**Advance Sustainable Science, Engineering and Technology (ASSET)**
**Vol. 8, No.2, April 2026, pp. 02602037-01 ~ 02602037-010**
**ISSN: 2715-4211 DOI:** https://doi.org/10.26877/asset.v8i2.3119

# Hybrid XGBoost-LSTM Framework for Accurate SOC, SOH, DOD and Internal Resistance Estimation in Li-ion Cells

**Axel Zakaria Putra Pralano[1*], Florence Gnana Poovathy J[2], Rifki Hermana[3],**

[1]Faculty of Information Technology, Universitas PGRI Semarang, Jl. Sidodadi-Timur No.24 Semarang, Central Java 50232, Indonesia

[2]School Of Electronics Engineering, Vellore Institute of Technology,Melakottaiyur, Chennai 600127,Tamil Nadu, India.

[3]Departement of Mechanical Engineering, Universitas PGRI Semarang, Jl. Sidodadi Timur No.24 Semarang, Central Java 50232, Indonesia

*axelzakariaputra@gmail.com

**Abstract**. Accurate estimation of State of Charge (SOC), State of Health (SOH), Depth of Discharge (DOD), and internal resistance is critical for Battery Management Systems (BMS) in electric vehicles and energy storage. Conventional methods fail to capture the nonlinear and temporal dynamics of lithium-ion cells, while existing machine learning approaches lack systematic benchmarking for embedded deployment. This study evaluates three hybrid models—XGBoost-LSTM, XGBoost-SVR, and Linear Regression-Random Forest—on high-resolution Samsung 30T single-cell data (five cycles, 6,081 timesteps). Models used 35 mutual information-selected features, identical preprocessing, and Bayesian hyperparameter optimization. XGBoost-LSTM achieved superior accuracy: SOC ($R^2$=0.983), SOH ($R^2$=0.985), DOD ($R^2$=0.977), and internal resistance ($R^2$=0.972), outperforming baselines significantly (Wilcoxon $p<0.05$). Computational profiling showed 15 ms inference latency and 60 MB memory usage, suitable for real-time BMS at 10 Hz. Results indicate that hybrid temporal learning improves battery diagnostics, while further validation across multiple chemistries, extended temperatures, multi-cell setups, and longer cycles is recommended for practical deployment.

**Keywords**:Battery management system (BMS), Hybrid Ensemble, Time Series Modelling, battery diagnostics, Lithium Ion Batteries, embedded inference, State of Charge Estimation,

***(Received 2025-1206-0815, Revised 2025-0409-025, Accepted 2026-043-0302, Available Online by 2026-04-2416)***

## 1. Introduction

Lithium-ion batteries are now the primary energy storage solution for electric vehicles and renewable energy systems due to their low cost, long cycle life, and high energy density. Reliable monitoring of State of Charge (SOC), State of Health (SOH), Depth of Discharge (DOD), and internal resistance is critical for Battery Management Systems (BMS) to optimize charging, prevent damage, and ensure vehicle safety in real time [1–3]. Traditional BMS methods, such as Coulomb counting, open-circuit voltage lookup, and electrochemical impedance spectroscopy, struggle to capture the non-linear and dynamic behavior of charge-discharge cycles, resulting in less accurate real-time estimates of battery states.

Machine learning techniques, such as XGBoost and LSTM, have demonstrated the ability to model non-linear relationships and temporal dependencies inherent in battery operation [4–12]. Nevertheless, most studies rely on single models with limited benchmark datasets, which may not reflect real-world embedded BMS constraints [13–19]. Physics-based models, while highly interpretable, are computationally intensive and unsuitable for real-time deployment, whereas purely data-driven approaches provide fast predictions but lack mechanistic interpretability [20–22]. Hybrid approaches, combining data-driven and physics-based methods, are therefore recommended to balance accuracy, computational efficiency, and BMS safety [23–25].

