Prediction of the Battery State Using the Digital Twin Framework Based on the Battery Management System

SADIQA JAFARI© AND YUNG-CHEOL BYUN©
Department of Computer Engineering, Jeju National University, Jeju 63243, South Korea
Department of Electronic Engineering, Institute of Information Science and Technology, Jeju National University, Jeju 63243, South Korea
Corresponding author: Yung-Cheol Byun (ycb@jejunu.ac.kr)

This work was supported in part by the Ministry of Small and Medium-Sized Enterprises (SMEs) and Startups (MSS), South Korea, under the Regional Specialized Industry Development Plus Program (Research and Development), supervised by the Korea Institute for Advancement of Technology (KIAT), under Grant S3246057; and in part by the Korea Institute for Advancement of Technology (KIAT) Grant through the Korea Government [Ministry of Trade, Industry, and Energy of Korea (MOTIE)]. The Establishment Project of Industry-University Fusion District, under Grant P0016977.

ABSTRACT Electric Vehicles (EVs) reliance on batteries, which currently have lower energy and power densities than liquid fuels and are prone to aging and performance degradation over time, restricts their mainstream adoption. With applications like electric vehicles and grid-scale energy storage, effective management of lithium-ion batteries is a vital enabler for a low-carbon future. Monitoring the battery’s condition of health and charge over the lifetime of an EV is, therefore, a highly pertinent issue. Battery Management Systems (BMS) are used during the operation of EVs to monitor, estimate and control battery states to ensure that batteries can function effectively and safely. Additionally, the materials composition, system design, and operating circumstances substantially impact a battery’s usable life, making it more challenging to govern and maintain battery systems. This work proposes the structure of a battery digital twin-based battery for the electronic vehicle, which has the potential to enhance BMS situational awareness greatly and enable the optimal functioning of battery storage units. Digitalization and Artificial Intelligence (AI) present an opportunity and offer. In this paper, a Digital Twin (DT) is proposed as a solution to the difficulty of onboard computation for the incremental State Of Health (SOH) and State Of Charge (SOC) by using Extreme Gradient Boost (XGBoost) model and Extended Kalman Filter (EKF) to predict the state estimate for the EV battery. The battery’s condition has been determined by using the EKF, which can provide vital information for maintenance. First, the battery’s usable life can be extended with an accurate estimate of the SOC to continue; then a learning-based prediction approach to gauge the battery’s health state is suggested in order to increase battery life. A SOC model is frequently retrained to depict the effects of aging, and a SOH model is often performed to foretell the reduction in the highest battery capacity. According to a result, DT models are useful for managing batteries, and full life cycle statistics are important for planning the battery’s upgrade path.

INDEX TERMS Battery management system, digital twin, the battery state, electric vehicle, XGBoost.

I. INTRODUCTION

Nowadayas, by improving energy storage technologies, battery systems have been known famous in the energy storage system to decrease global carbon emissions in fixed and mobile applications. Lithium-ion batteries are essential for net-zero carbon change because of their wide temperature range, increased energy density, and down memory development [1], [2]. Rising greenhouse gas emissions generate environmental change, so reducing them is an important priority. Lithium-ion batteries can also be used in cars to reduce CO₂ emissions [3]. A battery includes a combination...
of chemical, thermal, and mechanical, and the lifetime depends on the operational requirements and the system design. Considering the value of data to process, BMS can receive that it needs memory and processing power due to computationally centralized. A battery provides the power to steer the electric motor, while motors are the source of power for EVs [4]. A vehicle’s primary power source is the battery, which is crucial for EVs. A Li-ion battery is the most commonly used battery in electric vehicles. Lithium-ion batteries offer more energy density, extended life span, and power thickness than most valuable batteries. Still, they are not as durable or safe as other types of batteries. For optimum performance, li-ion batteries should be operated at temperatures and voltages within safe ranges. Batteries have internal reactions that can cause fires or other dangers if not properly managed due to lithium-ion batteries using solid chemicals [5]. BMS is usually used to quantify and predict battery performance to maximize battery life in real-world situations [6]. Its most important function is to accurately gauge the battery’s charge and health and provide information on its charge and aging levels. Battery life can be extended through maintenance or by adjusting operational techniques; the coulomb counting algorithm is an open-loop algorithm that does not require much computing power to implement [7].

Batteries can be monitored for voltage, current, and temperature, which can be used to estimate the batteries’ charge and health state based on those measurements. BMS also facilitates safe operation and prolongs battery life by regulating voltage and SOC through the cell balancing circuit. BMS has a variety of voltage, current, and temperature sensors to accomplish this. Batteries can be modeled digitally in several ways, but the models often lack real-world data to support their accuracy. The paper aims to determine the state estimation in BMS of EVs because many voltages, currents, and temperature sensors are established in BMS. Therefore, to ensure safety and reliability in automotive vehicles, DT architectures should be developed to monitor battery system state and maximize energy efficiency. A real-time battery monitoring, safety, and reliability, a self-reconfigurable battery digital twin was presented [8], [9]. Due to the improved availability of battery data and raised computer power, data-driven strategies for calculating SOC and SOH are growing in popularity. An overview of battery state estimation strategies based on ML techniques, including Feedforward Neural Networks (FNNs), Recurrent Neural Networks (RNNs), Support Vector Machines (SVM), Radial Basis Functions (RBF), and hamming networks, is presented to determine SOH and SOC in [10].

