International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group | Volume 7 Issue 1 PP 265-270, 2026
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V7I1P133

Original Article

Software-in-the-Loop Validation of Electric Vehicle Battery Systems Using AI-Augmented Digital Twins

Vijayachandar Sanikal
IEEE Senior Member, Independent Researcher, Michigan, USA.

Received On: 23/01/2026 ***Revised On: 24/02/2026*** ***Accepted On: 27/02/2026*** ***Published On: 04/03/2026***

Abstract - *The increasing reliance on software-controlled battery management and protection functions has made software-in-the-loop (SiL) validation a critical challenge for electric vehicle battery systems. Conventional battery validation approaches, which depend heavily on physical testing and late-stage verification, are costly, limited in coverage, and poorly suited to exploring rare or compounded operating conditions. While digital twins and data-driven models have been widely studied for battery modeling and health estimation, their use as validation-centric instruments for safety-critical battery software remains limited. This paper presents a software-in-the-loop validation framework based on AI-augmented digital twins, integrating physics-based electrothermal battery models with machine learning based residual and risk modeling. The framework enables scalable validation of battery control logic under both measured and synthetically generated operating scenarios, with validation objectives focused on limit enforcement, thermal derating, and fault-handling robustness. The approach is demonstrated using publicly available national-laboratory battery datasets, showing that synthetic scenario execution significantly expands validation coverage compared to experimental testing alone and reveals software-mediated risk conditions that are difficult to detect through conventional methods. The results demonstrate that AI-augmented digital twins can transform battery validation into a continuous, lifecycle-oriented process, supporting shift-left verification and improved safety assurance in software-defined electric vehicles.*

Keywords - *Digital Twin, Battery Software Validation, Electric Vehicle Batteries, AI-Augmented Modeling, Electro-Thermal Battery Modeling, Synthetic Scenario Generation, Software-In-The-Loop Validation, Safety-Critical Systems.*

1. Introduction

The rapid adoption of electric vehicles has elevated the battery system to one of the most safety-critical subsystems in modern automotive platforms. Beyond electrochemical performance, EV batteries are now tightly coupled with complex software functions, including battery management systems (BMS), thermal derating logic, fault detection, and protection strategies. Failures in these systems are frequently software-mediated, particularly under boundary and compounded operating conditions arising not from fundamental cell defects but from insufficient validation of control logic under extreme or unforeseen operating conditions [1]. Conventional battery validation relies primarily on extensive physical cycling tests, abuse testing, and hardware-in-the-loop verification. Although effective for component-level characterization, these approaches suffer from three fundamental limitations. First, they are prohibitively expensive and slow when scaled across multiple chemistries, variants, and software revisions. Second, safety constraints limit the exploration of rare but high-risk scenarios, such as combined high-load and high-temperature operation. Third, physical testing is typically performed late in the development lifecycle, reducing opportunities for early design correction [2].

Digital twins have emerged as a promising solution for addressing these challenges. Physics-based electro-thermal battery models and reduced-order representations are widely used for performance prediction, thermal analysis, and state estimation [3], [5]. In parallel, data-driven and machine learning approaches have demonstrated strong capabilities in predicting battery state of health (SOH), remaining useful life (RUL), and degradation trends from experimental data [6], [7]. However, most existing work treats digital twins as modeling or prediction tools, rather than as integrated validation systems capable of exercising and verifying battery software behavior.

This gap is particularly critical in the context of software-defined vehicles, where battery functionality is continuously updated through software releases and over-the-air updates. In such systems, validation must evolve from a one-time certification activity to a continuous, scalable, and predictive process. Achieving this requires a validation framework that combines the trustworthiness of physics-based models with the coverage and adaptability of artificial intelligence.

In this paper, we propose an AI-augmented digital twin framework explicitly designed for battery software validation. The framework integrates physics-based battery digital twins with machine learning–based residual and risk modeling, enabling software-in-the-loop validation under both measured and synthetically generated operating conditions. By grounding the approach in publicly available national-laboratory datasets, the proposed methodology emphasizes reproducibility and industrial relevance. The key contributions of this work are as follows:

- A validation-centric digital twin architecture for EV battery systems that explicitly targets software behavior and safety limits.
- An AI augmentation strategy that enhances validation coverage by learning residual dynamics and amplifying rare, high-risk scenarios.
- A reproducible case study demonstrates how open battery datasets can be leveraged for scalable battery software validation.

