluation of DMU is indicated by "o".

When discussing the requirements of an effective EBM model, Tone and Tsutsui mentioned the following requirements: $0 \leq P(c,d) = P(c,d) \leq 1/20 \leq \ldots 1/2$ and $0 \leq Q(c,d) = 1 - 2P(c,d)\ 0 \leq 1$. The analyzed data with low and high scattering are demonstrated in Figure 2

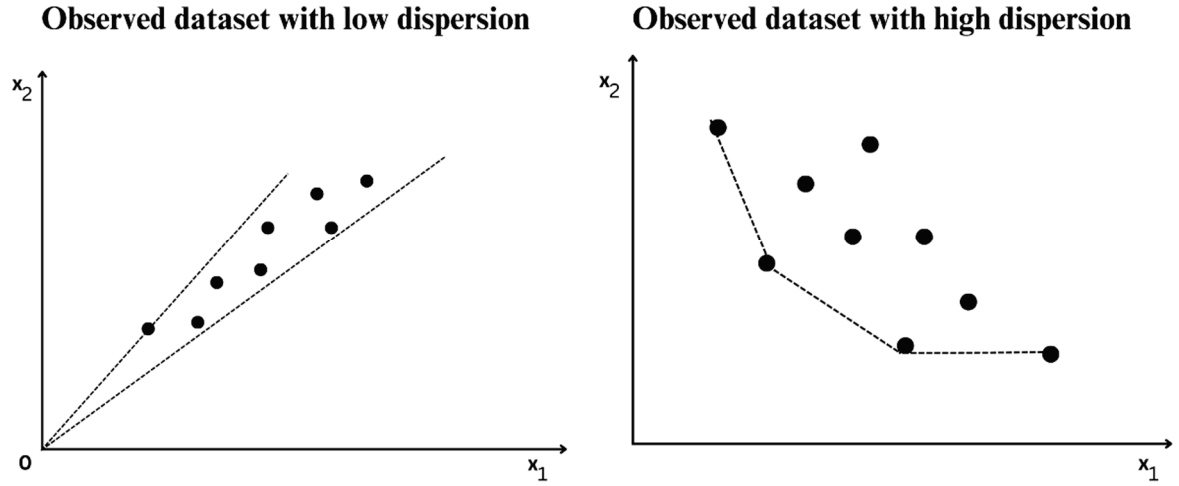


Figure 2. Low and high scattering of the obtained statistic.

4. Research's Empirical Findings

4.1. Obtaining Data

In this paper, we evaluate and rate the productivity and efficiency of the twelve LIB manufacturers throughout the four years 2018–2021. The data for the study were collected from the financial database of twelve LIB companies. Table 1 shows the list of DMUs and market cap up to 2023.

Table 1. List of DMUs and their market cap (data up to 7 April 2023).

DMUs	Company Name	Market Cap (USD)
A1	Contemporary Amperex Technology Co., Ltd.	282.79 B
A2	BYD Company Ltd.	138.64 B
A3	Panasonic Corporation	29.39 B
A4	Samsung SDI Co., Ltd.	40.14 B
A5	SK Innovation Co., Ltd.	21.07 B
A6	Toshiba Corporation	18.16 B
A7	LG Chem Ltd.	48.07 B
A8	Tesla Inc.	746.27 B
A9	Johnson Controls	49.25 B
A10	Nio Inc.	95.77 B
A11	Albemarle Corp.	24.77 B
A12	Sociedad Quimica y Minera De Chile SQM SA	9.35 B

In Data Envelopment Analysis (DEA), the Malmquist model stands as a vital tool for evaluating the performance of Decision-Making Units (DMUs) [17,40,41]. A critical aspect of this model is the selection of the inputs and outputs used to measure the DMUs' performance. Drawing upon a comprehensive review of the relevant literature spanning several decades, we have identified the most pertinent financial variables to be included in the model. According to the "thumb rule" suggested by Golany and Roll (1989) and Homburg (2001), to ensure there are enough DMUs for a meaningful and robust comparison, the number of DMUs should be at least twice the total number of inputs and outputs [32,42]. After examining various prior studies, we have arrived at a well-supported choice of three inputs: total assets, liabilities, and selling general and administrative (SG&A) expenses, and two outputs: revenue and gross income. These inputs and outputs have been carefully selected based on their significance in capturing the underlying performance of DMUs.

Table 2 demonstrates an overview of the various input and output variables employed in prior studies.

Table 2. Previous studies' input and output criteria.

Author/Reference	Inputs/Criteria	Outputs/Responses	Methodologies	Applied Areas
[43]	"Net fixed assets, Salaries and wages, Operating expenses, Current liabilities"	"Operating income"	"CCR, BCC"	Logistics
[44]	"Operating expenses, Liability, Equity, Employee"	"Net income, Net sales, Intangible value, Market value"		E-retailing
[45]	"Assets, Capital, Number of employees"	"Total revenue"	"CCR, BCC, Malmquist"	Logistics
[46]	"Business cost, Total assets, Employees number"	"Business income, Gross profit, Return on equity"	"CCR, BCC"	Real estate
[47]	"Management fees, Operating expenses, Interest expenses"	"Total assets, Net assets value, Total revenue"	"Malmquist model"	
[48]	"Assets, Equity, Employees, Expense"	"Revenue, Profit"	"CCR, Malmquist	Banking
[49]	"Assets, Capital, Operating cost, Employees number"	"Revenue, Gross profit, Return on equity"	SBM, Regression model"	Real estate
[50]	"Assets, Liabilities, Operating Expenses"	"Revenue, Gross Profit"	"DEA Malmquist, DEA Window"	Aviation
This paper	"Assets, Liabilities, SG&A Expenses"	"Revenue, Gross Income"	"DEA Malmquist, EBM"	Energy

The rigorous statistical analysis that has guided the selection of these inputs and outputs is meticulously documented in Table 3. This descriptive input and output data presentation further substantiates our thorough approach to constructing a robust DEA Malmquist model. By incorporating these carefully chosen variables, we have ensured that this model is well-founded and capable of providing valuable insights into the performance of DMUs.

Table 3. Statistical data of LIB companies' input and output (2018–2021).

Year	Statistics	Assets	Liabilities	SG&A Expense	Revenue	Gross Income
2018	Max	48,800	32,235	13,519	58,331	16,520
	Min	2759	1513	96	752	(92)
	Average	22,909	13,924	3271	18,811	4135
	SD	14,808	10,479	3675	16,913	4403
2019	Max	43,926	28,719	13,519	58,478	15,936
	Min	2094	2338	94	1138	(250)
	Average	23,387	14,074	3372	19,282	3964
	SD	13,469	8749	3655	16,288	4260
2020	Max	52,150	29,669	12,645	54,750	14,840
	Min	4097	2314	91	1666	(251,984)
	Average	26,440	15,266	3257	19,025	(16,918)
	SD	14,393	8942	3412	14,835	70,986
2021	Max	62,130	30,550	11,257	53,820	13,610
	Min	7153	3981	102	2523	960
	Average	33,242	18,922	3596	24,236	5449
	SD	16,761	9952	3097	16,464	4326

Table A1, Appendix A, displays the predicted Pearson correlation coefficient values for 2018, 2019, 2020, and 2021. The Pearson test coefficients show positive correlations, which are between 0 and 1. The research findings are incredibly delicate in selecting the input and output variables and parameters.

4.2. The Malmquist Model's Findings

4.2.1. Technical Efficiency Change

The catch-up index, as illustrated in Figure 3 and Table 4, reflects the variation in the operational effectiveness of the DMUs across periods. Figure 3 shows the development of

catch-up indexes across the entire range of LIB producers, and Table 4 gives specific values for "catch-up." The catch-up index's worth of more significant than 1, less than 1, and 1 reflects the advancement or decrease of the DMUs' technical effectiveness. Based on the table, all DMUs have achieved a progressive technological efficiency on average (average indexes for catch-up >1) throughout 2018–2021. This advanced efficiency resulted in an average catch-up score of 1.0925. In this group, the three DMUs, A10 (1.5509), A8 (1.2874), and A2 (1.1890), are the three DMUs that have achieved the most outstanding performance in terms of technical efficiency from 2018 to 2021. Meanwhile, A5 (0.8326), A3 (0.9268), and A11 (0.9278) are among the most efficient players based on the average.

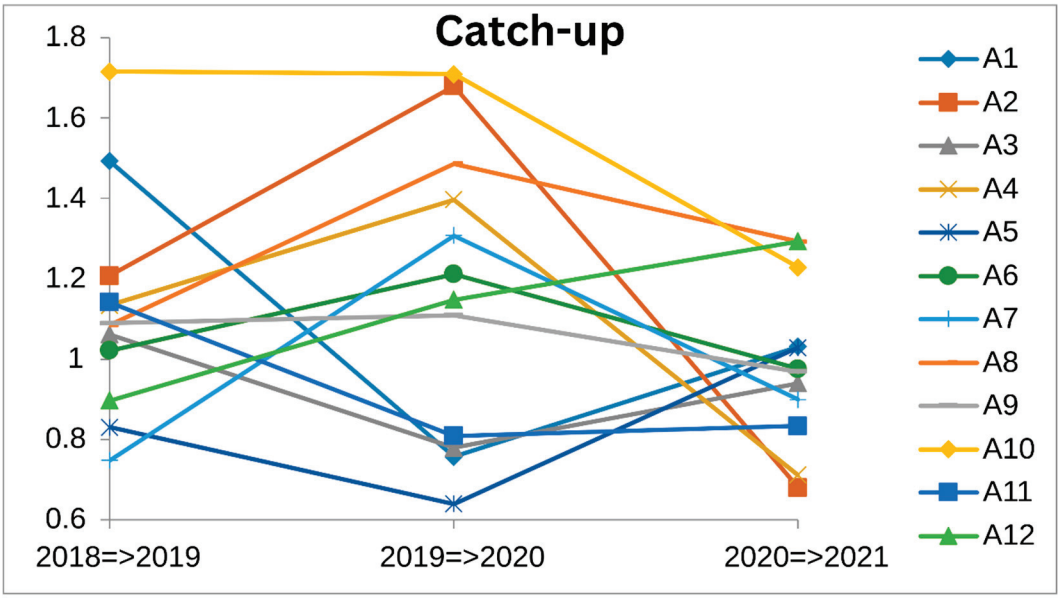


Figure 3. Changes in technical efficiency (catch-up).

Table 4. Changes in technical efficiency for the period 2018–2021.

Catch-Up	2018 to 2019	2019 to 2020	2020 to 2021	Average
A1	1.4926	0.7567	1.0311	1.0935
A2	1.2073	1.6793	0.6804	1.1890
A3	1.0615	0.7790	0.9401	0.9268
A4	1.1338	1.3968	0.7115	1.0807
A5	0.8305	0.6389	1.0283	0.8326
A6	1.0214	1.2111	0.9754	1.0693
A7	0.7476	1.3073	0.8989	0.9846
A8	1.0839	1.4861	1.2921	1.2874
A9	1.0891	1.1091	0.9687	1.0556
A10	1.7153	1.7090	1.2283	1.5509
A11	1.1408	0.8089	0.8336	0.9278
A12	0.8964	1.1477	1.2932	1.1124
Average	1.1183	1.1692	0.9901	1.0925
Max	1.7153	1.7090	1.2932	1.5509
Min	0.7476	0.6389	0.6804	0.8326
SD	0.2679	0.3646	0.2028	0.1892

Figure 3 shows that A10-Nio, Inc. has the highest performance, 1.7153 from 2018 to 2019 and 1.7090 for 2019–2020, before a decrease to 1.2283 between 2020 and 2021. There was a lot of fluctuation in the performance throughout the research, just as with A2. Notably, A2, A6, A7, and A8 reached their peak technical performance during 2019–2020 and fell in 2020–2021.

In contrast, A2 experienced a slump in the efficiency of its technology during 2020–2021 and a catch-up index of 0.6804 and was the most inefficient manufacturer then. Among twelve manufacturers, A7 was the first to have the poorest catch-up index (0.7476). In 2018–2019, they surpassed seven companies and climbed fifth in the rankings (1.3073). However, they had the fourth lowest catch-up rate in 2020–2021 (0.8989). The last time we looked at it, A7 had the highest percentage of manufacturers showing an increase in the efficiency of their technical changes. At the same time, A12, A3, and A11 significantly improved the performance of technology changes. Only A12 exhibits a steady growth trend in technical efficiency in the three intervals. It had dramatically improved its catch-up score, going from a mere 0.8963 from 2018 to 2019 to 1.1147 during 2019–2020, before it reached its peak efficiency at 1.2932 from 2020 to 2021.

4.2.2. Efficiency Frontiers

Frontier-shift indexes are an essential tool for assessing an organization's technical advancement over time, and the conclusions reported in this study are persuasive. Table 5 displays the frontier-shift values for each DMU, indicating that just one-sixth of the organizations evaluated obtained progressive average frontier-shift scores. The data suggest that most DMUs (nine out of twelve) stagnate and become poor technological performers.

Table 5. Technological change for the period 2018–2021.