Accurate SOH calculation reduces loss risk and increases dependability in EV applications by foretelling battery health situations. An accurate SOH estimate is difficult due to the EVs’ unknown dynamic operating conditions and the lithium-ion battery’s complicated nonlinear electrochemical properties. This study has proposed an Artificial Neural Network (ANN) classifier for lithium-ion battery SOH estimate. Training has been carried out using parameters such as the relative voltage SOC, State Of Energy (SOE) across a buffer, and the instantaneous SOC and SOE based on the SOH description [11]. It is essential to properly manage and administer battery-based energy storage systems, such as microgrids and electric cars, to accurately estimate the SOH of lithium-ion batteries. The model described in the paper for estimating the SOH is straightforward but effective. It is based on the feature extraction of voltage, current, and temperature patterns during the charging process, enabling per-cycle estimations using ANNs [12]. In order to address the issue of traditional estimation methods’ poor estimation accuracy, a SOH estimation method based on an improved ant lion optimization algorithm and support vector regression has been suggested. SOH estimation is crucial for maintaining the safe and stable operation of lithium-ion BMS Improved Ant Lion Optimizer-Support Vector Regression (IALO-SVR). Four health variables strongly correlated with SOH deterioration have been chosen as the input for the SVR model after geometric analysis of the battery charge and discharge data. After training has been utilized in a battery training set, the SOH estimation model has been created employing the IALO strategy to optimize the kernel parameters of the SVR [13]. The SOC of the battery is one of the parameters in the BMS for Battery Electric Vehicle (BEV); SOC has recently attracted a lot of attention; however, there are still modeling issues due to the nonlinear lithium-ion battery structure. SOC can provide accurate modeling and state estimation to increase the stability of a BMS. It has suggested current trends for SOC estimation in lithium-ion BEV [14].

Digital twin technology can represent a physical system virtually, using artificial intelligence, the internet of things, and big data. The BMS is updated through the cloud using machine learning and DT integration to maximize performance and offer battery longevity as the cells deteriorate. We have proposed to utilize machine learning for the digital twin. DT of the BMS’s performance can be improved by utilizing the proposed system, as the user can know the battery state in advance. This feature will help monitor the battery condition and allow the user to take appropriate measures.

In this paper, we proposed a DT model for BMS to estimate and predict the battery states, then analyze the correlation between features. Since lithium-ion batteries are always partially charged and discharged. Our suggested model uses voltage sensors and temperature data to determine estimates and predictions. Besides, the proposed model is helpful for battery performance degradation and evaluation and improves battery management’s reliability, optimization, and accuracy. Because aging significantly impacts the SOC model, the DT includes a SOC model that is periodically retrained utilizing data supplied from the EV. DT comprises a model of the SOC that is periodically retrained on data transferred from the EV since the model of the SOC is significantly influenced by age. The model of SOH by the digital twin was trained once on data from a full charge cycle.
and is based on historical data. In order to offer a workable solution for an accurate assessment of battery dynamics and aging effects, this study explores the difficulties involved in building a digital battery twin and suggests a unique architecture of models for SOC and SOH. Our contributions include the following under blow:

- Identify and analyze the relationship between the parameters of an electric vehicle battery;
- The suggested method to determine BMS is to predict the SOH and SOC based on the battery for an EV.
- Generated the digital twin structure based on the NASA dataset and performance evaluation.

The remainder of the methodology is structured as follows: section 2 describes the brief literature study of the offered design framework. Then we explain the proposed method in sections 3 and 4. Sections 5 and 6 supply the implementation process and the performance evaluation in section 7. Section 6 describes the result; finally, we complete this paper in the conclusion section.

II. RELATED WORK

DT technology is used for BMSs, which have been evaluated, and their functionality is outlined in existing studies. A digital twin-based battery diagnosis is only in its early stages, though the literature has extensively discussed the BMS of EVs. Developing and managing batteries is a critical component of new energy vehicles. Battery storage creation has been significantly aided by lithium-ion batteries, which have an excellent energy density, extended cycle lives, and a low self-discharge rate [15]. DT guides to multiphysics, the multiscale simulation that reflects the life of its physical twin. By connecting the best physical models with sensor data [16]. An analysis of unanticipated disturbances during a system run provides a diagnosis of the DT’s three main functions: prediction - execution of studies before the system run; safety - continuous monitoring and control during the system run; diagnosis - analysis of unanticipated disturbances during the system run. Models are simplified representations of natural systems based on their structure and processes [17]. A computer model is built to bring a digital representation closer to reality. Simulations and A computer model is built to bring a digital representation closer to reality. Simulations and interdisciplinary collaborations can produce multiscale models more accurately than some extreme simplifications of the actual system [18]. BMS are becoming increasingly required to maximize lithium-ion batteries’ performance. The safety and cycle life of the system is essential. BMS needs to get more information from actual batteries in order to accurately gauge their age and safety, necessitating a complete life cycle management system. Furthermore, the onboard BMS cannot adequately handle real-time data and has poor data utilization rates when processing large amounts of data during vehicle operation [19], [20].

