

Article

Smart Lithium-Ion Battery Monitoring in Electric Vehicles: An AI-Empowered Digital Twin Approach

Mitra Pooyandeh and Insoo Sohn *

Division of Electronics & Electrical Engineering, Dongguk University, Seoul 04620, Republic of Korea; mitra.p@dgu.ac.kr

* Correspondence: isohn@dongguk.edu

Abstract: This paper presents a transformative methodology that harnesses the power of digital twin (DT) technology for the advanced condition monitoring of lithium-ion batteries (LIBs) in electric vehicles (EVs). In contrast to conventional solutions, our approach eliminates the need to calibrate sensors or add additional hardware circuits. The digital replica works seamlessly alongside the embedded battery management system (BMS) in an EV, delivering real-time signals for monitoring. Our system is a significant step forward in ensuring the efficiency and sustainability of EVs, which play an essential role in reducing carbon emissions. A core innovation lies in the integration of the digital twin into the battery monitoring process, reshaping the landscape of energy storage and alternative power sources such as lithium-ion batteries. Our comprehensive system leverages a cloud-based IoT network and combines both physical and digital components to provide a holistic solution. The physical side encompasses offline modeling, where a long short-term memory (LSTM) algorithm trained with various learning rates (LRs) and optimized by three types of optimizers ensures precise state-of-charge (SOC) predictions. On the digital side, the digital twin takes center stage, enabling the real-time monitoring and prediction of battery activity. A particularly innovative aspect of our approach is the utilization of a time-series generative adversarial network (TS-GAN) to generate synthetic data that seamlessly complement the monitoring process. This pioneering use of a TS-GAN offers an effective solution to the challenge of limited real-time data availability, thus enhancing the system's predictive capabilities. By seamlessly integrating these physical and digital elements, our system enables the precise analysis and prediction of battery behavior. This innovation—particularly the application of a TS-GAN for data generation—significantly contributes to optimizing battery performance, enhancing safety, and extending the longevity of lithium-ion batteries in EVs. Furthermore, the model developed in this research serves as a benchmark for future digital energy storage in lithium-ion batteries and comprehensive energy utilization. According to statistical tests, the model has a high level of precision. Its exceptional safety performance and reduced energy consumption offer promising prospects for sustainable and efficient energy solutions. This paper signifies a pivotal step towards realizing a cleaner and more sustainable future through advanced EV battery management.



Citation: Pooyandeh, M.; Sohn, I. Smart Lithium-Ion Battery Monitoring in Electric Vehicles: An AI-Empowered Digital Twin Approach. *Mathematics* **2023**, *11*, 4865. <https://doi.org/10.3390/math11234865>

Academic Editors: Marina Anca Cidota and Clara Maathuis

Received: 1 November 2023

Revised: 30 November 2023

Accepted: 1 December 2023

Published: 4 December 2023

Keywords: digital twin; state of charge; lithium-ion battery; time-series GAN

MSC: 68T07



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Electric vehicles (EVs) have become a symbol of hope in the face of the energy crisis and the impending global greenhouse effect. There has been notable progress in EV technology sectors in 2023, including those of battery technology, autonomous driving, and other innovative technologies that are driving the industry forward, thus signaling their indispensable role in shaping our sustainable future. Lithium-ion batteries (LIBs) are at the forefront of this revolution, and they are known for their remarkable energy density,

efficiency, and life cycle. There is enormous significance for LIBs not only in the realm of EVs, but also in the realm of renewable energy storage and portable electronic devices. The state of charge (SOC) of LIBs is critical, as it serves as a guide for users and is a crucial component of battery management systems. As we enter a new era of renewable energy in which mitigating climate change is paramount, accurate SOC estimation for LIBs is becoming increasingly important. Battery energy storage systems are emerging as key players in this transformation, paving the way for economic, environmental, and social sustainability. This pursuit is driven by high-fidelity digital twin (DT) models, which provide insights into the complexities of battery system performance across various domains, including the fight against climate change. The efficient management of these batteries is paramount to their longevity and optimal performance. However, accurately predicting the SOC in these batteries remains a complex challenge.

In this paper, we delve into the realm of digital twin technology, a cutting-edge approach that promises to revolutionize SOC forecasting. By creating virtual replicas of real-world batteries, we aim to enhance our understanding of their behavior and provide more accurate SOC predictions. This research holds the potential to transform battery management systems, prolong battery life, and enable smarter energy consumption.

EVs need a reliable battery management system (BMS) to monitor the battery state. The SOC is a crucial factor of a BMS that determines the remaining battery energy and the time that it can last before charging. SOC estimation is complicated due to the complex dynamics of LIBs and changing external conditions. There are three primary categories of SOC estimation methods in the literature: physics-based electrochemical models [1], electrical equivalent circuit models [2,3], and data-driven models based on artificial intelligence algorithms, such as neural networks [4], support vector machines [5], random forests [6], and many others. The choice of model depends on factors such as accuracy requirements, data availability, computational resources, and real-world application needs, and the strengths and limitations of each category are considered. Physics-based models are highly detailed and complex, as they consider electrochemical reactions, diffusion, and heat transfer phenomena. Accurate predictions of the SOC can be made with these models, but they need a significant number of parameters related to the battery's materials, geometry, and operating conditions. Proper calibration is necessary for optimal results. Electrical equivalent circuit models simplify the complex electrochemical processes inside a battery into an electrical circuit. These methods use empirical data to estimate circuit parameters and are computationally efficient, which makes them appropriate for real-time applications. For certain applications, physics-based models may have a higher level of accuracy than that of other models.

Data-driven methods have become popular in recent years due to their ability to estimate the SOC using only battery measurement data. Machine learning algorithms such as support vector machines (SVCs) [7], artificial neural networks [8], and fuzzy logic [9] have been used for SOC estimation. According to [10], there are three types of neural network methods: feed-forward neural networks, deep learning neural networks, and hybrid neural networks. The authors discussed the accurate estimation of a lithium-ion battery's state of charge for high-level electric vehicles to support carbon neutrality and emission peak policies. The focus is on neural network techniques that provide precise SOC estimates without considering the battery's internal electrochemical state. However, these methods have limitations in accurately estimating the SOC due to the diversity and complexity of LIBs. Significant amounts of human labor and expertise are required for their training, and the learning architectures that they employ are relatively shallow. In response to these issues, the deep neural network (DNN) has proven to be a viable solution. By using a single DNN layer, it is possible to automatically learn a novel representation of the input. Additionally, the multi-layered structure of the network allows for the extraction of intricate feature information, which can be built from the original input. Currently, most researchers propose the estimation of batteries' SOC using DNN-based techniques, such as long short-term memory (LSTM) and a gated recurrent unit (GRU) [11]. However, deep

learning algorithms may overfit if they are not trained on high-quality datasets. To tackle this problem, regularization methods such as L1/L2 regularization and dropout layers are employed. Nevertheless, implementing these techniques necessitates specialized hardware, considerable computational resources, energy, and extensive fine-tuning. To mitigate these challenges, hybrid strategies that combine deep learning with other techniques are often employed. On the other hand, the growing complexity of battery systems coupled with the need for safety, optimization, sustainability, and cost-effectiveness has driven the demand for digital twins in the battery industry. To achieve energy autonomy and sustainability, renewable energy sources, storage systems, and energy management systems [12] must be efficiently integrated. In this context, our paper suggests a framework for digital twins that has the potential to transform battery management systems. This framework has the ability to extend battery life, enhance energy consumption intelligence, and align with evolving energy management technologies.