Frontier	2018 => 2019	2019 => 2020	2020 => 2021	Average
A1	0.6962	1.0012	1.1840	0.9605
A2	0.7878	0.7662	1.3373	0.9638
A3	0.9876	1.0566	0.9477	0.9973
A4	0.8374	0.8130	1.4029	1.0177
A5	0.8715	0.7407	1.1971	0.9364
A6	1.0033	0.9216	0.9185	0.9478
A7	0.9183	0.7827	1.2583	0.9864
A8	0.9285	0.8318	1.1574	0.9726
A9	0.9443	0.9499	1.1775	1.0239
A10	0.8336	0.6646	1.1515	0.8833
A11	0.6845	0.9957	1.2728	0.9843
A12	0.7292	0.9426	1.1699	0.9472
Average	0.8519	0.8722	1.1812	0.9684
Max	1.0033	1.0566	1.4029	1.0239
Min	0.6845	0.6646	0.9185	0.8833
SD	0.1097	0.1221	0.1392	0.0383

Figure 4 depicts the evolution routes of technological efficiencies for all DMUs, demonstrating a variety of shifting patterns. Interestingly, while some organizations endure a decline in technological progress during one time, they can achieve their peak during the next, emphasizing the need for ongoing innovation and adaptation. It is worth mentioning that A4 stands out as a remarkable performer, with the highest score of 1.4029 achieved over the research period. This number starkly contrasts with A3, which peaked at 1.0566 in 2019–2020 before plummeting to the second-worst score of 0.0977 in 2020–2021.

Only one-sixth of the DMUs in Table 5 attain the increasing average frontier-shift values, including A4 (1.0177) and A9 (1.0239). Meanwhile, A1 (0.9605), A2 (0.9638), A3 (0.09973), A5 (0.9364), A6 (0.9478), A7 (0.9864), A8 (0.9726), A10 (0.8822), A11 (0.9843), and A12 (0.9472) lag and become poor technological achievers. Most producers (9 among 12 units) had nonprogressive average frontier-shift indexes (FSI), with an average frontier-shift score of 0.9684 for the observed time. Figure 4 shows more fluctuating patterns of the DMUs, demonstrating the stable progression of their technological performances compared to Figure 3 (catch-up). Among these, A4 has the best score of 1.4029 in 2020–2021.

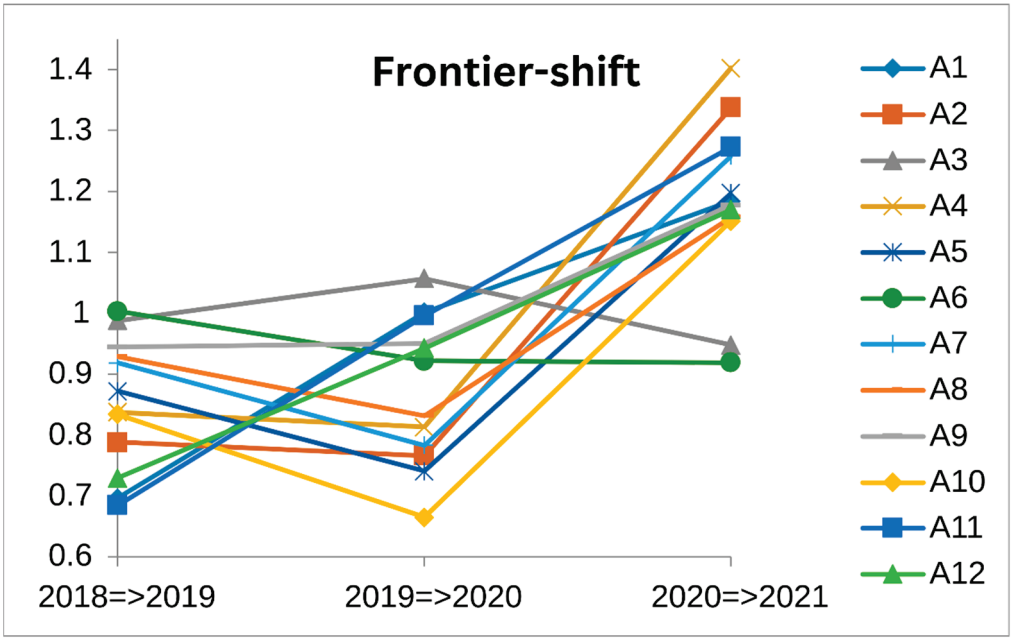


Figure 4. Technological change (Frontier-shift).

A2, A4, A5, A7, A8, and A10 indicate a decline in 2019–2020 and a peak in technical progress in 2020–2021. A3 had a maximum score of 1.0566 in 2019–2020, which was more significant than other corporations, but subsequently dropped to the second-worst score (0.0977) in 2020–2021.

In a continually changing corporate scene, the findings reported in this study highlight the significance of continuous technological innovation and adaptability. The frontier-shift indices offer a holistic view of technological advancement by considering various elements such as competitiveness, regulatory and political settings, and inventions. As a result, firms should take these findings to heart and seek to enhance their technological performance to remain competitive and thrive in the long run.

The "frontier-shift" chart depicts the changes in the technical efficiency and productivity of firms or industries over time relative to the production frontier.

4.2.3. Results of Malmquist Productivity Indexes

To obtain the result for the Malmquist Productivity Indexes (MPIs) of the twelve LIB producers, we must solve Equation (4). The granular MPI values are presented in Table 6, and Figure 5 illustrates how the MPIs have changed over time for each company. We note that an efficiency index of MPI = 1 represents the status of staying unchanged productivity, that an efficiency index of MPI > 1 reflects an efficiency increase, and that an efficiency index of MPI <1 correlates to a decrease in total productivity.

As seen in Table 6, all manufacturers operated effectively on a typical basis, except for A3, A7, and A11, which only achieved 0.8825. This result stands out since A3, A7, and A11 all display declining performance in terms of technical and technological efficiency.

In addition, the average MPI of all producers is greater than 1 (1.0320), which indicates a development in the overall productivity growth of the companies during the research period and shows an improvement in the manufacturers' efficiency. The three businesses that have had the most significant increase in productivity are A10 (1.3267), A8 (1.2460), and A12 (1.0828). These DMUs achieved exceptionally high technical efficiency levels, allowing them to compensate for a modest loss of ground in technological advancement. Figure 5 shows that, in certain instances, the DMUs almost follow the same trends as those shown in Figure 3 (the catch-up). A1, A3, A5, and A7 have the least stable performance for 2018–2021. These organizations' trends of catch-up efficiency are illustrated in (catch-up),

and these organizations' patterns are identical. These firms' catch-up efficiency trends are shown in Figure 3 (catch-up), and their patterns are similar.

Table 6. Total productivity change from 2018 to 2021.

Malmquist	2018 to 2019	2019 to 2020	2020 to 2021	Average
A1	1.0392	0.7576	1.2208	1.0059
A2	0.9512	1.2866	0.9099	1.0492
A3	1.0483	0.8231	0.8909	0.9208
A4	0.9494	1.1355	0.9981	1.0277
A5	0.7238	0.4733	1.2310	0.8093
A6	1.0247	1.1162	0.8958	1.0122
A7	0.6866	1.0233	1.1311	0.9470
A8	1.0065	1.2361	1.4954	1.2460
A9	1.0284	1.0536	1.1406	1.0742
A10	1.4300	1.1358	1.4144	1.3267
A11	0.7809	0.8054	1.0611	0.8825
A12	0.6536	1.0818	1.5130	1.0828
Average	0.9435	0.9940	1.1585	1.0320
Max	1.4300	1.2866	1.5130	1.3267
Min	0.6536	0.4733	0.8909	0.8093
SD	0.2131	0.2342	0.2241	0.1446

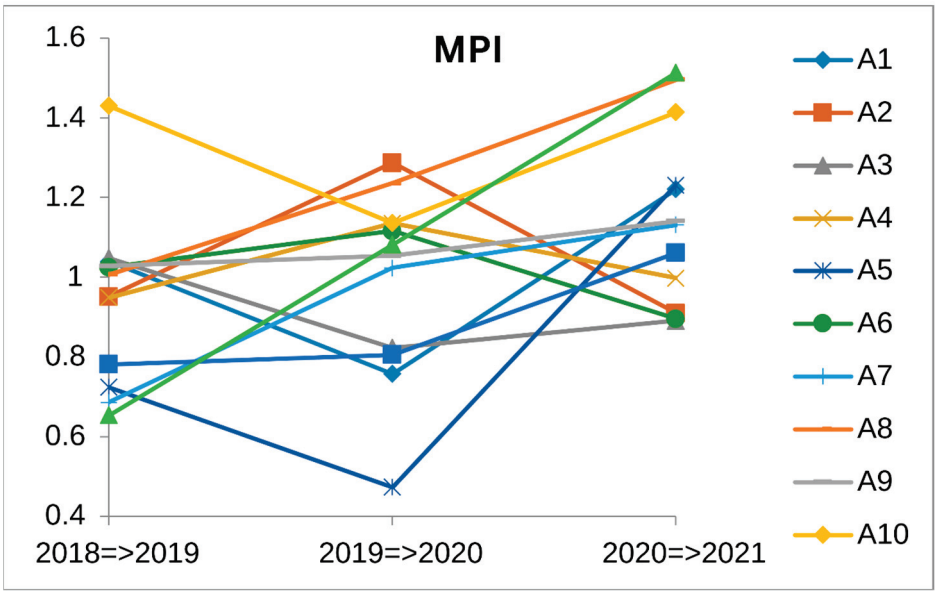


Figure 5. Malmquist Productivity Index.

4.3. Results of Epsilon-Based Measure Efficiency

The Malmquist approach results reflect the current operational picture of the world's leading twelve LIB manufacturers after calculating the total productivity via technical efficiency change (catch-up index) and investment in technology impacts (frontier-shift index). By applying the EBM approach in this research's third phase, we have data on twelve DMUs and may evaluate them according to their efficiency or inefficiency.

The EBM-I-C model, an input-oriented technique with a continuous restoration for scaling, is intentionally applied in the current investigation. The precision and relevance of the source to the topic at hand were the criteria that led to this selection. By constructing an affinity matrix from the observed input and output variables, we comprehensively investigate the diversity of production alternatives defined by EBM approaches. The diversity and affinity index matrix generated from the EBM methodology for 2018–2021 is presented in Tables A2 and A3, which may be seen below. According to the findings

of this research, the values of the diversity matrix and affinity matrix range from 0 to 0.2628 and 0.4744 to 1, which meet the basic requirements of the EBM model. Thus, we can utilize EBM to access and rank DMUs based on their efficiency scores. Table 7 then computes the EBM method input/output value and epsilon. The EBM model's epsilon, which integrates radial and non-radial features, is constantly positive during the period under assessment (range: 0.4136 to 0.4447).

Table 7. EBM Approach's Weight of Input and Epsilon (2018–2021).

Year	Weight to Input			Epsilon for EBM
	Assets	Liabilities	SG&A Expense	
2018	0.3338	0.3410	0.3252	0.4357
2019	0.3538	0.3406	0.3056	0.4136
2020	0.3444	0.3363	0.3192	0.4257
2021	0.3323	0.3245	0.3432	0.4447

EBM model epsilon and input/output weight are employed to determine the comparative effectiveness and inefficiencies of twelve DMUs for 2018–2021, as illustrated in Table 8 and Figure 6. The results show that enterprises A3, A5, and A12 are very efficient, with a score of 1 and a deficit of zero for 2018–2021. For the whole time of 2018–2021, A10 has the worst efficiency ratings. LIB manufacturers with less than one EBM efficiency rating (especially A1, A2, A4, A6, A8, A9, and A10 with the lowest score) do not operate at their full potential. For LIB companies whose EBM efficiency score reaches 1, this shows the potential for more profitable investment.

Table 8. The efficiency score of the EBM model (2018–2021).

DMUs	2018	2019	2020	2021	Average
A1	0.6588	0.915	0.7228	0.7545	0.7628
A2	0.5724	0.6756	1	0.7283	0.7441
A3	1	1	1	1	1.0000
A4	0.7095	0.8333	0.9711	0.7607	0.8187
A5	1	1	1	1	1.0000
A6	0.767	0.8129	0.9461	0.9027	0.8572
A7	1	0.8576	1	0.9373	0.9487
A8	0.5841	0.6693	0.9055	1	0.7897
A9	0.7103	0.8679	0.9066	0.8451	0.8325
A10	0.1532	0.3161	0.3925	0.454	0.3290
A11	0.8799	0.9217	0.7552	0.6306	0.7969
A12	1	1	1	1	1

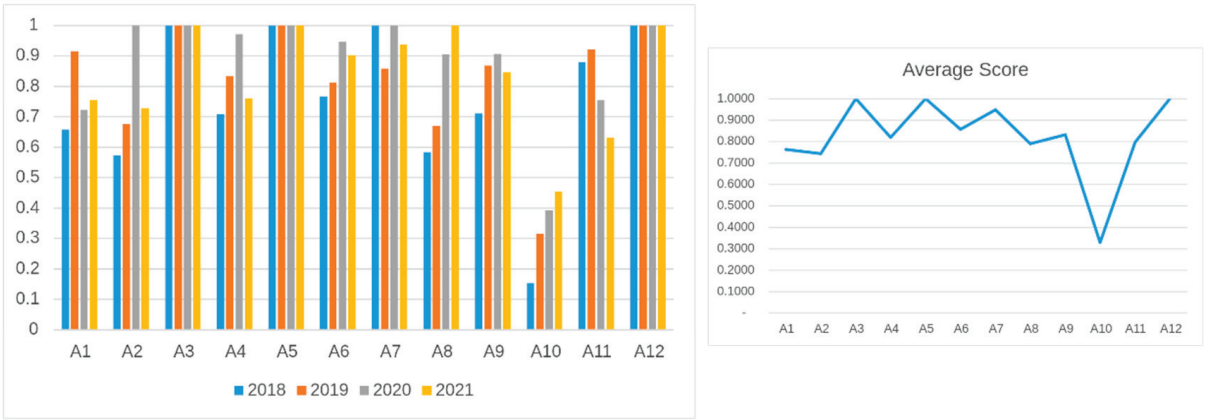


Figure 6. Ranking of LIB manufacturers.