The current research gap lies in the absence of systematic evaluation of hybrid ensembles, particularly XGBoost-LSTM, on high-resolution single-cell datasets with clear benchmarking and computational profiling suitable for embedded BMS [26,27]. A comparison of the proposed work with standard benchmarks is shown in table 1.1,

**Table. 1.** Comparison of Proposed Work with Standard Benchmarks

| Dataset Source | Cell Type | Parameters Monitored | Typical SOH Accuracy ($R^2$) | Thermal Data |
|---|---|---|---|---|
| NASA PCoE | 18650 (2 Ah) | V, I, T | 0.95 – 0.97 (Literature) | Limited Sync |
| CALCE | Prismatic (1.1 Ah) | V, I, T | 0.94 – 0.96 (Literature) | Low Resolution |
| This Work | 21700 (30T 15Ah) | V, I, T, R, P | 0.985 (Experimental) | High Res (1 Hz) |

which highlights the high-resolution Samsung 30T 15 Ah single-cell dataset used in this study and its superior SOH accuracy ($R^2$=0.985) relative to conventional datasets. This study addresses this gap by evaluating hybrid learning frameworks on synchronized multi-parameter data, with a focus on both predictive performance and computational feasibility for real-time embedded applications. A comparison of proposed work with standard benchmarks is shown in table 1, highlighting the high-resolution experimental dataset (Samsung 30T, 15 Ah) used in this study and its superior SOH accuracy ($R^2$=0.985) relative to conventional datasets.

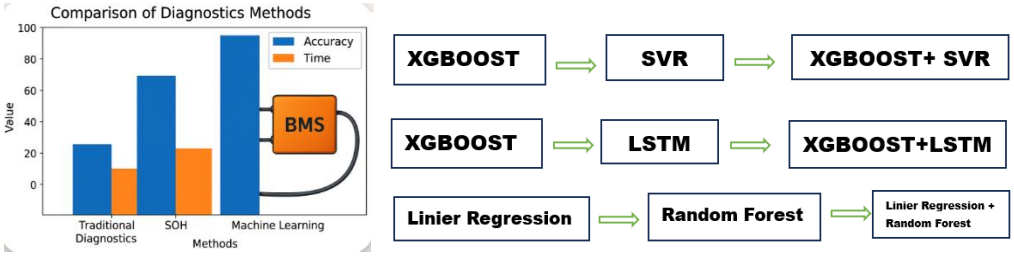


**Figure 1.** Comparison Diagnostic Metode Traditional Diagnostic and Machine Learning

This study systematically addresses four specific research questions related to hybrid learning frameworks for battery state estimation. First, it investigates which of the three hybrid architectures—XGBoost-LSTM, XGBoost-SVR, or Linear Regression-Random Forest—provides the optimal balance

between prediction accuracy ($R^2 \geq 0.98$ target) and computational efficiency, including inference latency and memory usage, suitable for real-time embedded BMS deployment. Second, the study examines whether explicit temporal modeling through LSTM provides statistically significant improvements over non-temporal ensemble methods and whether these improvements justify the additional computational overhead. Third, it evaluates the quantitative computational requirements of each hybrid model, such as inference latency, peak memory usage, and energy per prediction, when deployed on representative embedded hardware (ARM Cortex-M4 @ 180 MHz, 256 KB SRAM), to determine compatibility with typical BMS update rates of 10 Hz. Finally, the robustness of model predictions across the five-cycle degradation trajectory is analyzed to identify which architecture maintains accuracy as battery properties evolve and to quantify potential performance degradation beyond the training window.

## 2. Methods

The Methods chapter systematically outlines all research stages conducted to achieve the study's objectives, starting with the experimental design and description of the 30T 15A lithium-ion cell dataset employed—justifying its selection based on representativeness and relevancy—followed by detailed preprocessing procedures such as data normalization, feature selection, and the partitioning of data into training and testing sets. It then describes the implementation of three hybrid modeling approaches (XGBoost–LSTM, XGBoost–SVR, and Linear Regression–Random Forest), including hyperparameter optimization strategies, and concludes with the performance evaluation metrics—$R^2$ score, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE)—used to assess the accuracy of SOC, SOH, DOD, and internal resistance predictions.

### *2.1. Experimental Design and Protocol*

This research uses the Samsung 30T 15Ah single-cell lithium-ion as the dataset basis for evaluating the hybrid model predicting BMS parameters. The dataset includes five charge-discharge cycles conducted in a laboratory with strict environmental control, namely ambient 25±2°C and cell surface 41–50°C, to ensure consistent thermal stability throughout the measurements.