2. Related Work

This section reviews prior research relevant to battery digital twins, data-driven battery modeling, and validation practices, highlighting the limitations that motivate the proposed framework.

2.1. Battery Digital Twins and Physics-Based Modeling

Physics-based battery models, including equivalent circuit models and electro-thermal formulations, form the foundation of most battery digital twin implementations. Such models have been widely applied for thermal analysis, power capability estimation, and energy management in EV applications [3], [8]. Reduced-order models and functional mock-up units have further enabled integration into system-level simulations and control development workflows [5]. While these approaches provide strong physical interpretability, their fidelity degrades when operating conditions deviate significantly from calibration data, limiting their effectiveness for comprehensive validation.

2.2. Data-Driven and AI-Based Battery Modeling

Machine learning techniques have been extensively explored for battery health estimation, including SOH and RUL prediction using cycling data, impedance measurements, and temperature signals [6], [7], [9]. These methods excel at capturing nonlinear degradation trends and variability across cells. However, purely data-driven models often lack physical constraints and perform poorly when extrapolated to unobserved conditions, raising concerns for safety-critical validation [10]. As a result, recent research has emphasized hybrid approaches that combine physics-based models with data-driven correction terms [11].

2.3. Battery Testing and Validation Practices

Publicly available battery datasets, such as those released by NASA, Sandia National Laboratories, and academic consortia, have enabled significant progress in battery modeling and prognostics research [13], [14]. These datasets provide high-quality experimental data across multiple chemistries and operating conditions. Nevertheless, their use has largely focused on model development and health prediction rather than systematic software validation. Traditional validation workflows remain heavily dependent on physical testing and late-stage verification, with limited ability to assess software robustness under rare or extreme conditions [2], [4].

Summary of Gap:

Existing literature provides strong foundations in battery modeling and data-driven health estimation, yet lacks an integrated, validation-first framework that leverages digital twins and AI to systematically verify battery software behavior. Addressing this gap is essential for enabling scalable, safety-oriented validation in software-defined electric vehicles.

3. Dataset and Experimental Basis

Robust software validation for battery systems requires experimental data that captures electro-thermal behavior, degradation effects, and variability across operating conditions. However, access to proprietary automotive battery datasets is often restricted, limiting reproducibility and independent validation. To address this, the proposed framework is grounded in publicly available, national-laboratory-grade battery datasets, which provide both credibility and transparency.

3.1. Selection of Open-Source Battery Data

The framework prioritizes datasets that satisfy three key criteria:

- Coverage of multiple lithium-ion chemistries relevant to automotive applications
- Inclusion of time-resolved electrical and thermal measurements suitable for digital twin calibration.
- Well-documented experimental protocols enabling reproducibility.

Among publicly available resources, datasets released by national laboratories and academic institutions have been widely adopted for battery modeling and prognostics research [13], [14]. These datasets provide high-quality measurements under controlled conditions and span a range of temperatures, depths of discharge, and load profiles, making them suitable for validation-oriented studies.

3.2. Sandia National Laboratories Battery Dataset

This work primarily utilizes battery cycling data released by Sandia National Laboratories, which includes commercial 18650 lithium-ion cells spanning three representative chemistries: lithium iron phosphate (LFP), lithium nickel cobalt aluminum oxide (NCA), and lithium nickel manganese cobalt oxide (NMC) [14]. In total, the dataset comprises dozens of cells tested under systematically varied ambient temperatures, discharge rates, and depth-of-discharge windows.

The dataset provides in-cycle measurements of terminal voltage, current, and cell temperature, along with cycle-level summaries such as charge and discharge capacity and periodic impedance characterization. Cells are aged until defined end-of-life criteria, enabling analysis of both nominal performance and degradation-driven behavior. Data are distributed in standardized formats with accompanying metadata, facilitating integration into modeling and validation workflows.

3.3. Relevance to Software Validation

Although originally designed for aging and performance studies, the Sandia dataset is particularly well suited for battery software validation. The synchronized electrical and thermal measurements enable calibration of electro-thermal digital twins, while the diversity of operating conditions supports the exploration of control edge cases. Importantly, the dataset allows controlled extrapolation toward synthetic extreme scenarios, such as high-load and high-temperature combinations that are impractical or unsafe to test physically.

By grounding the proposed validation framework in this dataset, the study demonstrates how open experimental data can be repurposed to support scalable, software-in-the-loop validation of battery management and protection logic. This approach emphasizes reproducibility while maintaining relevance to safety-critical automotive applications.

