Hybrid XGBoost–LSTM Model for State of Health and Remaining Useful Life Prediction of Lithium-Ion Batteries

Amarjot Kaur[1], Shristi Chauhan[2], Dr. R. K. Pongiannan [3]

[1,] Department of Computing Technologies ,SRM Institute of Science and Technology, Kattankulathur, India-603203

Corresponding Author Email: ak9799@srmist.edu.in, sc2071@srmist.edu.in

How to Cite this Article:

Kaur, A. & Chauhan, S. (2026). Hybrid XGBoost–LSTM Model for State of Health and Remaining Useful Life Prediction of Lithium-Ion Batteries. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(05). https://doi.org/10.55041/ijcope.v2i5.018

https://doi.org/ 10.55041/ijcope.v2i5.018

Abstract—

For electric cars and energy storage devices to be more dependable, secure, and simple to maintain, it's crucial to possess the ability to precisely forecast the State of Health (SoH) and Lithium-ion battery Remaining Useful Life (RUL). Typical Structured battery data is a good fit for machine learning models. However, they don't always recognize how things evolve over time. Models for deep learning, like Long Short Term Memory (LSTM) networks, learn sequential patterns well, but they need a lot of data and processing power. This study shows a hybrid prediction system which combines XGBoost and LSTM models to use feature-based learning and temporal dependency modeling to guess the battery life. SoH. Experimental evaluations are performed using publicly available NASA lithium-ion battery discharge datasets. We use the Leave-One-Battery-Out strategy to measure how well the model works using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). The results show that the proposed hybrid model works better and more consistently than individual models on a number of battery datasets. The predicted SoH also makes it possible to use data to guess RUL, which is useful for real-world battery health monitoring.

Keywords— Lithium-ion battery; State of Health (SoH); Remaining Useful Life (RUL); XGBoost; Long Short-Term Mem ory (LSTM); Hybrid machine learning model; Battery health prediction

I. INTRODUCTION

In modern energy storage systems such as electric vehicles, renewable energy systems, and portable electronic devices, lithium-ion batteries play an important role. Accurate prediction of health parameters of batteries such as state-of-health (SoH) and remaining useful life (RUL) is required for their safe, reliable and efficient operation. SoH describes how a battery is performing compared to its original capacity, and RUL gives an estimate of how many more complete discharge/recharge cycles are left before reaching end-of-life.

Traditional physics-based and empirical approaches for battery health estimation often rely on complex electrochemical degradation models [1]. Although these methods provide interpretability, they are difficult to design, computationally expensive and highly sensitive to operating conditions. As a result, data-driven techniques using machine learning (ML) and deep learning (DL) have gained significant attention.

Batteries have been shown by numerous studies [2], [3] to exhibit nonlinear relationships between input features (e.g. voltage, current, temperature) and various indicators (e.g. capacity, internal resistance, cycle count) of a battery's state-of-health, through the use of machine learning models (e.g. support vector machines, random forests, XGBoost).

However, conventional ML models generally treat battery data as independent observations and therefore do not account for the influence of time on the degradation of batteries due to wear and tear. In contrast, deep learning models (e.g. Long Short-Term Memory (LSTM) networks) are designed to take time-series data and can effectively capture sequential dependencies [4], [5], [6]. However, LSTM-based models typically require large amounts of labelled data, precise hyperparameter tuning, and significant computational resources, which can restrict the capability of these models to generalize in real-world applications [7][8].

To address these limitations, this paper proposes a hybrid prediction framework that combines the strengths of both XGBoost and LSTM models. The XGBoost model captures complex feature-level relationships from battery sensor data, while the LSTM-based model capture temporal degradation trends. The goal of the proposed hybrid model is to achieve improved accuracy, robustness, and consistency across different battery datasets. In this work, SoH is directly predicted, and the estimated SoH trajectory is further utilized for data-driven RUL estimation.

The proposed approach is evaluated using publicly accessible lithium-ion battery discharge datasets, with voltage, current, temperature, and time as input features. The performance of models is compared using standard evaluation metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). Furthermore, a Leave-One-Battery-Out Cross-Validation (LOBO-CV) strategy is used to assess the generalization capacity of the proposed system across multiple battery cells.

