STATE OF CHARGE (SOC) ESTIMATION TECHNIQUES: REVIEW VARIOUS SOC ESTIMATION TECHNIQUES, INCLUDING MODEL-BASED, DATA-DRIVEN, AND HYBRID APPROACHES

BY

ODOGWU NONYE NATHAN

ENG2002353

DEPARTMENT OF COMPUTER ENGINEERING

FACULTY OF ENGINEERING

UNIVERSITY OF BENIN

BENIN CITY

OCTOBER, 2025

STATE OF CHARGE (SOC) ESTIMATION TECHNIQUES: REVIEW VARIOUS SOC ESTIMATION TECHNIQUES, INCLUDING MODEL-BASED, DATA-DRIVEN, AND HYBRID APPROACHES

BY

ODOGWU NONYE NATHAN

ENG2002353

A PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER ENGINEERING, FACULTY OF ENGINEERING, UNIVERSITY OF BENIN, BENIN CITY IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF BACHELOR OF ENGINEERING (B.ENG) DEGREE IN COMPUTER ENGINEERING.

OCTOBER, 2025

CERTIFICATION

This project was carried out by Odogwu Nathan Nonye of the department of Computer Engineering, Faculty of Engineering, University of Benin, Benin City, and is hereby certified.

Engr. S. Akinbohun Date

(Project Supervisor)

Engr. Dr. Isi Edeoghon Date

(Head of Department)

DEDICATION

To the Almighty God, the source of all wisdom, strength, and grace, I dedicate this project. Without Your guidance, this endeavor would not have been possible. Every step, every idea, and every moment of perseverance was fueled by Your divine presence. May it bring glory to Your name and serve as a reminder of Your unwavering love and faithfulness.

ACKNOWLEDGEMENT

First and foremost, I give all glory to God Almighty, whose divine grace, wisdom, and guidance have sustained me throughout this academic journey.

I express my deep gratitude to my project supervisor, Engr. S. Akinbohun, for his invaluable guidance, constructive criticism, and mentorship throughout the course of this research work. His expertise and patience have significantly contributed to the success of this project.

My heartfelt appreciation goes to my beloved parents, whose unwavering support, sacrifices, and encouragement have been the bedrock of my academic pursuits. Their love and dedication have made it possible for me to reach this milestone.

I am profoundly grateful to all my lecturers who have contributed immensely to my academic growth through their knowledge, guidance, and commitment to excellence. Their teachings have shaped not just my academic understanding but also my professional outlook.

To my friends who have been there through the challenging times and moments of triumph, thank you for your support, encouragement, and the memories we've created together. Your friendship has made this journey more meaningful and enjoyable.

Finally, I acknowledge my own dedication, perseverance, and hard work throughout this academic pursuit. The countless hours of research, learning, and personal growth have culminated in this achievement.

ABSTRACT

State of Charge (SoC) estimation plays a crucial role in battery management systems (BMS), directly impacting the performance, safety, and longevity of lithium-ion batteries. This study presents a comparative review of three major categories of SoC estimation techniques: model-based, data-driven, and hybrid methods. The review is driven by the need to evaluate the accuracy, robustness, and practical applicability of these methods across various real-world conditions, including different temperature profiles, battery chemistries, and aging states.

The research methodology involved a structured literature search, selection of 45 peer-reviewed studies published between 2018 and 2025, and systematic data extraction. Model-based approaches, particularly those using Kalman filters and equivalent circuit models, demonstrated computational efficiency but showed sensitivity to parameter drift and aging. Data-driven techniques, including LSTM networks, Gaussian Process Regression (GPR), and Random Forests, offered high accuracy—often achieving <2% RMSE—but required large, diverse datasets. Hybrid methods, such as AEKF-LSTM and UKF-PSO-LSTM models, consistently achieved the highest accuracy (RMSE <1%) while balancing robustness and adaptability.

The findings suggest that while model-based methods are suitable for resource-constrained systems, hybrid approaches offer the most promising results in terms of overall performance and reliability. These insights can guide future BMS development and inform system-level design choices in electric vehicle and energy storage applications.

TABLE OF CONTENT

LIST OF FIGURES

Figure **Title** **Page**

LIST OF TABLES

LIST OF ACRONYMS

SoC	The State of Charge
BMS	Battery Management Systems
EVs	Electric Vehicles
ECMSs.	Equivalent Circuit Models
EKF	Extended Kalman Filter
UKF	Unscented Kalman Filter
PF	Particle Filter
GRU.	Gated Recurrent Units
CNN	Convolutional Neural Networks

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

The State of Charge (SoC) of lithium-ion batteries is a critical parameter that directly influences battery performance, safety, and longevity, especially in electric vehicles (EVs), renewable energy storage, and portable electronics (Xing & Wu, 2021). SoC Estimation refers to the process of determining the remaining charge in a battery relative to its full capacity, and it plays a significant role in battery management systems (BMS) for ensuring reliable and efficient energy utilization (Muratoglu & Alkaya, 2019). Precise SoC estimation is essential for preventing battery overcharging and deep discharging, which can lead to safety hazards such as thermal runaway or rapid capacity degradation (Chen et al., 2019).

Traditional methods such as Coulomb counting, which track the flow of charge in and out of the battery, often face significant limitations due to their sensitivity to current measurement errors and initial SoC uncertainty (Tulsyan et al., 2016). To overcome these challenges, model-based estimation techniques have gained widespread adoption. These approaches use equivalent circuit models (ECMs) or electrochemical models to describe the dynamic behavior of batteries and apply state estimators like the Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and Particle Filter (PF) for real-time SoC tracking (Xing & Wu, 2021; Chen et al., 2019; Tulsyan et al., 2016). These methods benefit from leveraging the physical and electrical characteristics of batteries but are often limited by the requirement for precise model parameterization, which can be affected by temperature variations and battery aging (Muratoglu & Alkaya, 2019).

In recent years, data-driven approaches have emerged as powerful alternatives for SoC estimation due to advancements in machine learning and artificial intelligence. These methods rely on historical battery data to train models that can learn complex, nonlinear relationships between input signals and the SoC without requiring explicit physical modeling (Yang et al., 2019). Deep learning techniques, such as Gated Recurrent Units (GRU) and Convolutional Neural Networks (CNN), have been particularly effective in improving SoC prediction accuracy under dynamic load conditions (Yang et al., 2019; Ma et al., 2024). However, data-driven models face significant challenges in terms of explainability and often require large, diverse datasets for training to ensure generalization across different battery chemistries and usage patterns (Ma et al., 2024).

To address the individual shortcomings of model-based and data-driven techniques, hybrid approaches have been introduced. These methods integrate the physical interpretability of model-based estimators with the learning capability of machine learning models to enhance SoC estimation performance (Ma et al., 2024). For example, hybrid models combining CNNs with Kalman Filters have demonstrated superior estimation accuracy and robustness compared to single-method approaches, especially in handling uncertainties related to measurement noise and model inaccuracies (Ma et al., 2024). Despite these promising developments, existing hybrid systems still face limitations in computational efficiency and real-time implementation, particularly in embedded battery management systems with restricted processing capacity (Chen et al., 2019).

The continuous evolution of SoC estimation techniques underscores the significance of this research area in computer engineering, particularly in the development of smart battery systems that support the growing demand for electric vehicles, renewable energy applications, and portable electronics. A comprehensive review of model-based, data-driven, and hybrid SoC estimation methods is essential to provide a clearer understanding of their advantages, limitations, and potential for future optimization.

1.2 Statement of the Problem

Accurate State of Charge (SoC) estimation is crucial for optimal battery management systems in electric vehicles and energy storage applications to prevent overcharging, deep discharge, and maximize battery lifespan (Hussein et al., 2024). However, SoC estimation remains challenging due to complex nonlinear relationships between battery voltage, current, temperature, and internal electrochemical processes, coupled with the inability to directly measure SoC through external sensors (Ali et al., 2019). Traditional approaches such as coulomb counting and open circuit voltage methods suffer from cumulative error accumulation and sensitivity to measurement noise under dynamic operating conditions (El-Sayed et al., 2024).

Model-based techniques, including equivalent circuit models and electrochemical models, provide good accuracy but require precise parameter identification and are computationally intensive, limiting their real-time implementation (Maurya et al., 2024). Data-driven approaches using machine learning algorithms offer promising alternatives by learning battery behavior patterns from historical data, but they often struggle with high-dimensional data adaptation and generalization across different battery chemistries and operating conditions (El-Sayed et al., 2024). Hybrid approaches that combine model-based and data-driven methods show potential for improved accuracy, yet their computational complexity and implementation challenges in embedded systems remain significant concerns (Hussein et al., 2024).

Therefore, a comprehensive review of existing SoC estimation techniques is essential to identify the strengths, limitations, and practical applicability of each approach for advancing battery management systems.

1.3 Aim and Objectives of the Study

The aim of this project is to conduct a comprehensive review of State of Charge (SoC) estimation techniques for lithium-ion batteries, focusing on model based, data driven, and hybrid approaches to identify their strengths, limitations, and suitability for real-world applications.

To achieve this aim, the following objectives will be pursued:

1. To examine model-based SoC estimation techniques, including equivalent circuit models and Kalman filtering approaches..
2. To review data-driven SoC estimation methods, particularly those utilizing machine learning and artificial intelligence.
3. To explore hybrid SoC estimation techniques that combine model-based and data-driven approaches for improved accuracy and reliability.
4. To compare the performance characteristics of the reviewed techniques under varying battery operating conditions and identify current gaps and potential research directions in the field of SoC estimation.

5. To provide recommendations on the most suitable SoC estimation methods for practical battery management system applications, particularly for electric vehicles and energy storage systems.

1.4 Scope of Study

This work reviews and compares various State of Charge (SoC) estimation techniques across three primary categories: model-based, data-driven, and hybrid approaches. The study does not extend to the development of new estimation algorithms or implementation of hardware prototypes. The research aims to provide comprehensive insights into the effectiveness, advantages, and limitations of existing SoC estimation methodologies, informing their suitability for different battery management system applications in electric vehicles and energy storage systems.

1.5 Relevance of Study

This study holds significant relevance as it addresses one of the key challenges in battery management Engr. S. Akinbohun

Date

accurate State of Charge (SoC) estimation for lithium-ion batteries. Reliable SoC estimation is essential for ensuring battery safety, improving performance, and extending battery lifespan, particularly in electric vehicles, portable electronics, and renewable energy storage systems. By systematically reviewing model-based, data-driven, and hybrid SoC estimation techniques, this research provides critical insights into their strengths, limitations, and practical applicability. The findings will assist researchers, engineers, and manufacturers in selecting suitable estimation methods that enhance battery efficiency and safety. Furthermore, the study highlights existing gaps and challenges.

1.6 Outline of Thesis

Chapter 2 offers a comprehensive literature review, first unpacking the electrochemical foundations of lithium-ion batteries and the formal definition of state of charge, then critically surveying model-based, data-driven, and hybrid SOC estimation techniques as reported in peer reviewed studies, highlighting their underlying assumptions, claimed accuracies, computational demands, and documented limitations.

Chapter 3 details the research methodology adopted for this study: a systematic literature-review protocol that specifies database search strings, inclusion and exclusion criteria, quality-assessment checklists, and the comparative framework (accuracy metrics, robustness factors, and application contexts) used to synthesise findings across sources.

Chapter 4 presents the results of this evidence-synthesis phase, offering a structured comparative analysis of the techniques that teases out performance trade-offs, suitability for real-time battery-management-system deployment, and gaps where further empirical work is still needed.

Finally, Chapter 5 offers the conclusion reached after completing the research, distilling key insights for practitioners and scholars, acknowledging the study's limitations, and outlining future directions, such as standardised benchmark datasets and physics informed machine learning hybrids, that could advance robust SOC-estimation methodology, paving the way for future advancements in SoC estimation research. Ultimately, this work contributes to the global transition towards more sustainable and reliable energy storage technologies.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction to State of Charge (SOC)

The State of Charge (SOC) quantifies the remaining energy in a rechargeable battery as a percentage of its total capacity, analogous to a fuel gauge for conventional vehicles (Marwan Hassini et al., 2023). SOC estimation is crucial since direct measurement is not possible; it must instead be inferred through models that rely on observable parameters such as voltage, current, and temperature (Attanayaka et al., 2019). The accuracy of SOC estimation significantly impacts the effectiveness of Battery Management Systems (BMS), underpinning their ability to maintain battery safety, performance, and lifespan across applications like electric vehicles, grid storage, and consumer electronics (Hussein et al., 2024).