A digital twin is a virtual model that replicates a physical object or system in great detail and in real time. It captures the characteristics, behavior, and performance of its real-world counterpart and is utilized for analysis, monitoring, and simulation [13]. Digital twins are employed in multiple industries, such as manufacturing [14] and healthcare, to enhance comprehension, efficiency, and decision making. Digital twins are crucial for advancing battery technology and addressing challenges related to energy storage, control, and management. In a recent study [15], researchers introduced a control technique called the deep deterministic policy gradient (DDPG) that used reinforcement learning to improve the performance of maximum-power-point tracking (MPPT) controllers in photovoltaic (PV) energy systems. By utilizing a digital twin for simulation training, the RL approach was able to operate effectively under different weather conditions. The study found a significant improvement in the real-time total power output, with a 51.45% increase and 24.54-times-faster settling time compared to conventional P& O controllers. Digital twins offer several benefits for predicting the SOC. Firstly, they can generate synthetic data that accurately reflect the battery's behavior, even in extreme or rare situations. Secondly, they can speed up the development of SOC prediction models by providing a virtual platform for training and testing. Moreover, they can improve the reliability of SOC prediction models by creating synthetic data that cover a wider range of operating conditions. Lastly, digital twins can produce a realistic dataset without compromising sensitive information. To ensure authenticity, statistical distributions or generative models can be used to manage the random processes involved in data generation.

Additionally, digital twins can enhance the reliability of prediction models by producing synthetic data that cover a broader range of operating conditions while maintaining the privacy of sensitive information. Our digital twin framework incorporates the use of a TS-GAN to produce synthetic battery data, which greatly improves the capabilities of our system. GANs are a valuable tool for addressing data limitations in machine learning models. They have the ability to enhance small datasets, protect privacy, produce outliers, balance class distributions, manage missing data, create realistic testing environments, reduce expenses, and overcome data bias. By generating synthetic data that closely resemble real-world data, TS-GANs are able to increase diversity and robustness. They also aid in balancing class distributions, filling in missing data points, and creating realistic test environments. GANs also reduce the need for exhaustive data collection, making them a cost-effective and time-saving solution. Ultimately, TS-GANs improve the performance and capabilities of machine learning models. This innovative approach helps us overcome the challenge of limited and unreliable real-world battery data. By creating synthetic data that closely resemble real-world behavior, our training dataset is expanded, resulting in more accurate predictions of the SOC and improved battery management. Briefly, this paper proposes a digital twin framework for SOC prediction in EVs using a TS-GAN. The framework can generate synthetic battery data that accurately reflect the battery's behavior. In comparison with previous non-digital-twin methodologies, the accuracy of estimating the SOC of batteries improved with the integration of digital twin technology

and the TS-GAN. The method used in our study outperforms previous methods by utilizing high-quality data and a combination of the digital twin framework and a TS-GAN. This integration not only addresses the inherent limitations of utilizing limited and unreliable real-world battery data but also leads to a substantial improvement in the performance and capabilities of machine learning models. Beyond overcoming data constraints, the TS-GAN plays a pivotal role in augmenting the diversity, robustness, and accuracy of SOC predictions. It provides a practical solution to the pressing need to estimate the SOC more accurately in EVs and battery management systems. In addition to marking a significant advancement in the field, this work also paves the way for future developments at the intersection of digital twin technology and sustainable energy solutions. This study is structured as follows: In Section 1, the research background and motivation are presented through an introduction, Section 2 discusses related work, Section 3 outlines the system framework encompassing offline modeling and a digital twin, Section 4 presents the outcomes of the simulation, an in-depth discussion is provided in Section 5, and Section 6 presents concluding remarks.

1.1. Motivation

LIBs are the most common type of battery used in EVs. However, they are susceptible to degradation over time, which can lead to decreased performance, greater safety risks, and a shorter lifespan. It is possible for a battery to degrade in various ways, including by experiencing low capacity, overheating, unstable voltage, high internal resistance, the risk of overcharging or over-discharging, reduced lifespan, and thermal runaway. However, these risks can be minimized by employing battery management systems that constantly monitor and regulate the battery's condition through predictive algorithms, ensuring its safe operation despite its degradation. An EV requires a BMS to ensure safe and efficient operation by continuously monitoring and managing the battery's health. The BMS prevents overcharging, over-discharging, overheating, and imbalances. Additionally, it predicts maintenance requirements, thus greatly contributing to the overall safety and performance of EVs. Therefore, it is essential to develop effective methods for monitoring and managing lithium-ion batteries in EVs.

Traditional BMSs are typically based on simple heuristics and do not take the complex electrochemical processes that govern battery behavior into account. As a result, they can be inaccurate and unreliable, especially under challenging conditions such as high temperatures or fast charging. Recent advances in machine learning and artificial intelligence (AI) have made it possible to develop more sophisticated methods for monitoring and managing LIBs. These methods can learn the complex relationships between battery parameters and use this knowledge to make accurate predictions about battery behavior. In [16], a BPNN was trained using three distinct classifications of input features, along with random numbers of weights, biases, and hidden neurons. The resulting model was able to attain an average SOC error of 3.8% during the US06 (one of the commonly used drive cycles assessing the capabilities of vehicles—especially electric ones; this is known as a standardized testing method) drive cycle at a temperature of 25 °C. An RNN based on the NARXNN architecture was proposed in [17] for the accurate estimation of the lithium-ion battery SOC. The NARXNN technique emphasizes global feedback and backpropagation learning in order to improve robustness and computational intelligence. LiFP and lithium titanate (LiTO) were used in dynamic charge and discharge current profiles to experimentally test the effectiveness of the method in SOC estimation. The results showed that NARXNN was accurate and had a low computational cost. The integration of digital twins and AI into battery management systems can become a game changer in battery management intelligence. Digital twin integration can revolutionize the real-time monitoring and optimization of lithium-ion batteries by leveraging state-of-the-art techniques for accurate SOC estimation by creating virtual replicas that mirror real-world batteries. Digital twins can predict the SOC, health, and performance in real time to improve energy efficiency, extend battery life, and prevent failures. In recent studies, digital twins were utilized to

optimize energy consumption and facilitate proactive maintenance. The main objective of the authors of [18] was to create a digital twin specifically for hydropower turbines. They emphasized the importance of dynamic modeling, data interfaces, and adaptive learning for accurately representing system dynamics. The researchers utilized the recursive least squares algorithm as an adaptive learning method to develop hydropower turbine models for the DT. The successful application of this approach was demonstrated through its implementation in a pilot system at the Norwegian University of Science and Technology, which yielded positive adaptive learning and validation outcomes. To ensure accurate predictions and enable responsive battery management, one of the key research questions in this area was that of how to effectively model and simulate complex battery behaviors within a digital twin.

It is important to note, however, that the low availability of real-time data poses one of the major challenges of utilizing digital twin technology for BMSs. A digital twin's accuracy is limited in many cases due to the inability to collect data directly from a physical battery in real time; this improves battery management but hinders timely decisions.

Various methods have been proposed to supplement digital twins for lithium-ion batteries, including the use of historical battery data and the creation of synthetic data through computer simulations. Developing efficient battery management systems that conserve energy is crucial for the advancement of the time-series analysis of lithium-ion batteries. The battery industry faces growing complexities and demands for enhanced safety, optimization, sustainability, and cost-effectiveness. Developing batteries with higher energy density—particularly for electric vehicles—is a significant challenge that requires advanced materials and improved battery chemistry. Digital twins serve as virtual replicas that are capable of propelling battery technology forward by improving SOC anticipation and creating realistic datasets without compromising sensitive information. Incorporating a TS-GAN into the SOC prediction process can address the challenge of limited access to real-world battery data by generating synthetic data that closely mimic real-world behavior. This comprehensive SOC prediction system bridges the gap between physical and digital realms to contribute to the optimal performance, safety, and longevity. It has the potential to revolutionize the way in which lithium-ion batteries are managed in electric vehicles, thus ensuring a more sustainable and effective future for electric vehicles due to its advantages, such as precise SOC prediction, real-time monitoring, lifespan prediction, improved battery management, synthetic data production, and privacy concerns.

