Privacy-Preserving State Of Health Prediction for Lithium-ion Batteries in electric vehicles using Federated Learning

Luan L. Santos [1], Guilherme A. Thomaz[1], Lucas A. C. Souza[1], Marcelo L. D. Lanza[1], Matteo Sammarco[2], Anna Glownia[2], Robson F. S. Dias[1], Luís Henrique M. K. Costa[1], Miguel Elias M. Campista[1]

***Abstract*—Accurate State of Health (SOH) prognosis for electric vehicle (EV) batteries is critical, but training centralized models on sensitive operational data creates significant privacy risks. Federated Learning (FL) offers a solution but faces major challenges from the statistically heterogeneous (Non-IID) data inherent to real-world EVs. This paper contributes to the state-of-the-art on SOH prognosis by allowing privacy-preserving training with scarce and heterogeneous data from a small number of distributed clients. We propose a novel SOH multi-step predictor based on recurrent signal embedding (SOH-MPRSE) model and validate the FL setup on a real-world dataset including 20 EVs. Our analysis confirms the data is Non-IID, with up to 85% of client pairs differing statistically. Despite this heterogeneity, results show the federated model successfully converges. The model trained on heterogeneous data achieves a final Mean Squared Error (MSE) of 0.29%. Furthermore, we observe that increasing clients yields statistically indistinguishable MSEs, demonstrating that the model can be trained with a small number of clients. This confirms the SOH-MPRSE framework is a privacy-preserving solution for SOH prognosis in resource-constrained EVs.**

I. Introduction

The global transition toward sustainable transportation is heavily reliant on the widespread adoption of electric vehicles (EVs) [1]. At the core of EV viability and reliability lies the performance of their lithium-ion battery packs. Consequently, precise battery management is essential not only for ensuring operational safety, but also for optimizing vehicle performance and enabling effective fleet operations [1]. Among the critical tasks in battery health monitoring, accurately estimating the State of Health (SOH) stands as a fundamental requirement. SOH provides a quantitative measure of a battery's condition relative to its fresh state, serving as a vital indicator for remaining useful life, optimizing charging strategies, and planning cost-effective pack replacements [2].

The primary research problem addressed in this work is the online estimation of SOH using only operational data measured by the vehicle's onboard Battery Management System (BMS), such as voltage, current, and mean temperature profiles during charging [3]. This strategy avoids the need for additional onboard diagnostic hardware. However, it introduces two critical and interconnected challenges.

The first one concerns data privacy and security. The operational data streams generated by a vehicle are highly sensitive, containing detailed information about locations and charging habits. This data is vulnerable to malicious actors both within the vehicle's internal units and in cloud infrastructures, where data may be aggregated [4]. Transmitting raw data to a central server for conventional model training, therefore, creates substantial privacy vulnerabilities and cybersecurity risks [5]. The second challenge is the limited availability of local data. While keeping data local mitigates privacy risks, the battery aging data collected by a single vehicle is too scarce to train reliable, data-driven SOH models.

Federated Learning (FL) has emerged as a promising architectural solution to these challenges [6]. FL enables collaborative training across a distributed fleet of vehicles while ensuring that battery usage data never leaves the vehicle [7]. In this paradigm, each client (vehicle) trains a local model using its own data and only transmits updated model parameters to a central server. The server aggregates these parameters to produce an improved global model, which is then redistributed to clients for the next training round [6]. Despite its advantages, standard FL frameworks struggle with non-independent and identically distributed (Non-IID) data [8]. In real EV fleets, diverse operational conditions, including distinct charging protocols, ambient temperatures, and driver behaviors, produce local data distributions that differ substantially from one client to another. This statistical heterogeneity is known to degrade the convergence speed and final performance of FL models [9].

Although recent works have explored FL for battery applications, they have focused almost exclusively on SOH diagnostics, i.e., estimating the current battery SOH [9], [7], [10]. While useful, real operational needs for fleet managers and safety systems revolve around SOH prognosis, which is the ability to accurately predict the future SOH trajectory over the next several charging cycles. This work addresses this gap by proposing a novel federated learning framework designed specifically for multi-step SOH prognosis. We introduce the State of Health Multi-step Predictor from Recurrent Signal Embedding (SOH-MPRSE). The SOH-MPRSE framework first extracts deep temporal feature embeddings from recent BMS charging-cycle data (voltage, current, and temperature). It then employs a specialized temporal dependency model to capture relationships across these sequential embeddings, enabling accurate prediction of both the current SOH and its evolution over subsequent charging cycles.