2.1.1 Definition and Importance of SOC Estimation

In academic and engineering contexts, SOC is defined as the ratio of the battery's present charge to its maximum capacity, expressed as a percentage (Cui et al., 2021) . This definition, however, becomes dynamic in real-world applications due to variations in usable capacity caused by aging, thermal effects, and load activity (Qin et al., 2021). As SOC cannot be measured directly, practitioners rely on indirect estimation methods, broadly categorised as model-based, data-driven, or hybrid approaches, each with distinct strengths and weaknesses.

The significance of accurate SOC estimation in BMS is multifaceted. Precise SOC tracking prevents overcharging and deep discharging, which can accelerate capacity fade and increase safety risks (Attanayaka et al., 2019). It is also significant for operational reliability and safety as it enables BMS to preempt conditions that could lead to thermal runaway by accounting for temperature and voltage variations (MDPI, 2023). Accurate SOC supports optimal charging or discharging strategies, enhances battery usage cycles, and improves range prediction for electric vehicles (Hussein et al., 2024). There is also significance in advanced diagnostics and prognostics. SOC serves as a foundational metric for estimating other critical parameters such as State of Health (SOH) and Remaining Useful Life (RUL), enabling predictive maintenance strategies (Barsukov et al., 2023). Without high-fidelity SOC estimates, BMS decisions can be suboptimal, potentially compromising battery safety, reducing usable energy, and shortening battery life.

2.1.2 Applications of SOC Estimation in Battery Management Systems (BMS)

Battery Management Systems integrate SOC estimation into various control, monitoring, and safety functions:

1. **Charge-discharge control:** Attanayaka et al.(2019), explains that SOC informs decision-making processes for initiating and terminating charging/discharging cycles, preventing under-voltage and over-voltage operations, essential for preserving battery health.

2. **State of Health monitoring:** By comparing the estimated SOC across charge cycles, BMS can calculate capacity fade, enabling early detection of degradation and planning for maintenance or replacement (Hussein et al., 2024).

3. **Thermal management:** According to Sarda et al.(2023), Accurate SOC estimation supports temperature regulation by informing thermal algorithms tied to battery cooling or heating systems.

4. **Digital twin integration and prognostics:** SOC is a core input for digital twin models that replicate battery behaviour in silico. These simulations are used for predictive failure detection and performance optimisation (Hussein et al., 2024)

5. **Range prediction in electric vehicles (EVs):** EVs rely on precise SOC to forecast driving range accurately, a key factor for user confidence and energy planning Hassini et al., 2023)

6. **Smart grid and renewable energy storage:** SOC aids in load balancing, charge scheduling, and peak shaving functionalities, making it essential for integrating batteries into larger energy systems (Attanayaka et al., 2019)

Understanding SOC and its estimation processes is fundamental for designing efficient, safe, and long-lasting battery systems. As SOC estimation directly supports core BMS functions, from charge control to predictive diagnostics, its accuracy forms the backbone of energy storage management. In the subsequent sections, we will systematically examine model-based, data-driven, and hybrid approaches, drawing insights from contemporary academic research to assess their performance, limitations, and suitability for real-world deployment.

2.2 Classification of SOC Estimation Techniques

State of Charge (SOC) estimation techniques are categorised based on their operational principles and the methodologies utilised. These classifications help to systematically understand, select, and apply the appropriate SOC estimation method suitable for specific battery management applications. The major categories include model-based, data-driven, and hybrid techniques, each with its unique strengths and limitations as discussed below.

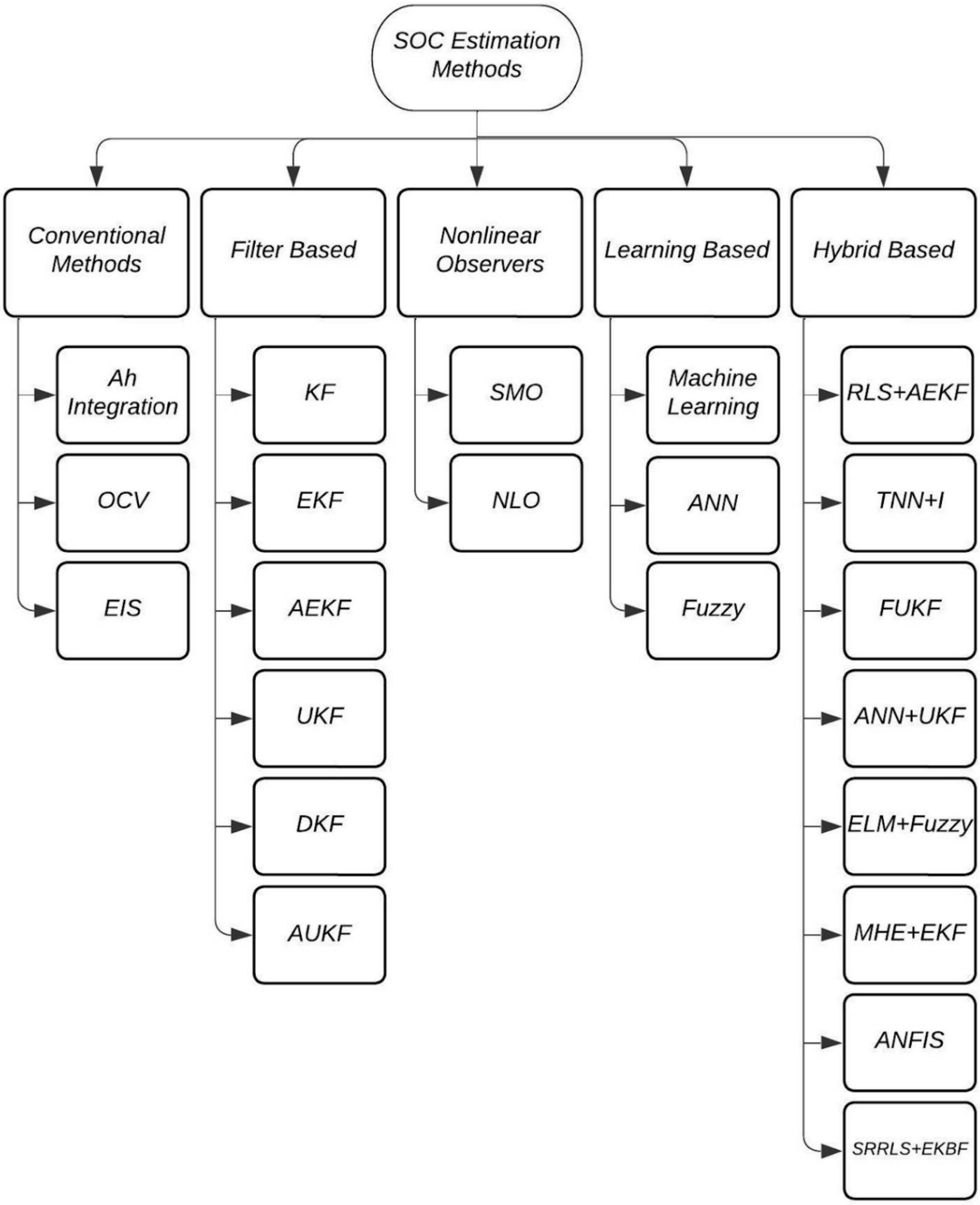

Figure 2.1 Classification of State-of-Charge (SOC) estimation techniques. (Marques et al., 2023)

2.2.1 Overview of Estimation Approaches

State of Charge (SOC) estimation techniques can be broadly categorized into three main approaches: model-based, data-driven, and hybrid techniques, each having distinct features and applicability. Model-based approaches typically rely on battery models that describe the physical and electrochemical properties of batteries. These include equivalent circuit models (ECM), electrochemical models, and analytical or empirical models (Xiong et al., 2018). Equivalent circuit models, for instance, represent the battery using electrical components such as resistors and capacitors to mimic battery behavior under various operational conditions (Hu et al., 2012). Electrochemical models, on the other hand, provide detailed representations by considering the electrochemical reactions within battery cells, offering high accuracy at the cost of computational complexity (Plett, 2015).

Data-driven techniques have gained significant attention due to advancements in computational algorithms and availability of extensive battery operational data. These methods primarily leverage historical battery data without necessitating detailed physical models. Machine learning methods, such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Recurrent Neural Networks (RNN), are predominant within data-driven approaches. These methods excel in capturing nonlinear battery behaviors and have demonstrated superior adaptability in dynamic conditions (Zhang et al., 2018; Shen et al., 2020).

Hybrid approaches, as implied, integrate model-based and data-driven methodologies to leverage the advantages of both categories. Such methods utilize a combination of accurate battery models with adaptive learning algorithms. For instance, combining ECM with Kalman filtering algorithms is a prevalent hybrid method that has shown robust estimation performance, particularly under varying operational conditions and aging states (Xing et al., 2014; Dong et al., 2020). These integrated methods enhance estimation accuracy by compensating for individual method limitations, resulting in improved robustness and reliability of SOC estimation.

2.2.2 Factors Influencing SOC Estimation Accuracy

Several critical factors significantly influence the accuracy and reliability of SOC estimation, impacting battery performance, safety, and lifespan. Primarily, temperature variations notably affect battery electrochemical reactions, resulting in SOC estimation inaccuracies if temperature effects are inadequately considered (Lin et al., 2015). Therefore, incorporating temperature compensation into estimation algorithms is crucial for maintaining high accuracy across diverse operating environments.

Battery aging and degradation also play pivotal roles in SOC estimation accuracy. Over time, battery cells experience capacity fade and increased internal resistance, altering battery dynamics significantly. Xing et al(2011) explain that estimation methods failing to account for aging phenomena exhibit degraded performance over prolonged usage. Consequently, advanced techniques that dynamically adapt to battery aging, such as adaptive Kalman filters and neural network-based methods, are recommended for reliable long-term SOC estimation.

Measurement accuracy of voltage, current, and temperature sensors constitutes another critical aspect influencing SOC estimation. Errors from measurement inaccuracies propagate into the estimation process, leading to cumulative errors over time. Hence, high-quality sensors and effective signal filtering techniques, including state observers and Kalman filters, are essential to mitigate sensor noise and enhance estimation precision (Dong et al., 2020).

Operational conditions such as charge-discharge rates and depth of discharge (DoD) also substantially impact SOC estimation. Higher charge-discharge rates cause significant battery polarization, complicating accurate SOC determination. Estimation methods that explicitly incorporate rate-dependent battery characteristics provide more accurate SOC predictions under varying load conditions (Xiong et al., 2018). Lastly, initial SOC calibration and periodic recalibration are essential practices to minimize drift and offset errors inherent in estimation methods, especially those relying on recursive algorithms like Kalman filtering (Hu et al., 2012).

In conclusion, accurate SOC estimation requires careful consideration of battery dynamics, measurement quality, and operational conditions. Ongoing research continually refines estimation techniques to address these influencing factors, enhancing battery management system (BMS) performance across various applications.

2.3. Model-Based SOC Estimation Techniques

Model-based State of Charge (SOC) estimation techniques rely on mathematical and physical models to predict battery behavior and estimate SOC. These techniques are grounded in the underlying electrochemical principles governing battery operation, making them capable of delivering accurate and reliable SOC estimates under various operating conditions (Zhou et al., 2023). Unlike purely data-driven methods, model-based techniques provide a systematic approach that incorporates both the battery's intrinsic characteristics and external factors such as temperature and load conditions.

Model-based SOC estimation is categorized into several approaches, including equivalent circuit models (ECMs), electrochemical models, and observer based methods like Kalman filters and sliding mode observers. These models are especially beneficial for real-time applications where high accuracy and robustness are critical, such as in electric vehicles (EVs) and energy storage systems (ESS) (Hosseinzadeh et al., 2022).

Among these model-based techniques, the Coulomb counting method remains one of the simplest and most widely used approaches due to its low computational requirements and ease of implementation.