1.2. Contribution of Our Work

We outline the key contributions of our work in this section. Our work makes significant contributions to the field of battery management systems by integrating digital twins and AI. This integration brings about a new era of energy-efficient smart battery management and addresses crucial research questions related to energy consumption. Our system utilizes advanced techniques for accurate SOC estimation by creating virtual replicas that predict the SOC, health, and performance in real time. This leads to improved energy efficiency, extended battery life, and the prevention of failures. We also introduce methods for generating supplementary data to enhance the accuracy and timeliness of our digital twin's predictions, which aids in responsive battery management and extends the capabilities of our system in optimizing energy consumption. Our system has the potential to revolutionize the management of LIBs in EVs, thus contributing to optimal energy performance, enhanced safety, and prolonged battery longevity. We also utilize TS-GANs to generate synthetic data that closely mimic real-world behavior to expand our training dataset and lead to more accurate predictions of the SOC. Our approach aligns with the goals of sustainable transportation and reduced greenhouse gas emissions, and promotes responsible battery usage, financial savings, and environmental advantages. Overall, our contributions advance our understanding of LIB management in EVs, offering an innovative, comprehensive solution that fosters a more sustainable and energy-efficient future for electric vehicles and the broader energy landscape.

2. Related Work

This section provides a summary of previously published literature reviews on the estimation of the SOC of LIBs and the utilization of digital twins that are integrated into battery management systems for SOC estimation. Due to its critical role in minimizing energy consumption in electric vehicles, the accurate estimation of the SOC has gained significant attention in recent years. A major goal of this research area is to provide manufacturers and BMS designers with the tools that they need to create high-performance batteries. While advances have been made in estimation methodologies for the SOC, generative adversarial networks (GANs) stand out as a promising approach. Through this innovative technique, real-time data are created from synthetic time-series data that are designed to simulate battery behavior in real-life settings, thus ultimately improving the accuracy of SOC estimations. Achieving precise estimation not only facilitates the production of low-energy batteries but also allows the fine-tuning of battery management strategies and the extension of battery life. The purpose of this section is to contribute to a greater understanding of BMSs through the synthesis of existing research and the use of this novel approach to the estimation of SOC.

2.1. SOC Estimation Based on LSTM Networks in BMSs

LSTM networks are increasingly used to estimate the SOC in LIBs due to their ability to capture complex battery behavior in real time. Their expertise lies in creating models that capture the impacts of variables such as voltage, current, and temperature on nonlinear relationships. Combining LSTM networks with other techniques further enhances accuracy. This approach is vital for optimizing battery performance in electric vehicles and renewable energy systems. A new model for SOC estimation in LIBs was proposed in [19]. LSTM cells were used to model state and process information together. A two-stage pre-training strategy was used to improve the feature-learning capabilities and resolve dynamic differences between loading profiles and sampling frequencies. Despite the variable sampling frequencies and unknown loading profiles, the proposed method achieved high accuracy in two cases with different batteries.

In [20], an encoder–decoder framework with bidirectional LSTM networks was used to capture sequential patterns in battery behavior. The bidirectional LSTM networks improved the model's ability to capture long-term dependencies from both past and future data points, resulting in improved estimation accuracy. The method was evaluated on publicly available battery datasets—particularly those with dynamic loading profiles—and demonstrated precise SOC estimation across diverse temperature conditions with mean absolute errors as low as 1.07%. This approach significantly enhanced the reliability and applicability of battery management systems for real-world scenarios with varying ambient conditions. A more comprehensive evaluation of the practical usefulness of this method can be achieved by examining its applicability to a broader range of battery types, as battery characteristics can differ significantly. The authors of [21] presented a new neural network called the uncorrelated sparse autoencoder with long short-term memory (USAL) for estimating the SOC in battery-powered machines over a long period of time. USAL combined the benefits of sparse autoencoders and LSTM networks by using a multi-task learning approach to penalize for high multi-collinearity between encodings and identify long and short temporal correlations between them. Using three publicly available accelerated aging datasets, the network outperformed existing models after training on five initial charge–discharge cycles. Overall, USAL showed promise for identifying patterns relevant to long-term SOC estimation that are typically missed by other methods, even with limited initial battery history. In [22], a novel technique was recommended for determining the SOC of large-scale LIB storage systems. The method employed an LSTM neural network model that could handle a nonlinear battery model and the uncertainties involved in the estimation process. The LSTM model surpassed the feed-forward neural network (FFNN) and deep-feed-forward neural network (DFFNN) models on three datasets. Real-world data from the Al-Manara PV power plant were used to train the model, which consistently

produced precise SOC calculations with a maximum standard error (MSE) of less than 0.62%. In contrast, the FFNN and DFFNN models exhibited MSEs ranging from 5.37% to 9.22% and 4.03% to 7.37%, respectively. In order to enhance the accuracy of SOC estimation in LIBs, a new technique called PSO-LSTM was introduced [23]. This approach combined an LSTM neural network with particle swarm optimization (PSO). The PSO algorithm was utilized to optimize the LSTM parameters to align them with the unique characteristics of a battery for optimal SOC prediction. Moreover, random noise was added at the input layer to improve the network's robustness against interference. The experimental results demonstrated that the PSO-LSTM framework consistently achieved a small error margin of only 0.5% when predicting the actual state of charge. The authors of [24] introduced an optimization technique for enhancing SOC estimation using LSTM networks. The novel approach incorporated fractal derivatives—specifically, the improved Borges derivative (a local fractional-order derivative that can be used to describe the dynamic behavior of complex systems)—into LSTM parameter optimization. This integration extended the concept to adaptive momentum estimation (Adam) algorithms, replacing integer-order derivatives with improved Borges derivatives. A comparative analysis of the improved Borges derivative's speed relative to the conventional integer-order derivative was conducted. That study also proposed an order-tuning method to effectively adjust the parameter training speed. The authors of [25] employed an LSTM recurrent neural network architecture to directly estimate the SOC of a battery by observing variables such as temperature, current, and voltage. This approach eliminated the need for utilizing inference systems from the Kalman filter family. The results demonstrated that the average mean absolute error (MAE) was consistently reduced to less than 1% across multiple test scenarios, underscoring the potential of deep learning in SOC estimation. This achievement was significant within the domain of cell state-of-charge estimation, as it replaced conventional estimation methods that were reliant on cell circuit modeling and fine-tuning SOC estimation accuracy through the use of Kalman filter family techniques. Additionally, the researchers made the battery test data publicly available to facilitate further research. The SOC estimation methods discussed in this section are all mostly effective.

However, our proposed method has the following advantages:

1. It uses a more robust LSTM architecture with four hidden layers. This allows it to capture more intricate temporal dependencies in the data, resulting in more accurate predictions;
2. It uses three different optimizers (SWATS, Adam, and SGD) during training. This allows the model to select the most effective optimizer for the data, resulting in faster convergence and better generalization;
3. It is less sensitive to noise and outliers in the data;
4. It compares the performance of LSTM models trained with different optimizers against a GRU network. This provides a more comprehensive understanding of the strengths and weaknesses of each architecture, allowing the researcher to choose the best one for the task at hand.

As a result, the proposed method is more robust and accurate in general. Battery management systems and other applications that require accurate SOC estimation could benefit from this approach.