To validate our contribution, we utilized real charging data of 20 EVs assesing prognostic performance of SOH-MPRSE in the FL setting. We also evaluate the framework's robustness to Non-IID scenarios which reproduce real-world fleet heterogeneity. Our experiments demonstrate that SOH-

[1]Universidade Federal do Rio de Janeiro (UFRJ), Brazil luansantos@lafae.ufrj.br
[2]Stellantis, France matteo.sammarco@stellantis.com

MPRSE achieves stable convergence under heterogeneous charging and temporal patterns across different battery packs. Furthermore, models trained without access to initial life-cycle data generalize effectively to early-cycle conditions, indicating strong robustness to different battery aging stages. These results highlight the method's prognosis precision for real-world battery packs.

This work is organized as follows. Section II presents the methodological framework, including the dataset, preprocessing steps, and FL implementation. Section III analyzes the experimental findings, evaluating the impact of key hyperparameters and data scenarios. The obtained results are discussed in Section IV. Finally, Section V summarizes the contributions and outlines directions for future work.

II. SOH-MPRSE Model Design and Implementation

This study evaluates FL health prognostics for electric vehicles. Our approach is based on the principle that historical operational data contains latent patterns indicative of battery degradation. By leveraging deep FL models, we can learn these patterns and project current and future degradation in a privacy-preserving context. This includes the ability to estimate SOH, and potentially Lithium-Ion batteries capacity regeneration, without requiring a deep understanding of underlying electrochemical processes.

A. Dataset and Experimental Scenarios

The dataset used in this study comprises operational data from a fleet of 20 commercial EVs [11]. Collected over several years, the data encompasses a wide range of real-world operating conditions and temperatures. Figure 1 displays the degradation trajectories for all 20 battery packs, from the first operational cycle to the end of the acquisition period.

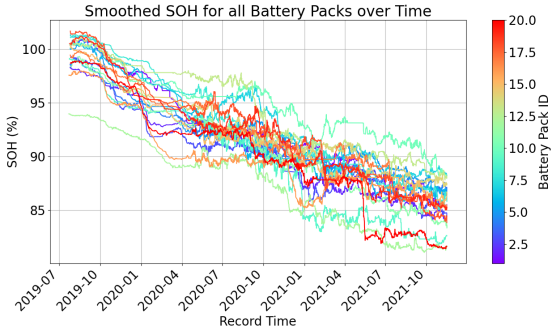


Fig. 1: Aging trajectories of all 20 EV Battery packs.

To evaluate the SOH estimator's robustness to incomplete data, we established two primary data conditions: full Sequence and trimmed Sequence. The trimmed sequence condition simulates data acquisition beginning only at battery's mid-life, while full sequence encompasses the whole aging trajectory. The trimmed sequence is generated using a trimming procedure, where for each battery's aging data, all charging cycles prior to a randomly selected SOH level (between 90-95%) are discarded. The model is only provided with the data from that point until the final collected sequence.

We then partitioned both sequences of all EVs into training and test datasets. This process results in four distinct scenarios detailed in Table I. The model is trained and evaluated using combinations of these scenarios to investigate the impact of partial data availability on estimation performance.

TABLE I: Dataset scenarios for model training and tests.

	Full Sequence	Trimmed Sequence
Training Set	Scenario 1	Scenario 2
Test Set	Scenario 3	Scenario 4

B. Data Preprocessing and Ground-Truth Generation

The raw time-series data is first segmented into discrete operational cycles. A new cycle boundary is identified when the time delta between consecutive records exceeds 10 seconds. Following segmentation, a curation step filters anomalous cycles to enhance data quality. Cycles are removed if they exhibit an insufficient State of Charge (SOC) change ($< 1.0\%$), have SOC values outside the valid [0%, 100%] range, or show an available capacity variation of less than 1.0 Ah. This results in the input data illustrated in Figures 2a–2c.