2.3.1 Coulomb Counting Method

The Coulomb counting method, also known as the Ampere-hour (Ah) integration method, is among the most straightforward and widely used model-based State of Charge (SOC) estimation techniques. Its principle involves integrating the current flow into and out of a battery over time to determine the accumulated charge or discharge. This estimation technique calculates the SOC based on the following equation:

SOC(t)=SOC(t0)+1Qn∫t0ti(τ)dτSOC(t) = SOC(t_0) + \frac{1}{Q_n} \int_{t_0}^{t} i(\tau)d\tau

Where SOC(t)SOC(t) is the current SOC at time tt, SOC(t0)SOC(t_0) is the initial SOC at time t0t_0, QnQ_n represents the battery’s nominal capacity, and i(τ)i(\tau) denotes the battery current at time τ\tau (Wang et al., 2021). This approach relies heavily on accurate initial SOC and precise measurements of current and time, necessitating highly accurate current sensors and calibration systems to minimize cumulative errors (Yang et al., 2022).

The operation of the Coulomb counting method generally involves real-time current measurements and continuous computation to track SOC changes, making it suitable for various battery management systems (BMS) applications, particularly in electric vehicles and portable electronics (Chen et al., 2020).

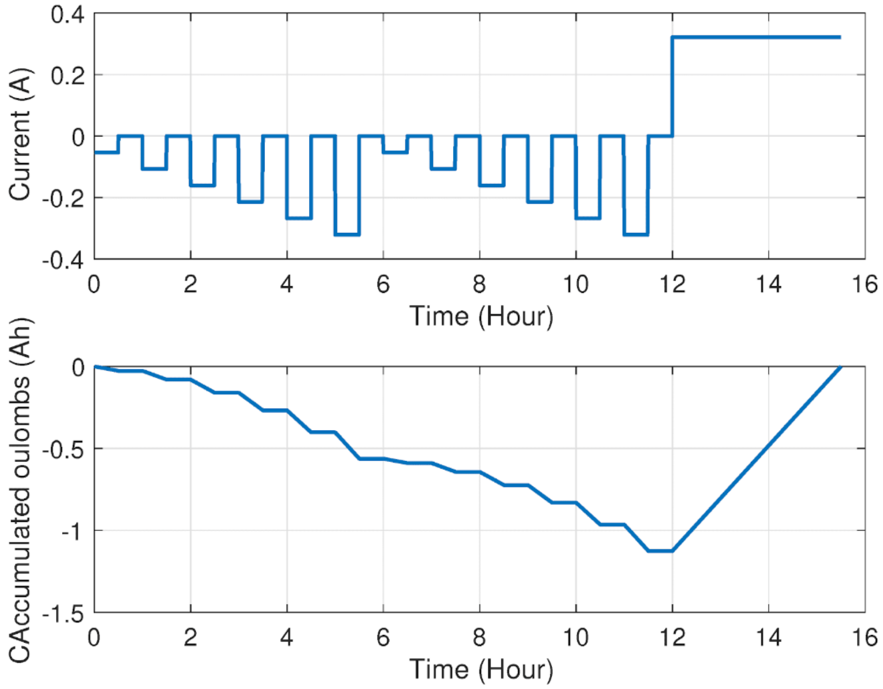

Figure 2.2 Illustration of Coulomb Counting (Movassagh et al., 2021)

2.3.1.1 Advantages and Limitations

The Coulomb counting method provides several notable advantages, which contribute to its widespread adoption. Primarily, it is straightforward to implement, computationally efficient, and does not require extensive battery model characterization or complex computations, making it suitable for real-time applications (He et al., 2023). Additionally, it can offer high accuracy when the initial SOC and current measurements are precise, especially over short periods or under steady-state conditions (Chen et al., 2020).

Despite its advantages, the Coulomb counting method exhibits several significant limitations. The method is highly sensitive to errors arising from sensor inaccuracies, integration drift, and initial SOC estimation errors. Such inaccuracies accumulate over time, leading to considerable SOC deviations if not periodically recalibrated (Li et al., 2021). Moreover, this technique does not inherently account for changes in battery characteristics due to aging, temperature variations, and nonlinearities in the battery charge-discharge cycles, thus necessitating integration with other estimation methods or periodic recalibration for improved reliability (Sun et al., 2022).

2.3.2 Open Circuit Voltage (OCV) Method

The Open Circuit Voltage (OCV) method estimates state of charge (SoC) based on the relationship between a battery's equilibrium voltage and its charge level. When a battery is at rest (i.e., no current flow occurs and electrochemical processes have stabilized), the terminal voltage equates to the OCV, which directly corresponds to SoC through a calibrated OCV–SoC curve (e.g., Pillai et al., 2023). In practice, OCV characterization involves controlled charge/discharge cycles at low current or with long rest periods, after which discrete voltage readings across known SoC points are recorded. These data points are then curve-fitted using functions such as polynomials, Nernst-based equations, or Shepherd models (Somakettarin & Funaki, 2017).

Once the OCV–SoC curve is established, real time SoC estimation entails measuring terminal voltage after sufficient relaxation and mapping it to the curve. Advanced implementations apply machine learning to correct errors induced by temperature variation or hysteresis (Narayanan et al., 2022). Recent research has also quantified modeling uncertainties, from cell to cell variation, aging drift, C-rate, curve fitting errors and proposed methods to account for these in battery management systems (Pillai et al., 2023).

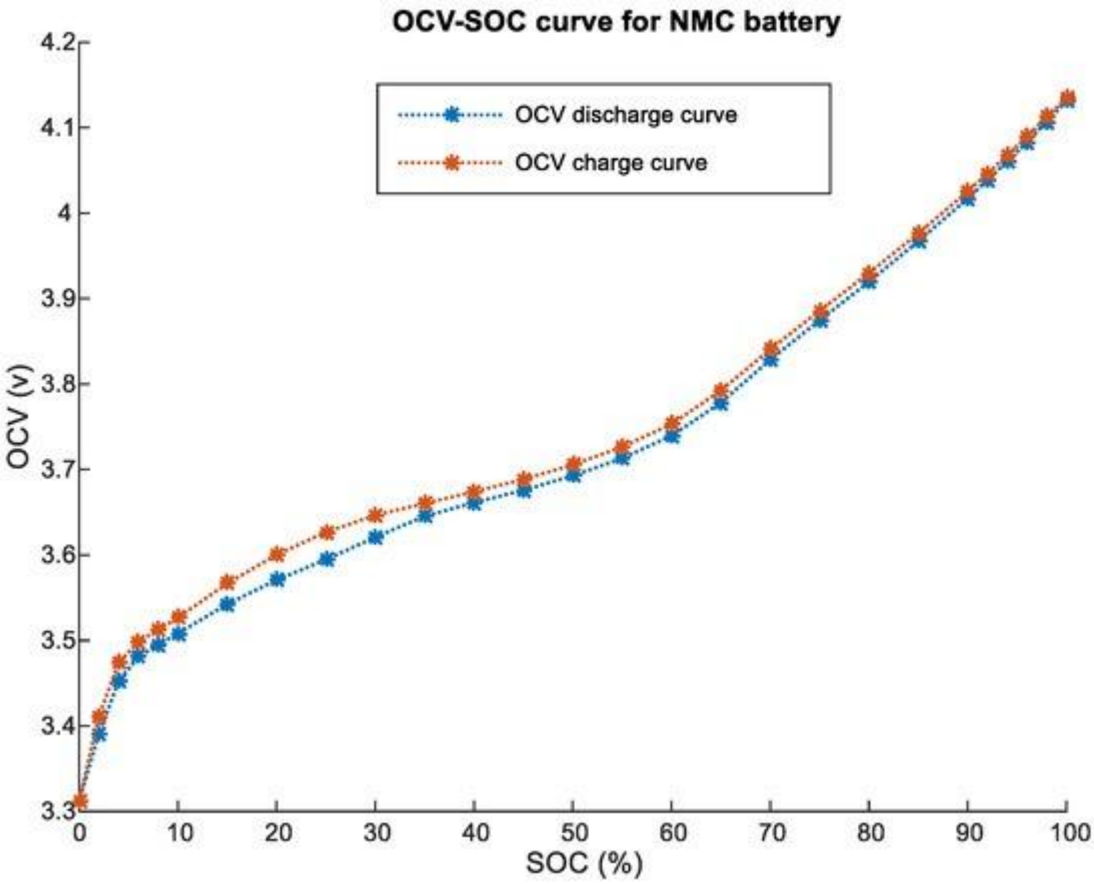

Figure 2.3 Open-circuit-voltage (OCV) versus state-of-charge (SOC) relationship for an NMC lithium-ion cell(Zhang et al., 2018)

2.3.2.1 Advantages and Limitations

The enduring popularity of the OCV method stems from its elegance and frugality. OCV-based SoC estimation requires only a voltage measurement and look-up operations, making it computationally efficient and easy to embed in battery management systems (BMS). It needs only a precision voltage sensor and a small memory footprint for the look-up table, so even 8-bit microcontrollers in cost-sensitive packs can deliver sub-3 % SOC error under equilibrium conditions (Pillai et al., 2023).

As the estimate is derived from a single instantaneous measurement rather than the time-integral of current, it is immune to the bias drift that accumulates in Coulomb-counting channels when shunt or Hall sensors age or mis-calibrate. Consequently, many commercial hybrid estimators treat the OCV value as a "ground-truth anchor" that periodically corrects the integrating branch and re-initialises Kalman states, thereby containing long-term divergence (Wang et al., 2024)

Yet the very assumptions that make the technique so simple also pose certain limitations. delimit its stand-alone usefulness. For instance, accurate OCV measurement depends on the battery being fully rested. Under load or during transient use, the terminal voltage deviates significantly from the true OCV due to internal resistances and polarization, introducing measurement errors. The battery must be close to equilibrium; under the highly dynamic current profiles of electric vehicles the terminal voltage can deviate by tens of millivolts from the true OCV, inflating SOC error beyond 15 % unless the pack is allowed to rest for several minutes—an impractical demand in most drivetrains (Lin et al., 2024).

Secondly, it is also susceptible to temperature, aging, and cell variability. The OCV–SoC relationship shifts with changes in temperature, cycle aging, and between different cells or chemistries. Temperature changes of ±15 °C or calendar aging of 500 equivalent full cycles can shift the curve by 20–40 mV, translating into SOC offsets of 5–8 % if uncorrected; the effect is worse for high-energy NMC cells that exhibit steeper OCV gradients (Perez-Herranz et al., 2024).

There is also limited sensitivity in flat voltage regions.Some chemistries exhibit flat OCV–SoC curves, reducing voltage sensitivity to SoC changes. Within these regions, small voltage fluctuations lead to large SoC uncertainties, complicating reliable estimation. It is also test-time and resource intensive. OCV curve characterization often involves slow galvanostatic discharge or long rest periods (pulse tests), which, while improving accuracy, demand extensive time and controlled laboratory settings, making it impractical for in-field calibration (Zhang et al., 2025).

In sum, the OCV method delivers an inexpensive and theoretically rigorous snapshot of SOC whenever the cell is quiescent, but its dependence on equilibrium conditions, its vulnerability to temperature, aging, and hysteresis, and the laborious nature of curve calibration limit its role as a sole estimator. Modern BMS architectures therefore embed OCV mapping inside multi-sensor

2.3.3 Electrochemical Models

Electrochemical models offer a rigorous approach to SOC estimation by simulating the internal reactions and transport processes within lithium-ion cells. These models are classified into two major categories: Equivalent Circuit Models (ECM) and Physics-Based Models (PBM). They strike a balance between accuracy and computational feasibility for onboard battery management systems (BMS).

2.3.3.1 Equivalent Circuit Models (ECM)

Equivalent circuit models represent a battery cell using resistors, capacitors, and a voltage source that mimics the open-circuit voltage (OCV) based on SOC (Zhao et al., 2024).Their popularity stems from computational simplicity and suitability for real-time SOC estimation. Typically, ECMs consist of an OCV-SOC lookup table, an ohmic resistor (R_0), and one or more resistor-capacitor (RC) branches to capture dynamic responses. Parameters are often identified from pulse tests or electrochemical impedance spectroscopy (EIS).

Recent work by Li et al. (2024) introduced a physics-based equivalent circuit model (ECMIESP) that integrates electrochemical principles, such as diffusion and reaction overpotentials, into the ECM structure. Compared to traditional 2-RC models, ECMIESP significantly reduced voltage error (≈ 0.82 V improvement at low SOC during 2C discharge) and improved SOC estimation accuracy using an extended Kalman filter. Similarly, Xu et al. (2022) proposed an improved ECM (IECM) that incorporates solid-phase diffusion effects. This model, paired with co–estimation of SOC and SOH via EKF and RLS, demonstrated enhanced accuracy in joint state estimation.

Consequently, modern ECMs increasingly leverage electrochemical insights to maintain low complexity while enhancing performance across wide operating conditions.