4. Problem Formulation and Validation Objectives

Battery-related failures in electric vehicles are increasingly driven by complex interactions between electro-thermal dynamics and embedded control software, rather than by isolated hardware defects. As battery management systems grow in complexity, validation challenges shift from verifying nominal performance to ensuring robust software behavior across a wide range of operating and environmental conditions. This work formulates battery validation as a software-centric problem, focusing on the verification of control logic, limit enforcement, and fault-handling behavior under both expected and extreme scenarios.

5. Proposed Digital Twin–Based Battery Software Validation Framework

This section presents the proposed AI-augmented digital twin framework designed explicitly for software validation of EV battery systems. Unlike conventional digital twin implementations focused on performance prediction or state estimation, the proposed framework treats the digital twin as an active validation instrument that continuously exercises battery software logic under both measured and synthetic operating conditions.

5.1. Framework Overview

The overall framework consists of three tightly coupled layers:

- A physics-based battery digital twin.
- An AI-based residual and risk modeling layer.
- A software-in-the-loop (SiL) validation orchestration layer.

Fig. 1. Show AI-augmented digital twin architecture for software-in-the-loop validation of electric vehicle battery systems, integrating physics-based electro-thermal modeling, data-driven residual learning, and battery control software execution. At a high level, experimental and operational data are first used to calibrate a physics-based electro-thermal battery model. Machine learning models are then trained to learn residual dynamics and degradation-sensitive behaviors that are insufficiently captured by physics alone. These combined models are executed in a SiL environment where battery management and thermal control software can be systematically validated across a broad range of scenarios.

5.2. Physics-Based Battery Digital Twin

The foundation of the framework is a physics-based battery digital twin that captures the dominant electro-thermal dynamics relevant to safety-critical operation. The model represents energy throughput, heat generation, and temperature evolution as functions of load, ambient conditions, and state variables. Reduced-order formulations are employed to ensure computational efficiency, enabling execution at scales compatible with cloud-based validation workflows [3], [5].

This physics-based layer provides two essential properties for validation. First, it ensures physical consistency and interpretability, which is critical for safety-related analysis. Second, it establishes a baseline representation of battery behavior against which deviations and anomalies can be identified. However, as observed in prior work, purely physics-based models struggle to maintain accuracy under aging effects, manufacturing variability, and complex operating regimes [8].

5.3. AI-Augmented Residual and Risk Modeling

To address these limitations, the digital twin is augmented with AI-based residual and risk models. Rather than replacing the physics-based model, machine learning is used to learn residual dynamics, the discrepancy between measured behavior and physics-based predictions under varying conditions.

The residual models capture effects associated with degradation acceleration, thermal imbalance, and nonlinear aging behavior that are difficult to parameterize explicitly. In addition, learned features derived from temperature gradients, energy throughput, and cycling history are used to construct risk indicators that correlate with elevated safety margins or impending limit violations [6], [11].

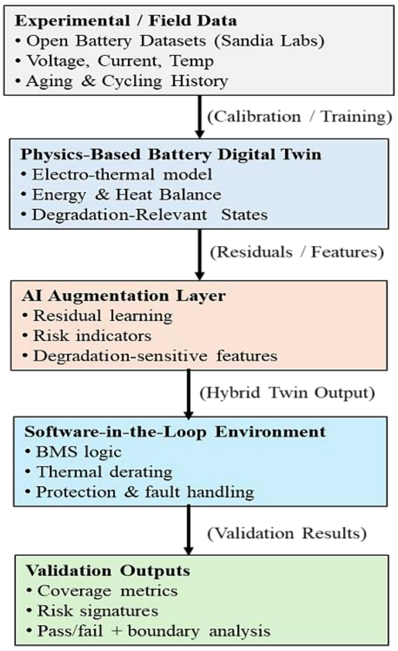

Fig 1: AI-Augmented Digital Twin Architecture for Software-In-The-Loop Validation

This hybrid approach combines the trustworthiness of physics with the expressiveness of data-driven learning, enabling improved extrapolation compared to purely data-driven models while preserving interpretability for validation purposes.

5.4. Software-in-the-Loop Validation Workflow

The calibrated hybrid digital twin is integrated into a software-in-the-loop validation environment, where battery control software interacts with the virtual battery in closed loop. Key software functions exercised within this environment include thermal derating logic, limit enforcement, fault detection, and protective shutdown strategies.