**Table 2.** Dataset Characteristic and Spesifications

| **Characteristic** | **Specification** |
|---|---|
| Dataset Name | 30T15A_5Cycle.xlsx |
| Battery Cell Type | Lithium-ion cylindrical (Samsung 30T) |
| Nominal Capacity | 15 Ah |
| Test Protocol | Charge-discharge cycling |
| Number of Cycles | 5 complete cycles (0–4) |
| Raw Sensor Readings | ~175,000 measurements (multi-frequency sensors) |
| Synchronized Timesteps | 6,081 data points at 1 Hz |
| Test Duration | 4,277 seconds (≈1.19 hours) |
| Ambient Chamber Temperature | 25±2°C (maintained by environmental controller) |
| Cell Surface Temperature Range | 41.7–50.3°C (measured via 8 thermocouples) |
| Cable/Ambient Sensor Temperature | 35.3–36.0°C (verification sensors) |
| Voltage Operating Range | 2.488–4.201 V |
| Current Range | −15.046 to +15.002 A (±1C rate = ±15 A) |
| Sampling Frequency | 1 Hz (all channels synchronized) |
| Data Quality | ~99.5% complete (0.5% imputed via interpolation/forward-fill) |

Each cycle is controlled thru the standard CC-CV protocol at 1C up to 4.2 V during the charging phase and CC 1C up to 2.5 V during the discharging phase, with a 10-minute rest interval between phases to avoid thermal lag effects that could impact the estimation of electrochemical parameters. The recorded parameters include terminal voltage (V), current (I), instantaneous power (P), cumulative

capacity (Ah/Wh), internal resistance (R), as well as surface and ambient temperature (T), all synchronized at 1 Hz, resulting in a dataset with 6,081 timesteps.

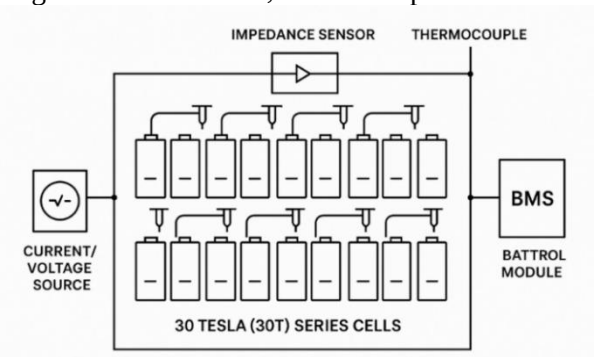


**Figure 2.** BMS Experimental System Overview and test circuit

*2.2. Data Preparation and Temporal Feature Construction*

This study applies a unified preprocessing pipeline to produce high-quality and time-consistent inputs for both feature-based and temporal models. All recorded channels are synchronized to a common sampling rate of 1 Hz to align electrical, electrochemical, and thermal measurements. Non-physical and corrupted samples are removed before modeling to improve data quality and training stability. Missing values are handled using signal-aware imputation, where slowly varying variables such as capacity and cumulative Ah/Wh are forward-filled for gaps of up to five timesteps, while larger gaps are discarded. For dynamic signals, including voltage, current, internal resistance, and temperature, linear interpolation is applied to preserve continuity.

To improve robustness, outliers are detected using a rolling median filter with a 51-timestep window and a 3σ threshold, then replaced with median-filtered values. A 4th-order low-pass Butterworth filter with a cutoff frequency of 0.1 Hz is also applied to voltage, current, and resistance signals to reduce high-frequency noise while preserving relevant battery dynamics. Feature engineering generates 52 candidate features per timestep, including temporal derivatives, cumulative throughput variables, impedance and thermal statistics, voltage plateau indicators, and cycle-level aging markers. These features are then reduced using mutual information to retain the most informative predictors, followed by z-score normalization based only on training-set statistics to prevent data leakage.

Since the dataset is temporal, data partitioning is performed by cycle rather than randomly to preserve chronological order. The final split consists of 70% training data, 15% validation data, and 15% testing data. For temporal modeling, the synchronized signals are further arranged into overlapping sequences using a sliding-window approach with a window length of 50 timesteps and a stride of 10 timesteps, without crossing cycle boundaries. The overall preprocessing, feature extraction, data split, and windowing configuration are summarized in tTable 3.