To establish a ground-truth degradation signal, the SOH is derived from the processed cycles. First, the absolute capacity ($C_{\text{abs},c}$) for each cycle c is estimated by normalizing the measured charge throughput ($Q_{\text{integral},c}$) by the corresponding fractional change in State of Charge (ΔSOC_c), as formulated in (1):

$$C_{\text{abs},c} = \frac{Q_{\text{integral},c}}{\Delta\text{SOC}_c} \times 100. \tag{1}$$

The SOH for cycle c is then defined as the ratio of its absolute capacity to the initial capacity of the first cycle ($C_{\text{abs},1}$), shown in (2),

$$\text{SOH}_c = \frac{C_{\text{abs},c}}{C_{\text{abs},1}}. \tag{2}$$

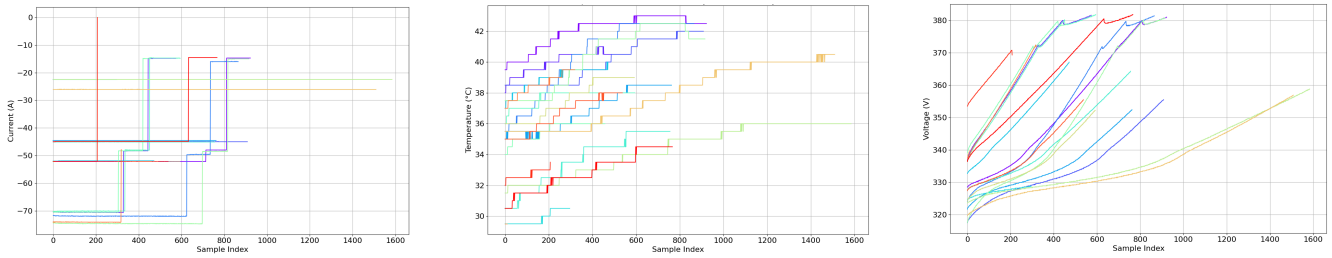
Finally, to remove noise from measurement and operational variations, the raw SOH_c signal is smoothed using a rolling median filter ($W = 50$), as described in Equation (3). This smoothed SOH curve serves as the primary target variable for the degradation models,

$$\text{SOH}_{\text{s},c} = \text{median}\left\{\text{SOH}_i \middle| i \in \left[c - \left\lfloor \frac{W}{2} \right\rfloor, c + \left\lfloor \frac{W}{2} \right\rfloor\right]\right\}. \tag{3}$$

C. Proposed SOH-MPRSE Model

We propose a model called SOH-MPRSE (SOH – Multi-step Prediction from Recurrent Signal Embedding), designed to capture hierarchical temporal dependencies from battery cycle data, as illustrated in Figure 3.

The first stage (intra-cycle) employs a Gated Recurrent Unit (GRU) to process the raw measurement arrays (voltage,



(a) Charging current (A). (b) Mean temperature (^{o}C). (c) Pack voltage (V).

Fig. 2: Input data for the first charging cycle (spanning both columns). Each color represents a cell ID.

current, and temperature) from a single charging cycle. These input arrays are typically masked or padded to ensure uniform sequence length. The resulting GRU output sequence is then processed by a multilayer perceptron (MLP) to generate a fixed-length time embedding that summarizes the state of that specific cycle.

The second stage (inter-cycle) uses a Long Short-Term Memory (LSTM) network to process a sequence of these cycle embeddings from multiple preceding cycles. This stage captures the long-term, inter-cycle degradation dependencies. A final MLP maps the LSTM output to a vector corresponding to predictions for multiple cycles in the future.

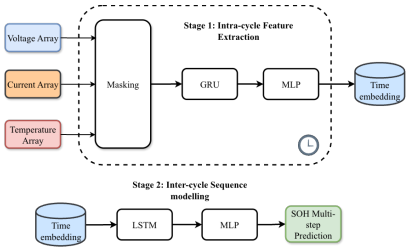


Fig. 3: SOH-MPRSE: State-of-Health multi-step predictor from recurrent signal embedding.

For feature engineering, we utilize a sliding window approach to generate overlapping input-output sequences. The SOH-MPRSE model uses a lookback window of 5 historical charging cycles (capturing voltage, current, and temperature data) as input. The model is trained to predict the SOH for the current cycle and the subsequent 2 cycles, resulting in a 3-cycles prediction horizon. This structure enables the model to learn temporal dependencies from recent degradation patterns.

D. Federated Learning Framework

The FL framework facilitates the collaborative training of the global battery aging model without sharing raw client data. The process follows an iterative cycle of local model training on client devices and parameter aggregation on a central server [6], illustrated in Figure 4 for vehicular applications.