5. Discussion and Conclusions

This article examines the productivity and efficiency of twelve lithium-ion battery (LIB) manufacturers from 2018 to 2021. Considering the Malmquist model's findings, we find that the efficiency of LIB producers has changed over time, but only in just a handful of distinct patterns (see Figure 5). However, in this increasingly computerized world, it is notable that technical developments have more unpredictable trends, which indicates how innovation in technology affects the performance of LIB manufacturers and the global lithium-ion batteries supply chain. On average, all DMUs advanced technologically throughout the research period. Looking at the catch-up index, which depicts the variation in operational efficiency, three DMUs (A10, A8, and A2) stand out for their remarkable technological efficiency. However, the performance of some DMUs has varied. For example, A2, A6, A7, and A8 technically peaked in 2019–2020 but fell in 2020–2021. Particularly noteworthy is A2, which declined in technological efficiency in the latter years and emerged as the least efficient producer. Only one-sixth of the evaluated firms attained the progressive average frontier change threshold, which suggests a lack of technical development in most DMUs. A4 shines as an exceptional processor, obtaining the highest score over the study period, but A3 witnessed a fall in technological progress in 2020–2021.

Our research emphasizes the significance of continuous technological innovation and flexibility in today's continually shifting corporate world. Frontier-shift indices provide insight into the many aspects impacting technological advancement. Companies must seek to improve their technical performance to remain competitive and flourish in the long run. Table 6 displays the results of the EBM model, which suggest that most enterprises have reached their maximum performance levels. Based on the results of this research, A10 is an interesting case. Since A10 is the best manufacturer in the catch-up index (the average catch-up index was 1.5509 in 2018–2021), it performs the best in implementing new technologies, methods, and processes that significantly impact its overall growth and productivity. In contrast, A10 has a meager efficiency score in the EBM model (an average of 0.3290 in 2018–2021) (Figure 6). Considering financial performance, management practices, strategic alignment, and the external environment, A10 shows low allocative efficiency or ineffective management practices.

Most significantly, the worldwide economic downturn caused by COVID-19 prevented most enterprises from obtaining efficiency scores in 2018–2021 (nine out of twelve DMUs). Among LIB manufacturers with EBM efficiency scores less than 1 (A1, A2, A4, A6, A7, A8, A9, A10, A11), companies A1, A2, A4, A6, A8, A9, and A10 have an MPI larger than 1. It means that these companies show a good performance of improvements in productivity and efficiency with technological and technical aspects but have a terrible performance in achieving their objectives and maximizing their potential with given available resources.

A3, A5, and A12 are three organizations that gain the highest efficiency score in the EBM model. While A3 (0.9208) and A5 (0.8093) have an average MPI score of less than 1, A12 (1.0828) has an average MPI score more significant than 1. As a result, A3 and A5 should put more effort into adopting new production techniques, better management practices, and implementing new technologies to improve total productivity change as it enables organizations to produce more outputs with the same or fewer inputs.

Overall, A12 is the company that has a high-efficiency score in the EBM model and shows a progressive performance in total productivity change for the 2018–2021 period. The development of lithium-ion batteries is moving at a breakneck pace, and numerous types of chemistry are successfully being made available. They are the batteries used in mobile smartphones, laptops, and other portable electronic gadgets [5,51,52]. Larger projects, such as energy storage, either partially or entirely electric motors, industrial vehicles, lifts, harbors and cranes, mining vehicles, boats, and submarines, are currently in the development stage [53,54]. This research contributes to the green energy market and gives a practical and detailed approach to determining how effective LIB enterprises are at achieving their goals. The hybrid method of DEA Malmquist and the techniques of

Epsilon-Based Measure (EBM) delivers an efficient and fair assessment framework that can be used to evaluate the performance of a firm and its growth in any direction.

These findings are helpful in that they can assist seaport owners in better understanding critical indicators for the growth and operation of LIB manufacturers, which in turn can lead to improved technological and technical element efficiency. Technology has become the dominant force in any sector because of the intensifying rivalry [55]. Because of this, there is a pressing need for a strategy that can be maintained to contribute to the construction of a more robust and resilient system.

The innovative comparative evaluation of LIB firms is one of the significant achievements of this research. This evaluation combines the DEA Malmquist and EBM models to determine the efficiency or inefficiency of DMUs by considering proportionate changes in inputs and outputs and the emergence of slacks. This method indicates the variety or dispersion of the data and the possibility of improving the parameters for information to the less efficient DMUs that span numerous times as well as multiple output and input variables. Additionally, this method reveals the possibility of improving parameters for inputs to the less efficient DMUs.

6. Limitations and Potential Further Research

Although DEA EBM and the DEA Malmquist model are valuable instruments for assessing performance and efficiency, their limits must be acknowledged. Subjectivity can bring bias into input and output choices in DEA EBM. Furthermore, it lacks benchmarking tools, making performance comparisons difficult. Malmquist's paradigm implies that technology is the primary driver of productivity change, disregarding other contributing factors. Furthermore, its success depends on proper period selection to avoid incomplete or misleading outcomes. Due to data availability constraints, our analysis used a four-year time range from 2018 to 2021.

The combination of DEA EBM and the Malmquist model addresses these constraints while providing comprehensive assessment benefits. This integration makes input–output selection more objective, minimizing subjectivity in DEA EBM. The Malmquist model is used to solve the DEA EBM scalability assumption by considering fluctuations in scaling efficiency over time. Furthermore, the Malmquist model allows for performance measurement and comparison, making it simple to find areas for improvement. Combining these models provides a comprehensive analysis considering technological development and relative efficiency.

To summarize, notwithstanding its shortcomings, the combination of DEA EBM and the DEA Malmquist model provides a solid evaluation framework. This integration enhances input and output selection, addresses scaling assumptions, allows benchmarking, and offers complete performance and efficiency analysis.

The subsequent studies must consider input and output variables to provide more accurate and reliable results. Moreover, approaches to multiple-criteria decision-making (MCDM), such as TOPSIS, AHP, VIKOR, WASPAS, and COCOSO, might provide more effective answers to the problem of ordering the business units [50]. Researchers can develop more exact approaches by comparing the outcomes using the ranking similarity coefficients. The hybrid approach, which combines the DEA Malmquist and EBM models, provides a more effective and transparent evaluation process to measure companies' performance and progress across all dimensions. This approach allows the creation of longer-term plans that contribute to the overall system's resilience.

Author Contributions: Conceptualization, C.-N.W. and V.T.T.N.; methodology, C.-N.W.; software, N.T.M.V.; validation, V.T.T.N.; formal analysis, V.T.T.N.; resources, N.T.M.V.; data curation, V.T.T.N. and N.T.M.V.; writing—original draft, N.T.M.V.; writing—review and editing, V.T.T.N. and F.-C.Y.; supervision, F.-C.Y. and C.-N.W.; project administration, F.-C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study did not receive any financial support from outside sources. Batteries Editorial Board funded APC.

Data Availability Statement: Provided when requested.

Acknowledgments: The authors would like to thank Ministry of Science and Technology, Taiwan. We also would like to thank the National Kaohsiung University of Science and Technology, Industrial University of Ho Chi Minh City, and Thu Dau Mot University for their assistance. Additionally, we would like to thank the reviewers and editors for their constructive comments and suggestions to improve our work.

Conflicts of Interest: The authors declare no conflict to interest.

Appendix A

Table A1. Input–output correlation (2018–2021).

Period	Inputs/Output	Assets (TOA)	Liabilities (LIA)	Selling, General & Administrative Expenses (SG&AE)	Revenue (REV)	Gross Income (GI)
2018	**TOA**	1	0.9474	0.7923	0.8191	0.8297
	LIA	0.9474	1	0.8583	0.808	0.8493
	SG&AE	0.7923	0.8583	1	0.8152	0.9802
	REV	0.8191	0.808	0.8152	1	0.8579
	GI	0.8297	0.8493	0.9802	0.8579	1
2019	**TOA**	1	0.9541	0.788	0.8501	0.8145
	LIA	0.9541	1	0.7799	0.8535	0.7865
	SG&AE	0.788	0.7799	1	0.8374	0.9861
	REV	0.8501	0.8535	0.8374	1	0.8587
	GI	0.8145	0.7865	0.9861	0.8587	1
2020	**TOA**	1	0.9764	0.7184	0.8553	0.021
	LIA	0.9764	1	0.7813	0.9161	0.034
	SG&AE	0.7184	0.7813	1	0.9031	0.2018
	REV	0.8553	0.9161	0.9031	1	0.0994
	GI	0.021	0.034	0.2018	0.0994	1
2021	**TOA**	1	0.9718	0.6803	0.9165	0.808
	LIA	0.9718	1	0.6299	0.8653	0.7216
	SG&AE	0.6803	0.6299	1	0.7942	0.9335
	REV	0.9165	0.8653	0.7942	1	0.8597
	GI	0.808	0.7216	0.9335	0.8597	1

Table A2. EBM model's Diversity matrix (2018–2021).

Period	Inputs/Output	Assets (TOA)	Liabilities (LIA)	Selling, General & Administrative Expenses (SG&AE)
2018	**TOA**	-	0.1977	0.2377
	LIA	0.1977	-	0.2187
	SG&AE	0.2377	0.2187	-
2019	**TOA**	-	0.1404	0.2218
	LIA	0.1404	-	0.2628
	SG&AE	0.2218	0.2628	-
2020	**TOA**	-	0.1782	0.2195
	LIA	0.1782	-	0.2422
	SG&AE	0.2195	0.2422	-
2021	**TOA**	-	0.2478	0.2008
	LIA	0.2478	-	0.2192
	SG&AE	0.2008	0.2192	-

Table A3. EBM model's Affinity matrix (2018–2021).

Period	Inputs/Output	Assets (TOA)	Liabilities (LIA)	Selling, General & Administrative Expenses (SG&AE)
2018	**TOA**	1.0000	0.6045	0.5247
	LIA	0.6045	1.0000	0.5627
	SG&AE	0.5247	0.5627	1.0000
2019	**TOA**	1.0000	0.7192	0.5565
	LIA	0.7192	1.0000	0.4744
	SG&AE	0.5565	0.4744	1.0000
2020	**TOA**	1.0000	0.6437	0.5610
	LIA	0.6437	1.0000	0.5156
	SG&AE	0.5610	0.5156	1.0000
2021	**TOA**	1.0000	0.5044	0.5985
	LIA	0.5044	1.0000	0.5616
	SG&AE	0.5985	0.5616	1.0000

References

1. Bridge, G.; Faigen, E. Towards the lithium-ion battery production network: Thinking beyond mineral supply chains. *Energy Res. Soc. Sci.* **2022**, *89*, 102659. [CrossRef]
2. Lithium-Ion Battery Market Size, Share & COVID-19 Impact Analysis. Available online: https://www.fortunebusinessinsights.com/industry-reports/lithium-ion-battery-market-100123 (accessed on 20 March 2023).
3. Chordia, M.; Nordelöf, A.; Ellingsen, L.A.-W. Environmental life cycle implications of upscaling lithium-ion battery production. *Int. J. Life Cycle Assess.* **2021**, *26*, 2024–2039. [CrossRef]
4. Duan, J.; Tang, X.; Dai, H.; Yang, Y.; Wu, W.; Wei, X.; Huang, Y. Building safe lithium-ion batteries for electric vehicles: A review. *Electrochem. Energy Rev.* **2020**, *3*, 1–42. [CrossRef]
5. Kurland, S.D. Energy use for GWh-scale lithium-ion battery production. *Environ. Res. Commun.* **2019**, *2*, 012001. [CrossRef]
6. Chapman, A.; Arendorf, J.; Castella, T.; Thompson, P.; Willis, P.; Espinoza, L.T.; Klug, S.; Wichmann, E. *Study on Critical Raw Materials at EU Level*; Oakdene Hollins: Buckinghamshire, UK, 2013.
7. Zhang, X.; Li, Z.; Luo, L.; Fan, Y.; Du, Z. A review on thermal management of lithium-ion batteries for electric vehicles. *Energy* **2022**, *238*, 121652. [CrossRef]
8. Cairns, E.J.; Albertus, P. Batteries for electric and hybrid-electric vehicles. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 299–320. [CrossRef]
9. Velázquez-Martínez, O.; Valio, J.; Santasalo-Aarnio, A.; Reuter, M.; Serna-Guerrero, R. A critical review of lithium-ion battery recycling processes from a circular economy perspective. *Batteries* **2019**, *5*, 68. [CrossRef]
10. Ding, Y.; Chen, X.; Wang, J. Deep Reinforcement Learning-Based Method for Joint Optimization of Mobile Energy Storage Systems and Power Grid with High Renewable Energy Sources. *Batteries* **2023**, *9*, 219. [CrossRef]
11. Mohanty, D.; Chen, S.-Y.; Hung, I.-M. Effect of Lithium Salt Concentration on Materials Characteristics and Electrochemical Performance of Hybrid Inorganic/Polymer Solid Electrolyte for Solid-State Lithium-Ion Batteries. *Batteries* **2022**, *8*, 173. [CrossRef]
12. Dunn, J.B.; Gaines, L.; Sullivan, J.; Wang, M.Q. Impact of recycling on cradle-to-gate energy consumption and greenhouse gas emissions of automotive lithium-ion batteries. *Environ. Sci. Technol.* **2012**, *46*, 12704–12710. [CrossRef]
13. Emilsson, E.; Dahllöf, L. *Lithium-Ion Vehicle Battery Production-Status 2019 on Energy Use, CO_2 Emissions, Use of Metals, Products Environmental Footprint, and Recycling*; IVL Svenska Miljöinstitutet: Stockholm, Sweden, 2019.
14. Yu, M.L.; Jin, H.; Zhang, H.J.; Chong, A.Y.L. ICT, Financial Development and Renewable Energy Consumption. *J. Comput. Inf. Syst.* **2023**, *63*, 190–203. [CrossRef]
15. Saranga, H. The Indian auto component industry–Estimation of operational efficiency and its determinants using DEA. *Eur. J. Oper. Res.* **2009**, *196*, 707–718. [CrossRef]
16. Tone, K.; Tsutsui, M. An epsilon-based measure of efficiency in DEA—A third pole of technical efficiency. *Eur. J. Oper. Res.* **2010**, *207*, 1554–1563. [CrossRef]
17. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [CrossRef]
18. Chen, K.; Ren, X.T.; Yang, G.L.; Qin, H.B. The other side of the coin: The declining of Chinese social science. *Scientometrics* **2022**, *127*, 127–143. [CrossRef]
19. Chen, X.; Liu, Z.; Saydaliev, H.B.; Abu Hatab, A.; Fang, W. Measuring Energy Efficiency Performance in China: Do Technological and Environmental Concerns Matter for Energy Efficiency? *Front. Energy Res.* **2021**, *9*, 779032. [CrossRef]
20. Chen, Y.W.; Yang, R.; Wong, C.W.Y.; Ji, J.W.; Miao, X. Efficiency and productivity of air pollution control in Chinese cities. *Sustain. Cities Soc.* **2022**, *76*, 103423. [CrossRef]