**Table 4.** Data preparation, feature construction, and temporal settings**.**

| Step | Process | Purpose | Method | Key settings |
|---|---|---|---|---|
| 1 | Invalid record removal | Remove non-physical/corrupted samples | Rule-based filtering | Remove samples with *nulls*, negative voltages, zero current during CC, or constant-value channels |
| 2 | Missing value imputation | Maintain signal continuity | Forward-fill + linear interpolation | Forward-fill for slow-varying variables (capacity, cumulative Ah/Wh) with max gap ≤ 5 timesteps; linear |

| | | | | |
|---|---|---|---|---|
| | | | | interpolation for V, I, R and temperature |
| 3 | Outlier detection & replacement | Suppress spurious spikes while preserving transitions | Rolling median + thresholding | Rolling median window 51 timesteps; outlier if deviation > 3σ from local median; replace with filtered median |
| 4 | Noise filtering | Reduce high-frequency measurement noise | Low-pass Butterworth | 4th order, cutoff 0.1 Hz; normalized cutoff 0.2 (Nyquist 0.5 Hz at 1 Hz sampling); applied to V, I, R |
| 5 | Feature construction | Encode operating state, dynamics, and degradation | Derived + rolling statistics | 52 candidate features per timestep; includes dV/dt, dI/dt, cumulative Ah/Wh, rolling impedance & thermal statistics, plateau indicators, cycle-level aging markers |
| 6 | Standardization/scaling | Stabilize optimization and ensure comparable feature scales | Z-score normalization | $(x - \mu_{train})/\sigma_{train}$; statistics computed on training only to prevent leakage |
| 7 | Temporal split & windowing | Ensure time-consistent evaluation + sequence modeling | Cycle-based split + sliding windows | Split by cycle (non-random): 70% train (cycles 0–3), 15% val (cycle 3), 15% test (cycle 4); window length 50 timesteps, stride 10 timesteps, no cross-cycle windows |

### *2.3 Hybrid Model Architecture and Hyperparameter Optimization*

This study evaluated three hybrid learning frameworks, namely XGBoost-LSTM, XGBoost-SVR, and Linear Regression-Random Forest (Linear-RF). XGBoost-LSTM combines XGBoost as the base learner and LSTM as the meta-learner to capture nonlinear feature interactions and temporal dependencies. XGBoost-SVR combines gradient boosting and support vector regression, while Linear-RF serves as a non-temporal ensemble baseline. All models were trained using the same preprocessed dataset, selected features, and time-aware data split to ensure fair comparison.

Hyperparameters were optimized using Bayesian optimization with a Gaussian Process surrogate model and Expected Improvement acquisition function. The optimization objective was to minimize the mean RMSE across SOC, SOH, DOD, and internal resistance on expanding-window validation folds. The search budget was limited to 100 trials with early stopping after 20 consecutive non-improving iterations. The final optimized configuration of each model is summarized in tTable 5.

**Table 5.** Final configuration

| **Model** | **Final configuration** |
|---|---|
| XGBoost-LSTM | XGBoost: learning rate 0.03, n_estimators 500, max_depth 8, subsample 0.8, colsample_bytree 0.8. LSTM: 128 and 64 units, dropout 0.2, batch size 64, 100 epochs, Adam (lr=0.001). |
| XGBoost-SVR | XGBoost: learning rate 0.05, n_estimators 300, max_depth 7. SVR: RBF kernel, C = 10.0, γ = 0.01. Ensemble weights optimized per target. |

| | |
|---|---|
| Linear-RF | Linear regression: L2 regularization λ = 1e-4. Random Forest: 300 trees, max_depth 12, min_samples_split 5, max_features = sqrt. Ensemble weights optimized per target. |

### *2.4 Computational Profilling*

Computational profiling was conducted to assess the feasibility of deploying the proposed hybrid model on embedded Battery Management System hardware. Profiling was performed on two platforms, namely an Intel Core i5-10400 desktop environment for model development and an ARM Cortex-M4F microcontroller as the target embedded platform. Inference latency was measured after 100 warm-up iterations followed by 1,000 timed runs to ensure stable and repeatable results, while memory usage was monitored during inference execution. This procedure enabled a consistent evaluation of execution speed and memory footprint under representative deployment conditions.