The concept of DT can be achieved via various advanced technologies, not just one specific technology. Consequently, other research areas demand different levels of clarity and ideas. A brief concept has also been presented for lithium-ion batteries; however, most researchers have emphasized utilizing DT to create high-precision models, and the estimated SOC and SOH in lithium-ion batteries [21], [22]. Currently, battery DT research focuses on problems encountered with the current BMS, including the limited computational capacity of embedded systems, difficulty sharing data between manufacturers, and data storage limitations. A hybrid HIF-PF online technique is suggested for predicting SOC under the Beijing Bus Dynamic Stress Test (BBDST) experimental setting. That design is based on power lithium BMS to monitor and visualize real-time voltages and currents to estimate SOC in [23]. An EKF has been suggested for estimating the battery pack’s parameters and SOC that has been utilized state-space model of an EV battery. Also, SOC has been compared with the present situation and previous situations, and then it has simulated battery parameters received from a characteristic EV battery pack in [24]. The current conversion of car batteries to second-life storage systems has been characterized by time- and money-consuming activities, including disassembly to the module level, and manual SOH evaluations have been used in expensive diagnostic equipment. The target of the study has been to present a brand-new cloud-connected battery management method to estimate the residual value of car batteries concerning several prospective second-life operations. The approach to electric-thermal modeling and the techniques has been utilized for dynamic electric parameter estimation are the main topics of paper [25]. An estimated e-golf battery state method has been suggested by applying diagnostic data and a DT to create a digital battery twin with a data pipeline to track the battery’s state, including SOC and SOH. Which has been searched for Universal Data Structure (UDS) messages containing both battery pack and cell-individual data based on reverse-engineering the diagnostics interface [26].

III. BATTERY MANAGEMENT SYSTEM

BMS regulates many functions needed to keep the batteries of EVs functioning correctly and safely. A BMS tracks voltage, current, and temperature, optimizes battery performance, detects and prevents failures, and collects and analyzes battery data. BMS must be programmed to meet the specific operating conditions per kind of battery cell. The range of ambient temperatures in which a power supply or any other electrical equipment may function is known as the operational temperature. Outside of these limits, the power supply may fail; these range from a minimum working temperature to a peak or maximum operating temperature. The safe operating voltage ranges in the lithium-ion battery have started from 3.2 °C to 3.6 °C. However, the lithium-ion battery can operate normally in temperatures between 20 °C and 60 °C as the safe operating temperature ranges but performance suffers when the temperature drops below 0°C and the discharge capacity is reduced [27]. An electric vehicle’s battery pack comprises a series of interconnected cells. Temperature changes and increased charge and discharge cycles decrease the capacity of...
individual battery cells. Managing and controlling individual cells’ charging and discharging conditions becomes crucial when connected in series or parallel, as in a battery pack. Monitoring, controlling, and balancing the pack are done through a BMS. As shown in Figure 1, a BMS serves several vital functions. In addition to risking unnecessary damage, a battery that is not balanced will also not function optimally. In a battery pack, preventing significant differences in cell charge status is very effective because the worst cell limits the performance of all cells, which depends on the functionality and intelligence of a BMS to define its cost and complexity. Nowadays, many different methods have been applied to predict accurate SOC indications based on BMS for battery-powered applications [28], [29]. Thermal management systems in battery packs are vital since battery cell performance varies with temperature. It provides electrical and thermal balance for all cells, developing their lifetime. Thermal management systems can transfer heat either through the air or through the liquid. A vehicle can integrate with an electronic system that consumes little power and does not add much mass to the vehicle. Thermal management systems can use passively or actively to achieve their performance needs. Some battery packs may be adequately controlled by passive systems that use only the ambient environment, while others may demand active control.

Figure 1 shows an efficient BMS for EVs. An electric vehicle’s BMS contains several circuits, features, power electronics, sensors, actuators, diodes, capacitors, inductors, transformers, switching, converters, and safety equipment controlled with several algorithms, models, and control signals.

The battery pack, measurement, cell balancing, capability estimations, SOC, SOH estimations, thermal management, and Controller Area Network (CAN) controllers are some of the parts that make up a BMS. As shown in Figure 2, a BMS is divided into blocks [30]. BMS can determine if there is a dangerous imbalance in the battery pack based on the difference between the minimum and maximum cell voltages. A battery pack’s charging and discharging current must be carefully controlled to avoid overheating and eventual loss. Over voltage and under voltage of the battery can occur due to improper control of the charging and discharging current. Batteries’ SOC refers to their remaining capacity as a percentage since some batteries cannot be discharged below a specific rate for many reasons, including the depth of discharge becoming too deep. This can cause some batteries to become damaged or lose capacity, which can be detected by monitoring their voltage and current values [31]. As the battery degrades faster at high temperatures, the temperature directly impacts the battery life. Batteries and individual cells also have a temperature function. The BMS and the battery thermal management system can cool the battery and keep its temperature within a specified range. In addition, it is essential to know the temperature of each cell, so it can determine if a hot spot exists, which indicates a possible failure. The intake and output coolant temperatures can choose a battery pack’s temperature.

The actual battery capacity, normalized to the initial battery capacity, is used to calculate battery capacity, which decreases with each charge or discharge. Measuring the battery’s SOH is a method of estimating a battery’s lifetime situation [32], [33]. Balancing the cells of the battery is one of the essential elements in the BMS for enhancing battery life. Each battery pack cell is balanced at a different voltage through the cell balancing block. In the course of charging and discharging, the various cells in a battery pack will naturally exhibit differences in the SOC due to manufacturing faults. In the cell balancing block, external balancing circuits ensure equal voltage and the SOC for the cells [34]. Each battery pack cell is balanced at a different voltage through the cell balancing block. After reaching the measurement block, capability estimation, SOC, SOH, and thermal management blocks, all data is sent to the measurement block. A thermal management block in the battery can measure temperature when the temperature is high or low. Considering (high or low) different temperatures can enable the cooling fan or heater. The battery’s performance depends on the temperature since to keep the temperature between limitations related to activating heaters and coolers; it is required to maintain the temperature of each cell of the battery. For estimating the highest charge current and the latest current charge in the battery, the capability estimate block can support it [35], [36].