1) Data Distribution and Partitioning: As partitioning strategies for EV battery aging are not well-established, we implement a two-stage approach. First, we separate 16 cells for training, where each cell represents a distinct client. This way,

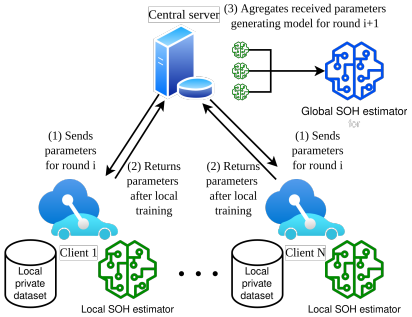


Fig. 4: Vehicular FL scenario.

the experiments have at most 16 clients, and, as the number of clients increases, the total number of data used for training also increases. Second, to simulate the data-scarce environment that motivates FL, we drastically limit the number of batches each client can access from its partition. This methodology allows for a controlled evaluation of model convergence based on the number of batches used in local training, local epochs, and learning rate.

We analyze the clients' datasets to verify if they are IID. First, we use the Covariance Distance [12] test for independence, followed by the Kolmogorov-Smirnov (K-S) test [13] for goodness of fit. The K-S test's null hypothesis is that the two distributions are the same; if the p-value is below 0.05, we reject the hypothesis and consider the data non-IID.

2) FL Implementation: The server and clients are implemented using PyTorch (version 2.9.0). The Flower framework (version 1.11.1) is used to manage all inter-process communication, parameter handling, and the Federated Averaging (FedAvg) aggregation algorithm.

The FL cluster specifications are summarized in Table II. When a client process starts, it is allocated to the next free GPU. If all GPUs are occupied, the next client shares the first GPU, following a round-robin allocation strategy [14]. While GPUs handle model training, all parameter-handling operations are performed by the CPU and RAM. The global model performance is evaluated on a dedicated test set after each aggregation round.

TABLE II: Hardware specifications of the FL cluster.

Machine	CPU/DRAM	GPU
Server	Core i9-10900/32 GB	–
Client 1	EPYC 7452/378 GB	Tesla V100S/32 GB
Client 2	EPYC 7452/378 GB	Tesla V100S/32 GB
Client 3	Core i7-13700K/64 GB	GeForce RTX 4070/12 GB
Client 4	Core i7-13700K/64 GB	GeForce RTX 4070/12 GB
Client 5	Core i5-9600K/32 GB	GeForce RTX 2080 Ti/12 GB

TABLE III: Number of non-IID datasets, comparing pairwise clients.

Metric	Trimmed	Untrimmed
Number of non-IID pairs	84	102
Fraction of non-IID pairs	70%	85%

III. Experimental Results

This section presents the experimental results in two main parts. First, we conduct a formal statistical analysis to quantify the degree of data heterogeneity (non-IID) across the client datasets, establishing the foundational challenge for the FL framework. Second, we evaluate the performance of our proposed federated prognostics model. This evaluation includes studies on key hyperparameters, an analysis of scalability by varying the number of clients, and an investigation into the model's robustness against temporal data shifts using the trimmed data scenario.

A. Data Heterogeneity Analysis

The input data consists of multivariate time series. As the statistical tests described in Section II-D1 are univariate, the non-IID analysis is decoupled and applied to each feature dimension independently. We perform pairwise comparisons (K-S test and covariance distance) across the 16 training clients, resulting in 120 unique pairs. Table III summarizes the total number of client pairs identified as statistically non-IID for each of the four training data scenarios.

The results confirm that the dataset is highly heterogeneous, and the scenario for trimmed data represents the most heterogeneous case, with 102 out of the 120 client pairs (85%) rejecting the IID hypothesis. A dimensional analysis revealed that the charge current is the primary source of heterogeneity, identified as non-IID in 90% of the heterogeneous cases. The mean temperature is the second-largest contributor, accounting for 25% of the cases.

B. Hyperparameters Optimization

A hyperparameter optimization (HPO) is employed to determine what configuration of the deep learning model minimizes the MSE on the validation dataset. This approach tailors the model configuration in a centralized manner using 50 trials of training/validation fold. The search space used for the SOH-MPRSE model, as well as the selected optimal hyperparameters, are detailed in Table IV. Notably, the dataset is big enough to make the model converge before the end of the first epoch.