The key contributions of this work are summarized as follows:

1) Development of a hybrid XGBoost–LSTM framework for lithium-ion battery SoH estimation.
2) Comparative analysis of ML, DL, and hybrid models using identical input features.
3) Validation of robustness and consistency across multiple batteries using LOBO-CV.

II. LITERATURE REVIEW

Lithium-ion batteries are widely used in Electric Vehicles (EVs), renewable energy systems and portable electronics. But long term use leads to battery degradation over cycles that affects to performance, safety and reliability. State of Health (SoH) indicates battery capacity degradation and Remaining Useful Life (RUL) indicates operational cycles before end of life of the batteries. Accurate SoH and RUL prediction helps in preventing unexpected battery failures and malfunctions, it also helps in optimizing maintenance scheduling, it improves safety in EVs and enhances battery management systems (BMS).

There are many challenges that interferes with the accurate prediction of SoH and RUL. Predictions fails to capture non-linear and temporal degradation behaviour, feature variation across battery cells, scarcity of reliable battery data, poor generalization across batteries.

In 2021, Mamo and Wang [4] proposed a hybrid LSTM model integrated with attention mechanism and optimized it using Differential Evolution (DE) algorithm. They used multiple datasets such as NASA, CALCE, Oxford and Toyota battery datasets. The model accurately predicted the degradation in terms of MAPE, RMSE. However, it requires significant amount of computational resources, extensive hyperparameter tuning, large training data. Also it also doesn't integrate structure based learning methods .

In 2024, Minghu WU et al. proposed a paper where a multiscale prediction framework using Ridge Regression is modeled for the purpose. The dataset used are NASA and Oxford battery datasets. It gives high accuracy with State of Health MAE and RMSE with value less than 0.2 and also Remaining Useful Life (RUL) MAE within two cycles. But the model heavily relies on linear regression models and cannot capture nonlinear regression degradation behaviour. Also it is sensitive to operating conditions of the battery [9], [10].

In another paper authored by J. N. Chandra Sekhar et al. in 2023 [11], used Support Vector Regression and other data driven approaches on NASA dataset. It improved the non-linear degradation tracking, however, it did not explicitly models temporal degradation dependencies and focused solely on RUL rather than RUL and SoH relationship .

In a study by Faith Durmus and Serap Karagol in year 2024 [7], Genetic Algorithm (GA), Convolutional Neural Network, Recurrent Neural Network and Back Propagation (BP) were used as a hybrid model to predict battery degradation. It improved convergence speed and accuracy. It comparatively performs better that traditional Neural Networks. But it requires large dataset and high computational complexity.

Another work done by Jian Liu and Ziqiang Chen in 2019 [12] proposed a Gaussian Process Regression model using public Li-ion battery dataset. It demonstrates good prediction accuracy, but requires high computational cost and has scalability issues. It also doesn't capture long term sequential dependencies.

After reviewing various works it is evident that there are many research gaps. These limitations are also highlighted in comprehensive battery surveys [1], [13]. Such as Traditional physics based models proved interpretability but are not flexible and suffer high complexity and sensitivity to operating conditions. It also fails to explicitly model and capture long term dependencies. Whereas, deep learning has better sequential degradation behaviour and achieves improved accuracy but requires large training, extensive tuning, high computational resources. They have strong non-linear regression but not temporal dependencies, also they focus only either on RUL or SoH not on a hybrid framework.

III. METHODOLOGY

This section presents the methodology for lithium-ion battery State of Health (SoH) and Remaining Useful Life (RUL) prediction using individual ML and DL models, as well as the proposed hybrid XGBoost-LSTM framework. The implementation is carried out in Python using libraries such as Scikit-learn, XGBoost, TensorFlow/Keras, NumPy, Pandas and other relevant visualization libraries.