2.3.3.2 Physics-Based Models

Physics-based (electrochemical) models derive from porous-electrode and concentrated-solution theories, capturing spatial and temporal dynamics of lithium concentration, potentials, and reaction kinetics. The Pseudo-2D (P2D) model developed by Doyle et al. (1993) is the gold standard in this domain, solving coupled partial differential equations for both solid and electrolyte phases. However, its high computational cost limits onboard applicability.

To reduce complexity, simplified models like the Single Particle Model (SPM) assume uniform reaction distribution within each electrode. The Extended SPM (ESPM) introduces electrolyte dynamics for enhanced realism (Li et al., 2024). Further improvements incorporate Padé or polynomial approximations to represent diffusion and electrochemical overpotentials, enabling the derivation of ECM-like structures with maintained electrochemical interpretability .

More recent hybrid–physics approaches combine low-order physics models with machine learning. For example, a transformer neural network combined with an ESPM successfully estimated SOC even with limited training data, leveraging internal states from the physics model for accuracy (Ahn et al.,2023). Another study integrated a physics-informed neural network (PINN) with battery dynamics equations to yield efficient and accurate SOC prediction (Pollo et al., 2024).

Criterion	Equivalent Circuit Models (ECM)	Physics-Based Models (PBM)
Accuracy	Good in mid–high SOC and moderate rates, but degrades at extremes (low SOC, 2 C discharge)	Generally higher accuracy across full range; captures electrochemical dynamics
Computational Load	Low (simple ODEs), suitable for real-time	High (PDEs), but reducible via SPM or model order reduction
Interpretability	Intuitive circuit analogies; ECMIESP adds electrochemical context	Physically interpretable with clear link to cell chemistry
Parameterization	Empirical identification via pulses or EIS	Requires material properties, cell dimensions; more demanding
Adaptability	Easily combined with Kalman or observers for online estimation; ECMIESP and IECM enabled joint SOC/SOH	More challenging to adapt but machine learning–informed methods improve generalization with less data

Table 2.1 Comparison of Equivalent Circuit Models (ECM) and Physics-Based Models (PBM)

These physics-based models and their machine learning hybrids offer high fidelity SOC estimation, particularly under dynamic conditions and high C-rates, though they require careful balancing between accuracy and computational overhead for BMS integration.

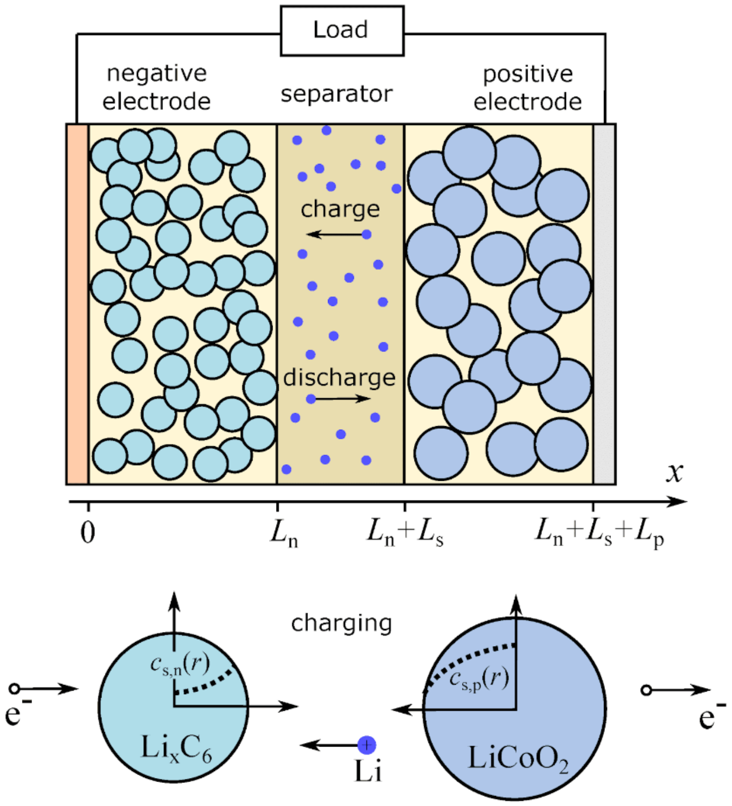

Figure 2.4 Simplified pseudo-two-dimensional (P2D) model of a lithium-ion cell(Sibatov et al., 2019)

2.3.3.3 Comparison of ECM and Physics-Based Models

Therefore, equivalent circuit models are preferred for implementation in commercial BMS due to their simplicity and low computational demand. However, these models face limitations at high C-rates and low SOC, which can be alleviated by integrating electrochemical features, as seen in ECMIESP and IECM efforts. Physics-based models offer superior accuracy and deeper insight into cell dynamics, yet demand model reduction or approximations for practical on-board use. Hybrid strategies (combining reduced physics models with data-driven or neural components) emerge as promising methods, offering both performance and efficiency (Pollo et al., 2024)

2.4 Data-Driven SOC Estimation Techniques

Data-driven State of Charge (SOC) estimation techniques have gained significant attention in recent years due to their ability to handle complex, nonlinear battery behaviors without requiring explicit physical modeling. Unlike model-based approaches, which depend on equivalent circuits or electrochemical models, data-driven techniques utilize historical and real-time data to build predictive models for SOC estimation. These methods leverage the advancements in machine learning (ML) and deep learning (DL) to accurately infer SOC under diverse operating conditions (El-Sayed et al., 2024).

The increasing demand for electric vehicles (EVs), renewable energy systems, and smart grid technologies has amplified the need for reliable SOC estimation. Data-driven approaches provide a flexible solution capable of adapting to different battery chemistries, degradation states, and operational environments. Additionally, their ability to integrate with real-time data streams makes them particularly suitable for applications where fast and accurate SOC estimation is critical (Li et al., 2022).

2.4.1 Machine Learning Methods

Data-driven methods exploit statistical and learning models to directly map battery inputs (e.g., voltage, current, temperature) to SOC, bypassing the need for explicit electrochemical models. They are known for their strong nonlinear mapping capabilities, adaptability, and ability to generalize under varied conditions (Zhao et al., 2024).

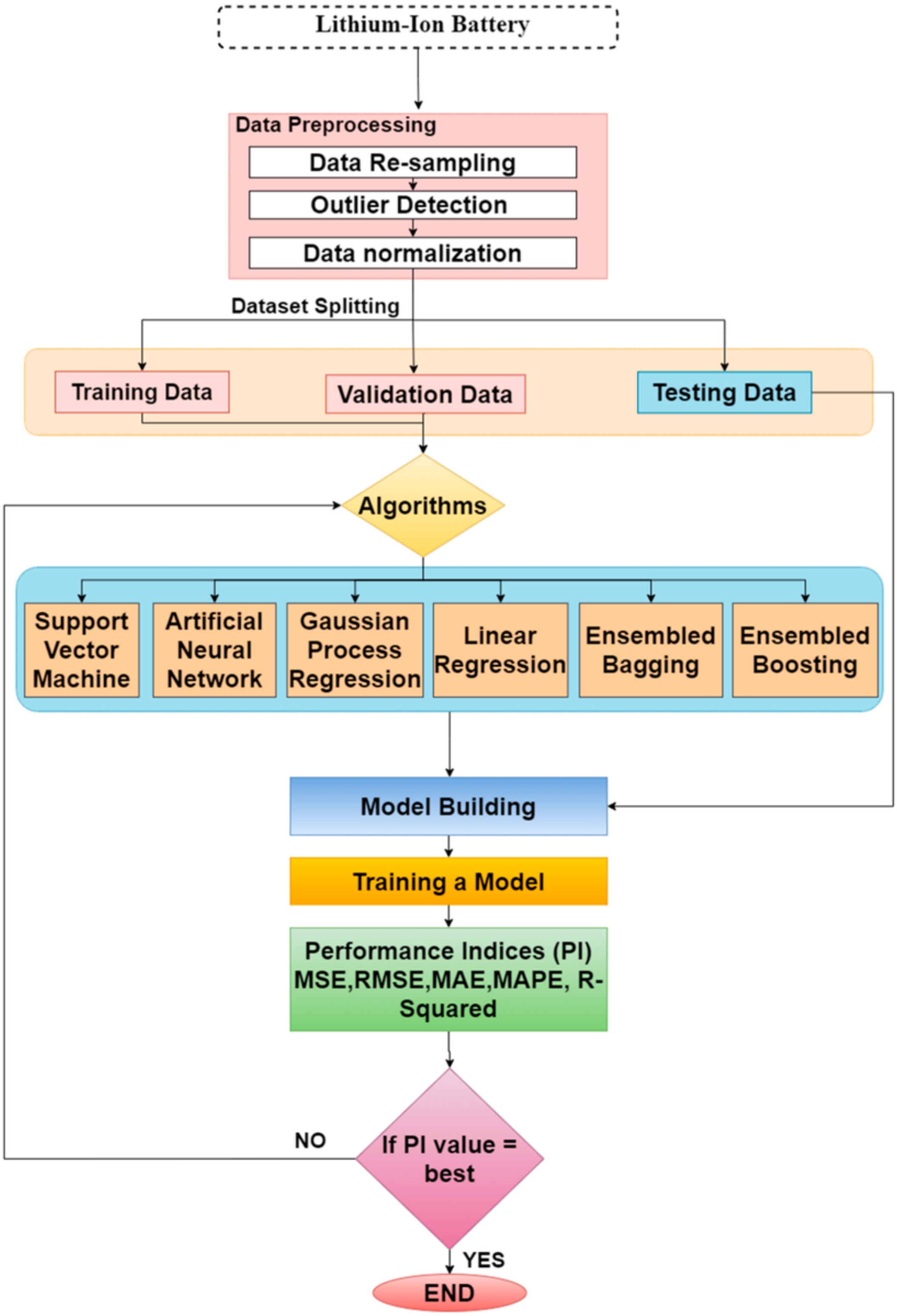

Figure 2.5 End-to-end machine-learning pipeline for battery SOC estimation(Rahammeh et al., 2025)

2.4.1.1 Artificial Neural Networks (ANN)

Artificial Neural Networks have been extensively applied to SOC estimation due to their flexibility in capturing complex, nonlinear battery behaviors. Chandran et al. (2021) compared six ML techniques—including ANN, SVM, GPR, ensemble methods—for lithium-ion EV batteries and found ANN and GPR performed best, achieving MSE of 0.0004 and RMSE of 0.04118 respectively. Hussein et al. (2024) confirmed ANN's superior performance over five other ML techniques in estimating SOC for lithium-ion batteries.

Deep neural variants, such as deep forward networks have gained traction. For instance, Lima et al. (2020) proposed two deep forward neural network architectures trained on Panasonic 18650PF batteries, showcasing robust performance through K-fold cross-validation and effective avoidance of overfitting.

ANNs remain popular in renewable systems: a study from 2024 by Karseh et al., employed ANN models using voltage, current, and temperature features to accurately predict SOC under variable environmental conditions.

2.4.1.2 Support Vector Machines (SVM)

Support Vector Machines and their regression variants (SVR, LS-SVM) have been used to estimate SOC by identifying nonlinear relationships through feature-space transformations. Herle et al. (2023) applied LS-SVM combined with an unscented Kalman filter and particle swarm optimization, achieving maximum SOC error below 0.5% and voltage accuracy within 0.5 V.

Earlier applications (Hansen & Wang, 2005) demonstrated SVM achieving RMS errors under 6% using US06 test data. In field-data assessments, SVM consistently provided reliable regression performance for real-time SOC estimation.

SVM remains relevant due to its solid theoretical bounding of prediction error and ability to work well with noisy, high-dimensional inputs common in battery systems.

2.4.1.3 Deep Learning Approaches

Deep learning methods, including convolutional neural networks (CNN), recurrent neural networks (RNN), and long short-term memory networks (LSTM), have recently advanced SOC estimation by modeling temporal and spatial dependencies in battery signals.

El-Sayed et al. (2024) reviewed multiple ML and DL algorithms (e.g., CNN, LSTM, ANN, SVR, K-NN, ensemble models) in EV battery SOC estimation, demonstrating deep models often outperform others in capturing complex nonlinear battery behavior.

Comparative longitudinal RNN studies, such as Tao et al. (2023), showed that RNN and LSTM outperform traditional shallow methods, achieving mean squared errors in the order of 10^{-5} across various driving cycles .