IV. STATE ESTIMATION AND BATTERY DIGITAL TWIN

An artificial intelligence-based machine learning method is operated to create the BMS-DT model. The DT model for lithium-ion batteries focuses on estimating battery performance states. In the proposed DT of the EV battery storage system, the estimation of SOC and SOH is the considerable influential factor affecting the EV driving knowledge and battery life. This section describes how a machine learning method creates a DT model for the BMS. The scheme of the proposed DT model approach used to build the DT will be presented first.

A. FRAMEWORK FOR THE DIGITAL TWIN OF BATTERY

An intelligent local energy system can be represented as a digital twin in a way that can be updated and dynamically developed in real-time. Hierarchical, bidirectional, and self-evolving are the main features of digital twins.

To ensure that the DT self-evolves, the model parameters are updated using a rolling learning technique as battery capacity declines. State estimation, such as RUL, prediction, and energy management, can all be addressed with different models. Besides, multi-dimensional data should still be able to interact between hierarchical structures. The main goals of this study are the SOC, and SOH measured variables analyzed by employing the DT to identify hidden transformation laws. The battery DT framework, demonstrated in Figure 3, is suggested as the foundation for the entire system to achieve a robust design. Digital twins operate as virtual physical entities on digital platforms. The DT platform has different
functions depending on the application scenario and physical entity.

B. BATTERY MODELING
Modeling battery systems is an essential aspect of digital twins, which is vital for accurately diagnosing batteries. The three main types of battery cell models are Equivalent Circuit Models (ECM), electrochemical models, and machine learning models [37], [38], [39]. According to our evaluation, each model has strengths and limitations and will play a different role in the battery system DT. A digital twin is used to accurately monitor the battery cell, the SOC, and SOH using ECMs. Numerous ECMs developed in the past years have been used to estimate the battery state in
A framework for developing battery digital twins.

FIGURE 4. Model of ECM structure.

The paper uses an equivalent circuit model, as shown in Figure 4. Due to its convenience for modeling and calculating, the ECM is a widely used Thevenin model. An extended Thevenin model is suggested that balances model accuracy with computation time. A model of battery dynamics is presented below:

\[
\begin{align*}
\dot{v}_1 &= -\dot{v}_1 + i/r_1c_1 + i/c_1 \\
\dot{v}_2 &= -\dot{v}_2 + i/r_2c_2 + i/c_2 \\
V_t &= V_{oc} - v_1 - v_2 - ir_0
\end{align*}
\]

where \( i \) denotes the input current in Equation 1, \( V_{oc} \) is the open-circuit voltage denotes in Equation 2. The model consists of SOC-controlled OPC (open circuit voltage), where \( V_t \) denotes terminal voltage, which is nonlinear to battery SOC. Also, \( r_0 \) denotes resistance for each circuit RC. Here \( c_{1,2} \) denote resistance and capacitance for each RC circuit. \( v_{1,2} \) denotes a voltage drop over each RC circuit in the Equation 1 and 2. Thus, dynamic changes in these parameters, such as the SOC state, temperature, and aging state, are to blame for the estimation error. Nevertheless, the offered algorithm could recompense these errors within the DT.
C. EXTENDED KALMAN FILTERING ALGORITHM

The EKF is one of the KF variants that can only be used with linear systems. To use a KF algorithm for the nonlinear battery method, the EKF is proposed via a linearization process. To provide the most accurate and unbiased estimation of the state variables, the conventional KF is essentially only suited for linear dynamic systems. However, because the battery is a nonlinear system, it must be linearized and resemble a linear system. This method is one of the best opportunities for accurate SOC prediction since the feedback is modified at each time step in light of the present circumstance and the likelihood of adoption within BMS. EKF has been regarded as a typical choice in the nonlinear state estimation for lithium-ion batteries, which are dynamic nonlinear systems. A battery model is initially created to provide internal state variables and state-space equations to develop the mathematical model in order to apply the EKF. Additionally, the battery model can calculate the internal state variables such as internal resistance, electromotive force, capacitance, and SOC depending on external factors such as terminal voltage, current, and temperature [40], [41]. Lithium-ion batteries are typical nonlinear systems in which noise significantly affects state estimation that the EKF algorithm is commonly used to optimize the state of the system because of its excellent noise-filtering performance. In EKF, the covariance estimation and the current mean value are linearized concerning the covariance estimation. Also, state-space equations for batteries can be obtained by discretizing the equivalent circuit model in Equation 3 and 4 that $M(n)$ is the terminal voltage $V_t$; $W(n)$ is the process noise, $U(n)$ is the current and $V(n)$ is the measurement noise.

$$X(n) = RX(n - 1) + ZU(n - 1) + W(n - 1)$$
$$M(n) = f(X(n)), U(n)) + V(n)$$

where Equation 5 is shows a prediction of the state variable at time $n$, and Equation 6 is shown system state error matrix 6 below:

$$X_p(n | n - 1) = RX(n - 1 | n - 1) + Zu(n - 1)$$
$$X_p(n | n - 1) = R(n)P(n)R(n)^T + \Phi$$

Equation 7,8 employs $G(n)$ as its state matrix, $\Phi$ and $E$ as its observation noise, and system noise, respectively. In addition, $n$ is its gain factor [42].

$$G(n) = \delta f(X(n), U(n))/\dot{X}(n)$$
$$n_f = \frac{X_p(n | n - 1)G(n)}{G(n)X_p(n | n - 1)G(n)^T} + E$$