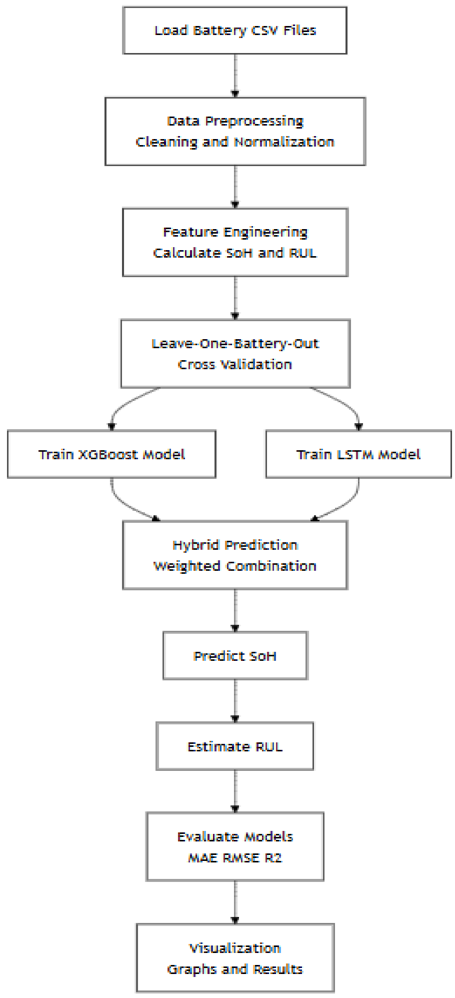

Figure. 1. Proposed Workflow

A. *Dataset and Feature Engineering*

Experiments were conducted on four lithium-ion batteries cells (B0005, B0006, B0007, B0018) from the NASA battery aging discharge dataset[6].

State of Health or SoH is computed as the ratio of discharge capacity at each cycle to the maximum observed capacity of the corresponding battery:

$$SOH = C_t / C_{max}$$

Remaining Useful Life or RUL is derived as the number of remaining cycles until end-of-life:

$$RUL = N - i - 1$$

where N denotes the total cycle life and i represents the current cycle index.

For model development, voltage, current, temperature, and time are used as input features, while cycle index is additionally included for XGBoost model.

Temporal sequences of fixed length are constructed for LSTM training, and Min-Max normalization is applied to ensure stable learning across batteries.

B. *Model Architectures*

This paper employs two models: XGBoost and LSTM, to predict SoH from discharge measurements.

The XGBoost model is a gradient-boosted decision tree method which effectively captures nonlinear relationships from structured battery discharge data.

Another model is LSTM which is a recurrent neural network designed for sequential time-series modeling. The proposed LSTM consists of stacked LSTM layers which is followed by a dense output layer. Since battery degradation evolves gradually over charge-discharge cycles, the LSTM model captures temporal dependencies using fixed-length input sequences.

XGBoost provides strong feature-based learning but does not explicitly model long-term temporal degradation. LSTM, on the other hand, captures sequential aging trends but require large datasets and high computational cost. These complementary characteristics motivate the hybrid learning framework.

C. *Proposed Hybrid XGBoost-LSTM Framework*

To achieve more consistent performance battery health prediction, this work proposes a hybrid framework that integrates the outputs of both XGBoost and LSTM models.

In the proposed approach, XGBoost generates SoH predictions by learning nonlinear relationships from structured discharge features, whereas the LSTM network estimates SoH by capturing sequential degradation patterns over multiple cycles. Since both models provide independent predictions, a weighted fusion strategy is applied to obtain a single hybrid estimate.

The hybrid prediction is computed as a weighted combination of the two outputs:

$$\hat{y}_{hyb} = w_{xgb}\hat{y}_{xgb} + w_{lstm}\hat{y}_{lstm}$$

where the weights are automatically determined based on the Mean Absolute Error (MAE) of each model, such that models with lower error contribute more significantly to the final prediction.

This hybrid architecture enhances generalization across different battery cells and provides more stable performance compared to using either XGBoost or LSTM alone. The overall workflow of the framework is illustrated in Figure.1

D. *Training and Validation*

The dataset is divided into training and testing sets following the LOBO-CV protocol.

The proposed framework is trained using a supervised manner, where discharge measurements are mapped to corresponding SoH targets. All input features are normalized using Min-Max scaling to ensure stable convergence particularly for the LSTM network.

The XGBoost model is trained using cycle-wise structured feature vectors, while LSTM model is trained on fixed-length temporal sequences of length 20 of past discharge cycles. The network was trained using the Adam optimizer with mean squared error loss.

To evaluate robustness and generalization across different battery cells, a Leave-One-Battery-Out Cross-Validation (LOBO-CV) strategy is adopted. In each fold, one battery is held out as unseen test set, while the remaining are used for training. This ensures that the model performance reflect its ability to generalize to new battery conditions rather than memorizing a single battery's aging pattern.
Aligned predictions from both models were combined through the proposed hybrid weighing scheme.