2.2. Integration of Digital Twins into BMSs

In recent years, the use of cyber-physical systems—also known as digital twins—has become more widespread due to the availability of low-cost sensors and the deployment of Internet of Things (IoT)-enabled devices. To create a virtual representation of a physical system, remote sensing is combined with cloud-based models. Incorporating cloud technology into a BMS facilitates the real-time monitoring and control of battery performance, thus improving energy storage efficiency and providing data-driven insights.

The authors of [26] tackle the difficulties of managing batteries by combining edge and cloud technologies to monitor and regulate charging and discharging. This approach

resolved problems that arise with solely edge-based systems, such as limited computational power, speed, and data storage capacity. The system involved a web interface for remote monitoring and control, and the prototype successfully executed battery control commands while ensuring precise data transmission between the edge and cloud, leading to the comprehensive storage of historical data.

Digital twins of batteries can be used to develop multi-scale intelligent management systems. In addition, there are challenges such as the need for multiphysics models, nano-/microscale characterization, and low-latency communication networks. Additionally, effective data pre-processing and increased data security must also be considered. Battery control and lifetime estimation have shifted from an empirical approach to model-driven techniques, with data-driven and machine learning (ML) approaches gaining popularity. For instance, the authors of [27] highlighted the significance of effective lithium-ion battery management for a low-carbon future, particularly in applications such as electric vehicles and grid-scale energy storage. The article suggested the possibility of improving real-world control by combining knowledge of battery degradation, modeling tools, diagnostics, and new machine learning techniques to create a digital twin of a battery. This cyber–physical system allowed for the closer integration of the physical and digital aspects of batteries, resulting in smarter control and a longer lifespan and providing a useful framework for future intelligent and interconnected battery management. Nonetheless, challenges in real-world applicability persist. Digital twin technology in EV battery management systems offers advantages such as the real-time monitoring, analysis, and simulation of battery behavior, which enhance the SOC estimation accuracy. Factors such as temperature variations, aging effects, and load fluctuations are incorporated, contributing to prolonged battery life, optimized energy management, and increased confidence in EV range estimation. An efficient and informed decision-making process is enabled by integrating digital twin and AI technologies into BMS SOC estimation, thus promoting a more reliable and efficient EV ecosystem. The authors of [28] suggested a combination of the digital twin approach and parameter estimation methods to estimate the online SOC of a battery pack. They devised a unique approach that combined offline parameter estimation with recursive least squares for online updates to estimate battery parameters. Monitoring the SOC and offering diagnostics are possible using this method. The results obtained from EV battery packs demonstrated that this approach was effective in accurately estimating parameters and the SOC. The authors of [29] proposed a digital twin platform for the degradation assessment of lithium-ion battery packs in spacecraft. The platform was composed of visual software and an assessment unit. Through remote sensing links, the visual software received and analyzed real-time data from the battery pack. Through the use of models and algorithms, the assessment unit determined the battery pack's state of charge (SOC), state of health (SOH), and remaining useful life (RUL). This study used a Kalman filter–least squares support vector machine (KF-LSSVM) for SOC estimation and an autoregressive particle filter (AR-PF) for the evaluation of the SOH. Based on the results, the platform was capable of accurately estimating battery packs' SOC and RUL. The authors of [30] described a cloud-based management system for battery systems. By seamlessly transmitting battery data to the cloud, the system aimed to enhance batteries' computational power and data storage capacity. Afterward, battery diagnostic algorithms were used to determine the charge level and aging of batteries. Using equivalent circuit modeling and cloud-based SOC and SOH estimations, a digital twin of the battery system is created. An adaptive extended H-infinity filter was proposed for accurate state-of-charge estimation for lithium-ion and lead–acid batteries. In addition, a particle swarm optimization algorithm was used for state-of-health estimation to monitor the capacity and power fading during aging. The authors of [31] recommended a performance degradation evaluation model for LIBs in dynamic operating conditions. Using a digital twin, the model calculated the actual capacity of the battery. The digital twin model was generated using the LSTM algorithm, which utilized a health indicator (HI) as a temporal measurement. The HI was derived from measurable parameters to represent the battery's performance degradation.

The results of their experiments demonstrated that the proposed model could precisely estimate the actual capacity of the battery under dynamic operating conditions. The authors of [32] discussed challenges such as range anxiety and slow charging. A special emphasis was placed on the incorporation of digital twin technology in order to identify critical gaps and highlight advanced technologies. They suggested a thorough plan for creating an effective BMS, facilitating live monitoring, and tackling battery recycling in a complete and unified system. The authors of [33] investigated the feasibility of using a digital twin for battery cells. They used a systematic approach and applied it to a Doyle–Fuller–Newman model. There are several benefits of using a battery DT, including improved representation, performance estimation, and behavioral predictions based on real data. To update the battery model parameters, they used PyBaMM, a Python package. Using PyBaMM to develop a digital twin of a battery can be effective, especially for accurately updating parameters during battery cycles. In [34], the authors suggested a digital twin model for BMSs that can forecast battery characteristics such as the temperature, current, and state of charge by solely measuring voltage. They employed linear and multi-linear regression models to make predictions. The experimental outcomes indicated that this method attained a high level of prediction accuracy (over 90%) for variables such as the current, maximum cell voltage, and state of charge. However, the results of maximum/minimum cell temperature prediction were not as impressive. To ensure the dependability and safety of lithium-ion batteries, it is essential to have precise SOH estimation. However, the existing methods that used digital twins necessitated complete charge/discharge cycles, which is not practical for situations with partial discharges. To overcome this challenge, in [35], a new digital twin framework that allowed on-the-go battery SOH sensing and updates to the physical battery model was proposed. This framework included energy-discrepancy-aware cycling synchronization, a time-attention SOH estimation model, and a data reconstruction method based on similarity analysis. Comprehensive benchmark tests showed that this solution provided real-time SOH estimation accuracy within a 1% error range for most sampling instances during ongoing cycles. The critical role of machine learning (ML) models in optimizing battery thermal management system (BTM) design was emphasized by the authors of [36], as they provided a thorough review of the literature on ML-based BTMs. A digital-twin-based method for enhancing BTMs was also proposed.

Our approach successfully integrated and monitored battery management in a holistic and comprehensive manner. By incorporating digital twin technology, we surpassed existing methodologies and achieved greater robustness and real-time accuracy and created a comprehensive framework for EV battery management systems.

2.3. GANs for Generating Synthetic Data

Artificial intelligence has made great strides toward producing data that mimic real-life patterns. One of the most important techniques in this area is the use of GANs, which allow for the creation of synthetic data that are very similar to actual datasets. The potential of GANs is vast, as they can generate a wide variety of datasets that accurately represent different industries, such as finance, healthcare, and manufacturing, without compromising sensitive information. Additionally, they provide a unique opportunity to delve into the complexities that drive real-world phenomena. Digital twins use GANs to obtain high-fidelity, real-time data in order to perform precise simulations and predictions, which rely on accurate and up-to-date data. Digital twin simulations and predictions are made more accurate by combining the generator and discriminator components of GANs. This approach also enables digital twins to operate with up-to-date, high-quality data even in scenarios where access to real-time data is limited.