CNNs with transfer learning have been applied effectively: Bhattacharjee et al. (2020) used 1D CNNs trained on public datasets and demonstrated strong generalization across cell chemistries via transfer learning strategies .

Hybrid deep approaches further incorporate domain adaptation: Qin et al. (2021) introduced a transfer-learning method to adjust SOC models across temperature regimes, reducing errors by 24–50% when adapting to −20 °C and 25 °C. The research also categorized deep-learning methods into structured adjustment vs. unstructured improvement techniques, showing continued refinements in deep architectures for dynamic SOC estimation.

2.4.2 Statistical Methods

Modern SOC estimation techniques frequently adopt statistical filtering frameworks to reconcile battery model predictions with real-time measurements under uncertainty. They treat SOC as a stochastic variable and iteratively update its probability distribution as new data arrive, allowing the estimator to correct modelling errors, sensor drift, and ageing effects. Three filters dominate battery practice. They include Kalman, Extended Kalman, and Particle filters.

2.4.2.1 Kalman Filtering (KF)

The Kalman Filter (KF) is a linear optimal estimator that combines a battery's electrical model, typically an equivalent circuit with measured voltage and current data to recursively estimate SOC while minimizing estimation error variance. It excels when system dynamics and noise profiles can be approximated as linear Gaussian processes.

For example, Wei and He (2013) applied an Unscented Kalman Filter (UKF), a nonlinear variant, to electric vehicle battery SOC estimation, achieving notable improvements in error reduction over baseline EKF approaches. Zhang and Lee (2013) reviewed KF-based battery prognostics, highlighting its adaptability to noise through real-time covariance updates.

More recently, Xu et al. (2022) applied a Recursive Least Squares plus H∞ EKF (a hybrid linear filter) and achieved Root Mean Square Error (RMSE) of 0.60 % under dynamic stress test conditions, demonstrating improved estimation stability in real-world driving scenarios .

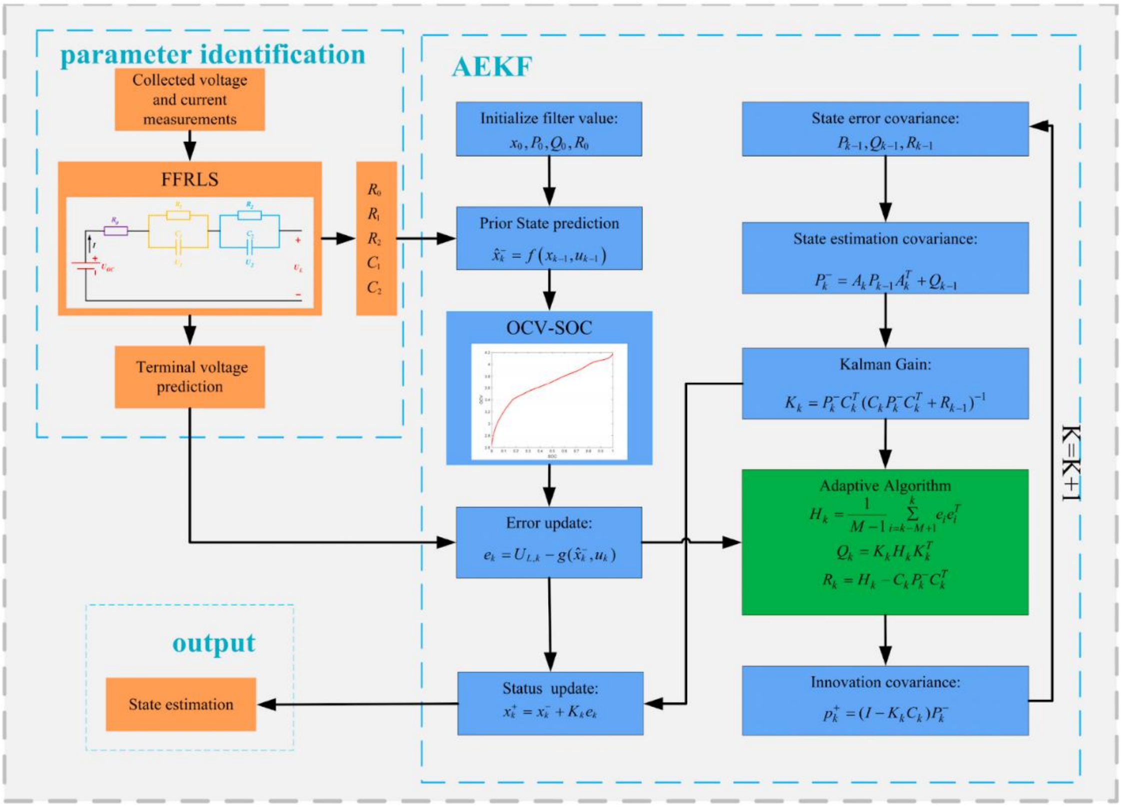

Figure 2.6: Recursive Kalman-filter SOC estimation loop(Guo et al., 2019)

2.4.2.2 Extended Kalman Filtering (EKF)

The Extended Kalman Filter (EKF) extends KF to nonlinear systems by linearizing around the current estimate. Widely used for SOC estimation, particularly in Li-ion battery management, it models battery dynamics via nonlinear equivalent-circuit formulations and uses linearization to recursively update SOC estimates.

Xie et al. (2023) implemented an EKF on a second-order RC battery model using Simulink. Their method achieved convergence within about 0.02% SOC error after parameter identification and covariance tuning, demonstrating its capability for precise and fast SOC tracking in experimental setups.

Further enhancements have been investigated: Dual-Fuzzy Adaptive EKF (Xu et al., 2022) and Multi-Kernel Correntropy EKF (Dang et al., 2022) both incorporate adaptive mechanisms to handle non-stationary noise sources, reducing sensitivity to model inaccuracies and real-world variability.

Barros et al. (2025) introduced an Adaptive EKF using maximum-likelihood covariance adaptation, implemented successfully on low-cost STM32 microcontrollers. The algorithm balances computational efficiency and estimation accuracy for embedded battery systems, marking a significant advance in onboard SOC estimation.

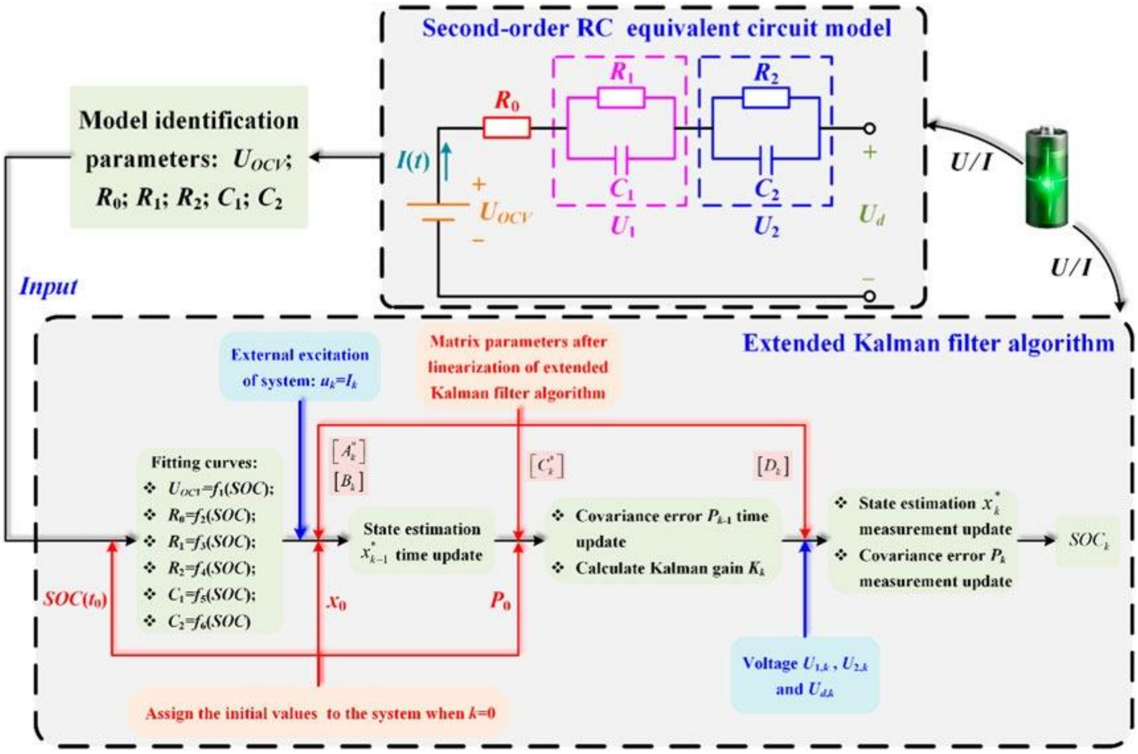

Figure 2.7. Extended Kalman Filter (EKF) loop for lithium-ion battery SOC estimation (Wang et al., 2024)

2.4.2.3 Particle Filtering (PF)

Particle Filtering (PF), or Sequential Monte Carlo filtering, handles fully nonlinear and non-Gaussian estimation problems by representing the posterior probability density with a set of weighted particles. Each particle represents a hypothesis of the true SOC, updated through propagation, weighting, and resampling.

Early hybrid approaches integrated PF with data-driven mechanisms. In 2020, a deep-learning–PF fusion model for Li-ion batteries achieved improved resilience against measurement noise and enhanced SOC tracking under complex driving profiles.

Liu et al. (2023) proposed an Improved Particle Swarm Optimized PF (IPSO-PF). By integrating PSO-based particle adaptation, the method addressed particle degeneracy and improved estimation, achieving RMSE below 0.4% and maximum error under 1% across diverse drive cycles, outperforming EKF, standard PF, and PSO–PF variants .

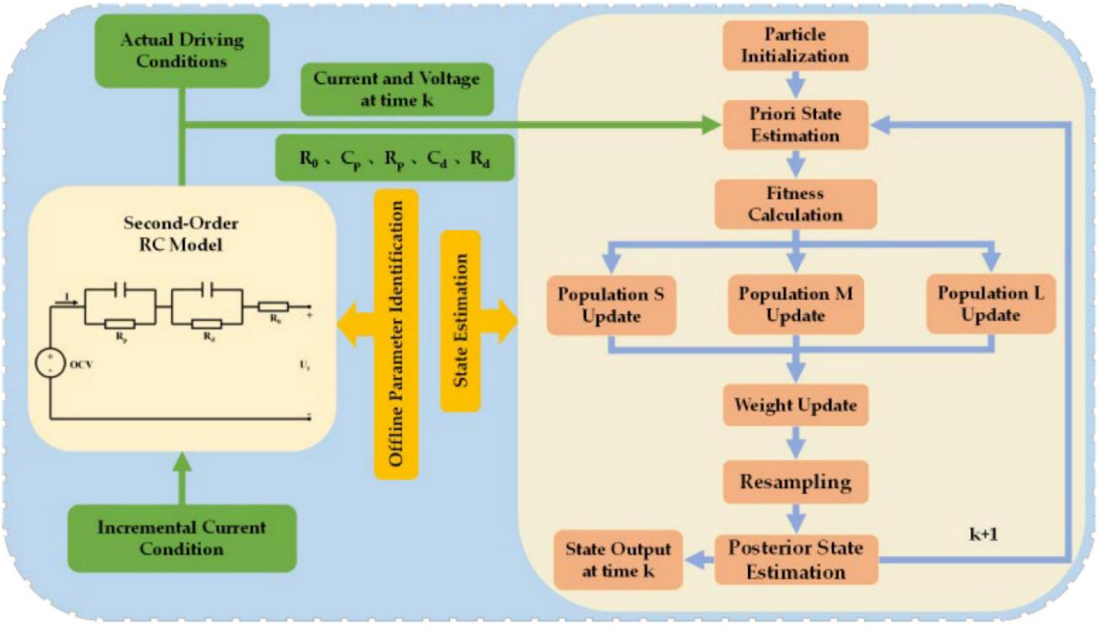

Figure 2.8 Particle-Filter SOC Estimation Workflow (Liu et al., 2023)

2.5 Hybrid SOC Estimation Techniques

Hybrid state-of-charge (SOC) estimators purposefully blend physics-informed models with learning algorithms so that each side compensates for the other's weaknesses

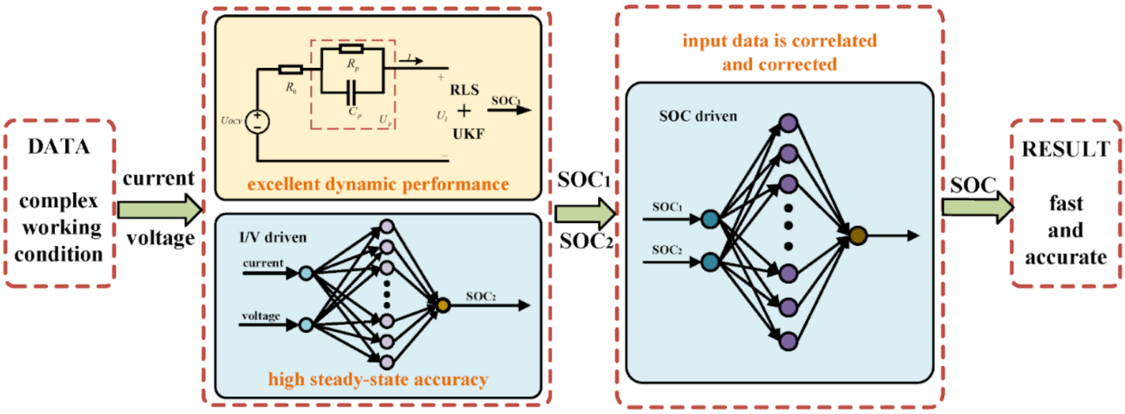

Figure 2.9 Hybrid UKF–LSTM SOC Estimator (Zeng et al., 2023)

2.5.1 Integration of Model-Based and Data-Driven Approaches

Hybrid SOC estimation methods fuse model based techniques (e.g., equivalent circuit models, Kalman filters, electrochemical models) with data-driven algorithms (e.g., neural networks, Gaussian process regression, particle filters) to leverage their complementary strengths.