E. *Performance Evaluation*

To make sure that the proposed models are fully evaluated, following standard regression performance metrics are used:

- Mean Absolute Error (MAE): It tells you how far off the absolute difference between the predicted and actual values. A lower MAE value means that the prediction is more accurate.
- Root Mean Square Error (RMSE): This checks the average size of prediction errors by figuring out the mean of the squared differences between predicted and real values. Lower RMSE values show better model performance.
- Coefficient of determination (R^2): It shows the pro-part of the change in the dependent variable that is explained by the variables that are not dependent. R^2 values that are higher show that the model fits the data better.

For improved clarity, visual comparisons are also done. The actual and predicted SoH curves are shown over discharge cycles to see how well each model captures trends in degradation. Also, performance that has been combined. Charts are used to show how people act in general across many cells in a battery.

Also, estimating Remaining Useful Life (RUL) is derived from the estimated SoH trajectories, emphasizing the real-world usefulness of the suggested method for batteries predictions.

IV. RESULTS AND DISCUSSION

This part shows the results of the proposed hybrid XGBoost-LSTM framework for battery SoH prediction and estimating RUL. Model performance is evaluated under the LOBO-CV strategy using MAE, RMSE and R^2 metrics.

A. *Overall Model Performance*

To summarize the general prediction capability, the average performance across all test batteries as shown in Table I. Among the three, the hybrid model achieves the lowest MAE and RMSE, along with highest R^2 value, indicating improved accuracy and stability to the individual models.

Table I: Overall Performance Comparison

Model	MAE	RMSE	R^2
XGBoost	0.067742	0.072586	0.293885
LSTM	0.054066	0.062190	0.461967
Hybrid	0.052102	0.058615	0.529354

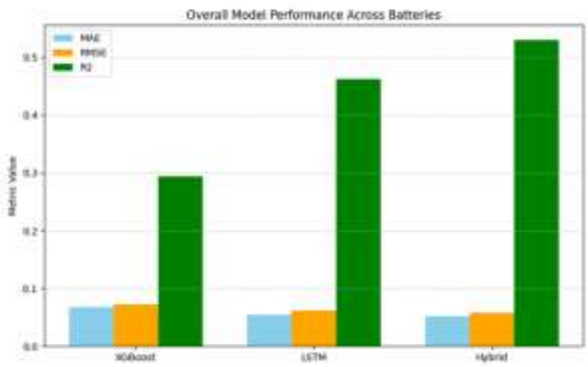

Figure 2. Overall Actual vs. Predicted across batteries

B. *SoH Prediction Curve Analysis*

To further evaluate tracking performance, Actual vs. Predicted SoH curves are plotted for all models. The hybrid model demonstrates more stable prediction trend across cycles, through deviations from the true SoH curve are still observed, especially in later degradation stages.

In addition, detailed SoH prediction tables and plots are represented to illustrate cell-level behaviour. For Battery B0007, all models perform well, with LSTM giving the most accurate SoH tracking. The hybrid approach provides stable and consistent predictions across cycles as shown in Table II . Battery B0018 exhibits more complex degradation behaviour, making accurate tracking more challenging. The hybrid approach improves robustness compared to XGBoost alone as shown in Table III.

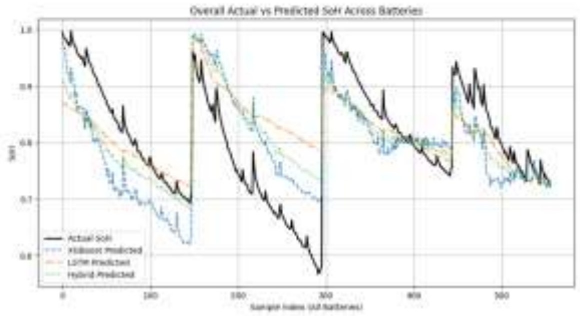

Figure 3. Overall Actual vs. Predicted SoH across batteries

Table II: Performance Metrics - Battery B0007

Model	MAE	RMSE	R^2
XGBoost	0.044451	0.049862	0.585529
LSTM	0.035483	0.046304	0.642567
Hybrid	0.039358	0.046425	0.640703