Validation scenarios are executed by subjecting the digital twin to both experimentally observed inputs and synthetically generated stress conditions. Software responses are monitored for violations of safety constraints, delayed reactions, or unstable control behavior. By executing these tests early and repeatedly, the framework enables shift-left validation, reducing reliance on late-stage physical testing and enabling repeatable regression testing across software revisions. [2], [15].

6. Synthetic Scenario Generation and Case Study

This section demonstrates how the proposed framework expands battery software validation coverage through synthetic scenario generation and presents a case study grounded in open experimental data to illustrate practical validation outcomes.[12]

6.1. Synthetic Scenario Generation for Extreme Conditions

Physical battery testing is inherently constrained by cost, safety, and time, limiting the exploration of rare or compounded operating conditions. To address this limitation, the proposed framework incorporates a synthetic scenario generation mechanism that systematically extends beyond the experimental envelope represented in measured data.

Synthetic scenarios are generated by perturbing validated operating variables such as load profiles, ambient temperature, cooling effectiveness, and sensor characteristics within physically plausible bounds defined by the calibrated digital twin. The physics-based model ensures energy and thermal consistency, while the AI-based residual models capture nonlinear effects associated with degradation acceleration and thermal imbalance. This combination enables safe exploration of conditions such as sustained high-power operation at elevated ambient temperatures, partial cooling degradation, and delayed sensor feedback. Fig. 2. Synthetic scenario generation and validation workflow, illustrating physics-constrained stress amplification and objective-driven exploration of software decision boundaries.

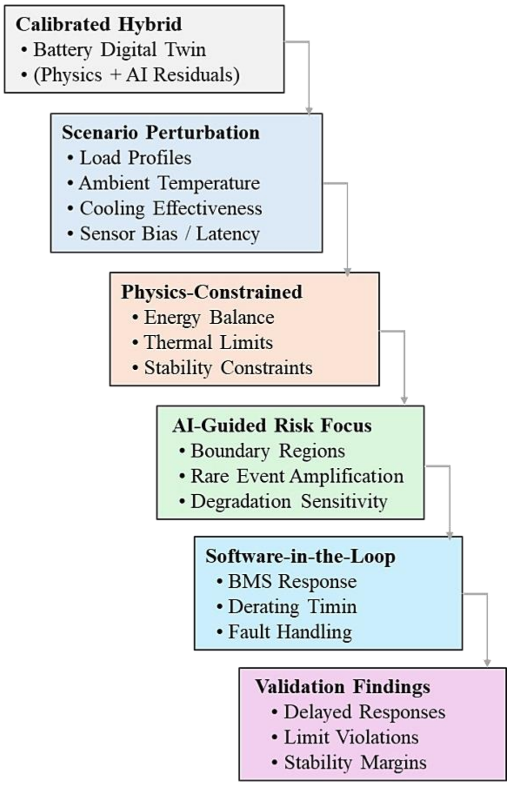

Fig 2: Synthetic Scenario Generation and Validation Workflow

6.2. Case Study Description

The proposed framework is evaluated using publicly available battery cycling data from Sandia National Laboratories, encompassing lithium-ion cells of multiple chemistries tested across varied temperatures, discharge rates, and depth-of-discharge ranges [16]. A subset of the data is used to calibrate the physics-based electro-thermal digital twin, while the remaining data supports validation and residual learning.

Battery management and thermal protection logic are exercised in a software-in-the-loop environment, interacting with the hybrid digital twin under both measured and synthetic operating scenarios. Validation metrics focus on software behavior rather than model prediction accuracy, including timing of derating actions, stability of control responses, and adherence to thermal and operational safety limits.[17]

6.3. Validation Outcomes and Observations

The case study demonstrates that synthetic scenario execution significantly expands the range of validated operating conditions compared to measured data alone. Compounded stress scenarios combining elevated ambient temperature with high discharge rates exposed delayed or marginal software responses that were not observed during nominal validation runs. These behaviors were traced to threshold interactions within thermal derating logic rather than deficiencies in the underlying battery model.

Across scenarios, the hybrid digital twin enabled consistent reproduction of thermal trends while highlighting residual-driven risk indicators associated with accelerated temperature rise. Compared to baseline validation using only experimental cycles, the proposed approach achieved substantially broader coverage of software decision boundaries with minimal additional computational cost.[18]

These results illustrate how AI-augmented digital twins can serve as practical validation tools, uncovering software-mediated risk conditions that would be difficult or unsafe to evaluate through physical testing alone.