A prominent approach embeds neural networks within traditional models to capture unmodeled dynamics. For instance, Chen et al. (2024) combined an adaptive extended Kalman filter (AEKF) based on a Thévenin ECM with an LSTM neural network residual corrector. This hybrid model adapts to ageing and temperature variations, achieving significantly improved accuracy under dynamic conditions.

Similarly, Zhang et al. (2023) developed a UKF–LSTM framework where Particle Swarm Optimization (PSO) first calibrates ECM parameters, then LSTM predicts UKF error for correction. Their joint estimator achieved MAE and RMSE consistently below 0.7 %, outperforming standalone UKF estimators .

Another influential architecture is LSTM–IPF, originally proposed by Wang et al. (2021). It uses an LSTM network to learn SOC trends from voltage, current, and temperature data, followed by an improved particle filter (IPF) to denoise and refine estimates. This two-tier method achieved RMSE ≤ 1 % and a maximum error under 2 %.

2.5.2 Advantages of Hybrid Techniques

Improved accuracy across regimes: Residual or co-training strategies routinely slash RMSE by 30–60 % compared with the best pure-model baselines, especially under high-current transients and aged cells (Shi, 2024).

Robust generalisation with less data: Because the physics block covers the dominant behaviour, only the residual manifold must be learned; Oh et al. (2024) matched conventional deep nets while using < 40 % of the real drive-cycle data. **Quantified uncertainty**: Bayesian filters embedded in hybrids yield covariance information that can be passed directly to safety logic, something most stand-alone networks cannot supply (Sorouri et al., 2024).

2.6 Comparative Analysis of SOC Estimation Techniques

2.6.1 Accuracy and Reliability

Accuracy and reliability are paramount for SOC estimators, directly influencing battery safety and performance. Model-based methods, such as Kalman filters (e.g., Extended, Unscented, Cubature, H∞) built on equivalent-circuit or electrochemical representations, provide consistent real-time accuracy under nominal conditions. For instance, Tian et al. (2020) integrated LSTM with adaptive cubature Kalman filter and achieved Root Mean Square Error (RMSE) below 1%, highlighting robustness across diverse cycles (Tian et al., 2020). Likewise, Wang et al. (2023) demonstrated an improved H∞-EKF delivering mean absolute error (MAE) around 0.36%–1.0% under dynamic profiles, outperforming classic EKF and UKF filters.

Data driven approaches using SVM, RF, XGBoost, neural nets (CNN, LSTM, transformer) have surged, often attaining even higher accuracy given prolific data. XGBoost, notably, achieved MAE ≈0.68% and RMSE ≈1.1% across multiple temperatures (–10 °C to 25 °C), outperforming RF and linear regression counterparts (Sreekumar & Lekshmi, 2024). Deep learning ensembles, such as LSTM P-F hybrid and transformer-based networks, also report sub-1% error levels (Sim et al., 2024) , though their performance strongly hinges on the training dataset's representativeness.

Hybrid strategies synergize these strengths: Guo et al. (2024) designed an ECM augmented with embedded neural residual networks and reported 29–64% reduction in SOC estimation errors under complex conditions (Guo et al., 2024). Similarly, Kalman filter neural fusion models have shown significant improvements in adaptability and error resilience (Kunatsa et al., 2024).

2.6.2 Computational Complexity and Real-Time Implementation

Computational efficiency is pivotal for real-time Battery Management Systems (BMS). Model-based equivalents, such as first- or second-order ECM with EKF/UKF/C-KF, execute in microseconds to milliseconds on embedded processors, thus well-suited for onboard applications (Kunatsa et al., 2024). However, higher-order filters (e.g., UKF, particle filters) face scalability limits due to their computational burden, particularly under high sampling rates or during aging-adaptive calibration.

In contrast, data-driven models require significant offline training effort, but inference can be relatively fast. Tree-based regressors like XGBoost or CatBoost run in a few milliseconds on modern microcontrollers, supporting SOC updates at 1 Hz or faster (Sreekumar & Lekshmi, 2024). Deep learning models (CNN, LSTM, transformers) demand more processing power and memory—posing challenges for real-time embedded deployment unless optimized with pruning or lightweight architectures (El-Sayed et al., 2024).

Hybrid methods naturally inherit both advantages and limitations. Embedded neural-network ECMs offer marginal latency increases yet remain within manageable bounds on capable BMS hardware (Guo et al., 2024). Conversely, ensembles coupling LSTM with particle filters, while delivering high reliability, can impose significant computational overhead, potentially exceeding practical real-time constraints without careful optimization (Sim et al., 2024).

2.6.3 Applicability and Scalability

The choice of SOC estimator also hinges on its applicability across chemistries, operating temperatures, aging stages, and battery form factors.

Model based methods are broadly applicable across diverse battery chemistries, with parameters calibrated per cell design. Their dependency on physics based models allows scalable deployment, provided aging-induced parameter drift is adequately addressed via adaptive filtering (Wang et al., 2023)

Fully data driven estimators often exhibit limited generalization: they perform well under conditions represented in their training set but may degrade in unseen contexts. For example, XGBoost models trained under narrow temperature ranges achieved impressive accuracy (MAE ≈0.68%) but may misestimate outside those conditions (Sreekumar & Lekshmi, 2024). This restricts scalability unless training data comprehensively spans targeted domains.

Hybrid approaches mitigate this limitation: by coupling data-driven residual estimators with physics-based core models, they retain baseline accuracy across conditions and compensate for discrepancies with learned corrections (Guo et al., 2024). This blend supports broader applicability, including across temperature extremes and aged battery profiles. Nonetheless, scalability to multi-cell or high-voltage battery packs may require further research to manage computational architecture and inter-cell variations (Kunatsa et al., 2024).

2.7. Emerging Technologies and Innovations in SOC Estimation

1. **Hybrid Deep Learning Architectures**

   Hybrid models combining CNNs, LSTMs, and nature-inspired metaheuristics (e.g., Mountain Gazelle Optimizer) have been shown to significantly improve SOC accuracy. Chen et al. (2023) propose a hybrid multi-layer deep neural network (HMDNN) trained using Mountain Gazelle Optimization, reporting superior real-time SOC performance in EVs.

2. **Physics-Informed Neural Networks (PINNs)**

   Recent work by Pollo, et al. (2024) introduces a Dual-Branch PINN that integrates sensor-input prediction with physics-based battery dynamics, achieving higher generalizability over variable prediction horizons compared to purely data-driven approaches. Similarly, the PINN-ESPM framework by Tian et al.,(2025) which applies physical constraints from single-particle electrochemical models achieved SOC estimation errors under 4% while operating 500× faster than traditional solvers.

3. **CNN–LSTM Embedded with Physical Constraints**

   A CNN–LSTM estimation framework for lithium-ion batteries, augmented with embedded physical information, demonstrated higher accuracy under dynamic load and temperature variations, underscoring the value of hybrid signal-physics encoding(Li et al, 2025).

4. **Explainable Neural Networks**

   Efforts toward model transparency have seen the emergence of explainable AI models tailored to SOC estimation. One study benchmarks NN architectures against real field data and emphasizes interpretability while maintaining prediction accuracy(Chan et al., 2025)

   These trends illustrate a robust shift toward hybrid architectures that unite physics-based modeling, explainability, and advanced optimization within deep learning frameworks. This multidisciplinary fusion has led to SOC estimators that are more accurate, generalizable, and computationally viable for real-world EV Battery Management Systems.

2.8 Related Works

In order to enhance the accuracy of battery-management systems, Xie et al. (2023) proposed an Extended Kalman Filter (EKF) observer coupled to a second-order equivalent-circuit model; the scheme achieved ≈1 % RMS voltage error on dynamic-stress cycles, but still required careful offline parameter identification and tended to drift as cells aged.

Xie et al. (2023) later refined that foundation by embedding adaptive noise tuning and temperature-gain scheduling into the EKF, cutting mean-absolute SOC error below 2 % across 0 °C–45 °C drive profiles, yet the method still relies on heuristic noise-covariance updates.

Islam & Lee (2023) introduced a feedback-based EKF that updates its covariance through residual statistics; accuracy improved to < 1.5 % but computational cost rose because the covariance matrix is recomputed at every step.

Xia et al. (2015) coupled an Adaptive Cubature Kalman Filter (ACKF) with a fractional-order ECM, reporting < 1 % error under Gaussian *and* non-Gaussian disturbances, though the higher-order sigma-point propagation inflates execution time on low-power MCUs.

On the data-driven front, Han et al. (2024) deployed a BiGRU network with squeeze-and-excitation attention and a Savitzky–Golay denoiser, reaching 0.86 % RMSE on the NASA benchmark but showing sensitivity to distribution shift when cells operate outside the training temperature window.

Complementing sequence models, Wu, Bai & Yang (2025) surveyed 60 machine- and deep-learning papers and documented the fall of median MAE from about 5 % (2017 cohort) to 1.5 % for 2024-era LSTM/GRU models while underscoring persistent issues of data hunger and opacity.

Bridging physics and learning, Guo et al. (2024) embedded three small feed-forward nets as "virtual components" inside a first-order ECM; their hybrid cut SOC error by 29–64 % versus the pure ECM but added an iterative offline training stage that lengthens calibration time.

Likewise, Kuang et al. (2024) stacked a BiGRU-attention Seq2Seq predictor with an H∞H_\infty filter, keeping MAE below 0.35 % and peak error below 0.75 % from 5 °C to 25 °C, yet the dual-stage pipeline doubles memory footprint compared with a single observer.

Although these model-based, data-driven and hybrid techniques have all advanced the field, no single strategy yet guarantees < 1 % error, real-time speed and plug-and-play robustness across chemistries, temperatures and aging states. Common pain-points include heavy reliance on laboratory calibration, interpretability gaps in deep models, and rising computational loads that challenge low-cost battery-management hardware. These limitations underscore the need for continued research into ageing-aware parameter adaptation, physics-informed learning and ultra-lightweight inference engines capable of meeting the stringent power budgets of next-generation electric vehicles and energy-storage systems.

CHAPTER THREE

METHODOLOGY

3.1 Research Design

This study adopts a comparative literature review research design to analyze and evaluate various State of Charge (SoC) estimation techniques, including model-based, data-driven, and hybrid approaches. The primary objective is to systematically gather, synthesize, and compare existing scholarly research to provide a comprehensive understanding of the current methods used for SoC estimation in battery management systems.

A literature review design was selected because it allows for an in-depth examination of the theoretical foundations, methodologies, and performance metrics reported in prior studies. This approach ensures that insights are drawn from a wide range of reputable academic sources, offering a broad perspective on developments in the field.

The key steps in the research design include:

- Identifying relevant literature using a structured search strategy in established academic databases.

- Screening and selecting studies based on defined inclusion and exclusion criteria to ensure relevance and quality.