TABLE IV: Hyperparameter search space for SOH-MPRSE model.

Hyperparameter	Type	Range / Choices	Selected
LSTM/GRU Units	Integer	32 to 128	64
Dense Units	Integer	16 to 64	32
Local Epochs	Integer	1 to 20	1
Batch Size	Categorical	{32, 64, 128}	32

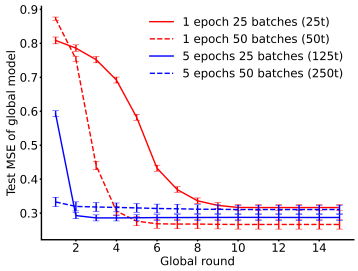


Fig. 5: Impact of local epochs (e_{local}) and batches (B) on model convergence (MSE). Relative computational time per round is shown between parentheses.

C. Federated Learning

Since the data subset and parameters are initialized randomly, the experiments must be repeated one hundred times to account for random uncertainty. After statistically analyzing the reported logs generated by the scripts, the results are presented as the mean value, with the 95% confidence intervals (CIs) depicted by the error bars in the corresponding plots. We use $\eta = 10^{-4}$ as the learning rate, $K = 2$ as the number of clients, and $E = 15$ as the number of global epochs/rounds, unless otherwise stated.

1) Influence of Local Epochs and Batches: The results are shown in Figure 5. The parentheses after the labels indicate the average time for local training in a single global round, where t is the time for the client's hardware to compute a gradient over a batch of 32 samples, and $B \cdot e_{\text{local}}$ is the number of local gradient computations. The model converges in fewer rounds when the number of local epochs is increased, at the cost of a substantially higher training time, as indicated in the label. Although having more batches and epochs, the setting ($e_{\text{local}} = 5, B = 50$) has a final error which is statistically indistinguishable from other settings.

Thus, we select ($e_{\text{local}} = 1, B = 25$) for all subsequent experiments for three primary reasons: (i) it provides the lower computational overhead, leading to faster experimentation; (ii) it better represents scenario with time, computation, and data volume constrains, which aligns with FL premises; and (iii) its smoother convergence profile facilitates clearer comparisons in subsequent experiments.

2) Influence of the Number of Clients: As depicted in Figure 6, aggregating models from multiple clients ($K > 1$) leads to a lower final MSE compared to a single-client baseline ($K = 1$), validating the collaborative approach. Nonetheless, increasing the number of clients beyond $K = 2$ (i.e., $K = 4$)

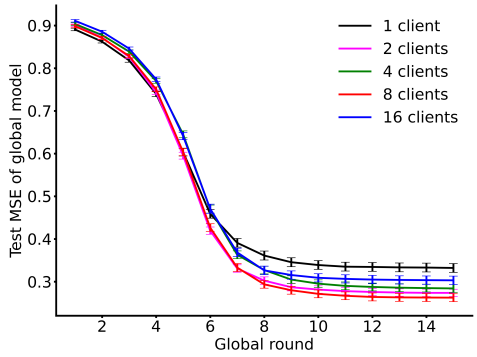


Fig. 6: Impact of the number of participating clients (K) on final model MSE.

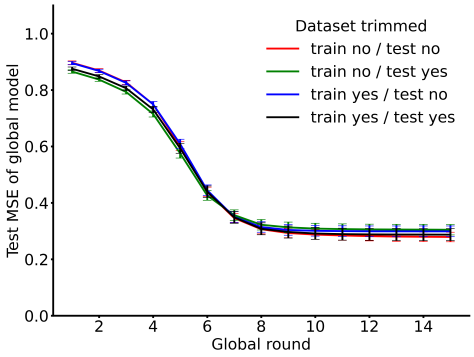


Fig. 7: Effect of time-series trimming (simulating data heterogeneity) on model performance, using $K = 4$ clients.

yields results that are statistically indistinguishable. This suggests that the learning task is not overly complex and that a robust model can be trained even with data from a small number of clients.

3) Federated vs. Centralized Performance: Table V presents the average MSE for centralized ($K = 1$) and federated ($K \geq 1$) learning. FL achieves a statistically higher performance than the centralized benchmark, since CIs for the two methods do not overlap.

TABLE V: MSE of Centralized and Federated approaches.