7. Discussion and Implications

The results presented in this study highlight a fundamental shift in how battery systems can be validated in software-defined electric vehicles. Rather than treating validation as a late-stage, test-centric activity, the proposed framework demonstrates how AI-augmented digital twins can enable continuous, scalable, and safety-oriented validation of battery software behavior.

A key observation from the case study is that many safety-relevant issues emerge not from inaccurate physical modeling, but from software-mediated interactions between control thresholds, thermal dynamics, and operating variability. Conventional validation approaches, which rely heavily on predefined physical test matrices, are ill-suited to uncover such interactions particularly under compounded or rare conditions. By contrast, the proposed framework systematically exposes these behaviors through objective-driven synthetic scenario generation, expanding validation coverage without incurring the cost or risk of additional physical testing.

From an engineering perspective, the hybrid digital twin architecture offers an effective balance between trust and coverage. Physics-based models provide interpretability and enforce physical constraints, while AI-based residual and risk modeling enhances sensitivity to degradation-driven and nonlinear effects. This combination is particularly important for safety-critical applications, where purely data-driven extrapolation may be unacceptable, yet purely physics-based approaches lack sufficient adaptability.

The framework also has important implications for development workflows. By enabling software-in-the-loop validation early and continuously, the approach supports shift-left validation, reducing late-stage rework and accelerating design convergence. Moreover, the validation artifacts produced such as coverage metrics and risk signatures offer a structured basis for traceability and decision-making, which is increasingly relevant in regulatory and compliance-driven contexts.

Despite these advantages, several limitations must be acknowledged. The fidelity of validation outcomes depends on the quality and representativeness of the underlying datasets, and synthetic scenario generation must be carefully constrained to avoid unrealistic extrapolation. Additionally, the current implementation focuses on cell-level behavior; extension to pack-level and vehicle-level interactions will require additional modeling and data integration.

Overall, the findings suggest that AI-augmented digital twins can play a central role in the evolution of battery validation practices, particularly as EV platforms continue to adopt software-defined architectures and over-the-air update capabilities.

8. Conclusion and Future Work

This paper presented an AI-augmented digital twin–based framework for software validation of electric vehicle battery systems, addressing critical limitations of traditional, test-centric validation approaches. By integrating physics-based electro-thermal battery models with data-driven residual and risk modeling, the proposed framework enables scalable, software-in-the-loop validation of battery control logic under both measured and synthetic operating conditions. The results demonstrate that such an approach can significantly expand validation coverage, particularly for rare and compounded scenarios that are impractical or unsafe to evaluate through physical testing alone.

A key contribution of this work is the reframing of battery validation as a continuous, lifecycle-oriented process rather than a one-time certification activity.[19] The case study using open, national-laboratory battery datasets illustrates how digital twins can serve as trusted validation instruments, uncovering software-mediated safety risks that may otherwise remain undetected until late-stage testing or field operation. This validation-first perspective is especially relevant for software-defined vehicles, where battery functionality is increasingly shaped by frequent software updates.

Future work will focus on extending the framework from cell-level validation to pack- and vehicle-level digital twins, incorporating interactions with cooling systems, power electronics, and vehicle energy management. Additional research will explore closed-loop integration with fleet data to enable continuous model refinement and postdeployment validation.[20] Finally, alignment with emerging safety and certification frameworks represents an important direction for translating digital twin–based validation into formal evidence for regulatory and compliance processes.

References

[1] A. Barré, B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, and D. Riu, "A review on lithium-ion battery ageing mechanisms and estimations for automotive applications," *Journal of Power Sources*, vol. 241, pp. 680–689, Nov. 2013, doi: 10.1016/j.jpowsour.2013.05.040

[2] S. Saxena, C. Hendricks, and M. Pecht, "Cycle life testing and modeling of graphite/$LiCoO_2$ cells under different state of charge ranges," *Journal of Power Sources*, vol. 327, pp. 394–400, Sept. 2016. doi: 10.1016/j.jpowsour.2016.07.057

[3] M. Doyle, T. F. Fuller, and J. Newman, "Erratum: Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell [*J. Electrochem. Soc.*, vol. 140, p. 1526, 1993]," *Journal of The Electrochemical Society*, vol. 165, no. 11, p. X13, Sept. 2018, doi: 10.1149/2.1181811jes.