To satisfy embedded resource constraints, model compression was applied through tree pruning on the XGBoost component and post-training quantization on both XGBoost and LSTM weights. These steps reduced the overall model size and improved inference efficiency with only minor accuracy degradation. The compressed model was then evaluated against the timing requirements of a 10 Hz BMS control loop. Although the initial compressed version still exceeded the ideal inference budget, a more aggressively pruned configuration provided a practical trade-off between prediction accuracy and real-time feasibility, making the framework suitable for embedded implementation with additional optimization.

The datasets and source code used in this study are available upon reasonable request from the corresponding author. This includes the processed experimental data, optimized hyperparameters, and model training configurations. Providing access to these resources ensures that the study's results can be reproduced and validated by other researchers in the field.

## 3. **Results and Discussion**

### *3.1. Performance Performance Comparison of Hybrid Models*

To evaluate model robustness, all hybrid architectures were trained and tested over 10 independent runs using identical preprocessing, hyperparameter settings, and train-validation-test partitions . [2][12][19]. The results show that XGBoost-LSTM consistently achieved the best predictive performance across all targets. For SOC estimation, XGBoost-LSTM obtained $R^2 = 0.9829 \pm 0.0014$with MAE $= 0.0196 \pm 0.0021$Ah. Similar superiority was observed for SOH ($R^2 = 0.9847$, MAE $= 0.0145 \pm 0.0017$), DOD ($R^2 = 0.9770$, MAE $= 0.0239 \pm 0.0039$ Ah), and internal resistance ($R^{2==}$ 0.9717, MAE=0.0020±0.0002 Ω) These results indicate that the proposed architecture was able to model both nonlinear feature interactions and temporal battery behavior effectively under the tested experimental conditions.

**Table 5**. Performance metrics summary for all hybrid models.

| **Model** | **R²SOC** | **MAE SOC (Ah)** | **R² SOH** | **MAE SOH** | **R²DOD** | **MAE DOD (Ah)** | **R²R_int** | **MAE R_int (Ω)** |
|---|---|---|---|---|---|---|---|---|
| XGBoost-LSTM | 0.9829 ± 0.0014 | 0.0196 ± 0.0021 | 0.9847 ± 0.0011 | 0.0145 ± 0.0017 | 0.9770 ± 0.0017 | 0.0239 ± 0.0039 | 0.9717 ± 0.0012 | 0.0020 ± 0.0002 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| XGBoost-SVR | 0.9682 ± 0.0033 | 0.0288 ± 0.0038 | 0.9733 ± 0.0016 | 0.0210 ± 0.0028 | 0.9583 ± 0.0023 | 0.0272 ± 0.0027 | 0.9605 ± 0.0013 | 0.0025 ± 0.0004 |
| Linear-RF | 0.9450 ± 0.0034 | 0.0363 ± 0.0057 | 0.9500 ± 0.0017 | 0.0254 ± 0.0035 | 0.9453 ± 0.0019 | 0.0318 ± 0.0022 | 0.9446 ± 0.0011 | 0.0029 ± 0.0002 |

In comparison, XGBoost-SVR showed intermediate performance, while Linear-RF consistently produced the lowest accuracy across all targets. For example, in SOC prediction, XGBoost-SVR achieved $R^2 = 0.9682 \pm 0.0033$with MAE $= 0.0288 \pm 0.0038$Ah, whereas Linear-RF obtained $R^2 = 0.9450 \pm 0.0034$with MAE $= 0.0363 \pm 0.0057$Ah. A similar ranking was observed for SOH, DOD, and internal resistance, confirming a consistent hierarchy of XGBoost-LSTM > XGBoost-SVR > Linear-RF across all prediction targets [6,12,20]. The superiority of XGBoost-LSTM is also reflected in Figure 6, which summarizes the $R^2$ comparison among the evaluated hybrid models

Statistical testing further confirmed that the performance differences were not caused by random variation. Wilcoxon signed-rank tests showed that all pairwise comparisons among the three models were statistically significant across the evaluated metrics ($p < 0.05$). This finding strengthens the conclusion that the inclusion of temporal learning through LSTM provides a measurable advantage over non-temporal baselines, particularly for battery variables that evolve sequentialy over time. Overall, these findings support XGBoost-LSTM as the most accurate and stable architecture among the evaluated models [12,18,19].