Method	**MSE**	**95% CI**
Centralized (Baseline)	0.3329	[0.3201, 0.3457]
Federated	0.2917	[0.2687, 0.3147]

4) Effect of Time-Series Trimming (Data Heterogeneity): Figure 7 presents the model's performance (using $K = 4$ clients) under data heterogeneity, which is simulated by trimming the time-series data in for training/testing datasets, producing four scenarios. The red curve assumes that train and test datasets contain untrimmed time-series, leading to the same result as presented in the last experiment. The gray curve consider trimmed time-series in both datasets, and, despite the higher degree of heterogeneity in this scenario (as demonstrated in Subsection III-A), the model still converges to a comparable final MSE value. This indicates that removing initial data points, and thus increasing inter-client data variance, does not significantly impair the model's generalization capability in this application.

The goal of the other two curves is to evaluate how a model trained without data from the initial battery life performs in early-cycle data (trimmed train dataset and untrimmed test dataset, represented in blue) and vice-verse (untrimmed train dataset and trimmed test dataset, represented in green). The statistical equivalency in all results confirms that ability of all models to generalize to situations never seen before.

IV. Discussions and Analyses

The experimental results provide several key insights into the practical application of Federated Learning for EV battery SOH prognosis. This section discusses the primary implications of our findings.

- **Convergence under heterogeneous data:** the charging and discharging temporal patterns of different battery packs are very heterogeneous. This holds true even in the case where the packs being compared include data from the same period of the battery life-time. At first, such initial results could demotivate using FL due to known issues with convergence under non-IID settings. Yet, our experiments demonstrate that convergence still holds, which is achieved by this work in terms of encouraging the usage of FL for this task.
- **Processing requirements:** by adopting a GPU cluster, we were able to reduce the duration of a global round to less than 5 seconds, which allows more repetitions and variations for the experiments. However, just a small number of local computations is enough for training a federated model. Therefore, the requirements for computing power and energy needed to train the proposed model locally may not be prohibitive for next-generation EV embedded computing using cheaper hardware, such as CPU. This suggests that shifting from the traditional cloud-centric ML approach to FL is an economically viable solution for increasing vehicle owners' privacy and trust in the manufacturers.
- **Data collection requirements:** using a dataset of $N_{local} = 800$ local samples (25 batches/round $\times$ 32 samples/batch) was enough for training a high-performance model. An interesting conclusion is that the requirements in terms of data volume needed to train a federated model are practical for real-world vehicular users.
- **Scalability and availability:** selecting just a small fraction of the vehicular fleet to participate in FL is enough to achieve excellent results. Then, the server can select only clients with stable wireless connections and idle com-

puting and memory resources, increasing fault-tolerance and reducing costs with expensive cloud networking infrastructure.

- **Incomplete SOH history:** a model trained without data from initial cycles generalize well to data from the start of the battery lifetime. This encourages the aggregation of models trained in vehicles with different moments of the battery life-time.

Nonetheless, some research gaps remain:

- **Limited access to signals:** modern electronic control units (ECUs) may not report in the controller area network (CAN) electric measurements from the high-voltage battery. Instead, the battery management system (BMS) computes the SOC and overall electrical power using simple estimators, and only reports this high-level data instead of raw voltage and current information. Therefore, its paramount to investigate strategies to reduce the number of input variables for our model and to verify if we can transfer the knowledge from our trained model to new domains, with different features, resolutions, and sampling rates.
- **Privacy evaluation:** it is still unclear if just using FL is enough to protect user privacy in the EV prognosis task, since model parameters are vulnerable to model inversion attacks that reverse engineer sensitive data [15]. Data anonymization, differential privacy [16], mixed federated learning, client/data sub-sampling, and trusted execution environments (TEEs) [17] may mitigate these privacy breaches. Evaluating if the computational overhead introduced by such techniques is worth the privacy gain is still not addressed.

V. Conclusions

This work investigates privacy-preserving EV SOH battery prognosis. Unlike previous work, we perform extensive experiments with a real-world vehicular dataset under realistic conditions, such as limited computational power, and limited data from heterogeneous battery life time periods. The proposed SOH-MPRSE model achieves low error prediction for state-of-health for 3 charging cycles in the future when trained with centralized data, and FL provide further reduce the error by 12% when compared to centralized benchmark. As future work, we aim to tackle the challenges concerning limited access to features provided by modern electric vehicles and the evaluation of the trade-off between complementary privacy-enhancing mechanisms for distributed training and their performance overhead.