21. Mykhalovskiy, E.; Weir, L. The problem of evidence-based medicine: Directions for social science. *Soc. Sci. Med.* **2004**, *59*, 1059–1069. [CrossRef]
22. Li, Y.; Chiu, Y.-h.; Liu, Y.; Lin, T.-Y.; Chang, T.-H. The impact of the media and environmental pollution on the economy and health using a modified meta 2-stage EBM Malmquist model. *Inq. J. Heal. Care Organ. Provis. Financ.* **2020**, *57*, 0046958020921070. [CrossRef]
23. Rusli, N.A.E.; Ramli, N.A.; Shariff, S.S.R.; Zahid, Z.; Hussin, S.A.S. Evaluating the efficiency and productivity of Malaysian logistics companies using epsilon-based measure and Malmquist index during the Covid-19 pandemic. *J. Ind. Eng. Manag.* **2022**, *15*, 521–537.
24. Qin, Q.D.; Li, X.; Li, L.; Zhen, W.; Wei, Y.M. Air emissions perspective on energy efficiency: An empirical analysis of China's coastal areas. *Appl. Energy* **2017**, *185*, 604–614. [CrossRef]
25. Yang, W.P.; Zhao, B.Y.; Zhao, J.K.; Li, Z.D. An Empirical Study on the Impact of Foreign Strategic Investment on Banking Sustainability in China. *Sustainability* **2019**, *11*, 181. [CrossRef]
26. Cheng, S.X.; Xie, J.H.; Xiao, D.; Zhang, Y. Measuring the Environmental Efficiency and Technology Gap of PM2.5 in China's Ten City Groups: An Empirical Analysis Using the EBM Meta-Frontier Model. *Int. J. Environ. Res. Public Health* **2019**, *16*, 675. [CrossRef] [PubMed]
27. Lu, L.; Zhang, J.J.; Yang, F.; Zhang, Y.J. Evaluation and prediction on total factor productivity of Chinese petroleum companies via three-stage DEA model and time series neural network model. *Sustain. Comput. Inform. Syst.* **2020**, *27*, 100397. [CrossRef]
28. Nguyen, P.H.; Nguyen, T.L.; Nguyen, T.G.; Nguyen, D.T.; Tran, T.H.; Le, H.C.; Phung, H.T. A Cross-Country European Efficiency Measurement of Maritime Transport: A Data Envelopment Analysis Approach. *Axioms* **2022**, *11*, 206. [CrossRef]
29. Färe, R.; Grosskopf, S. Malmquist productivity indexes and Fisher ideal indexes. *Econ. J.* **1992**, *102*, 158–160. [CrossRef]
30. Golany, B.; Roll, Y. An application procedure for DEA. *Omega* **1989**, *17*, 237–250. [CrossRef]
31. Färe, R.; Grosskopf, S.; Lindgren, B.; Roos, P. Productivity developments in Swedish hospitals: A Malmquist output index approach. In *Data Envelopment Analysis: Theory, Methodology, and Applications*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 253–272.
32. Bai, H.; Fan, Y.; Wang, L.; Vo, N.V.T.; Nguyen, T.V.T. Research on Battery Characteristics and Management System of New Energy Vehicle Based on BMS System Design and Test. *Comput. Aided Des. Appl.* **2023**, *20* (Suppl. S3), 200–212. [CrossRef]
33. Kulshreshtha, M.; Parikh, J. Study of efficiency and productivity growth in opencast and underground coal mining in India: A DEA analysis. *Energy Econ.* **2002**, *24*, 439–453. [CrossRef]
34. Ahn, T.; Charnes, A.; Cooper, W.W. Efficiency characterizations in different DEA models. *Socio-Econ. Plan. Sci.* **1988**, *22*, 253–257. [CrossRef]
35. Tavana, M.; Mirzagoltabar, H.; Mirhedayatian, S.M.; Saen, R.F.; Azadi, M. A new network epsilon-based DEA model for supply chain performance evaluation. *Comput. Ind. Eng.* **2013**, *66*, 501–513. [CrossRef]
36. Xiao, Q.W.; Tian, Z.; Ren, F.R. Efficiency assessment of electricity generation in China using meta-frontier data envelopment analysis: Cross-regional comparison based on different electricity generation energy sources. *Energy Strategy Rev.* **2022**, *39*, 100767. [CrossRef]
37. Yao, L.M.; Shuai, Y.H.; Chen, X.D.; Xiao, A.R. A two-stage EBM-based approach to evaluate operational performance of unattended convenience store. *Int. J. Retail. Distrib.* **2020**, *48*, 609–627. [CrossRef]
38. Ting, I.W.K.; Chen, F.C.; Kweh, Q.L.; Sui, H.J.; Le, H.T.M. Intellectual capital and bank branches' efficiency: An integrated study. *J Intellect. Cap.* **2021**, *23*, 840–863. [CrossRef]
39. Wang, C.N.; Day, J.D.; Nguyen, T.K.L. Applying EBM Model and Grey Forecasting to Assess Efficiency of Third-Party Logistics Providers. *J. Adv. Transp.* **2018**, *2018*, 1212873. [CrossRef]
40. Agostino, M.; Nifo, A.; Ruberto, S.; Scalera, D.; Trivieri, F. Productivity changes in the automotive industry of three European countries. An application of the Malmquist index decomposition analysis. *Struct. Chang. Econ. Dyn.* **2022**, *61*, 216–226. [CrossRef]
41. Pan, Z.W.; Tang, D.C.; Kong, H.J.; He, J.X. An Analysis of Agricultural Production Efficiency of Yangtze River Economic Belt Based on a Three-Stage DEA Malmquist Model. *Int. J. Environ. Res. Public Health* **2022**, *19*, 958. [CrossRef]
42. Homburg, C. Using data envelopment analysis to benchmark activities. *Int. J. Prod. Econ.* **2001**, *73*, 51–58. [CrossRef]
43. Zhou, G.G.; Min, H.; Xu, C.; Cao, Z.Y. Evaluating the comparative efficiency of Chinese third-party logistics providers using data envelopment analysis. *Int. J. Phys. Distrib. Logist. Manag.* **2008**, *38*, 262–279. [CrossRef]
44. Lu, W.M.; Hung, S.W. Exploring the efficiency and effectiveness in global e-retailing companies. *Comput. Oper. Res.* **2011**, *38*, 1351–1360. [CrossRef]
45. Park, H.G.; Lee, Y.J. The efficiency and productivity analysis of large logistics providers services in Korea. *Asian J. Shipp. Logist.* **2015**, *31*, 469–476. [CrossRef]
46. Chen, Q.; Li, F. Empirical analysis on efficiency of listed real estate companies in China by DEA. *Ibusiness* **2017**, *9*, 49. [CrossRef]
47. Ahmed, A.A.; Mohamad, A. Data envelopment analysis of efficiency of real estate investment trusts in Singapore. *Int. J. Law Manag.* **2017**, *59*, 826–838. [CrossRef]
48. Shah, A.A.; Wu, D.; Korotkov, V. Are sustainable banks efficient and productive? A data envelopment analysis and the Malmquist productivity index analysis. *Sustainability* **2019**, *11*, 2398. [CrossRef]
49. Atta Mills, E.F.E.; Baafi, M.A.; Liu, F.; Zeng, K. Dynamic operating efficiency and its determining factors of listed real-estate companies in China: A hierarchical slack-based DEA-OLS approach. *Int. J. Financ. Econ.* **2021**, *26*, 3352–3376. [CrossRef]

50. Wang, C.N.; Yang, F.C.; Vo, N.T.M.; Nguyen, V.T.T. Wireless Communications for Data Security: Efficiency Assessment of Cybersecurity Industry-A Promising Application for UAVs. *Drones* **2022**, *6*, 363. [CrossRef]
51. Meng, F.; McNeice, J.; Zadeh, S.S.; Ghahreman, A. Review of lithium production and recovery from minerals, brines, and lithium-ion batteries. *Miner. Process. Extr. Metall. Rev.* **2021**, *42*, 123–141. [CrossRef]
52. Schnell, J.; Nentwich, C.; Endres, F.; Kollenda, A.; Distel, F.; Knoche, T.; Reinhart, G. Data mining in lithium-ion battery cell production. *J. Power Sources* **2019**, *413*, 360–366. [CrossRef]
53. Dai, Q.; Kelly, J.C.; Gaines, L.; Wang, M. Life cycle analysis of lithium-ion batteries for automotive applications. *Batteries* **2019**, *5*, 48. [CrossRef]
54. Hüger, E.; Riedel, L.; Zhu, J.; Stahn, J.; Heitjans, P.; Schmidt, H. Lithium Niobate for Fast Cycling in Li-ion Batteries: Review and New Experimental Results. *Batteries* **2023**, *9*, 244. [CrossRef]
55. Balan, S.; Conlon, S. Text Analysis of Green Supply Chain Practices in Healthcare. *J. Comput. Inf. Syst.* **2018**, *58*, 30–38. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

batteries

MDPI

Article

Deep Reinforcement Learning-Based Method for Joint Optimization of Mobile Energy Storage Systems and Power Grid with High Renewable Energy Sources

Yongkang Ding [1], Xinjiang Chen [1] and Jianxiao Wang [2,*]

[1] Department of Industrial Engineering and Management, Peking University, Beijing 100871, China
[2] National Engineering Laboratory for Big Data Analysis and Applications, Peking University, Beijing 100871, China
* Correspondence: wang-jx@pku.edu.cn

Abstract: The joint optimization of power systems, mobile energy storage systems (MESSs), and renewable energy involves complex constraints and numerous decision variables, and it is difficult to achieve optimization quickly through the use of commercial solvers, such as Gurobi and Cplex. To address this challenge, we present an effective joint optimization approach for MESSs and power grids that consider various renewable energy sources, including wind power (WP), photovoltaic (PV) power, and hydropower. The integration of MESSs could alleviate congestion, minimize renewable energy waste, fulfill unexpected energy demands, and lower the operational costs for power networks. To model the entire system, a mixed-integer programming (MIP) model was proposed that considered both the MESSs and the power grid, with the goal of minimizing costs. Furthermore, this research proposed a highly efficient deep reinforcement learning (DRL)-based method to optimize route selection and charging/discharging operations. The efficacy of the proposed method was demonstrated through many numerical simulations.

Keywords: renewable energy; battery energy storage system; machine learning; deep reinforcement learning; data-driven optimization; cost minimization

1. Introduction

Renewable energy sources such as wind, water, and solar energy have considerable capacity to lower energy costs and carbon emissions while playing a crucial role in creating low-carbon and sustainable energy systems, as highlighted in [1–3]. However, the intermittency and variability of renewable energy pose a great challenge to the safe and economic operation of power systems in regions with an abundance of renewable energy potential [4,5]. To address this challenge, energy storage systems (ESSs) are a promising technology for the integration of renewable energy and the reduction in renewable curtailment.

Stationary energy storage systems (SESSs), the most conventional application mode of BES, have enabled energy conversion and capacity sharing during a limited time period [6,7]. Specifically, given the predicted renewable energy generation curve, SESSs, renewable energy, and thermal units were scheduled collaboratively to integrate renewable energy. However, SESS was inflexible, as it relies on large-capacity and long-distance transmission lines. Meanwhile, the utilization efficiency of SESSs also depended on its location, as it was difficult for the owners of a SESS to determine the optimal location because numerous constraints had to be considered, such as political, economic, social, and geographical factors [8–10]. Therefore, mobile energy storage systems (MESSs), as an ongoing development application mode of BES that couples energy and transportation systems, incorporate vehicles (e.g., electric vehicles (EVs), trains, ships, etc.), batteries,

power converters, and transformers. As compared to SESSs, MESSs enable energy conversion and the capacity sharing of energy storage over longer time periods through their transportation network and considerably improve the utilization rate of battery assets and the emergency dispatching capability of the power system. MESSs had more flexible deployment capabilities and had a more remarkable capacity in energy arbitrage [11,12], renewable energy integration [13–15], peak-shaving [16,17], grid-congestion relief [18,19], and grid-investment deferral [20–22].