Table III: Performance Metrics - Battery B0018

Model	MAE	RMSE	R^2
XGBoost	0.047046	0.056066	0.292399
LSTM	0.035669	0.045189	0.540323
Hybrid	0.040295	0.047763	0.486459

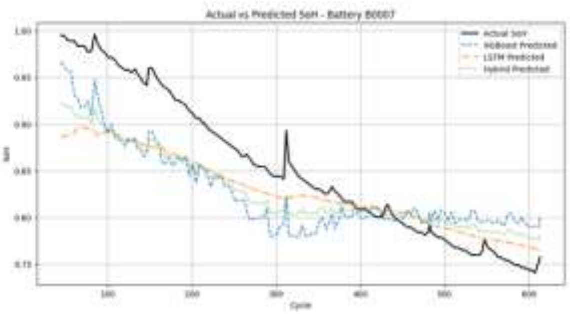

Figure 4. Actual vs. Predicted SoH (B0007)

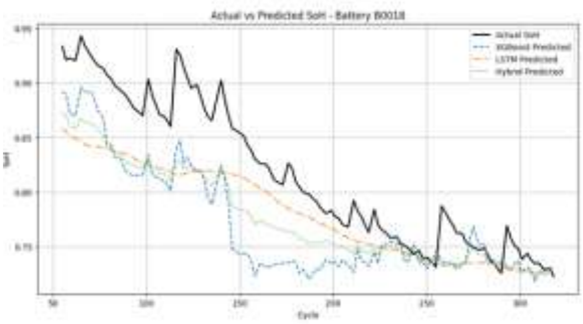

Figure 5. Actual vs. Predicted SoH (B0018)

C. *Strengths and Limitations*

The suggested hybrid framework combines the best parts of XGBoost and LSTM by mixing learning based on features with modeling the degradation over time. This integration makes it possible the model to take into account both structured feature relationships and behavior of degradation over time. This improves the model's performance more consistent and applicable across various battery cells when using the Leave-One-Battery-Out (LOBO) validation strategy.

But the method has some problems. Preparing data in a certain order is necessary for the LSTM part, and it costs more to run than other machine learning models. Also, the estimate for Remaining Useful Life (RUL) comes from predicted SoH values, not from learning them as a separate prediction target. This indirect estimation might not be as accurate in different real-life situations.

D. *Brief RUL Estimation Results*

In addition to estimating SoH, we also look at derived RUL trends to get a rough idea of how many operational cycles are left. The hybrid model makes RUL trajectories that are smoother and more consistent than those made by individual predictors. This shows that it can be used to monitor the health of batteries in real life.

E. *Improvements and Future Enhancements*

The suggested hybrid XGBoost-LSTM model works well and consistently on several battery cells, but there are ways to make it better in the future. Instead of getting it from SoH estimates, one way to go is to directly predict RUL as a different learning goal. This might help make predictions more accurate.

Combining more and more varied battery aging datasets that take into account different operating conditions and looking into advanced deep learning architectures like attention mechanisms or transformer-based models could also make predictions more accurate and generalize better. These enhancements would help make the framework more reliable when it is used in real-world battery management systems.

V. CONCLUSION

The proposed hybrid XGBoost-LSTM framework effectively monitors the health of lithium-ion batteries by integrating feature-based regression with the modeling of temporal degradation. Cross-validation of LOBO on multiple NASA battery discharge datasets demonstrates that the hybrid model more effectively predicts State of Health (SoH) in a consistent and generalized manner compared to individual models. The derived RUL trends also show how helpful this method is for planning maintenance. We will work on expanding the framework in the future to directly train RUL as a separate target. This will include adding more operational conditions, such as changing load and temperature profiles, and looking into physics-informed or attention-based architectures to make it easier to understand and more reliable for use in the real world.

REFERENCES

[1]. S. A. Hasib, S. Islam, R. K. Chakrabortty, M. J. Ryan, D. K. Saha, M. H. Ahamed, S. I. Moyeen, S. K. Das, M. F. Ali, M. R. Islam, et al., “A Comprehensive Review of Available Battery Datasets, RUL Prediction Approaches, and Advanced Battery Management,” IEEE Trans., vol. 9, pp. 86166–86193, 2021.