D. STATE OF CHARGE ESTIMATION

Battery voltage does not directly relate to battery capacity, so knowing what it does not suffice. As a percentage of the battery’s rated capacity, the SOC represents the amount of battery available. The SOC enables the BMS to assess the battery’s state and ensures safe battery operation by managing charge and discharge. The SOC is usually calculated by separating the battery’s current capacity from its specified capacity. ECMs and parameters can be identified through digital twins to estimate the battery’s SOC because the SOC cannot be directly measured. The effects of dividing the battery’s current capacity by its nominal capacity on SOC estimation are as follows:

$$\text{(SOC)}_T = 1 - \int_0^T \delta \frac{i}{c_n} dt$$

$\text{(SOC)}_T$ shows the battery at the initial time, respectively, $\delta$ shows coulomb efficiency, $c_n$ is the capacity and $i$ shows current in the Equation 9. Based on a nonlinear relationship, the open-circuit voltage $V_{oc}$ and its closed-circuit voltage, the SOC, are relatives. In addition, no direct measurement can be made of the voltage $V_{oc}$. Figure 5 shows the battery’s SOC during charge and discharge.

E. STATE OF HEALTH ESTIMATION

The battery’s SOH shows the battery’s age because of its capacity loss mechanisms; the remaining capacity of a battery reduces as it ages, resulting in the capacity fade. During that time, the battery’s power fades due to an increase in ohmic resistance and polarization resistance [43]. Additionally, as the capacity fades, there may be an increase in internal resistance and a decrease in capacity. In Equation 10, battery health is typically described as follows:

$$\text{(SOH)}_T = \frac{c_i}{c_n}$$

where Equation 10, $c_i$ indicates the current practical capacity and $c_n$ demonstrates the nominal battery capacity at time $T$, respectively. Besides, as shown in Equation 11, the battery’s inner resistance varies at the beginning of its life and present.

$$\text{(SOH)}_T = \frac{R_T}{R_{0,0}}$$

Based on a machine learning program, one can also determine the battery’s current capacity and ohmic resistance. By combining 10 and 11, the SOH can be calculated both from the perspective of capacity fade and resistance increase. As a result of newly manufactured batteries, a battery’s SOH estimation describes the battery’s current state.

V. EXPERIMENTAL VALIDATION AND DISCUSSION

The accuracy of the predictions is evaluated using two assessment indicators, the Mean Absolute Error (MAE)
and the Root Mean square Error (RMSE). To show the difference between the predicted value and the actual value in Equation 12 and 13, RMSE and MAE are calculated as follows:

\[
E_{\text{RMSE}} = \frac{1}{M} \sqrt{\frac{1}{M} \sum_{m=1}^{M} (\hat{Y}_m - Y_m)^2} \tag{12}
\]

\[
E_{\text{MAE}} = \frac{1}{M} \sum_{m=1}^{M} |\hat{Y}_m - Y_m| \tag{13}
\]

A. DATASET DESCRIPTION

Battery data from NASA’s prognostics dataset is used to evaluate the prediction model. A digital twin is validated using the case study’s data collected on Panasonic’s lithium-ion battery [44]. A total of 18,650 lithium-ion batteries were tested to identify their degradation behavior. Basic information about lithium-ion batteries based on Table 1, the lithium-ion battery used in the testing was able to discharge at a constant current of 2.2A, charge at a constant current of 2.2A, and charge at a constant voltage of 4.2V, continuously operated. We used the different battery cell types; also, the experiments are run at room temperature (24 °C) at a constant temperature.

The dataset comprises data on the battery’s current, voltage, temperature, relative time per reference discharge and charge cycle, and time throughout the measurements are monitored for reference and randomized processes. Periodically, reference charging and discharging cycles were carried out to estimate the battery’s total functional capacity, which determines the battery SOH at various periods over the battery’s lifetime.

B. DATA VISUALISATIONS

When tuning statistical models to choose which variables are essential, the visualizations of the data in Figure 6 are used to highlight any apparent trends between parameters that make it easier to understand the data.

C. DATA PROCESSING

In this study, at the measurements, we keep reference discharge data. We use all reference discharge data for this research to construct our dataset. Due to data quality, voltage data cannot be directly used in the XGBoost when transmitted from the BMS to the digital twin server. As a result, it is essential to have a data processing stage where various operations are carried out to increase the speed and usability of the data, such as data splitting, abnormal data deletion, and data sorting. Two subsets of the datasets are created: normal and charging.

The dataset is divided into two parts; the first is normal, and the second is the charging. Preferably, the battery pack is discharged to make the EV move; in the second step, EVs can park at stations and are recharged at charging stations. We cleaned up the dataset by eliminating any abnormal data. Then some data has invalided that according to the result of measurement and communication errors. In order to ensure that a dataset is normal of abnormal data, abnormal values are removed. During each reference discharge cycle, both the SOH and SOC models should be built, and the SOC model should be retrained when the SOH decreases. As shown in Figure 7, voltage, current, and temperature are input, followed by \( r_0, r_1, r_2, c_1 \) and \( c_2 \) according to the ECM parameters; the equation of state and measurement equation of the equivalent circuit model are then determined. Next, this work uses EKF and XGBoost to minimize the error and prediction of battery states.

In addition to being utilized to estimate the battery’s SOH, the detected parameters \( r_n, r_0, 1, 2 \) may also be fed back into SOC algorithms to increase estimation accuracy with precise model parameters.