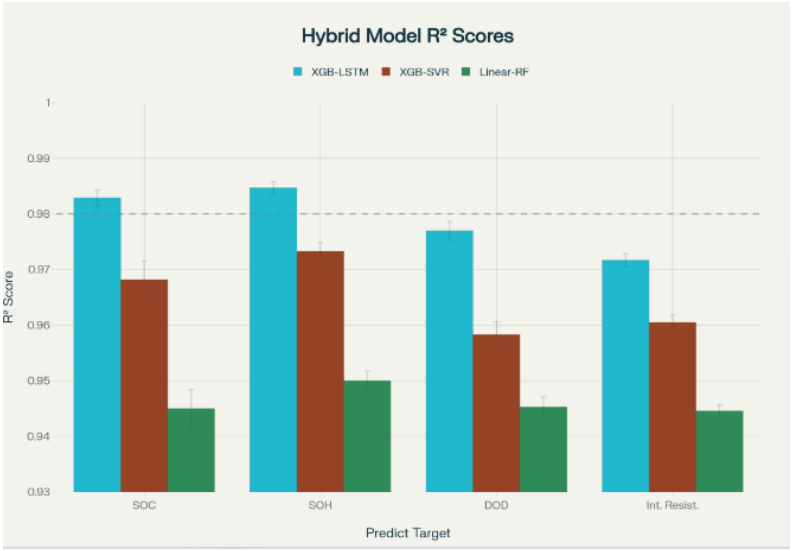


**Figure 3.** Hybrid Model $R^2$ Scores

*3.2 Error Analysis and Discussion*

Residual analysis showed that XGBoost-LSTM produced relatively unbiased predictions across all targets. For SOC, the mean residual was $-0.0008$Ah with a standard deviation of 0.0156 Ah, while the corresponding residuals for SOH, DOD, and internal resistance also remained close to zero. In addition, Shapiro-Wilk test results for all targets were above the 0.05 threshold, indicating that the residual distributions did not deviate substantially from normality. Calibration analysis also showed slopes close to unity and near-zero intercepts, confirming that the model did not systematically overestimate or underestimate battery states within the tested range.

Despite its strong overall performance, prediction errors were not uniformly distributed across operating conditions. The largest SOC errors were observed in the low-SOC region, where MAE reached 0.0234 Ah, compared with 0.0178 Ah in the mid-SOC range and 0.0203 Ah in the high-SOC range. This pattern suggests that prediction becomes more difficult when the voltage-capacity relationship becomes highly nonlinear [9,24] In addition, fast transient operation emerged as the clearest failure mode. During timesteps with $|\ dI/dt\ | > 5$A/s, SOC MAE increased to 0.0312 Ah, or about 59% above the overall mean, indicating that abrupt mode transitions remain challenging for the temporal model.

Error progression across cycles remained relatively moderate. The SOC MAE increased from 0.0191 Ah in Cycle 3 to 0.0196 Ah in the unseen Cycle 4, indicating only a modest extrapolation error under early degradation conditions. This suggests that the model retained acceptable generalization over the short degradation trajectory covered in this study. However, the experimental scope remains limited to a single Samsung 30T cell, a surface temperature range of 41.7–50.3°C, and five charge-discharge cycles. [5,20,27] Therefore, although the proposed framework is promising for battery diagnostics under controlled laboratory conditions, broader validation across wider temperature ranges, longer cycle life, and multi-cell configurations is still required before stronger generalization claims can be made.

## 4. 4. Conclusion

This study shows that the XGBoost-LSTM model achieved the best performance for estimating SOC, SOH, DOD, and internal resistance compared with XGBoost-SVR and Linear-RF. The results indicate that combining nonlinear feature learning and temporal sequence modeling improves battery state prediction under the tested conditions. In addition, statistical testing confirmed that the superiority of XGBoost-LSTM was significant, while error analysis showed that the model remained relatively stable, although larger errors still appeared in low-SOC conditions, fast transients, and unseen cycle progression. Overall, the proposed framework is promising for battery diagnostics in controlled laboratory settings. However, further validation is still needed across wider temperature ranges, longer cycle life, and multi-cell battery configurations before broader practical deployment can be considered.