[2]. D. Roman, S. Saxena, V. Robu, M. Pecht, and D. Flynn, “Machine learning pipeline for battery state-of-health estimation,” Nature Machine Intelligence, vol. 3, no. 5, pp. 447–456, Apr. 2021, doi: 10.1038/s42256-021-00312-3.

[3]. J. Wu, X. Cui, H. Zhang, and M. Lin, “Health Prognosis With Optimized Feature Selection for Lithium-Ion Battery in Electric Vehicle Applications,” IEEE Trans. Power Electron., vol. 36, pp. 12646–12655, 2021.

[4]. T. Mamo and F.-K. Wang, “Attention-Based Long Short-Term Memory Recurrent Neural Network for Capacity Degradation of Lithium-Ion Batteries,” Batteries, vol. 7, no. 4, p. 66, 2021, doi: 10.3390/batteries7040066.

[5]. J. Lu, R. Xiong, J. Tian, C. Wang, and F. Sun, “Deep learning to estimate lithium-ion battery state of health without additional degradation experiments,” Nature Communications, vol. 14, no. 1, p. 2760, May 2023, doi: 10.1038/s41467-023-38458-w.

[6]. G. Ding, W. Wang, and T. Zhu, “Remaining useful life prediction for lithium-ion batteries based on CS-VMD and GRU,” IEEE Access, vol.10, pp. 89402–89413, 2022.

[7]. F. Durmus and S. Karagol, “Lithium-Ion Battery Capacity Prediction with GA-Optimized CNN, RNN, and BP,”Applied Sciences, vol. 14, no. 13, p. 5662, 2024, doi: 10.3390/app14135662.

[8]. C. Zhang, S. Zhao, and Y. He, “An integrated method of the future capacity and RUL prediction for lithium-ion battery pack,” IEEE Trans. Veh. Technol., vol. 71, pp. 2601–2613, 2021.

[9]. M. Wu, C. Yue, F. Zhang, R. Sun, J. Tang, S. Hu, N. Zhao, and J. Wang, “State of Health Estimation and Remaining Useful Life Prediction of Lithium-Ion Batteries by Charging Feature Extraction and Ridge Regression,” Applied Sciences, vol. 14, no. 8, p. 3153, 2024, doi: 10.3390/app14083153.

[10].X. Wang, X. Wang, B. Ma, Q. Li, and Y.-Q. Shi, “High Precision Error Prediction Algorithm Based on Ridge Regression Predictor for Reversible Data Hiding,” IEEE Signal Process. Lett., vol. 28, pp.1125–1129, 2021.

[11].J. N. C. Sekhar, B. Domathoti, and E. D. R. Santibanez Gonzalez, “Prediction of Battery Remaining Useful Life Using Machine Learning Algorithms,” Sustainability, vol. 15, no. 21, p. 15283, 2023, doi:10.3390/su152115283.

[12]. J. Liu and Z. Chen, ”Remaining Useful Life Prediction of Lithium-Ion Batteries Based on Health Indicator and Gaussian Process Regression Model,” in IEEE Access, vol. 7, pp. 39474-39484, 2019, doi:10.1109/ACCESS.2019.2905740.

[13]. S. Zhao, F. Blaabjerg, and H. Wang, “An Overview of Artificial Intelligence Applications for Power Electronics,” IEEE Trans. Power Electron., vol. 36, pp. 4633–4658, 2020.

[14]. Z. Deng, X. Hu, P. Li, X. Lin, and X. Bian, “Data-Driven Battery State of Health Estimation based on random partial charging data,” IEEE Transactions on Power Electronics, vol. 37, no. 5, pp. 5021–5031, Dec. 2021, doi: 10.1109/tpel.2021.3134701.

[15]. J. Wang, H. Li, C. Wu, Y. Shi, L. Zhang, and Y. An, "State of Health estimations for Lithium-Ion batteries based on MSCNN," Energies, vol. 17, no. 17, p. 4220, Aug. 2024, doi: 10.3390/en17174220.

[16]. H. Chun, J. Kim, M. Kim, J. Lee, T. Lee, and S. Han, "Capacity Estimation of Lithium-Ion Batteries for Various Aging States Through Knowledge Transfer," IEEE Trans. Transp. Electrif., vol. 8, pp. 1758–1768, 2022