VI. STATE ESTIMATION WITH COLLABORATING EKF AND XGBoost

Through the construction of various machine learners, the formation and training of numerous weak learners, and the combination of multiple vulnerable learners using a combination technique, ensemble learning (or ensemble learning) is a model framework for a strong learner [45], [46]. The boosting method improves the performance of students who perform poorly by giving them feedback repeatedly. In order to make the iteration \( M \) times, the training effect of the previous learner is modified, and the new sample distribution is used to train the following learner. A group of weak learners is eventually merged into strong learners. The XGBoost does not resample the sample at sampling time, preventing a sample from being used repeatedly in each round of computation. XGBoost enables subsampling and allows partial sample sizes in each computation phase to lessen overfitting. Another highlight is that XGBoost enables column sampling and extracts a random number of features every cycle of computation for training, speeding up and minimizing overfitting [47]. The SOH prediction model chosen in this instance is the XGBoost model, and features retrieved using the technique above are utilized as input for model training. The trained model then predicts the SOH.

Different optimization strategies are employed to improve the super parameters of the XGBoost model since they significantly impact the model. Each model’s effectiveness is compared to choose the best model. The standard EKF-based SOC estimate approach offers the benefits of being straightforward and quick to react to suit the real-time demands of the DT systems [48]. We have experimented with a machine learning-based model in this study to accomplish the SOH and SOC estimates. This effort aims to define the architecture of our suggested DT and its workings.

---

**TABLE 1. A short summarization of the tested lithium-ion batteries.**

| Size       | 18,650     |
|------------|------------|
| Anode      | Li(NiCoMn)O2 |
| Cathode    | Graphite   |
| Rated capacity | 2200 mAh |
| Voltage range | 2.2 - 4.2 V |

---
The selection of the inputs for an estimating algorithm is a complex problem. However, as they can be directly monitored, current, temperature, and voltage have also been crucial factors in determining a battery’s status. Current, voltage, temperature, relative time, SOC, and SOH are the columns that make up our dataset. The first four are utilized as inputs to train the models in our suggested DT, while the last two are used as model outputs.

The XGBoost uses to estimate SOH and EKF to determine SOC. As a result, the ECM and XGBoost get these parameters as inputs in this investigation. Utilizing driving data from the battery’s initial few working cycles, the XGBoost is initialized and pre-trained to increase accuracy. This approach offers higher SOC estimate performance and reduces the initial battery condition data uncertainty compared to the conventional EKF estimation method. Due to the immense influence of battery aging on assessing the battery charge level, real-time updating is required as a battery DT. XGBoost is utilized in the cloud for retraining and calibration, and the XGBoost on the virtual end is updated when the cumulative runtime $t$ exceeds a certain threshold of time $T$. This is exactly how the battery DT works. Real-time data is continually gathered.

VII. EXPERIMENT RESULTS

The first model to be put into implemented is a KF. It was written up as a class, making it easy to tune once the model was built to determine a more precise estimate of the SOC. Noise from the Open-Circuit Voltage (OCV) readings was removed using the KF to determine a more accurate estimate of the SOC, and noise from the OCV voltage readings was extracted using the KF. The model was trained using the BMS circuit’s identical inputs. Figure 8 depicts the result graph, which contrasts predicted OCV versus observed OCV after the model has been trained on the dataset.

To demonstrate the efficacy of the applied improved XGBoost method based on the EKF to analyze the comprehensive effect and superiority of the SOC estimation of the built XGBoost model, only one XGBoost model with 80% train inputs and 20% test as output is used in this paper. Figure 9 illustrates how the EKF algorithm’s filtering impact might limit the peak inaccuracy of the SOC calculation. Only current and voltage are used as the input characteristic values for a single XGBoost, which results in poor model convergence. The trained XGBoost and EKF can estimate the actual SOC according to the chosen input physical quantity.

In this study, SOH prediction is made using the feature extraction and prediction model presented as shown in
To ensure the dependability and safety of electric and hybrid cars, monitoring and prognostication of cell deterioration in lithium-ion batteries are crucial. Compared to a brand-new battery, a battery’s SOH can gauge how well it can store and distribute electrical energy. The EKF-XGBoost method efficiently enhances the total state estimation (SOC and SOH) impact, as demonstrated by the experimental simulation findings in Figure 11, which also show improved generalization and resilience while retaining high estimate accuracy. The suggested XGBoost model based on the EKF algorithm uses the model trained using the EKF algorithm data to state estimate the battery because the choice of training data greatly influences the machine learning method’s prediction performance.

The majority of the existing work in the literature ignores the substantial correlation between the relative temporal discharge time and the related voltage, current, and temperature values. The estimation approach of appropriate SOC and SOH is implemented together. Considering the identical features operated to a train of the SOC has been utilized for training the SOH models after normalization. Table 2 shows the performance of SOH models trained that the estimation results for the EKF method and the modified EKF-XGBoost algorithm are compared in terms of RMSE and MAE. The trained XGBoost and EKF can forecast the actual according to the chosen input physical amount. Both

Figure 10. The SOH of the battery is predicted using the XGBoost model. Figure 10’s prediction results make it clear that the proposed method is performing well.
techniques successfully decrease estimation error, as shown by the comparison in Table 2. However, when given the right feature inputs, the XGBoost algorithm has superior machine-learning skills.

According to the experimental findings, the suggested EKF-XGBoost algorithm performs more comprehensively than the EKF and XGBoost algorithms. The estimated accuracy is significantly improved, and the estimation effect is enhanced overall while considering the approach’s generality and resilience. To evaluate the performance of the suggested approach with existing methods, shown in Table 3 that the error rate compared different models between other techniques. With attention to the error rate, the proposed method performs well compared to existing methods.