[4] V. Sanikal, "Digital twin architecture for continuous calibration of electric vehicle control systems: From development to production," Journal of Engineering and Applied Sciences Technology, June 2025, doi: 10.47363/JEAST/2025(7)322

[5] J. Marcicki, M. Canova, A. T. Conlisk, and G. Rizzoni, "Design and parametrization analysis of a reduced-order electrochemical model of graphite/$LiFePO_4$ cells for SOC/SOH estimation," *Journal of Power Sources*, vol. 237, pp. 310–324, Sept. 2013, doi: 10.1016/j.jpowsour.2012.12.120

[6] J. Remmlinger, M. Buchholz, M. Meiler, and K. Dietmayer, "State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation," Journal of Power Sources, vol. 196, no. 12, pp. 5357–5363, June 2011, doi: 10.1016/j.jpowsour.2010.08.035

[7] K. A. Severson, P. M. Attia, N. Jin, *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, Mar. 2019, doi: 10.1038/s41560-019-0356-8

[8] G. L. Plett, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 1. Background," *Journal of Power Sources*, vol. 134, no. 2, pp. 252–261, June 2004, doi: 10.1016/j.jpowsour.2004.02.03

[9] Y. Xing, E. W. M. Ma, K. L. Tsui, and M. Pecht, "Battery management systems in electric and hybrid vehicles," *Energies*, vol. 4, no. 11, pp. 1840–1857, Oct. 2011, doi: 10.3390/en4111840

[10] D.-I. Stroe and E. Schaltz, "Lithium-ion battery state-of-health estimation using the incremental capacity analysis technique," *IEEE Transactions on Industry Applications*, vol. 56, no. 1, pp. 44–52, Jan.–Feb. 2020, doi: 10.1109/TIA.2019.2955396

[11] Y. Ye, Y. Shi, N. Cai, J. Lee, and X. He, "Electro-thermal modeling and experimental validation for lithium-ion battery," *Journal of Power Sources*, vol. 199, pp. 227–238, Mar. 2012, doi: 10.1016/j.jpowsour.2011.10.027

[12] V. Sanikal, "Machine Learning Models for Predictive Maintenance of EV Thermal Systems Reducing Catastrophic Failure Risk." Journal of Artificial Intelligence & Cloud Computing SRC/JAICC-519. August 2025, doi: doi.org/10.47363/JAICC/2025(4)477

[13] G. dos Reis, C. Strange, M. Yadav, and S. Li, "Lithium-ion battery data and where to find it," *Energy and AI*, vol. 5, p. 100081, Sept. 2021, doi: 10.1016/j.egyai.2021.100081

[14] A. Garg, X. Peng, P. M. L. Le, C. Chin, *et al.*, "Design and analysis of capacity models for lithium-ion batteries," *Measurement*, vol. 120, pp. 114–123, Feb. 2018, doi: 10.1016/j.measurement.2018.02.003

[15] C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, "Degradation diagnostics for lithium-ion cells," *Journal of Power Sources*, vol. 341, pp. 373–386, Mar. 2017, doi: 10.1016/j.jpowsour.2016.12.011

[16] Xiong, J. Cao, Q. Yu, and F. Sun, "Critical review on the battery state of charge estimation methods for electric vehicles," *IEEE Access*, vol. 6, pp. 1832–1843, 2018, doi: 10.1109/ACCESS.2017.2780258

[17] G. Balan, P. Neninger, E. Ruiz Zúñiga, E. Serea, D.-D. Lucache, and A. Sălceanu, "A perspective on software-in-the-loop and hardware-in-the-loop within digital twin frameworks for automotive lighting systems," *Applied Sciences*, vol. 15, no. 15, p. 8445, Aug. 2025, doi: 10.3390/app15158445

[18] Y. Lu, C. Liu, K. Wang, H. Huang, and X. Xu, "Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101837, Feb. 2020, doi: 10.1016/j.rcim.2019.101837

[19] F. S. Schranner, A. Abassi Misheni, and J. Warnecke, "Deriving a representative variant for the functional safety development according to ISO 26262," *Reliability Engineering & System Safety*, vol. 213, p. 107436, Sept. 2021, doi: 10.1016/j.ress.2021.107436

[20] C. Duan, H. Cao, F. Liu, X. Duan, H. Pu, and J. Luo, "An interactive prognostics framework for lithium-ion battery remaining useful life based on neural networks and statistical processes," Reliability Engineering & System Safety, vol. 246, p. 111434, 2025, https://doi.org/10.1016/j.ress.2025.111434