Gene expression datasets in cancer research often have a small sample size due to privacy constraints, making it difficult to accurately classify cancer types. Data augmentation can increase the dataset size by generating synthetic data points. GANs are capable of generating synthetic data, and a modified generator GAN (MG-GAN) generates synthetic data that conform to a Gaussian distribution [37]. In comparison to traditional

data augmentation methods, a MG-GAN significantly improved cancer type classification accuracy for a breast cancer patient gene expression dataset. A significant reduction in error function loss was achieved with the MG-GAN (from 0.6978 to 0.0082), demonstrating its high sensitivity. In [38], a framework for generating synthetic data that mimicked the joint distribution of variables within original electronic health record (EHR) datasets while managing ambiguities in anonymization was proposed. The methodology leveraged conditional generative adversarial networks (CGANs) to synthesize data while prioritizing patient confidentiality, resulting in a model that was named ADS-GAN. ADS-GAN's ability to replicate joint distributions and uphold patient privacy was validated through rigorous evaluations against real datasets. In [39], synthetic data creation was emphasized as a growing necessity in financial services. The main areas of focus were the generation of genuine artificial datasets, assessment of the resemblance between actual and produced information, and maintenance of confidentiality throughout the creation procedure. The financial sector faces unique challenges related to regulations and privacy mandates. This paper addressed these challenges and aimed to establish a shared framework and terminology for generating synthetic financial data.

There are numerous studies proposing GAN-based algorithms for generating synthetic data for a variety of industrial applications, including healthcare and finance, but GAN applications in the realm of digital twins remain unexplored. An innovative approach within the digital twin domain is the generation of real-time synthetic data, which this study pioneers. The study also introduces a secure framework that combines Digital Twin technology with a cloud environment specifically for BMSs. The following section provides more details on the complexities of the proposed framework.

3. Proposed System Framework

Our proposed system provides a comprehensive solution for monitoring LIBs in EVs by addressing both physical and digital aspects. The system consists of two main components: the physical side, which encompasses offline modeling, and the digital side, which includes real-time monitoring and prediction (Figure 1).

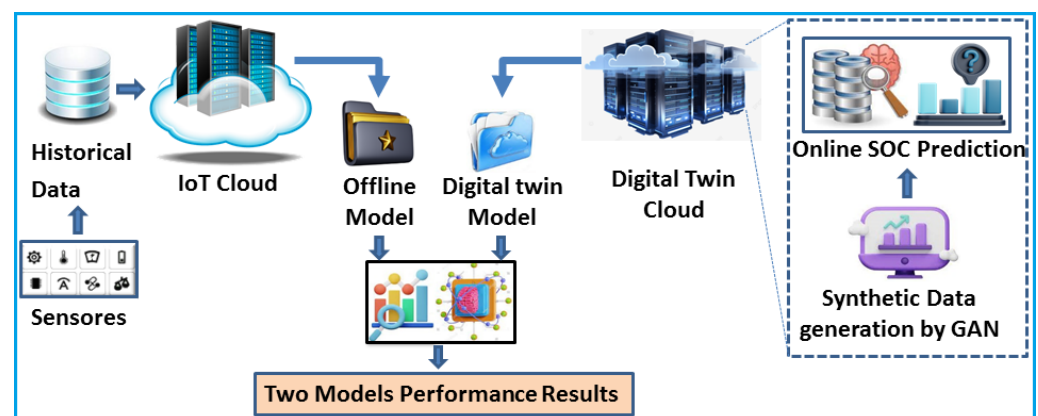


Figure 1. Incorporated digital twin and physical system architecture for lithium-ion battery monitoring.

Physical Side: The physical side of our system involves a cloud-based IoT network that connects and monitors all EVs equipped with LIBs on the road. According to [40], today, data from various sources and IoT devices are generated and transmitted over cloud-based networks. Virtualized resources, such as servers, storage, databases, and networking, are available on demand through cloud computing. We consider that each EV is equipped with a BMS responsible for monitoring key battery parameters, such as the voltage, temperature, and SOC. This ensures optimal battery performance, safety, and longevity. Additionally, we employ machine learning algorithms to train an offline model, which accurately represents a physical battery system. A powerful LSTM network is used to train the offline model and incorporates data from four types of LIBs (B0005, B0006, B0007, and B0018). The physical

battery system, along with the offline model, allows precise SOC prediction in the digital twin. This ensures that the virtual representation of the battery closely resembles the real one (Figure 2).

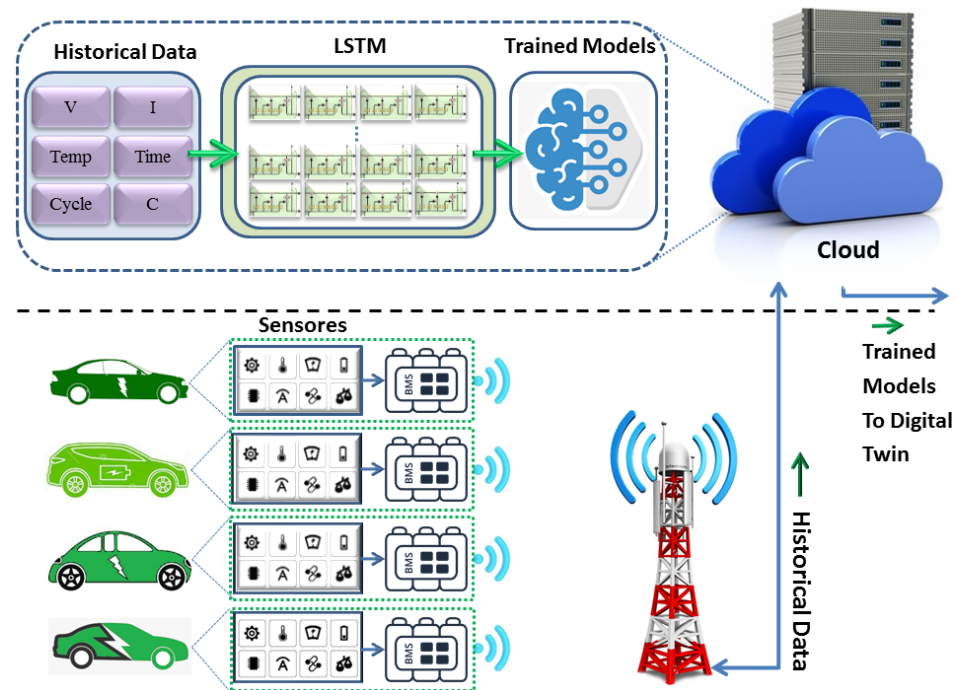


Figure 2. Physical system architecture for lithium-ion battery monitoring.

Digital Side: The digital aspect of our system involves a cloud-based digital twin that acts as a virtual version of a lithium-ion battery (Figure 3). This digital twin allows for the real-time monitoring and prediction of battery activity, thus facilitating efficient battery management. To address the challenge of limited access to real-time battery data, we employ a TS-GAN to generate synthetic data that mimic real-time battery performance. These synthetic data supplement the physical infrastructure and enhance the digital twin’s ability to accurately predict the SOC. It generates additional data samples that are similar to the real data, thus expanding the training dataset for improved model generalization. Additionally, it mitigates the issue of mode collapse by diversifying the generated samples and creating a more balanced dataset when combined with real data. Furthermore, the TS-GAN addresses class imbalance by generating synthetic samples for minority classes. Lastly, it enables the fine-tuning of pre-trained models with synthetic data, making it valuable when limited task-specific data are available, thereby reducing the need for collecting a large labeled dataset, which can be costly and time-consuming. The digital twin utilizes these synthetic data to provide real-time information in the virtual environment, thus improving its SOC prediction accuracy. The digital twin’s SOC prediction results are then compared with the offline model’s SOC prediction results to enable performance evaluation.

Combined System: By seamlessly integrating physical and digital components, our system offers an integrated approach to effective LIB monitoring in EVs. This framework enables the precise analysis and prediction of battery behavior, thus contributing to optimal performance, safety, and longevity. Our solution paves the way for a more sustainable and efficient future for electric vehicles. All of the processes mentioned above are demonstrated in Algorithm 1.