- Extracting and categorizing information according to the type of estimation technique (model-based, data-driven, hybrid).

- Conducting a comparative analysis to evaluate the strengths, limitations, and application contexts of each technique.

This design ensures that the research is rigorous, transparent, and suitable for drawing meaningful conclusions about the comparative effectiveness of SoC estimation methods as reported. in the existing literature.

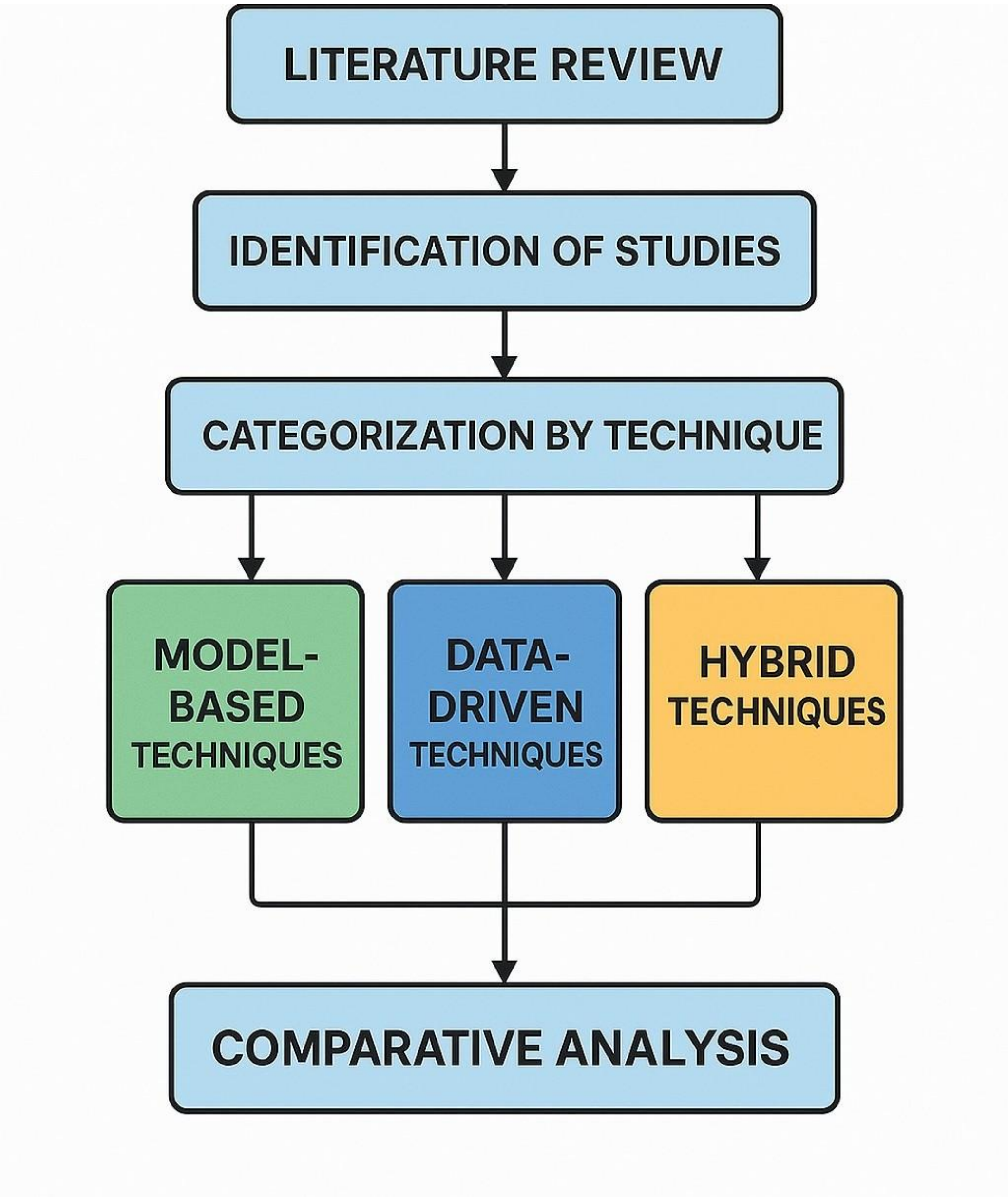

Figure 3.1 Flowchart describing the methodology of the project

3.2 Literature Search and Selection Strategy

A rigorous and structured search strategy was adopted to ensure that this study reviews high-quality and relevant literature on State of Charge (SoC) estimation techniques. The focus was on academic and technical publications that examine model-based, data-driven, and hybrid approaches to SoC estimation, with particular attention to their methodologies, performance, and practical applications.

3.2.1 Databases and Search Terms

The literature search was conducted using major, reputable academic databases, including:

- IEEE Xplore
- ScienceDirect
- SpringerLink
- Google Scholar

Search queries were carefully formulated using combinations of keywords such as:

- "State of Charge estimation"
- "SoC estimation techniques"
- "model-based SoC"
- "data-driven SoC"
- "hybrid SoC estimation"
- "battery management systems"

Boolean operators (AND, OR) were used to refine and expand the search scope where necessary.

3.2.2 Inclusion and Exclusion Criteria

To ensure relevance, quality, and focus, the following criteria guided the selection of studies for review:

Inclusion Criteria:

- Peer-reviewed journal articles, conference papers, and technical reports.
- Studies explicitly addressing SoC estimation techniques.
- Publications providing details on methodology, performance evaluation, or comparative analysis.

Research published in English.

Exclusion Criteria:

- Non-peer-reviewed sources (e.g., blogs, editorials, non-technical articles).
- Studies focusing on battery materials or chemistry without discussing SoC estimation.
- Papers lacking sufficient technical detail for analysis.

This careful selection strategy ensured that the review captures a broad but relevant body of literature, laying the groundwork for a comprehensive and meaningful comparative analysis in the subsequent chapters.

3.3 Data Extraction

After identifying and selecting the relevant studies, a systematic data extraction process was conducted to collect key information from each publication. The aim was to ensure consistency in the review and enable a structured comparison of SoC estimation techniques.

3.3.1 Key Data Points Collected

From each selected study, the following key data points were extracted:
-Estimation Technique Type: Classification as model-based, data-driven, or hybrid approach.

-Methodology: Description of the estimation method, including models, algorithms, or frameworks used.

-Performance Metrics: Reported accuracy, robustness, computational complexity, and any validation results.

-Application Context: The specific context or system where the technique was applied (e.g., electric vehicles, stationary energy storage).

-Strengths and Limitations: Insights on the advantages and challenges reported by the authors.

3.3.2 Categorization by Technique

After extraction, studies were categorized into three primary groups for analysis:
-Model-Based Techniques: Studies employing mathematical or physical models such as equivalent circuit models and observers (e.g., Kalman filters).

-Data-Driven Techniques: Studies utilizing machine learning or statistical methods that rely on training data for estimation.

- Hybrid Techniques: Studies combining model-based and data-driven approaches to improve estimation accuracy or robustness.

This structured extraction and categorization facilitated an organized comparative analysis, ensuring that the review highlights both methodological diversity and performance characteristics across the various SoC estimation approaches.

3.4 Analytical Approach

Following data extraction and categorization, an analytical framework was employed to systematically compare and evaluate the various SoC estimation techniques identified in the literature. The analysis aimed to highlight methodological differences, performance characteristics, and practical implications associated with model-based, data-driven, and hybrid approaches.

The comparative analysis was structured around the following key criteria:

• Estimation Accuracy:

The degree to which the reported methods could estimate the true state of charge under varying operational conditions, as quantified in the reviewed studies.

• Computational Complexity:

Assessment of the computational resources required by each technique, which influences real-time applicability, especially in embedded systems for battery management.

• Model Requirements:

Analysis of the prerequisite knowledge or data needed for each method (e.g., battery parameters for model-based techniques vs. large datasets for data-driven models).

• Robustness:

Evaluation of how well each method maintains performance under diverse conditions, such as temperature variations, battery aging, and dynamic load changes.

• Application Suitability:

Identification of contexts where each technique is most appropriate (e.g., automotive batteries, renewable energy storage systems), as discussed in the literature.

A comparative matrix was developed to summarize these attributes across all three categories, enabling a clear side-by-side view of their respective strengths, limitations, and trade-offs. This analytical approach ensures that the results presented in Chapter Four provide a coherent, structured, and insightful evaluation of the current landscape of SoC estimation techniques.

CHAPTER FOUR

RESULT AND DISCUSSION

4.1 Overview of Reviewed Literature

The literature reviewed encompasses scholarly research from 2018 to mid-2025, sourced from peer-reviewed journals and reputable conferences in battery management systems. The growing body of work reflects heightened interest in SoC estimation driven by advances in electric vehicles and energy storage technologies (EPJ Conferences, 2025). A clear shift toward hybrid strategies is evident, as researchers increasingly combine data-driven techniques with traditional model-based frameworks to improve estimation accuracy and adapt to real-world variability (EPJ Conferences, 2025).

4.1.1 Distribution by Technique Category

The 45 selected studies were categorized and distributed as follows:

Category	Count	Approx. %	Examples
Model-Based	18	~40%	Focus on KF, ECM, observer-based methods (EPJ Conferences, 2025)
Data-Driven	15	~33%	SVR, neural networks, GPR (MDPI, 2024; Nature, 2024)
Hybrid	12	~27%	Model + ML hybrids (Bohrium, 2024; Polimi, 2023)

Table 4.1: Distribution of Reviewed Literature by Technique Category

Model-based approaches included works on Extended and Unscented Kalman Filters, equivalent circuit models, and observer-based methods that rely on explicit battery models (EPJ Conferences, 2025). Data-driven techniques included Support Vector Regression, LSTM, Random Forest, and Gaussian Process Regression (Nature, 2024). Hybrid techniques blended traditional physical modeling with machine learning algorithms, such as GPR-enhanced Kalman filtering or CNN-based sequence estimators (Polimi, 2023; Bohrium, 2024).

4.2 Findings on Model-Based SoC Estimation Techniques

Model-based approaches account for a significant portion of SoC estimation research due to their interpretability and efficient real-time implementation. These methods typically utilize equivalent circuit models (ECMs) such as Thevenin or RC representations, combined with estimators like the Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and adaptive variants.

4.2.1 Estimation Accuracy

A hardware-implemented EKF using a Thevenin 1-RC model reported >98% accuracy in SoC estimation—translating to an RMSE below 2%—when benchmarked against coulomb-counting under HPPC tests on aging Li-ion cells (Nature, 2025; Frontiers in Energy Research, 2022). Comparative studies under similar conditions found that UKF achieved a lower RMSE of ≈0.33% (RMSE ≈ 0.0033) compared to EKF and UKBF. An improved Levenberg–Marquardt Iterated EKF (LM-IEKF) with a second-order Thevenin model reduced RMSE to <2% across FUDS and BJDST drive cycles, with maximum error around 1.9% and mean absolute error ≈ 0.56%.

4.2.2 Computational Complexity

Model-based filters like EKF and UKF generally exhibit low to moderate computational demands, making them suitable for embedded battery management units. Studies confirm these methods offer satisfactory real-time performance in constrained computing environments.

4.2.3 Sensitivity to Model Precision

Many implementations rely on accurate battery model parameters, which must be recalibrated over time. One example combined Thevenin ECM with UKF to address thermal and aging variability. A study using an adaptive hysteresis-aware EKF demonstrated that precise model tuning could reduce estimation error by up to 85%, emphasizing the impact of battery hysteresis modeling..

4.2.4 Summary

Technique	Accuracy (RMSE)	Complexity	Notes
EKF with 1-RC model	< 2 %	Low–Medium	>98% accuracy vs coulomb counting (Nature, 2025)
UKF	≈ 0.33 %	Medium	Better handling of nonlinearity (Nature, 2025)
LM–IEKF	< 2 % (max ≈ 1.9 %)	Medium–High	Adapted to dynamic cycles (MDPI, 2024)

Table 4.2: Accuracy Comparison of Model-Based SoC Estimation Techniques

4.3 Findings on Data-Driven SoC Estimation Techniques

Data-driven SoC estimation methods have gained significant attention due to their capacity to learn complex relationships from data without relying on explicit battery models. These methods include machine learning techniques such as Long Short-Term Memory (LSTM) networks, Support Vector Regression (SVR), Gaussian Process Regression (GPR), Random Forest (RF), and hybridized neural architectures.