The authors of [11] proposed the concept of a utility-scale MESS, which incorporated electric trucks, energy storage, and energy conversion systems; constructed an optimization model involving charge scheduling and route planning of MESSs; and used the proposed model in the energy arbitrage of a power grid in California. The results revealed that, as compared to the SESS, the MESS had considerable benefits over a life-cycle time period. In [13], the researchers integrated MESS into a large-scale power distribution system in order to integrate renewable energy in remote areas and proposed a particle-swarm optimization to streamline a scheduling model of MESSs. The results showed that the proposed models and algorithm could considerably reduce the operating cost of the power system. The authors of [14] developed a decision framework that used MESS to improve system resilience for the emergency dispatching of the power distribution system, and they verified that the proposed decision framework could effectively improve system response-and-recovery speed in the emergency scenarios. In [18], the researchers integrated large-scale EVs into the electric energy distribution system and proposed a mixed-integer linear programming (MILP) model to optimize the charging schemes of EVs. The results showed that the proposed model could substantially alleviate grid congestion and improve the performance of EVs and grid benefits. The authors of [20] developed a collaborative planning model for SESSs and MESSs to reduce the capital cost of transmission lines, energy storage systems, and the operating costs of power systems. Meanwhile, ref. [20] used the proposed model to determine the optimal region for SESS and MESS implementation for China's northwestern grid. The aging issue of mobile energy storage devices has also attracted attention. Reference [23] analyzed the influence mechanism of grid-connected operation on the life degradation of lithium-ion battery energy storage systems and established a cost accounting model for frequency regulation, considering the impact of battery life degradation. Reference [24] studied the frequency regulation problem of a power grid model that includes loads, traditional generators, and multiple electric vehicles, aiming to minimize the degradation of battery devices. Their proposed strategies demonstrated high effectiveness under actual operating conditions. In this paper, in order to emphasize the modeling of mobile energy storage systems and their joint optimization with the power grid, we have simplified the description of this part.

Recently, deep reinforcement learning (DRL) techniques have seen growing applications in battery management [25,26], grid management, and micro-grid management [27,28]. The authors of [25] presented a DRL-based method for a battery system in EVs, which consisted of a high-power battery pack in order to reduce energy loss and enhance thermal safety. The proposed strategy demonstrated advantages in terms of reducing calculation time and energy consumption. In [26], the researchers proposed a joint optimization framework of DRL and binary integer programming for the battery swapping–charging systems for EVs, which had favorable performance in a real-time demonstration, as well as improved computational efficiency and privacy preservation. A multi-agent system based on an RL method was applied to manage a stand-alone micro-grid including a power-production unit, a power-consumption unit, and a power-storage unit in [27]. The authors of [28] proposed an optimal control strategy based DRL, which had asynchronous advantages via its $actor - critic$ framework to manage and optimize an online energy system.A multi-agent DRL-based approach was developed by [29] for real-time route selection and dispatching of multiple coordinated MESSs, and it had the ability to handle a hybrid continuous-discrete action space and to strengthen resiliency after extreme events. In [30], the researchers proposed a DRL-based method that coordinated the scheduling of the

MESSs and the resource allocation of micro-grids in order to minimize the overall cost of the system. In the future, a significant amount of data is expected to be generated in the fields of energy and transportation [19], which could provide significant data support for solving MESS's challenges through the application of DRL-based methods.

We have summarized the commonly used forms of energy storage, including their optimization objectives, research methods, and integration with the power grid, as shown in Table 1.

Table 1. Summary of energy management problem and power grid.

	Authors	**[31]**	**[9]**	**[10]**	**[11]**	**[32]**	**[30]**	**Ours**
	Year	**2021**	**2020**	**2021**	**2021**	**2022**	**2020**	**2023**
Modality	SESSs	✓						
	MESSs		✓	✓	✓	✓	✓	✓
Objective function	Cost/benefit	✓	✓		✓	✓		✓
	Grid peak			✓				
	Resilience						✓	
Method	Math programming		✓		✓			
	Heuristic					✓		
	Control	✓		✓				
	Deep reinforcement learning						✓	✓
Coordinated with	Power grid	✓	✓	✓		✓	✓	✓
	Renewable energy							✓

Therefore, based on the aforementioned research, most of the methods adopted to solve joint optimizations in power systems, MESSs, and other renewable energy components have been commercial solvers, heuristics, and meta-heuristics [12,20]. More recently, DRL-based methods have shown strong abilities in resolving combinatorial optimization challenges,and all have achieved remarkable results. To the best of our knowledge, however, few studies have focused on applying DRL to solve the cooperative scheduling challenges of batteries and power systems. Therefore, we proposed a joint mixed-integer programming (MIP) model for MESSs and the power grid. To address the nonlinearity and the elimination of integer variables, a hybrid strategy was proposed. A MESS model, which included 0–1 integer variables, was resolved using a DRL-based approach. The output from the DRL-based method was used as a portion of the input for the power-grid model, which was a simple linear programming (LP) model and easy to solve. The MESSs were formulated as constrained Markov decision processes (CMDPs), which served as the foundation for the DRL design. To enable the agent to make simultaneous decisions about the destination and power, a hybrid discrete-continuous action space was designed. The effectiveness of the proposed method was demonstrated through numerical experiments. Our contributions were summarized in three key areas:

(1) We developed an MIP model for the joint system of the MESSs and power grids to minimize the total operating costs. The model included constraints on the output of thermal power and renewable energy sources, the energy transmission of the power grid and the MESSs, the locations of the MESSs, etc.;
(2) We presented a formulation of the MESS challenge as a CMDP and introduce a DRL-based algorithm to make decisions in a hybrid action space that combined both discrete and continuous variables;
(3) We proposed a new linear programming (LP) model that eliminated the constraints on the MESSs in the original model by using the DRL-based method, thereby eliminating the integer-variable constraints and significantly improving the solving time. There-

fore, we proposed an algorithmic framework that combined DRL-based methods with the solutions of new LP models.

The remainder of this paper is structured as follows. In Section 2, the scheduling optimization challenge of the MESSs and the power grid is formulated as a MIP model. Section 3 describes the proposed DRL-based method. In Section 4, simulation studies are performed to validate the efficacy of the proposed approach. Section 5 presents the conclusions.

2. Mobile Energy Storage Systems and Power Grid Model

The joint optimization model for the MESSs and power grids aimed to minimize costs while considering constraints on energy transportation and operation. The objective function was comprised of two components: thermal power cost and MESS transportation cost, as follows:

$$\min \sum_{g\in G}\sum_{u\in U_g}\sum_{t\in T} C^g_{u,t}P^g_{u,t} + c^{tra}\Delta t \sum_{g\in G}\sum_{f\in G}\sum_{t\in T}\gamma_{g,f,t}, \tag{1}$$

where G denotes the electricity consumption node set and g denotes a single node; U_g denotes the thermal power unit set in node g; T represents the decision cycle, and t represents time index, where time slot length δt is set to 1 hour. The variables $C^g_{u,t}$ and $P^g_{u,t}$ indicate the generation cost of thermal power units and output, respectively. In addition, c^{tra} denotes the transportation cost per unit of time, which was set to \$20/h in this study. The expression $\gamma_{g,f,t} \in 0,1$ represents whether the MESSs were traveling between nodes g and f (1 indicated traveling and 0 indicated not traveling).

The constraints of the joint model were the following:

$$P^U_{u,\min} \le P^U_{u,t} \le P^U_{u,\max}, \qquad \forall u \in U_g, g \in G, t \in T, \tag{2}$$

$$P^U_{u,t} - P^U_{u,t-1} \le RU_{u,t}, \qquad \forall u \in U_g, g \in G, t \in T, \tag{3}$$

$$P^U_{u,t} - P^U_{u,t-1} \ge -RD_{u,t}, \qquad \forall u \in U_g, g \in G, t \in T, \tag{4}$$

$$0 \le P^R_{r,t} \le P^R_{r,t,\max}, \qquad \forall r \in R_g, g \in G, t \in T, \tag{5}$$

$$-P^{TL}_{l,t,\max} \le P^{TL}_{l,t} \le P^{TL}_{l,t,\max}, \qquad \forall l \in TL, g \in G, t \in T, \tag{6}$$

$$\sum_{l\in TL} P^{TL}_{l,t} = 0, \qquad \forall t \in T, \tag{7}$$

$$\sum_{u\in U_g} P^g_{u,t} + \sum_{r\in R_g} P^R_{r,t} + \sum_{l\in TL} P^{TL}_{l,t} = P^L_{g,t} + \omega_{g,t}P^{MESS}_{g,t}, \quad \forall u \in U_g, r \in R_g, g \in G, t \in T, \tag{8}$$

$$0 \le P^{MESS}_{g,t} \le \omega_{g,t}P_{max}, \qquad \forall g \in G, t \in T, \tag{9}$$

$$\sum_{g\in G} \omega_{g,t} \le 1 - \sum_{f\in G} \gamma_{g,f,t}, \qquad \forall g \in G, t \in T, \tag{10}$$

$$\alpha_{g,t} - \beta_{g,t} = \omega_{g,t} - \omega_{t,(t-1)}, \qquad \forall g \in G, t \in T, \tag{11}$$

$$\sum_{g\in G} \left(\alpha_{g,t} + \beta_{g,t}\right) \le 1, \qquad \forall t \in T, \tag{12}$$

$$\sum_{f\in G} \gamma_{g,f,t} \ge \beta_{g,t}, \qquad \forall f \in G, t \in T, \tag{13}$$

$$\alpha_{f,t} - \theta_{f,t} = \sum_{g\in G} \left(\gamma_{g,f,t-1} - \gamma_{g,f,t}\right), \qquad \forall f \in G, t \in T, \tag{14}$$

$$\sum_{g\in G} \left(\alpha_{g,t} + \theta_{g,t}\right) \le 1, \qquad \forall t \in T, \tag{15}$$

$$\gamma_{g,f,t} \ge \gamma_{g,f,t-1} - \gamma_{g,f\left(t-H_{g,f,t}\right)}, \qquad \forall g \in G, f \in G, t \in T. \tag{16}$$

where $P^U_{u,\min}$ and $P^U_{u,\max}$ denote the minimum output and maximum output of thermal power units, respectively. The variables $RU_{u,t}$ and $RD_{u,t}$ denote the increased or decreased output of the thermal units, respectively, (with ramp constraints). In addition, R_g is the set

of renewable energy resources including WP, PV, and hydropower. The actual output of the renewable energy resources and the maximum output of the renewable energy resources are expressed as $P^R_{r,t}$ and $P^R_{r,t,max}$, respectively. Furthermore, $P^{TL}_{l,t}$ and $P^{TL}_{l,t,\max}$ represent the actual transmission power and maximum transmission power of tie-line l, respectively. The expression $l = (g, f) \in TL$ denotes the tie-line between nodes g and f. The load in node g is represented as $P^L_{g,t}$.

The constraints were divided into two categories. The first category included the constraints on the power grid, including inequalities, as expressed in (2)–(8). Constraint (2) included the upper and lower limits of the thermal unit output. Constraints (3) and (4) represented the ramp constraints for thermal power units. The output of renewable energy resources was constrained by inequality, as shown in (5). The constraints on the transmission power of the tie-lines were expressed in (6) and (7). Constraint (8) represented the balance of power output for the whole joint system of the MESSs and power grids.

The second category included constraints (9)–(16) of the MESSs, with respect to [11]. Constraint (9) denoted the limit of the power output of the MESSs. The storage could not stop at a node if it was traveling between nodes and could only appear at one node at one time, as expressed in constraint (10). Equations (11)–(15) modeled the transportation status between nodes g and f, where $\alpha_{g,t} \in \{0, 1\}$ and $\beta_{g,t} \in \{0, 1\}$ were both binary variables that denoted whether the MESSs were moving to node g at time t or whether the MESSs were moving from node g at time t, respectively.In addition, $\theta_{g,t} \in \{0, 1\}$ represented an auxiliary binary variable. Constraint (11) connected the change of the locator indicators $\omega_{g,t}$ with the arrival indicators $\alpha_{g,t}$ and depart indicators $\beta_{g,t}$. The arrival and departure could not occur simultaneously, as expressed in constraint (12). Constraint (13) managed the storage depart node. Constraints (14) and (15) ensure that the arrival indicators $\alpha_{g,t}$ are equal to 1 for the arrival times; otherwise, they were equal to 0. Constraint (16) calculated the transportation time $H_{g,f,t}$ from node g to node f. In this study, the traveling time was estimated as a square matrix for all nodes.

Model (1) contained a large number of integer variables, which are known to be computationally demanding. In this proposal, we considered a machine-learning approach to solve for the integer variables. This study formulated the storage movement process as a Markov decision process (MDP) and applied a DRL-based method to solve it.

3. Deep Reinforcement Learning-Based Method

This section begins with a concise overview of the evolution of DRL-based techniques. The MESSs challenge is then framed as an MDP. The agent in the DRL-based approach was represented by an electric truck and its operator. Next, we demonstrated the integration of the DRL-based technique into the MIP model. Finally, the formulation of the action space, state space, and reward function for the DRL-based method was established. The proximal gradient projection algorithm was employed to ensure the safe exploration by the agent.

3.1. Background of DRL

The basement of the reinforcement-learning phase was the Bellman equation [33]:

$$Q^*(s, a) = \mathbb{E}_{n' \sim \varepsilon}\left[r + \gamma \max_{a'} Q^*\left(s', a'\right) \mid s, a\right], \tag{17}$$

where the optimal value of the state sequence s, represented by $Q(s, a)$, is determined by selecting the action a that maximizes the expected value of $Q(s, a)$.