Algorithm 1: Battery Monitoring System

Input:

Historical data: X

Manufacturer’s SOC estimates: Y

Real-time battery data: X' (generated by the TS-GAN)

Output:

SOC prediction results from the offline model: Y (Trained Model)

Real-time SOC predictions from the digital twin: Y'

1: **Step 1: Offline Modeling**

2: Train an offline model (M) using historical data (X) and the manufacturer’s SOC estimates (Y).

3: **Input:** X, Y

4: **Output:** Offline model (M)

5:

6: **Step 2: Online SOC Prediction by the Digital Twin**

7: Generate real-time synthetic battery data using the time-series generative adversarial network (TS-GAN).

8: **Input:** X

9: **Output:** X' (Synthetic data generated by the TS-GAN)

10: Perform online SOC prediction using the digital twin.

11: **Input:** X'

12: **Output:** Y'

13:

14: **Step 3: Evaluation of Online Prediction in Comparison with the Offline Model**

15: Compare the digital twin’s SOC prediction results (Y') with the offline model’s SOC prediction results (Y) to evaluate performance.

16: **Input:** Y, Y'

17: **Output:** Performance evaluation metrics (e.g., accuracy, precision, recall)

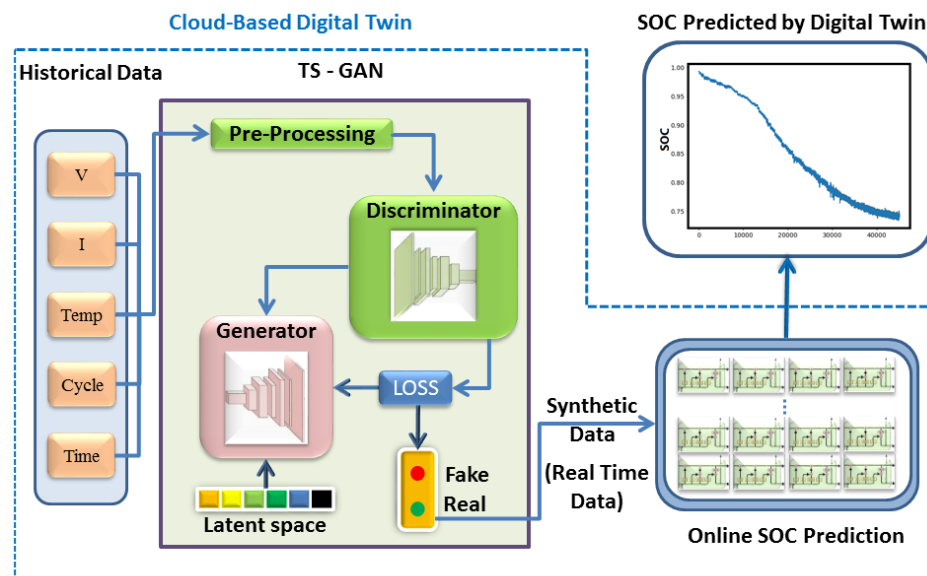


Figure 3. Digital twin framework for real-time lithium-ion battery monitoring.

3.1. Modeling of the Physical Battery System

This section provides an overview of the physical battery system and its offline model, which are important for ensuring that the battery performs properly, is safe to use, and lasts for a long time. By utilizing the offline model, it is possible to accurately monitor and predict battery behavior. The offline model utilizes historical data from battery

manufacturers and physical sensors to establish a comprehensive dataset that reveals the relationships between physical battery parameters and their associated behaviors. This dataset enables the offline model to accurately predict battery behavior and facilitate proactive decision making. To develop the offline model, data are collected from physical sensors and battery manufacturers. Manufacturer-provided information, including battery specifications and manufacturing details, offers invaluable insights into battery chemistry, electrode materials, cell composition, and other relevant parameters. A physical battery system can be represented using LSTM networks [41], which are a form of recurrent neural network (RNN), through the implementation of the offline model. LSTM networks excel at capturing temporal dependencies and patterns in time-series data, making them suitable for analyzing battery behavior.

During the training process, three distinct optimizers were examined, namely, SWATS (Switching from Adam to SGD), Adam, and SGD (stochastic gradient descent). Optimizers play a crucial role in updating a model's weights and biases. The SWATS technique [42] is an innovative optimization approach that merges adaptive techniques such as Adam with stochastic gradient descent (SGD) to tackle the problem of inadequate generalization during deep learning training. By utilizing a triggering condition linked to the projection of the steps of Adam on the gradient subspace, SWATS seamlessly transitions from Adam to SGD. Through experiments conducted on different benchmarks, SWATS was proven to significantly reduce the generalization gap between SGD and Adam in various tasks, including image classification and language modeling. Adam is an optimizer that adapts learning rates (LRs) based on gradient moments, while SGD updates parameters by using gradients from the loss function.

The performance of the LSTM-based offline model trained with these optimizers was compared with that of a model trained using GRU networks, another popular RNN variant. This comparative analysis revealed the strengths and weaknesses of each model architecture, aiding in accurate battery behavior forecasting.

LSTM Model Architecture

The offline model was trained on a separate server from the cloud-based IoT network that included all sensors. This saved time and resources by allowing the model to be trained offline. Additionally, a larger dataset of historical data was used to train the model, thus improving the prediction accuracy. An LSTM network was utilized to train the model offline. LSTM networks are a suitable choice for training offline models in lithium-ion battery monitoring due to the following:

1. Their ability to capture long-term dependencies;
2. Their handling of vanishing/exploding gradients;
3. Their preservation of memory and context;
4. Their accommodation of variable-length sequences;
5. Their provision of robustness against overfitting.

These characteristics enabled the model to effectively capture complex dynamics and temporal patterns in battery behavior, filter out irrelevant information, and adapt to varying data sampling rates. Leveraging these properties can enhance the accuracy and reliability of battery behavior forecasting in electric vehicles. Due to their struggle to retain input information from the past over extended periods of time, traditional recurrent neural networks have difficulty modeling long-term structures. To solve this issue, LSTM is a more effective solution. This challenge becomes even more difficult in conditional generative models, where predictions depend solely on the network's own generated inputs. Although injecting noise into predictions can be helpful, LSTM is specifically designed to enhance information storage and access in recurrent neural networks. It provides a longer-term memory that references past information, which improves predictions while also ensuring stability.

An LSTM unit comprises three main components: the input gate, forget gate, and output gate, which control the flow of information and allow the unit to selectively retain or

discard information over time. The memory cell is the internal memory of the LSTM unit, and it retains information over long sequences. The combination of the gates and memory cell enables the LSTM to capture long-term dependencies in sequential data, making it valuable for tasks such as time-series prediction and natural language processing. During training, the LSTM learns optimal gate parameters and memory cell values through backpropagation through time.

3.2. Proposed Digital-Twin-Based BMS Framework

The digital twin system is a key component of our proposed framework for monitoring LIBs in EVs. The digital twin system serves as a virtual model of the actual battery, allowing for the continuous observation and anticipation of battery performance.

In order to address the issue of limited access to real-time data, we integrated a TS-GAN into the digital twin system to generate synthetic data. The generation of synthetic data is important, as it offers a viable solution for improving the accessibility of data for real-time monitoring and prediction. If synthetic data are easily accessible and kept current, the digital twin can make accurate predictions even when there is a lack of real-world data.

Within this section, we explore the combination of a TS-GAN for the generation of artificial data and an LSTM-based SOC estimator. The battery monitoring and forecasting become more precise and reliable with these components. We will describe the implementation details, training processes, and evaluation results of the TS-GAN-based synthetic data generation and the LSTM-based SOC estimator.