### VIII. CONCLUSION AND FUTURE WORK
Enhancing the safety, dependability, and performance of the battery systems depends on effective battery management. We present an estimation and prediction model using a digital twin for the BMS. By examining the shortcomings of a single algorithm in state estimation, a novel model for predicting the state estimation of lithium-ion batteries extracts batteries’ SOC and health characteristics. An equivalent circuit model based on second-order RC is employed for parameter identification and modeling. The parameter values are updated online, improving the accuracy of the battery model. As a result, the low-current OCV test data is utilized to receive the relationship between SOC and OCV, which minimizes fitting error. In addition, we introduce a digital twin battery twin and online model refinement methods operating machine learning. A prediction model based on EKF-XGBoost is then constructed to predict SOC-SOH with the open-loop to the trained machine learning model.

Combining the EKF algorithm’s filtering capabilities with the XGBoost algorithm’s exceptional nonlinear appropriate regression prediction ability, the suggested algorithm can have the stability of estimation errors across the entire range of estimation errors while improving filtering accuracy. Simulation results demonstrate that the offered EKF-based XGBoost algorithm performs better in prediction. In addition to supplying the reliability and accuracy of state estimation, the proposed algorithm has clear benefits regarding generality and robustness. Future work will create a hierarchical structure of digital twins to maximize battery life and data collected from the cloud battery systems to develop machine learning algorithms for lifetime prediction and system optimization.

### REFERENCES

[1] Y. Wang and Z. Chen, “A framework for state-of-charge and remaining discharge time prediction using unscented particle filter,” Appl. Energy, vol. 260, Feb. 2020, Art. no. 114324.

[2] S. Jafari, Z. Shahbazi, and Y.-C. Byun, “Lithium-ion battery health prediction on hybrid vehicles using machine learning approach,” Energies, vol. 15, no. 13, p. 4753, Jun. 2022.

[3] C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, “Degradation diagnostics for lithium ion cells,” J. Power Sources, vol. 341, pp. 373–386, Feb. 2017.

[4] W. Wang, J. Wang, J. Tian, J. Lu, and R. Xiong, “Application of digital twin in smart battery management systems,” Chin. J. Mech. Eng., vol. 34, no. 1, pp. 1–19, Dec. 2021.

[5] L. Merkle, A. S. Segura, J. T. Grummel, and M. Lienkamp, “Architecture of a digital twin for enabling digital services for battery systems,” in Proc. IEEE Int. Conf. Ind. Cyber Phys. Syst. (ICPS), May 2019, pp. 155–160.

[6] Z. Chen, H. Zhao, Y. Zhang, S. Shen, J. Shen, and Y. Liu, “State of health estimation for lithium-ion batteries based on temperature prediction and gated recurrent unit neural network,” J. Power Sources, vol. 521, Feb. 2022, Art. no. 230892.

[7] S. Pillar, M. Perrin, and A. Jossen, “Methods for state-of-charge determination and their applications,” J. Power Sources, vol. 96, no. 1, pp. 113–120, Jun. 2001.

[8] A. Vezzini, “Lithium-ion battery management,” in Lithium-Ion Batteries. Amsterdam, The Netherlands: Elsevier, 2014, pp. 345–360.

[9] H. Gabbar, A. Othman, and M. Abdussamad, “Review of battery management systems (BMS) development and industrial standards,” Technologies, vol. 9, no. 2, p. 28, Apr. 2021.

[10] C. Vidal, P. Malysh, P. Rollmeyer, and A. Emadi, “Machine learning applied to electrified vehicle battery state of charge and state of health estimation: State-of-the-art,” IEEE Access, vol. 8, pp. 52796–52814, 2020.

[11] E. Ezemobi, M. Silvagni, A. Mozaffari, A. Tonoli, and A. Khajepour, “State of health estimation of lithium-ion batteries in electric vehicles under dynamic load conditions,” Energies, vol. 15, no. 3, p. 1234, Feb. 2022.

[12] L. Driscoll, S. De La Torre, and J. A. Gomez-Ruiz, “Feature-based lithium-ion battery state of health estimation with artificial neural networks,” J. Energy Storage, vol. 50, Jun. 2022, Art. no. 105584.

[13] Q. Li, D. Li, K. Zhao, L. Wang, and K. Wang, “State of health estimation of lithium-ion battery based on improved ant lion optimization and support vector regression,” J. Energy Storage, vol. 50, Jun. 2022, Art. no. 104215.

[14] I. B. Espedal, A. Jinasena, O. S. Burheim, and J. J. Lamb, “Current trends for state-of-charge (SoC) estimation in lithium-ion battery electric vehicles,” Energies, vol. 14, no. 11, p. 3284, Jun. 2021.

[15] V. Etacheri, R. Marom, R. Elazari, G. Salitra, and D. Aurbach, “Challenges in the development of advanced Li-ion batteries: A review,” Energy Environ. Sci., vol. 4, no. 9, pp. 3243–3262, Aug. 2011.

[16] M. Shafro, M. Comroy, R. Doyle, E. Glaesgen, C. Kemp, J. LeMoigne, and L. Wang, “Modeling, simulation, information technology & processing roadmap,” Nat. Aeronaut. Space Admin., vol. 32, no. 2012, pp. 1–38, 2012.

[17] M. Batty, “Digital twins,” Environ. Planning B, Urban Anal. City Sci., vol. 45, no. 5, pp. 817–820, 2018.

[18] L. Wright and S. Davidson, “How to tell the difference between a model and a digital twin,” Adv. Model. Simul. Eng. Sci., vol. 7, no. 1, pp. 1–13, Dec. 2020.