Traditional reinforcement learning encounters the challenge of dimensionality. The authors of [34] addressed this issue by utilizing neural networks (NNs) to approximate the Q-value table. The Q-value was approximated using NNs, expressed as $Q(s, a; w)$, which was a close estimation of the true Q-value $Q(s, a)$. The weights of the NNs were represented

by the variable w. The weight parameters were updated through the loss function [35]:

$$L_t(w) = \left\{ Q(s_t, a_t; w) - \left[r_t + \gamma \max_{a' \in \mathcal{A}} Q\left(s_{t+1}, a'; w_t\right) \right] \right\}^2. \tag{18}$$

The maximization calculation in Equation (18) is computationally inefficient for continuous action spaces due to the non-convex nature of the function $Q(s, a; w)$ with respect to action a. This renders the maximization calculation NP-hard in the worst-case scenario [36]. To overcome this challenge, ref. [37] introduced the deterministic policy gradient (DPG) theorem which utilizes policy-based methods for continuous action spaces through the implementation of deterministic policies $\mu_\theta : S \to \mathcal{A}$. The objective of policy gradient methods is to discover a policy π_θ that optimizes the expected reward. The objective of the policy, $J(\pi_\theta)$, is then updated through gradient descent [38]:

$$\nabla_\theta J(\mu_\theta) = \mathop{\mathbb{E}}_{s \sim \rho^{\mu_\theta}} \left[\nabla_\theta \mu_\theta(s) \nabla_a Q^{\mu_\theta}(s, a)|_{a=\mu_\theta(s)} \right]. \tag{19}$$

3.2. Algorithm Framework

The *actor* − *critic* framework has proven to be effective in various domains, such as energy storage systems [26] and the game Go [39]. We adopted a similar structure in this proposal. The *Actor* network was a strategic network designed to determine the power parameters of the MESSs, which had a continuous feasible range. The *Critic* network assessed the actor–network's parameter choices and decided on other discrete variables, such as destination selection, charging and discharging selection, etc.

The hybrid action space $\mathcal{A}$ was defined according to [36], as follows:

$$\mathcal{A} = \{(k, x_k) \mid x_k \in \mathcal{X}_k \text{ for all } k \in K\}, \tag{20}$$

which consists of discrete action $k \in \mathcal{K}$ and continuous action $x_k \in \mathcal{X}_k$. The Bellman equation transforms into [36], as follows:

$$Q\left(s_t, k_t, x_{k_t}\right) = \mathop{\mathbb{E}}_{r_t, s_{t+1}} \left[r_t + \gamma \max_{k \in K} Q\left(s_{t+1}, k, x_k^Q(s_{t+1})\right) \mid s_t = s \right]. \tag{21}$$

The value-based network's weights are denoted as ω, and the policy-based network's weights are denoted as θ. In each step t of the updating process, ω_t and θ_t were updated, respectively. First, the weight parameter, ω, was estimated by employing the gradient descent to minimize the mean-squared Bellman error, which was similar to the DQN method. Then, ω and θ were fixed by maximizing $Q(s, k, x_k(s; \theta); \omega)$. The loss functions were formulated as follows [40]:

$$\ell_t^Q(\omega) = \frac{1}{2} \left[Q\left(s_t, k_t, x_{k_t}; \omega\right) - y_t \right]^2, \tag{22}$$

$$\ell_t^\Theta(\theta) = -\sum_{k=1}^{K} Q(s_t, k, x_k(s_t; \theta); \omega_t), \tag{23}$$

where $y_t = r_t + \gamma \max_{k' \in K} Q(s_{t'}, k, x_k(s_{t'}, \theta_t); \omega_t)$.

According to [41], there were three approaches for utilizing machine learning in combinatorial optimization challenges: using machine learning on its own, integrating machine learning with traditional optimization algorithms, and iteratively combining optimization and machine learning. For this proposal, we employed the DRL-based approach to efficiently eliminate integer variables in model (1). The objective function of model 1 served as the reward for reinforcement learning, and the DRL-based method was used to obtain the integer variables in constraints (9)–(16). The value of $\omega_{g,t}$ in constraint (8) was passed as an input to the new model.

The new model was described by the following:

$$\begin{aligned} &\min \sum_{g\in G}\sum_{u\in U_g}\sum_{t\in T} C^g_{u,t}P^g_{u,t} + C^{tra} \\ &s.t. \quad \text{constraints (2)} - \text{(8) in model (1)}, \end{aligned} \tag{24}$$

where C^{tra} is the total transportation cost of the MESSs.

The algorithm framework is depicted in Figure 1. State S_t (P1) consisted of the output of thermal units and renewable energy at grid nodes, while State S_t (P2) included the remaining energy of the MESSs, the cost of thermal power units, and the location information. State S_t was the combination of these five elements.

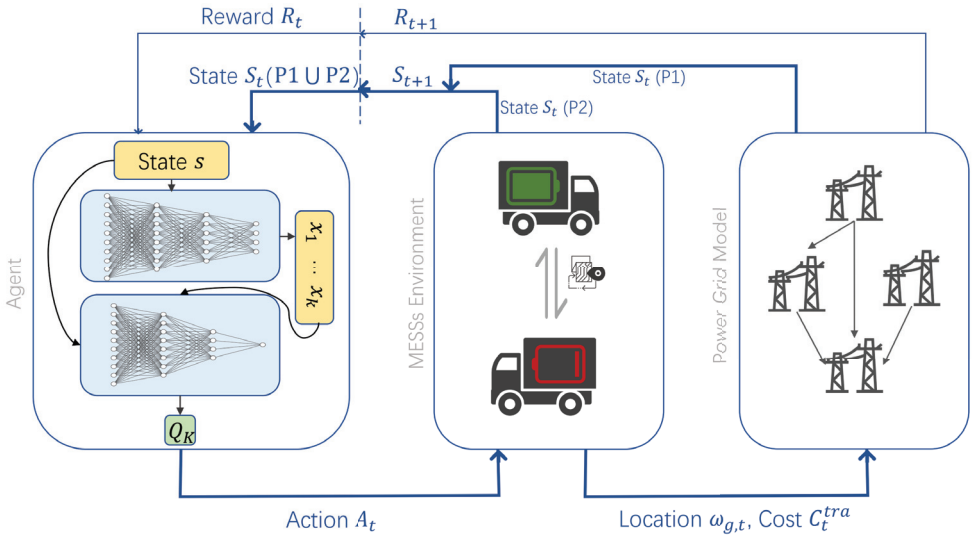


Figure 1. Framework of optimization method for the joint model of MESSs and power grids.

3.3. Constrained MDP Formulation

According to model (1), this study modeled the MESSs in this model as MDPs. However, due to the capacity constraint of the battery packs in the MESSs, the processes became CMDPs. To address this, we designed a hybrid action space and a state space for the CMDPs. A proximal gradient projection algorithm was proposed to ensure safe exploration by the agent.

3.3.1. Action Space $\mathcal{A}$, State Space $\mathcal{S}$, and Reward Function $\mathcal{R}$

Action space $\mathcal{A}$ was composed of two levels: The discrete action space that included a chosen charging/discharging/holding action and the chosen destination node; and the continuous action that included the output power parameters of the MESSs. At time t, the action a_t was represented by $\left((g_t, cdh_t)|P^{cdh}_{g,t}\right)$, where g is the chosen destination node, cdh represents the charging/discharging/holding choice, and $P^{cdh}_{g,t}$ is the corresponding power at node g at time t. The range of actions and their values were defined by the following:

$$\mathcal{A} = \begin{cases} g \in \{g_1, g_2, ..., g_N\} \\ cdh \in \{-1, 0, 1\} \\ P = \begin{cases} (0, P_{max}], cdh = -1 \text{ or } chd = 1 \\ 0, cdh = 0 \end{cases} \end{cases}, \tag{25}$$

where N is the node's number, and $cdh = -1$, 0, and 1 means discharging, holding, and charging, respectively.

The state space $\mathcal{S}$ consisted of two parts, as illustrated in Figure 1. The variable $\mathcal{S}_t(P1)$ described the state of the power grid, and $\mathcal{S}_t(P2)$ described the state related to the MESS. At time t, $\mathcal{S}$ could be defined as $\mathcal{S}_t = \left(P_{u,t}, RE_{g,t}, E_{g,t}, Load_{g,t}, g\right)$, where $RE_{g,t}$ is the renewable energy resources at node g at time t, $E_{g,t}$ is the remaining energy in the MESSs at node g at time t, and $Load_{g,t}$ represents the load at node g. There was a capacity constraint for the MESSs, $\underline{E} \leq E_{g,t} \leq \bar{E}$, which indicated the battery pack had an upper and lower limit on its capacity, represented by $\underline{E}$ and $\bar{E}$, respectively.

The reward $\mathcal{R}$ was based on the objective value of the linear programming (LP) model (24). The reward in RL was a cumulative value, as depicted in Equation (21), where $Q(s,a) = \sum_{t \in T} \gamma^t r_t$. In this study, the objective value of model (24) was used as the reward: $r_t = -Obj_{m2}$, where Obj_{m2} represents the optimal objective value of model (24).

Theorem 1. *In a finite MDP, if $\sum_{t \in T} -Obj_{m2,t}$ was the maximum when training converged with a policy, then each $-Obj_{m2,t}$ $t \in T$ was the maximum in the policy.*

Proof of Theorem 1. If ϕ is a policy, then

$$q_\pi(s^*, a^*) \geq v_\pi(s^*).$$

Suppose, then, that there is a policy $\phi^{'}$:

$$v_{\pi^{'}}(s) \geq v_\pi(s) \text{ for all } s,$$
$$\pi^{'}(a \mid s) = 1(s^* = s)1(a = a^*) + 1(s^* \neq s)\pi(a \mid s).$$

We considered $v^*(s) = \max_{\pi \in \Pi} v_\pi(s)$ and $q^*(s,a) = \max_{\pi \in \Pi} q_\pi(s,a)$. Now extrapolating backward, there existed s^*, resulting in the following:

$$\max_a q^*(s^*, a) > v^*(s^*).$$

Suppose there existed a policy $\pi^{''}$, as well as a^*:

$$q_{\pi^{''}}(s^*, a^*) = \max_a q^*(s^*, a)$$

while v^* was the maximum over all π, for $\pi = \pi^{''}$, we obtained the following inequality:

$$q_{\pi^{''}}(s^*, a^*) > v_{\pi^{'}}(s^*).$$

Therefore, we obtained the following:

$$v_{\pi^{'}}(s^*) \geq q_{\pi^{''}}(s^*, a^*) > v_{\pi^{'}}(s^*).$$

However, $q_{\pi^{''}}(s^*, a^*)$ was strictly larger than $v_\pi(s^*)$ for any policy π, so we obtained the following contradiction:

$$v_{\pi^{'}}(s^*) > v_{\pi^{'}}(s^*)$$

Therefore, if a policy was optimal, then it would be optimal for all states in all steps. □

3.3.2. State-Updating Process

Based on the design of $\mathcal{S}$ and $\mathcal{A}$, new states could be observed after taking actions via $\mathcal{A}$ in both the MESS environment and the power-grid environment. The state-updating equations for time-step t were described separately for the three cases of charging, discharging, and holding, as follows:

$$\text{charge}: \begin{cases} P_{u,t^{'}} = P^{LP}_{u,t} \\ RE_{g,t^{'}} = \mathcal{M}_{re}(RE_{g,t}) + E_{g,t} \\ E_{g,t^{'}} = E_{g,t} + P_{g,t}\Delta h \\ Load_{g,t^{'}} = \mathcal{M}_{load}(Load_{g,t}) \\ g^{'} = Action(g) \end{cases},$$

$$\text{discharge}: \begin{cases} P_{u,t'} = P_{u,t}^{LP} \\ RE_{g,t'} = \mathcal{M}_{re}(RE_{g,t}) - E_{g,t} \\ E_{g,t'} = E_{g,t} - P_{g,t}\Delta h \\ Load_{g,t'} = \mathcal{M}_{load}(Load_{g,t}) \\ g^{'} = Action(g) \end{cases},$$

$$\text{hold}: \begin{cases} P_{u,t'} = P_{u,t}^{LP} \\ RE_{g,t'} = \mathcal{M}_{re}(RE_{g,t}) \\ E_{g,t'} = E_{g,t} \\ Load_{g,t'} = \mathcal{M}_{load}(Load_{g,t}) \\ g^{'} = Action(g) \end{cases},$$

where $\mathcal{M}_{re}$ and $\mathcal{M}_{load}$ are the matrices that hold the renewable energy sources and loads, respectively.

3.3.3. Proximal-Gradient Algorithm

In accordance with our previous discussion in Section 3.3.1, we studied the MESSs challenge under the CMDP framework. Several approaches were available for ensuring safe exploration by the agent, including the primal-dual algorithm, the adaptive-penalty method, and the gradient-projection algorithm. However, for the purposes of this study, we deemed the proximal-gradient algorithm to be both readily implementable and computationally efficient.