TS-GAN-Based Real-Time Data Generation

In the field of digital twin research, not having prompt access to up-to-date information is a significant issue. There are several factors and challenges that restrict access to real-time data about EVs and their batteries. These include privacy and security concerns, ownership and control issues, regulatory and legal constraints, communication and connectivity challenges, data collection costs, infrastructure limitations, user consent and awareness issues, and technical compatibility. Obtaining real-time data from EVs can be difficult due to limitations and obstacles. However, a TS-GAN can generate synthetic data that closely resemble real-world data, allowing the digital twin system to accurately simulate and predict battery behavior even without real-time data. This approach provides a practical solution for overcoming data scarcity and ensuring continuous monitoring and precise forecasting of EV battery performance. However, we found a remedy by integrating a TS-GAN into the digital twin framework, which generated synthetic data. The generation of synthetic data is crucial as it allows for better access to data for predicting and monitoring things in real time. “Real-time monitoring” means continuously observing and analyzing data as they appear, without any delay. “Prediction” means using data to make educated guesses about future events or trends. In this section, we delve into the details of the TS-GAN-based synthetic data generation, including the implementation, training processes, and evaluation results. The use of synthetic data enhances the precision and reliability of monitoring and forecasting batteries.

The TS-GAN framework, which was presented at NeurIPS in December 2019 by Yoon, Jarrett, and van der Schaar, is a significant development in the field of synthetic time-series data generation [43]. TS-GANs are specifically designed to address the unique challenges associated with generating time-series data, which extend beyond creating cross-sectional distributions of features at each time point. Additionally, a TS-GAN aims to capture the temporal dynamics that govern the sequence of observations in time-series data, reflecting their autoregressive nature.

A TS-GAN is a specialized machine learning model that is intended to operate with time-series data. In order to learn the distribution of transitions between different points in time, it combines unsupervised and supervised learning techniques. Synthetic data that are indistinguishable from real data are the ultimate goal of a TS-GAN.

4. Simulation Results

4.1. Dataset Description

In this study, a dataset comprising four LIBs (labeled #5, #6, #7, and #18) was utilized. This type of data is known as historical or physical data, and the experimental setup closely followed the methodology outlined in [44]. At room temperature (25 °C), batteries were subjected to three distinct operational profiles: charge, discharge, and impedance measurements. Charging was accomplished by applying a constant current of 1.5 A until the voltage reached 4.2 V and then switching to a constant voltage until the charge current dropped to 20 mA. In order to discharge, a constant current of 2 A was applied until specific voltage thresholds were reached. In addition, electrochemical impedance spectroscopy was performed. During repeated cycles, batteries were exposed to accelerated aging, allowing for the observation of changes in internal battery parameters until the end-of-life criteria were met, which were defined as a 30% decrease in the rated capacity. Finally, the dataset included the battery capacity for discharge until 2.7 V, which was recorded and analyzed. The multivariate time-series data obtained from Li-ion batteries included 45,122 samples and eight features, namely, the ID cycle, measured voltage, measured current, measured temperature, capacity, charging current, charging voltage, and time. Additionally, temperature is a crucial factor that significantly affects the performance and overall health of battery systems. Our framework addressed this by incorporating temperature as an important aspect in the following ways.

- **Addition to the Feature Set:** Temperature was a fundamental feature in our dataset, and it was included as one of the key input parameters for our model. Our framework was able to consider the changing temperature levels that occur during battery operation thanks to the inclusion of these data in the dataset.
- **Feature Importance:** We recognize the significance of temperature in influencing battery behavior. In our feature engineering process, temperature was assigned an appropriate weight based on its impacts on various performance metrics, such as the state of charge (SOC), capacity, and voltage.
- **Dynamic Modeling:** Using our framework, we employed a dynamic temperature model that reflected real-time changes. This allowed the model to understand and represent the subtle impacts of temperature changes on the battery's functionality.
- **Temperature-Dependent Responses:** The model responses, including its state estimation and performance predictions, were inherently linked to the temperature feature. The model adjusted its predictions to maintain precision in varying temperature conditions.

4.1.1. Time-Series Data

Time-series data are a type of data that are collected over time at regular intervals or time steps, and they are commonly used in various fields to analyze and forecast trends, patterns, and behaviors that evolve over time. Key characteristics include the temporal order, time intervals, seasonality, trends, and seasonal and residual components. Analyzing and modeling time-series data is essential for tasks such as forecasting [45], anomaly detection [46], pattern recognition [47], time-series imputation, simulation, and digital twin applications. A time series is a set of observations that are arranged in chronological order and captured at regular intervals. Each data point in the series is influenced by its previous values, indicating a temporal correlation or dependence between consecutive observations. The joint distribution of the sequence of observations can be represented using the chaining product rule as follows:

$$p(x_1, x_2, x_3, \dots, x_T) = p(x_1) \prod_{t=2}^T P(x_t | x_1, x_2, \dots, x_{t-1}) \quad (1)$$

where x_t is a data point belonging to T , and the conditional probability $p(\cdot | \cdot)$ for each event signifies the relationship between the current state and its preceding ones in terms of time.

It represents a factorization of a joint probability distribution over a time series of variables $x_1, x_2, x_3, \dots, x_T$ into a product of conditional probabilities. In this research, a dataset consisting of multiple batteries' voltage, current, capacity, and temperature measurements was used. The study focused on four lithium-ion batteries, namely, B0005, B0006, B0007, and B0018. After 168 cycles of charging and discharging, B0005, B0006, and B0007 showed a 30% decrease in their rated capacity, which met the end-of-life (EOL) criteria. Similarly, B0018 reached the EOL criteria after 132 cycles of charging and discharging. This dataset is useful for various purposes, such as predicting and identifying anomalies and recognizing patterns. By utilizing chaining product rules, the researchers were able to effectively model the joint distribution of sequential observations and take advantage of the temporal correlations between these data points.

Figure 4 shows how the voltage of the battery decreased at various time intervals for B0005. As the battery was utilized, its voltage was reduced with every cycle. Moreover, the battery's ability to hold a charge also diminished with an increase in the number of cycles. With a higher cycle count, the battery reached its discharge voltage threshold, which was 2.7 volts, more quickly. Consequently, as the battery aged, its capacity diminished at an accelerated rate.

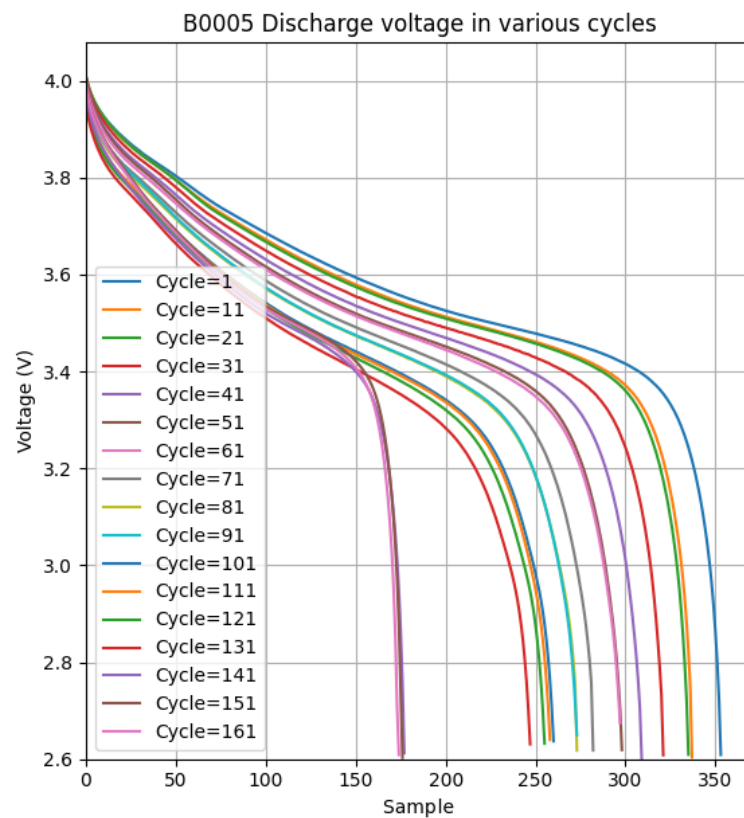


Figure 4. Battery voltage drop over different numbers of cycles for B0005.