**Declaration of AI and AI assisted technologies in the writing process**

During the preparation of this work, the authors used ChatGPT to assist in language refinement, grammar correction, sentence restructuring, and improving the readability of the manuscript. The tool was used only for limited editorial support and not for generating the core scientific content, analysis, or conclusions of the study. After using this tool/service, the authors carefully reviewed and edited the content as needed and take full responsibility for the content of the publication.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

The authors gratefully acknowledge Universitas PGRI Semarang for providing access to laboratory facilities and technical support throughout the experimental campaign. We also extend our sincere thanks to the School of Electronics Engineering at Vellore Institute of Technology, Chennai, for invaluable guidance on data acquisition and model development. Special appreciation goes to the technical staff and students who assisted with cell conditioning, thermocouple installation, and data preprocessing, as well as to our colleagues who offered critical feedback on feature engineering and

hyperparameter tuning. Finally, we thank the funding agencies and administrative teams whose timely assistance made this research possible.

**References**

[1] Zhu B, Jia L, Pan Q, Zhang H. Cross-domain battery SOH and RUL estimation via Domain-Adaptive Transformer. Energy 2025;341. https://doi.org/10.1016/j.energy.2025.139288.

[2] Cao C, Gao H, Liu F, Maysrah M, Zhou Y, Chen L. Early prediction of lithium-ion battery life using a hybrid deep neural network and ensemble learning approach. Computers and Chemical Engineering 2026;205. https://doi.org/10.1016/j.compchemeng.2025.109465.

[3] Park MS, Lee JK, Kim BW. SOH Estimation of Li-Ion Battery Using Discrete Wavelet Transform and Long Short-Term Memory Neural Network. Applied Sciences (Switzerland) 2022;12. https://doi.org/10.3390/app12083996.

[4] Chou JH, Wang FK, Lo SC. Predicting future capacity of lithium-ion batteries using transfer learning method. Journal of Energy Storage 2023;71. https://doi.org/10.1016/j.est.2023.108120.

[5] Severson KA, Attia PM, Jin N, Perkins N, Jiang B, Yang Z, et al. Data-driven prediction of battery cycle life before capacity degradation. Nature Energy 2019;4. https://doi.org/10.1038/s41560-019-0356-8.

[6] Song Z, Zhang H, Jia J. Data-Driven State of Health Interval Prediction for Lithium-Ion Batteries. Electronics (Switzerland) 2024;13. https://doi.org/10.3390/electronics13203991.

[7] Wu X, He T, Zhu W, Liao Y. State of health estimation of lithium-ion battery based on automatic feature extraction and BiLSTM-SA. Ionics 2025;31. https://doi.org/10.1007/s11581-025-06681-8.

[8] Islam S, Namkoong G. High-Fidelity SOH Prediction in Lithium-Ion Batteries Using Hybrid ML Networks. Journal of The Electrochemical Society 2025;172. https://doi.org/10.1149/1945-7111/adf35e.

[9] Lin SL. Deep learning-based state of charge estimation for electric vehicle batteries: Overcoming technological bottlenecks. Heliyon 2024;10. https://doi.org/10.1016/j.heliyon.2024.e35780.

[10] Fan Y, Lin Z, Wang F, Zhang J. A hybrid approach for lithium-ion battery remaining useful life prediction using signal decomposition and machine learning. Scientific Reports 2025;15. https://doi.org/10.1038/s41598-025-92262-8.

[11] Bockrath S, Lorentz V, Pruckner M. State of health estimation of lithium-ion batteries with a temporal convolutional neural network using partial load profiles. Applied Energy 2023;329. https://doi.org/10.1016/j.apenergy.2022.120307.

[12] Tianqi YU, Yongxu ZHU, Xianbin W, Delhi N, Gugulothu N, Tv V, et al. Analysis of degradation in residential battery energy storage systems for rate-based use-cases. Applied Energy 2019;PP.