In the normal gradient-descent process, the iterative equation was the following:

$$x_{t+1} = x_t - \alpha_t \cdot f_t(x_t), \tag{26}$$

where α_t is the step length in step t. When the solution $\widetilde{x}_t$ exceeded its feasible domain, we needed to project it back into the feasible domain. Therefore, the iterative equation was transformed into the following:

$$x_t = \mathcal{P}_X(\widetilde{x}_t) := \arg\min\Big\{\|x - \widetilde{x}_t\|^2 : x \in X\Big\}, \tag{27}$$

where $\mathcal{P}_X$ is the projection operator, which was chosen as the Euclidean distance function in our study, and $\widetilde{x}_t := x_{t-1} - \gamma_n \hat{\nabla}_x f(x_{t-1})$ denoted the original infeasible solution.

3.4. Learning Process

As mentioned in Section 3.2, this proposal employed the "$actor - critic$" (A-C) framework to train the policy NN and value NN, respectively. To ensure the network was convergent and stable, we used target networks that were added to each NN of the A-C structure and added an experience replay pool, such as the double-DQN(DDQN) [42]. The learning process was also similar to the update process of the DDQN.

The required inputs for the algorithm included the initialized parameters such as exploration parameter ϵ, mini-batch size B, a probability distribution ξ, etc. The capacity N of the experience replay memory $\mathcal{D}$; the parameters of the *Actor* network Θ and the *Critic* network Q; and their target networks $\hat{\Theta}$ and $\hat{Q}$, respectively, also needed to be initialized.

At the beginning of each episode $i \in I$, where I is the maximum training time, the agent acquired an initial state by observation and computed continuous action parameters through network Θ: $x_k \leftarrow x_k(s_t, \theta_t)$. Then, it selected an action $a_t = \left(k_t, x_{k_t}\right)$, according to the ϵ-greedy policy:

$$a_t = \begin{cases} \text{a sample from distribution } \xi \text{ with probability } \epsilon \\ \left(k_t, x_{k_t}\right) \text{such that } k_t = \arg\max_{k \in K} Q(s_t, k, x_k; \omega_t) \text{ with probability } 1 - \epsilon \end{cases}$$

in each decision time-step t. Then, we performed the action in the MESSs and power grid environment to define the new state s_{t+1} and the reward r_t, which was obtained by solving the model (24).

Next, the transition tuple $[s_t, a_t, r_t, s_{t+1}]$ was stored in $\mathcal{D}$. Through sampling a batch $\{s_b, a_b, r_b, s_{b+1}\}_{b\in[B]}$ of size B, the target y_b could be calculated via equation

$$y_b = \begin{cases} r_b & \text{if } s_{b+1} \text{ is the terminal state,} \\ r_b + \max_{k\in K} \gamma Q(s_{b+1}, k, x_k(s_{b+1}, \theta_t); \omega_t) & \text{if otherwise.} \end{cases}$$

Then, the stochastic gradients $\nabla_\omega \ell_t^Q(\omega)$ and $\nabla_\theta \ell_t^\Theta(\theta)$ were calculated through $\{y_b, s_b, a_b\}_{b\in[B]}$, as shown in Equations (22) and (23). The weights were updated by $\theta_{t+1} \leftarrow \theta_t - \beta_t \nabla_\theta \ell_t^\Theta(\theta_t)$. In the above processes, the proximal gradient projection mentioned in Section 3.3.3 was applied. We used a soft update (Polyak averaging) method to update the target networks $\hat{Q} = (1-\tau)Q + \tau\hat{Q}$ and $\hat{\Theta} = (1-\tau)\Theta + \tau\hat{\Theta}$, where $\tau \in [0,1]$ is a hyper-parameter, usually $\tau \ll 1$.

4. Case Studies

In order to verify the effectiveness of the proposed strategy, case studies were implemented for the integrated model, utilizing simulation data. Our comparison of the energy output after incorporating MESSs demonstrated that the proposed strategy was effective in reducing the proportion of thermal power output and enhancing the utilization of renewable energy sources.

4.1. Experiment Settings

The training process for the 10-node scenario was approximately 5 h in length, on a desktop computer with an NVIDIA GTX 3080 GPU and an Intel i7-13700KF CPU, and approximately 6 h on a server with an NVIDIA Tesla P100.The solver for the MIP and LP models was Gurobi Optimizer, version 9.11.

The MESSs comprised an electric truck, a battery pack, and an operator. The unit movement cost c^{TRA} incorporated all expenses associated with the movement, including the charging costs of the electric truck and the operator's wages. The experimental arithmetic was represented by a topological diagram comprising 10 nodes that were interconnected by electrical conduits. The hyper-parameters of the experiment were subject to multiple simulations to find a balance between convergence speed and stability. The main simulation parameters are shown in Table 2.

Table 2. Experiment Parameters

Parameter Names	Value
Maximum capacity $\bar{E}$ (MWh)	27
Maximum charge/discharge power pwr_{MAX} (MW)	27
Maximum number of nodes n	10
Maximum number of thermal power units	3
Charging and discharging efficiency η	0.95
Transportation cost per unit of time c^{TRA} (\$/h)	20
Random exploration parameter ε	0.9–0.1
Batch size B	128
Discount factor γ	0.95
Probability distribution in random exploration ξ	U(0,1)
Soft updating parameter τ	0.1
Sizes of three layers of neural networks	[256,128,64]
Learning rate in policy network lr_p	1×10^{-6}
Learning rate in value network lr_v	1×10^{-4}

4.2. Results Analysis

Figure 2 depicts the convergence of the cumulative rewards achieved by the proposed DRL-based algorithm during the training phase. The x-axis represented the number of iteration rounds, with each increment on the axis representing 100 iterations. The y-axis displayed the accumulated cost, which was presented as a negative value. The graph demonstrated that the training reached convergence after approximately 1500 iterations. The training process was approximately 5 h for 7000 iterations, whereas the execution time of the trained network was virtually insignificant. By utilizing a trained network, we observed a significant improvement in computational efficiency.

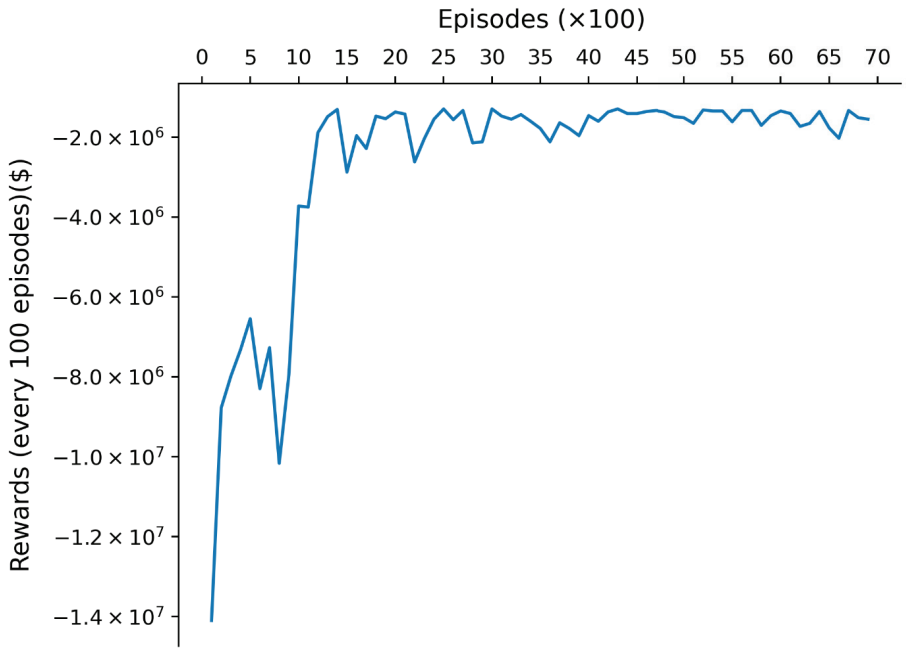
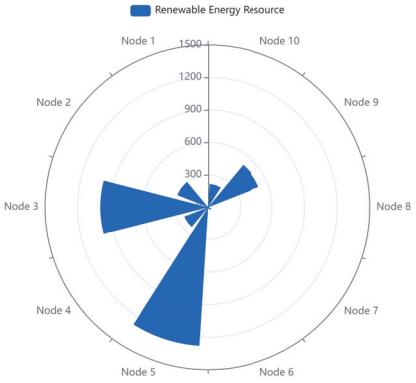


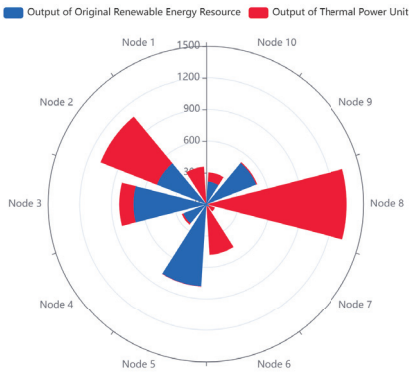
Figure 2. Episodic average reward in every 100 episodes in a scenario with 10 nodes.

Figure 3 shows the alteration in the distribution of renewable energy and thermal unit outputs within the system. As depicted in Figure 3a, certain nodes, such as nodes 3 and 8, exhibited a richer availability of renewable energy sources, whereas others, such as nodes 1, 6, and 7, exhibited virtually no presence of such resources. To cater to the electricity demand of these nodes with limited renewable energy, it was necessary to utilize thermal power units. Figure 3b presents a polar plot, demonstrating the distribution of renewable energy and thermal unit outputs in the absence of a connection to the MESS system. The figure clearly illustrates that some nodes, such as node 5, exhibited a predominant presence of renewable energy output, while others, such as node 8, were dominated by thermal power generation. Overall, the system exhibited a relatively substantial proportion of thermal power generation.

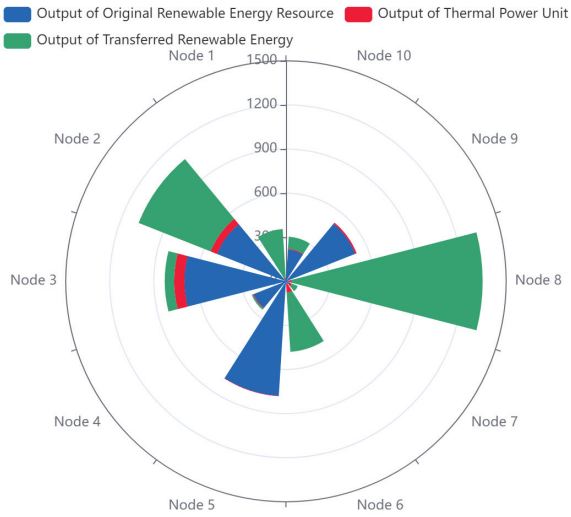
Figure 3c presents a polar plot that displays the optimal energy distribution among various nodes. The blue, red, and green segments in the plot depicted the local renewable energy sources, the energy generated by thermal units, and the energy transmitted through the MESSs, respectively. As shown in the plot, certain nodes, such as nodes 5 and 9, possessed abundant local renewable energy resources and, thus, exhibited a higher proportion of local energy output. These nodes were also the main charging nodes and were not typically involved in discharging the MESSs.Comparatively, nodes such as node 2 had both local renewable energy resources and energy transmitted by the MESSs contributing to their power supply. Furthermore, nodes such as node 8, which primarily relied on the energy transmitted by the MESSs, effectively reduced their share of the thermal power-unit output.



(**a**) Renewable Energy Resource at Each Node in a 10-Node Scenario



(**b**) Share of Energy Output of the Original System at Each Node in a 10-Node Scenario



(**c**) Share of Energy Output of the System after Integration of the MESSs at Each Node in a 10-Node Scenario

Figure 3. The Output of Thermal Power Units, Original Renewable Energy, and Transferred Renewable Energy at Each Node in a 10-Node Scenario.

Figure 4 presents a chord diagram that displays the cumulative energy transfer within the decision cycle of the MESSs. The diagram provided a visual representation of the flow of energy transfer and its accumulation during the decision cycle. The direction of the arrows represents the flow of energy transfer. As depicted in the figure, some nodes were designated as primary charging nodes, such as nodes 3 and 5, while others served as primary discharging nodes, such as nodes 1 and 8. This was because these nodes, such as 3 and 5, possessed abundant renewable energy sources, which was a result of their geographical locations and other contributing factors. However, when these nodes were not connected to the MESSs, the surplus of renewable energy resources could lead to issues such as excess light or wind. By incorporating these nodes into the MESSs,the proportion of thermal power output was significantly reduced, thus improving the utilization of renewable energy sources. Some nodes, such as nodes 3 and 8, served as both supply and demand nodes, which was due to the intermittent nature of renewable energy and the mismatch between the supply of renewable energy and the peak demand for electricity. For example, the power generation capacity of PV was related to factors such as sunlight duration and weather and reached its peak output during sunny afternoons, but the local demand may not be able to consume all the generated electricity during this time period.

Therefore, the excess electricity would then be transported to nearby nodes to meet the electricity demand of other nodes. In the evening, when the power output of PV decreased, a hydroelectric power supply from neighboring nodes could be required.



Figure 4. Optimal cumulative total renewable energy transfer in a 10-node scenario.

By conducting numerous experiments, we demonstrated that by integrating the MESSs into the power grid, it was feasible to transfer surplus renewable energy from nodes with an abundance of such resources, to nodes with limited resources. This resulted in a reduction in thermal power generation in the latter, thereby validating the proposed model and algorithm.