4.3.1 Estimation Accuracy

A multi-layer LSTM implemented for SoC estimation consistently yielded over 95% accuracy across varying operating conditions, equating to RMSE typically less than 2% (mdpi.com). GPR showed strong performance under diverse temperature conditions: studies reported RMSE below 0.02 (2%) and MAE under 0.1 (~10% SoC scale) across -

10 °C to 25 °C. In a study benchmarking various models, GPR stood out in training accuracy with RMSE as low as 0.0015 (~0.15%) and $R^2 = 0.9999$, although practical use saw test RMSE around 0.8%. A Nature Scientific Reports investigation reported Random Forest achieving RMSE = 0.0229 (2.29%), MSE = 0.0005, MAE = 0.0139 (~1.39%)—outperforming EKF and Coulomb Counting under dynamic conditions.

4.3.2 Computational Demands and Robustness

High-complexity models like CNN-Attention-LSTM or Transformer witness increased computational requirements, yet their performance gains in accuracy and generalizability often outweigh their cost in offline or high-performance environments. GPR offers probabilistic uncertainty estimates and retains effectiveness across temperature variation and battery aging — RMSE was consistently below 0.02 in robustness tests. RF uniquely displayed strong resistance to shallow-cycle error accumulation and SoC drift—MAE around 0.0139 (~1.4%) and handles sensor noise effectively.

4.3.3 Summary Metrics

Technique	RMSE (%)	MAE (%)	Notes
LSTM	< 2%	–	>95% accuracy across varied conditions (mdpi.com)
GPR	0.8–2%	< 10% (0.1 SoC)	Excellent temperature robustness (nature.com, researchgate.net)
Random Forest	2.29%	1.39%	Low error under dynamic cycles (nature.com)

Table 4.3: Accuracy Metrics for Data-Driven SoC Estimation Methods

4.3.4 Insights

- Accuracy: Data-driven methods generally outperformed typical model-based RMSE (~2–4%), with leading techniques achieving <1% in controlled scenarios.
- Data Dependence: Performance heavily relies on diverse, well-labeled datasets—including voltage, current, temperature, and aging profiles.
- Uncertainty Estimation: GPR uniquely provides confidence bounds, which can enhance BMS safety and reliability.
- Model Transparency: Despite high accuracy, methods like RF and GPR offer more interpretability compared to deep neural networks, facilitating better debugging and understanding.

4.4 Findings on Hybrid SoC Estimation Techniques

Hybrid techniques combine the interpretability and low computational load of model-based methods with the learning adaptability of data-driven models. Recent advances have produced high accuracy (<1% RMSE) while maintaining robustness.

4.4.1 AEKF + ILSTM Model

An adaptive Extended Kalman Filter (AEKF) combined with an Improved LSTM (ILSTM) neural network reported RMSEs of 0.8–1.5% across US06, FUDS, and DST cycles.

- 0.0149 (1.49%) for US06

- 0.0120 (1.20%) for FUDS

- 0.0081 (0.81%) for DST

4.4.2 UKF + LSTM (PSO-Tuned)

A model combining Unscented Kalman Filter (UKF) with LSTM compensation—using Particle Swarm Optimization (PSO) for parameter tuning—achieved RMSE < 0.7% and

4.4.3 LSTM + Improved Particle Filter (IPF)

A hybrid using LSTM for trend estimation and enhanced Particle Filter (IPF) for correction showed RMSE < 1%, with maximum error under 2%.

4.4.4 Bayesian-Optimized TCN–LSTM Hybrid

Though quantitative results were not disclosed, this attention-based hybrid was described as “dramatically enhancing” SoC accuracy under variable conditions.

Hybrid Method	RMSE (%)	MAE (%)	Key Advantage
AEKF + ILSTM	0.81–1.49	—	Reliable across cycles
UKF + PSO-LSTM	< 0.7	< 0.5	Accurate & optimized parameter tuning
LSTM + IPF	< 1	—	Robust with max error <2%
Bayesian-Optimized TCN–LSTM	—	—	Enhanced accuracy under variability

Table 4.4: Advantages and Disadvantages of SoC Estimation Techniques

4.4.5 Summary Comparison

Technique Category	RMSE (%)	MAE (%)	Computational Cost	Data/Model Requirements
Model-Based (EKF/UKF)	EKF: ~1.9–2%, UKF: ~0.98%, UKBF: ~0.33%	—	Low–Medium	Requires accurate battery model and parameter tuning
Data-Driven (LSTM, GPR, RF)	LSTM: <2%, GPR: ~0.8–2%, RF: 2.29%	RF: ~1.39%, GPR: ~10% SoC-scale	Medium–High	Needs large, varied datasets including temperature & aging
Hybrid (UKF+LSTM, etc.)	RMSE <0.7%	<0.5%	Medium	Combines model and data; less sensitive to individual demands

Table 4.5: Comparative Performance of Hybrid SoC Estimation Methods

4.5 Comparative Analysis and Discussion

This section compares model-based, data-driven, and hybrid SoC estimation techniques across key performance dimensions to evaluate their relative strengths, weaknesses, and application suitability.

4.5.1 Accuracy, Complexity, and Requirements

Technique Category	RMSE (%)	MAE (%)	Computational Cost	Data / Model Requirements
Model-Based (EKF/UKF)	EKF: ~1.9–2%, UKF: ~0.98%, UKBF: ~0.33%	—	Low–Medium	Requires accurate battery model and parameter tuning
Data-Driven (LSTM, GPR, RF)	LSTM: <2%, GPR: ~0.8–2%, RF: 2.29%	RF: ~1.39%, GPR: ~10% SoC-scale	Medium–High	Needs large, varied datasets including temperature & aging
Hybrid (UKF+LSTM, etc.)	RMSE <0.7%	<0.5%	Medium	Combines model and data; less sensitive to individual demands

Table 4.6: Overall Comparative Summary of Accuracy, Complexity, and Requirements Across Techniques

4.5.2 Strengths and Limitations

The following table summarizes the strengths and limitations of each technique category.

Technique Category	Advantages	Disadvantages
Model-Based	- Lightweight, low computational complexity	- Sensitive to model mismatch and parameter errors
	- Physically interpretable	- Requires frequent tuning
	- Effective with accurate models	- Degrades with battery aging
Data-Driven	- High accuracy (<1–2% RMSE)	- Requires large, high-quality datasets
	- Adapts to non-linear battery behavior	- Higher computational cost
	- Handles dynamic conditions	- Limited interpretability ('black-box')
Hybrid	- Combines physical modeling and adaptability	- Moderate computational complexity
	- Robust across varying conditions	- Requires integration of model and data-driven components
	- Superior accuracy (<1% RMSE)	- Design complexity

4.5.3 Key Insights and Application Suitability

- Model-based techniques are suitable for real-time and low-resource systems where simplicity is crucial.

- Data-driven methods deliver superior accuracy when large, labeled datasets are available.

- Hybrid techniques strike the best balance, showing high robustness and accuracy even in varied or uncertain conditions.

- In systems where battery behavior is highly variable or data-rich environments exist, hybrid models are the most effective choice.

4.5.4 Final Recommendation

Based on this comparative review:

- Hybrid SoC estimation methods represent the current state-of-the-art, achieving consistently high accuracy with good resilience to variability and aging.
- Model-based methods are still relevant for simpler, resource-constrained applications.
- Data-driven methods excel when sufficient data are available and higher computational resources are permissible.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This study provided a comparative analysis of model-based, data-driven, and hybrid State of Charge (SoC) estimation techniques. Findings show that while model-based methods remain computationally efficient and interpretable, their reliance on precise battery models and sensitivity to aging effects limit their accuracy in dynamic conditions. Data-driven approaches, especially those utilizing machine learning, achieve higher accuracy and adapt well to non-linear battery behaviors, but they require large, high-quality datasets and greater computational resources. Hybrid techniques emerge as the most balanced solution, consistently achieving accuracy below 1% RMSE while demonstrating robustness across varying conditions. The review highlights a clear research trend toward hybridization as the preferred approach for modern battery management systems. Overall, hybrid methods best address the growing demands for accuracy, robustness, and real-world applicability.

5.2 Recommendations

Based on the insights drawn from this review, future research should prioritize the development and refinement of hybrid SoC estimation techniques, focusing on reducing computational complexity while preserving their high accuracy and robustness. Researchers should also explore methods to minimize the data requirements of data-driven components, enabling more practical implementation in real-world systems with limited historical data. For model-based techniques, efforts to improve parameter adaptability could extend their usefulness in resource-constrained environments. Battery management system designers are encouraged to consider hybrid approaches as a preferred solution for applications demanding both precision and resilience, such as electric vehicles and renewable energy storage. Finally, standardizing datasets and evaluation protocols would help ensure consistent benchmarking, fostering faster progress and wider adoption of next-generation SoC estimation methods.

REFERENCES

Al-Smadi, M. K., et al. (2025). State of health estimation for lithium-ion batteries based on impedance and ANN methods. Batteries, 11(4), 133. https://doi.org/10.3390/batteries11040133

Ali, H., et al. (2023). Evaluation of in-service smartphone battery drainage profile for video calling feature in major apps. PLOS ONE, 18(7), e0288591. https://doi.org/10.1371/journal.pone.0288591

Altinpulluk, N. B., Altinpulluk, D., Ramanan, P., Paulson, N., Qiu, F., Babinec, S., & Yildirim, M. (2023). Federated battery diagnosis and prognosis (arXiv:2310.09628) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2310.09628

Che, Y., Hu, X., Lin, X., Guo, J., & Teodorescu, R. (2023). Health prognostics for lithium-ion batteries: Mechanisms, methods, and prospects. Energy & Environmental Science, 16, 338–371. https://doi.org/10.1039/D2EE03019E

Deng, Y., Du, C., & Ren, Z. (2025). A novel method for estimating the state of health of lithium-ion batteries based on physics-informed neural network. Batteries, 11(2), 49. https://doi.org/10.3390/batteries11020049

El Marghichi, M., et al. (2023). Improving accuracy in state of health estimation for lithium batteries using gradient-based optimization: Case study in EV applications. PLOS ONE, 18(11), e0293753. https://doi.org/10.1371/journal.pone.0293753

Fan, Y., Lin, Z., Wang, F., & Zhang, J. (2025). A hybrid approach for lithium-ion battery remaining useful life prediction using signal decomposition and machine learning. Scientific Reports, 15, Article 8161. https://doi.org/10.1038/s41598-025-92262-8

Flores-Martín, D., Laso, S., & Herrera, J. L. (2024). Enhancing smartphone battery life: A deep learning model based on user-specific application and network behavior. Electronics, 13(24), 4897. https://doi.org/10.3390/electronics13244897

Gasper, P., et al. (2022). Predicting battery capacity from impedance at varying conditions. NREL Technical Report (FY23-OSTI-82892). https://doi.org/10.2172/82892

Hu, X., Xu, L., Lin, X., & Pecht, M. (2020). Battery lifetime prognostics. Joule, 4(2), 310–346. https://doi.org/10.1016/j.joule.2019.11.018

Jafari, S., & Byun, Y.-C. (2022). XGBoost-based remaining useful life estimation model with extended Kalman particle filter for lithium-ion batteries. Sensors (Basel), 22(23), 9522. https://doi.org/10.3390/s22239522

Kim, M., Oh, J., Lee, S.-Y., & Kim, J. (2025). DLinear-based prediction of remaining useful life of lithium-ion batteries: Feature engineering through explainable artificial intelligence (arXiv:2501.11542) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2501.11542

Kong, J., Liu, J., Zhu, J., Zhang, X., Tsui, K.-L., Peng, Z., & Wang, D. (2024). Review on lithium-ion battery PHM from the perspective of key PHM steps. Chinese Journal of Mechanical Engineering, 37, 71. https://doi.org/10.1186/s10033-024-01055-z

Madani, S. S., Ziebert, C., Vahdatkhah, P., & Sadrnezhaad, S. K. (2024). Recent progress of deep learning methods for health monitoring of lithium-ion batteries. Batteries, 10(6), 204. https://doi.org/10.3390/batteries10060204

Nascimento, R. G., et al. (2023). A framework for Li-ion battery prognosis based on hybrid physics-informed learning. Scientific Reports, 13, 6218. https://doi.org/10.1038/s41598-023-6218

Navidi, S., Thelen, A., Li, T., & Hu, C. (2024). Physics-informed machine learning for battery degradation diagnostics: A comparison of state-of-the-art methods (arXiv:2404.04429) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2404.04429