[13] Li W, Lin C, Hosseininasab S, Bauer L, Pischinger S. Lithium-Ion Battery SOH Estimation Based on a Long Short-Term Memory Model Using Short History Data. IEEE Transactions on Power Electronics 2025;40. https://doi.org/10.1109/TPEL.2025.3532839.

[14] Shi Z, Xu J, Ibrahim AW, He Y, Liu S, Zeng L. Accurate state of health estimation based on the multi-feature extracted and improved MHA assisted CNN-BiGRU model for lithium-ion batteries. Journal of Energy Storage 2025;131. https://doi.org/10.1016/j.est.2025.117598.

[15] Ogundana AS, Terala PK, Amarasinghe MY, Xiang X, Foo SY. Electric Vehicle Battery State of Charge Estimation with an Ensemble Algorithm Using Central Difference Kalman Filter (CDKF) and Non-Linear Autoregressive with Exogenous Input (NARX). IEEE Access 2024;12. https://doi.org/10.1109/ACCESS.2024.3371883.

[16] Yang Y, Chen S, Chen T, Huang L. State of Health Assessment of Lithium-ion Batteries Based on Deep Gaussian Process Regression Considering Heterogeneous Features. Journal of Energy Storage 2023;61. https://doi.org/10.1016/j.est.2023.106797.

[17] Wang Z, Liu Y, Wang F, Wang H, Su M. Capacity and remaining useful life prediction for lithium-ion batteries based on sequence decomposition and a deep-learning network. Journal of Energy Storage 2023;72. https://doi.org/10.1016/j.est.2023.108085.
[18] Kim SW, Oh KY, Lee S. Novel informed deep learning-based prognostics framework for on-board health monitoring of lithium-ion batteries. Applied Energy 2022;315. https://doi.org/10.1016/j.apenergy.2022.119011.
[19] Xu J, Liu B, Zhang G, Zhu J. State-of-health estimation for lithium-ion batteries based on partial charging segment and stacking model fusion. Energy Science and Engineering 2023;11. https://doi.org/10.1002/ese3.1338.
[20] Yuan H, Liu J, Zhou Y, Pei H. State of Charge Estimation of Lithium Battery Based on Integrated Kalman Filter Framework and Machine Learning Algorithm. Energies 2023;16. https://doi.org/10.3390/en16052155.
[21] Fei C, Lu Z, Jiang W, Zhao L, Zhang F. Research on Lithium-Ion Battery State of Health Prediction Based on XGBoost–ARIMA Joint Optimization. Batteries 2025;11. https://doi.org/10.3390/batteries11060207.
[22] Alsuwian T, Ansari S, Ammirrul Atiqi Mohd Zainuri M, Ayob A, Hussain A, Hossain Lipu MS, et al. A review of expert hybrid and co-estimation techniques for SOH and RUL estimation in battery management system with electric vehicle application. Expert Systems with Applications 2024;246. https://doi.org/10.1016/j.eswa.2023.123123.
[23] Xiao Z, Jiang B, Zhu J, Wei X, Dai H. State of Health Estimation for Lithium-Ion Batteries Using an Explainable XGBoost Model with Parameter Optimization. Batteries 2024;10. https://doi.org/10.3390/batteries10110394.
[24] Wang D, Bao Y, Shi J. Online lithium-ion battery internal resistance measurement application in state-of-charge estimation using the extended kalman filter. Energies 2017;10. https://doi.org/10.3390/en10091284.
[25] Wang J, Zhang C, Meng X, Zhang L, Li X, Zhang W. A Novel Feature Engineering-Based SOH Estimation Method for Lithium-Ion Battery with Downgraded Laboratory Data. Batteries 2024;10. https://doi.org/10.3390/batteries10040139.
[26] University of Maryland. Center for Advanced Life Cycle Engineering. School of Engineering 2019.
[27] Ali S, Bianchi V, De Munari I. A Microcontroller Based Optimized Framework for the State of Charge Estimation of a Lithium Ion Battery. 2025 IEEE International Workshop on Metrology for Automotive, MetroAutomotive 2025 - Proceedings, 2025. https://doi.org/10.1109/MetroAutomotive64646.2025.11119267.