Acknowledgments

This work was partially sponsored by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Funding Code 001 and 88887.954253/2024-00, CNPq (310234/2025-5, 408255/2023-4, and 405940/2022-0), FAPERJ (E-26/200.380/2023 and E-26/204.122/2024), FAPESP (2023/00673-7 and 2023/00811-0), Development and Research Foundation - Fundep - Rota 2030, and our partners Stellantis and Mobway.

References

[1] R. Zaino, V. Ahmed, A. M. Alhammadi, and M. Alghoush, "Electric vehicle adoption: A comprehensive systematic review of technological, environmental, organizational and policy impacts," *World Electric Vehicle Journal*, vol. 15, no. 8, 2024. [Online]. Available: https://www.mdpi.com/2032-6653/15/8/375

[2] G. Ramkumar, I. Chandra, and N. V. Krishnamoorthy, "Experimental evaluation of improved battery management system design for electric vehicles to resist performance over long duration," in *Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, April 2024, pp. 1–8.

[3] S. Zhang, B. Zhai, X. Guo, K. Wang, N. Peng, and X. Zhang, "Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks," *Journal of Energy Storage*, vol. 26, p. 100951, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X19307340

[4] K. Tang, B. Luo, D. Chen, C. Wang, L. Chen, F. Li, Y. Cao, and C. Wang, "The state of health estimation of lithium-ion batteries: A review of health indicators, estimation methods, development trends and challenges," *World Electric Vehicle Journal*, vol. 16, no. 8, 2025. [Online]. Available: https://www.mdpi.com/2032-6653/16/8/429

[5] T. Aljohani and A. Almutairi, "A comprehensive survey of cyberattacks on EVs: Research domains, attacks, defensive mechanisms, and verification methods," *Defence Technology*, vol. 42, pp. 31–58, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221491472400151X

[6] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2017. [Online]. Available: https://arxiv.org/abs/1610.05492

[7] W. Fang, J. Zhang, X. Lin, J. Hu, L. Huang, and G. Yuan, "A study on data privacy protection in power battery SOH prediction based on federated learning," *Ionics*, 2025.

[8] Y. Liu, M. Ahmed, J. Feng, Z. Mao, and Z. Chen, "A federated transfer learning framework for lithium-ion battery state of health estimation based on fast-charging segments," *IEEE Transactions on Transportation Electrification*, 2025.

[9] B. Chen, L. Fan, X. Zeng, M. Gu, and J. Zhou, "A contribution-aware federated framework for electric vehicle batteries health estimation," *IEEE Internet of Things Journal*, vol. 12, no. 5, pp. 4605–4612, March 2025.

[10] L. Fan, B. Chen, X. Hu, and W. Niu, "PIFL: Physics-informed federated learning for progressive degradation estimation in energy storage devices," *Advanced Engineering Informatics*, vol. 69, p. 103908, 2026.

[11] Z. Deng, L. Xu, H. Liu, X. Hu, Z. Duan, and Y. Xu, "Prognostics of battery capacity based on charging data and data-driven methods for on-road vehicles," *Applied Energy*, vol. 339, p. 120954, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261923003185

[12] N. Blomqvist, "On a measure of dependence between two random variables," *The Annals of Mathematical Statistics*, pp. 593–600, 1950.

[13] F. J. Massey Jr, "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[14] T. Balharith and F. Alhaidari, "Round robin scheduling algorithm in CPU and cloud computing: A review," in *2nd International Conference on Computer Applications & Information Security (ICCAIS)*, May 2019, pp. 1–7.

[15] J. Geiping, H. Bauermeister, H. Dröge, and M. Moeller, "Inverting gradients-how easy is it to break privacy in federated learning?" *Advances in neural information processing systems*, vol. 33, pp. 16 937–16 947, 2020.

[16] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. Quek, and H. V. Poor, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3454–3469, 2020.

[17] F. Mo, H. Haddadi, K. Katevas, E. Marin, D. Perino, and N. Kourtellis, "PPFL: Privacy-preserving federated learning with trusted execution environments," in *19th Annual International Conference on Mobile Systems, Applications, and Services*, 2021, pp. 94–108.