[19] H. Dai, B. Jiang, X. Hu, X. Lin, X. Wei, and M. Pecht, “Advanced battery management strategies for a sustainable energy future: Multilayer design concepts and research trends,” Renew. Sustain. Energy Rev., vol. 138, Mar. 2021, Art. no. 110480.

[20] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, and Z. Chen, “A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems,” Renew. Sustain. Energy Rev., vol. 131, Oct. 2020, Art. no. 110015.

[21] X. Qu, Y. Song, D. Liu, X. Cui, and Y. Peng, “Lithium-ion battery performance degradation evaluation in dynamic operating conditions based on a digital twin model,” Microelectron. Rel., vol. 114, Nov. 2020, Art. no. 113887.
[22] W. Li, M. Rentemeister, J. Bededa, D. Jost, D. Schulte, and D. U. Sauer, “Digital twin for battery systems: Cloud battery management system with on-state-of-charge and state-of-health estimation,” J. Energy Storage, vol. 30, Aug. 2020, Art. no. 101557.

[23] H. Tang, Y. Wu, Y. Cai, F. Wang, Z. Lin, and Y. Pei, “Design of power lithium battery management system based on digital twin,” J. Energy Storage, vol. 47, Mar. 2022, Art. no. 103679.

[24] R. Ramachandran, D. Ganeshaperumal, and B. Subathra, “Parameter estimation of battery pack in EV using extended Kalman filters,” in Proc. IEEE Int. Conf. Clean Energy Efficient Power. Circuit Sustain. Develop. (INCCES), Dec. 2019, pp. 1–5.

[25] M. Baumann, S. Rohr, and M. Lienkamp, “Cloud-connected battery management for decision making on second-life of electric vehicle batteries,” in Proc. 13th Int. Conf. Ecol. Vehicles Renew. Energies (EVER), Apr. 2018, pp. 1–6.

[26] L. Merkle, M. Pöthig, and F. Schmid, “Estimate e-Golf battery state using diagnostic data and a digital twin,” Batteries, vol. 7, no. 1, p. 15, Feb. 2021.

[27] Y. Preger, H. M. Barkholtz, A. Fresquez, D. L. Campbell, B. W. Juba, H. S. Ramadan, M. Becherif, and F. Claude, “Extended Kalman filter for lithium-ion batteries for real applications,” in Proc. IEEE 5th Int. Conf. Condition Assessment Techn. Electr. Syst. (CATCON), Dec. 2021, pp. 310–311.

[28] J. Tian, R. Xiong, and W. Shen, “A review on state of health estimation for lithium-ion batteries in photovoltaic systems,” eTransportation, vol. 2, Nov. 2019, Art. no. 100028.

[29] P. W. Khan and Y.-C. Byun, “Blockchain-based peer-to-peer energy trading and charging payment system for electric vehicles,” Sustainability, vol. 13, no. 14, p. 7962, Jul. 2021.

[30] Z. B. Omariba, L. Zhang, and D. Sun, “Review of battery cell balancing methodologies for optimizing battery pack performance in electric vehicles,” IEEE Access, vol. 7, pp. 129335–129352, 2019.

[31] Z.-Y. Hou, P.-Y. Lou, and C.-C. Wang, “State of charge, state of health, and state of function monitoring for EV BMS,” in Proc. IEEE Int. Conf. Consumer Electron. (ICCE), Jun. 2016, pp. 1–6.

[32] D. N. How, M. A. Hannan, M. H. Lipu, and J. P. Ker, “State of charge estimation for lithium-ion batteries using model-based and data-driven methods: A review,” IEEE Access, vol. 7, pp. 136116–136136, 2019.

[33] T. J. Kim, J. Oh, and H. Lee, “Review on battery thermal management system for electric vehicles,” Appl. Thermal Eng., vol. 149, pp. 192–212, Feb. 2019.

[34] A. Farmann and D. U. Sauer, “Comparative study of reduced order equivalent circuit models for on-board state-of-available-power prediction of lithium-ion batteries in electric vehicles,” Appl. Energy, vol. 225, pp. 11102–11122, 2018.

[35] K. Liu, Y. Li, X. Hu, M. Lucu, and W. D. Widanage, “Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries,” IEEE Trans. Ind. Informat., vol. 16, no. 6, pp. 3767–3777, Jun. 2020.

[36] W. Li, D. Cao, D. Jost, F. Ringbeck, M. Kuipers, F. Frie, and D. U. Sauer, “Parameter sensitivity analysis of electrochemical model-based battery management systems for lithium-ion batteries,” Appl. Energy, vol. 269, Jul. 2020, Art. no. 115104.

[37] J. Du, Z. Liu, Y. Wang, and C. Wen, “An adaptive sliding mode observer for lithium-ion battery state of charge and state of health estimation in electric vehicles,” Control Eng. Pract., vol. 54, pp. 81–90, Sep. 2016.

[38] X. Lai, C. Qin, W. Gao, Y. Zheng, and W. Yi, “A state of charge estimator based extended Kalman filter using an electrochemistry-based equivalent circuit model for lithium-ion batteries,” Appl. Sci., vol. 8, no. 9, p. 1592, Sep. 2018.

[39] H. S. Ramadan, M. Becherif, and F. Claude, “Extended Kalman filter for accurate state of charge estimation of lithium-based batteries: A comparative analysis,” Int. J. Hydrogen Energy, vol. 42, pp. 29033–29046, Sep. 2017.

[40] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. Van Mierlo, and P. Van Den Bossche, “Critical review of state of health estimation methods of Li-ion batteries for real applications,” Renew. Sustain. Energy Rev., vol. 56, pp. 572–587, Apr. 2016.