Furthermore, our proposed mobile energy storage system-grid joint optimization model has the potential to address network contingency events. First, our study on coupling mobile energy storage systems with the power grid contributes to achieving higher renewable energy penetration levels [43,44]. In the event of network contingencies (e.g., equipment failure, fluctuations in power demand, etc.), mobile energy storage systems can quickly adjust their output to balance supply and demand imbalances in the grid. Moreover, mobile energy storage systems can transfer electrical energy from one location to another within the grid in a short period of time, alleviating local transmission congestion issues and improving the stability of the entire grid. Second, by optimizing the dispatch strategy of mobile energy storage systems, we can allocate renewable energy resources reasonably across different time periods to meet grid demand. This will help mitigate power supply shortages or overloads caused by contingency events, further enhancing the stability and reliability of the grid [29]. Lastly, our research also indicates that mobile energy storage systems can help reduce the operational costs of the entire power system [45]. By utilizing renewable energy resources more efficiently, we can decrease our reliance on conventional fossil fuel-based generation, thus lowering the cost losses associated with network contingencies.

5. Conclusions

In this study, we developed a mixed-integer nonlinear programming (MINP) model that coupled MESSs with a power grid to balance a region with an uneven distribution of renewable energy. We modeled the MESSs as constrained Markov decision processes (CMDPs) and proposed a framework based on a deep reinforcement learning (DRL) algorithm that considered the discrete-continuous hybrid action space of the MESSs. We solved the constraint on battery capacity in the CMDP by applying a proximal-gradient-projection algorithm. Based on this, we reformulated the original MINP model as a linear programming (LP) model to arrive at our solution framework. The case study showed that our proposed algorithm and framework effectively improved the utilization of renewable energy and reduced the generation costs of thermal power units.

The DRL-based algorithm we proposed has good scalability and application prospects in energy challenges with discrete-continuous hybrid decision-action spaces. Furthermore, our approach of using a DRL-based method to eliminate integer decision variables and nonlinear constraints provided insight for solving the MINP challenges. However, to highlight the main innovation points of this paper, we simplified some constraints, such as overlooking power-flow constraints and the aging characteristics of batteries. In future research, we will consider these additional constraints and objective functions more carefully, including the initial investment cost for implementing MESSs, and attempt to optimize scheduling decisions over longer time periods.

Author Contributions: Conceptualization, Y.D. and J.W.; methodology, Y.D. and J.W.; software, Y.D.; validation, Y.D.; formal analysis, Y.D., X.C. and J.W.; investigation, Y.D. and X.C.; resources, Y.D.; data curation, X.C.; writing—original draft preparation, Y.D.; writing—review and editing, Y.D., X.C. and J.W.; visualization, Y.D.; supervision, J.W.; project administration, Y.D. and J.W.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported, in part, by the National Key Research and Development Program, under Grant 2022YFB2405600; and, in part, by the National Natural Science Foundation of China, under Grants 72131001, T2121002, and 52277092.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the High-performance Computing Platform of Peking University.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MESSs	Mobile Energy Storage Systems
WP	Wind Power
PV	Photovoltaic
MIP	Mixed-Integer Programming
LP	Linear Programming
DRL	Deep Reinforcement Learning
ESSs	Energy Storage Systems
SESS	Stationary Energy Storage System
EV	Electric Vehicle
MILP	Mixed-Integer Linear Programming
MDP	Markov Decision Process
CMDP	Constrained Markov Decision Process
A-C	Actor-Critic
NNs	Neural Networks

References

1. Olabi, A.; Abdelkareem, M.A. Renewable energy and climate change. *Renew. Sustain. Energy Rev.* **2022**, *158*, 112111.
2. Elavarasan, R.M.; Shafiullah, G.; Padmanaban, S.; Kumar, N.M.; Annam, A.; Vetrichelvan, A.M.; Mihet-Popa, L.; Holm-Nielsen, J.B. A comprehensive review on renewable energy development, challenges, and policies of leading Indian states with an international perspective. *IEEE Access* **2020**, *8*, 74432–74457.
3. Qazi, A.; Hussain, F.; Rahim, N.A.; Hardaker, G.; Alghazzawi, D.; Shaban, K.; Haruna, K. Towards sustainable energy: A systematic review of renewable energy sources, technologies, and public opinions. *IEEE Access* **2019**, *7*, 63837–63851.
4. Cole, W.J.; Greer, D.; Denholm, P.; Frazier, A.W.; Machen, S.; Mai, T.; Vincent, N.; Baldwin, S.F. Quantifying the challenge of reaching a 100% renewable energy power system for the United States. *Joule* **2021**, *5*, 1732–1748.
5. Impram, S.; Nese, S.V.; Oral, B. Challenges of renewable energy penetration on power system flexibility: A survey. *Energy Strategy Rev.* **2020**, *31*, 100539.
6. Kebede, A.A.; Kalogiannis, T.; Van Mierlo, J.; Berecibar, M. A comprehensive review of stationary energy storage devices for large scale renewable energy sources grid integration. *Renew. Sustain. Energy Rev.* **2022**, *159*, 112213.
7. Park, S.; Ahn, J.; Kang, T.; Park, S.; Kim, Y.; Cho, I.; Kim, J. Review of state-of-the-art battery state estimation technologies for battery management systems of stationary energy storage systems. *J. Power Electron.* **2020**, *20*, 1526–1540.
8. Maestre, V.; Ortiz, A.; Ortiz, I. Challenges and prospects of renewable hydrogen-based strategies for full decarbonization of stationary power applications. *Renew. Sustain. Energy Rev.* **2021**, *152*, 111628.
9. Saboori, H.; Jadid, S. Optimal scheduling of mobile utility-scale battery energy storage systems in electric power distribution networks. *J. Energy Storage* **2020**, *31*, 101615.
10. Kucevic, D.; Englberger, S.; Sharma, A.; Trivedi, A.; Tepe, B.; Schachler, B.; Hesse, H.; Srinivasan, D.; Jossen, A. Reducing grid peak load through the coordinated control of battery energy storage systems located at electric vehicle charging parks. *Appl. Energy* **2021**, *295*, 116936.
11. He, G.; Michalek, J.; Kar, S.; Chen, Q.; Zhang, D.; Whitacre, J.F. Utility-scale portable energy storage systems. *Joule* **2021**, *5*, 379–392.
12. He, G.; Chen, X.; Yang, Y.; Wang, J.; Song, J. Hybrid Portable and Stationary Energy Storage Systems with Battery Charging and Swapping Coordination. In Proceedings of the 2022 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), Shanghai, China, 8–11 July 2022; pp. 1465–1470.
13. Abdeltawab, H.H.; Mohamed, Y.A.R.I. Mobile energy storage scheduling and operation in active distribution systems. *IEEE Trans. Ind. Electron.* **2017**, *64*, 6828–6840.
14. Nazemi, M.; Dehghanian, P.; Lu, X.; Chen, C. Uncertainty-aware deployment of mobile energy storage systems for distribution grid resilience. *IEEE Trans. Smart Grid* **2021**, *12*, 3200–3214.
15. Ebadi, R.; Yazdankhah, A.S.; Kazemzadeh, R.; Mohammadi-Ivatloo, B. Techno-economic evaluation of transportable battery energy storage in robust day-ahead scheduling of integrated power and railway transportation networks. *Int. J. Electr. Power Energy Syst.* **2021**, *126*, 106606.
16. Sexauer, J.M.; Mohagheghi, S. Voltage quality assessment in a distribution system with distributed generation—A probabilistic load flow approach. *IEEE Trans. Power Deliv.* **2013**, *28*, 1652–1662.
17. Eyer, J.; Corey, G. Energy storage for the electricity grid: Benefits and market potential assessment guide. *Sandia Natl. Lab.* **2010**, *20*, 5.
18. Abolhassani, M.H.; Safdarian, A. Electric Vehicles as Mobile Energy Storage Devices to Alleviate Network Congestion. In Proceedings of the 2019 Smart Grid Conference (SGC), Tehran, Iran, 18–19 December 2019; pp. 1–5.
19. Song, J.; He, G.; Wang, J.; Zhang, P. Shaping future low-carbon energy and transportation systems: Digital technologies and applications. *iEnergy* **2022**, *1*, 285–305.
20. Pulazza, G.; Zhang, N.; Kang, C.; Nucci, C.A. Transmission planning with battery-based energy storage transportation for power systems with high penetration of renewable energy. *IEEE Trans. Power Syst.* **2021**, *36*, 4928–4940.
21. Sun, Y.; Li, Z.; Shahidehpour, M.; Ai, B. Battery-based energy storage transportation for enhancing power system economics and security. *IEEE Trans. Smart Grid* **2015**, *6*, 2395–2402.
22. Kwon, S.Y.; Park, J.Y.; Kim, Y.J. Optimal V2G and route scheduling of mobile energy storage devices using a linear transit model to reduce electricity and transportation energy losses. *IEEE Trans. Ind. Appl.* **2019**, *56*, 34–47.
23. Yan, G.; Liu, D.; Li, J.; Mu, G. A cost accounting method of the Li-ion battery energy storage system for frequency regulation considering the effect of life degradation. *Prot. Control. Mod. Power Syst.* **2018**, *3*, 1–9.
24. Scarabaggio, P.; Carli, R.; Cavone, G.; Dotoli, M. Smart control strategies for primary frequency regulation through electric vehicles: A battery degradation perspective. *Energies* **2020**, *13*, 4586.
25. Li, W.; Cui, H.; Nemeth, T.; Jansen, J.; Uenluebayir, C.; Wei, Z.; Zhang, L.; Wang, Z.; Ruan, J.; Dai, H.; et al. Deep reinforcement learning-based energy management of hybrid battery systems in electric vehicles. *J. Energy Storage* **2021**, *36*, 102355.
26. Liang, Y.; Ding, Z.; Zhao, T.; Lee, W.J. Real-time operation management for battery swapping-charging system via multi-agent deep reinforcement learning. *IEEE Trans. Smart Grid* **2022**, *14*, 559–571.
27. Kofinas, P.; Dounis, A.; Vouros, G. Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids. *Appl. Energy* **2018**, *219*, 53–67.

28. Hua, H.; Qin, Y.; Hao, C.; Cao, J. Optimal energy management strategies for energy Internet via deep reinforcement learning approach. *Appl. Energy* **2019**, *239*, 598–609.
29. Wang, Y.; Qiu, D.; Strbac, G. Multi-agent deep reinforcement learning for resilience-driven routing and scheduling of mobile energy storage systems. *Appl. Energy* **2022**, *310*, 118575.
30. Yao, S.; Gu, J.; Zhang, H.; Wang, P.; Liu, X.; Zhao, T. Resilient load restoration in microgrids considering mobile energy storage fleets: A deep reinforcement learning approach. In Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2–6 August 2020; pp. 1–5.
31. Kebede, A.A.; Coosemans, T.; Messagie, M.; Jemal, T.; Behabtu, H.A.; Van Mierlo, J.; Berecibar, M. Techno-economic analysis of lithium-ion and lead-acid batteries in stationary energy storage application. *J. Energy Storage* **2021**, *40*, 102748.
32. Yu, Z.; Dou, Z.; Zhao, Y.; Xie, R.; Qiao, M.; Wang, Y.; Liu, L. Grid Scheduling Strategy Considering Electric Vehicles Participating in Multi-microgrid Interaction. *J. Electr. Eng. Technol.* **2022**, 1–16. https://doi.org/10.1007/s42835-022-01294-x.
33. Bellman, R. A Markovian decision process. *J. Math. Mech.* **1957**, *6*, 679–684.
34. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv* **2013**, arXiv:1312.5602.
35. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533.
36. Xiong, J.; Wang, Q.; Yang, Z.; Sun, P.; Han, L.; Zheng, Y.; Fu, H.; Zhang, T.; Liu, J.; Liu, H. Parametrized deep q-networks learning: Reinforcement learning with discrete-continuous hybrid action space. *arXiv* **2018**, arXiv:1810.06394.
37. Silver, D.; Lever, G.; Heess, N.; Degris, T.; Wierstra, D.; Riedmiller, M. Deterministic policy gradient algorithms. In Proceedings of the International Conference on Machine Learning, Beijing, China, 21–26 June 2014; pp. 387–395.
38. Sutton, R.S.; McAllester, D.; Singh, S.; Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. *Adv. Neural Inf. Process. Syst.* **1999**, *12*, 1–7.
39. Silver, D.; Schrittwieser, J.; Simonyan, K.; Antonoglou, I.; Huang, A.; Guez, A.; Hubert, T.; Baker, L.; Lai, M.; Bolton, A.; et al. Mastering the game of go without human knowledge. *Nature* **2017**, *550*, 354–359.
40. Bester, C.J.; James, S.D.; Konidaris, G.D. Multi-pass q-networks for deep reinforcement learning with parameterised action spaces. *arXiv* **2019**, arXiv:1905.04388.
41. Bengio, Y.; Lodi, A.; Prouvost, A. Machine learning for combinatorial optimization: A methodological tour d'horizon. *Eur. J. Oper. Res.* **2021**, *290*, 405–421.
42. Van Hasselt, H.; Guez, A.; Silver, D. Deep reinforcement learning with double q-learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; Volume 30.
43. Dugan, J.; Mohagheghi, S.; Kroposki, B. Application of mobile energy storage for enhancing power grid resilience: A review. *Energies* **2021**, *14*, 6476.
44. Wang, Y.; Rousis, A.O.; Strbac, G. Resilience-driven optimal sizing and pre-positioning of mobile energy storage systems in decentralized networked microgrids. *Appl. Energy* **2022**, *305*, 117921.
45. Ahmadi, S.E.; Marzband, M.; Ikpehai, A.; Abusorrah, A. Optimal stochastic scheduling of plug-in electric vehicles as mobile energy storage systems for resilience enhancement of multi-agent multi-energy networked microgrids. *J. Energy Storage* **2022**, *55*, 105566.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Batteries Editorial Office
E-mail: batteries@mdpi.com
www.mdpi.com/journal/batteries



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.





Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-8426-1