4.1.2. Preprocessing

In order to process the raw data for model training, we performed a series of preprocessing steps. To read and display information from the CSV file, we used the pandas library. Subsequently, we used the Matplotlib library to generate a correlation map. Figure 5 displays the connections among the variables based on a heatmap analysis that we conducted as part of our exploratory analysis of the essential features. Five critical features were identified in our analysis: the ID cycle, time, temperature, voltage, and current. These variables directly influenced the SOC. Through this analysis, we sought to quantify the impact of each feature on SOC prediction, thereby enhancing our comprehension of battery behavior and improving the predictive accuracy.

We then partitioned the data into training and testing sets, allocating the initial 35,000 samples for training and the remainder for testing. To ensure efficient calculations during our operations, we normalized the data, scaling the columns to a specific range (0 to 1). Lastly, accounting for the data series' length, we transformed them into fixed-length sequences or windows, creating windows of 1000 time steps each. This process yielded two arrays, X-train and X-test, with dimensions of (33,950, 100, 8) and (9073, 100, 8), respectively, using a window length of 1000. The sampling sequence for voltage, temperature, and current for battery B0005 is depicted in Figure 6.

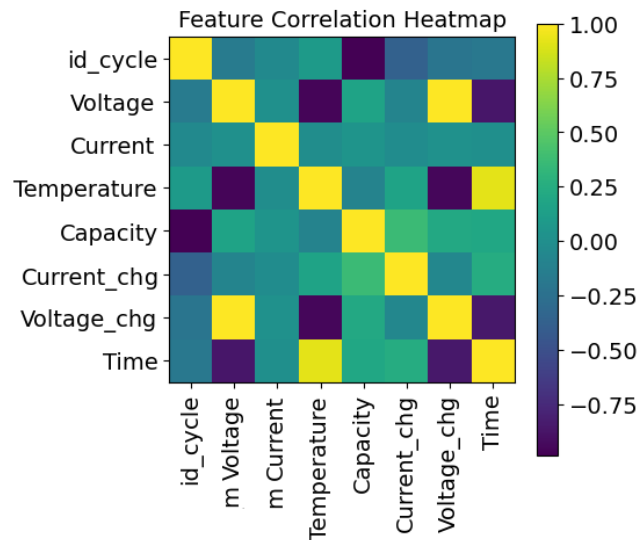


Figure 5. Feature correlation heatmap for SOC prediction in LIBs.

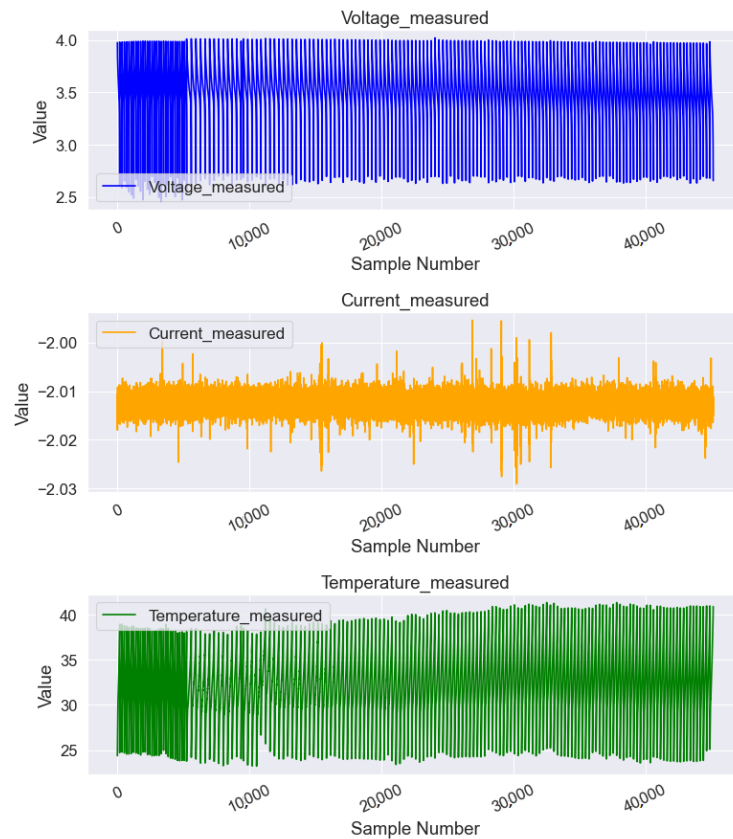


Figure 6. The temporal variations in voltage, current, and temperature for battery B0005 over 168 cycles of charging and discharging.

Additionally, Figure 7 displays the SOC graphs for B0005, B0006, B0007, and B0018 based on information provided by the manufacturer. The batteries were subjected to multiple cycles of charging and discharging until their SOC reached 72%. Batteries B0005, B0006, and B0007 underwent 166 cycles each, while battery B0018 underwent 132 cycles. It can be observed from the graph that the SOC of battery B0018 decreased at a faster rate.

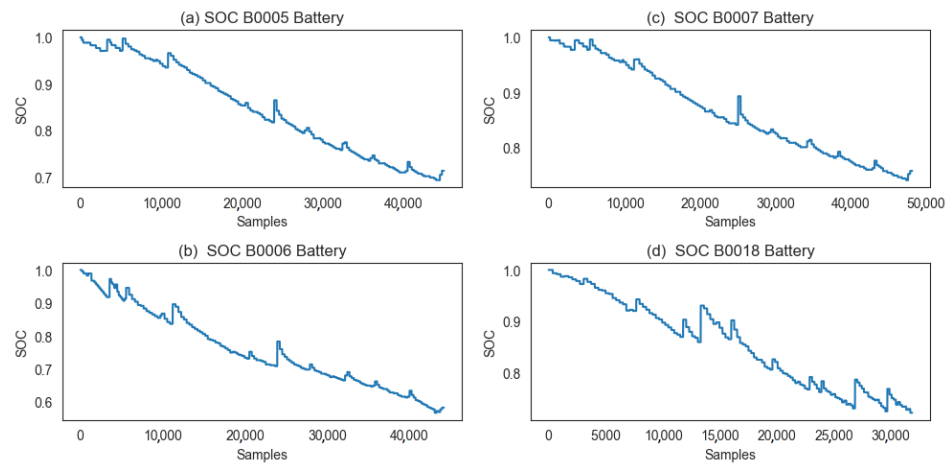


Figure 7. A comparison of the changes in SOC over 168 cycles of charging and discharging for battery B0005 and 132 cycles of charging and discharging for batteries B0006, B0007, and B0018.

4.1.3. Problem Formulation with Physical Batteries and Simulation Results

Overall, the SOC is an important factor when it comes to overseeing and controlling the energy of a battery, as it indicates the present amount of charge that it holds. The SOC can be obtained for time t as follows:

$$SOC_t = SOC_{t_0} - \frac{1}{C} \int_{t_0}^t CE \cdot I_t dt \tag{2}$$

where SOC_t and SOC_{t_0} are the state of charge of the battery at the starting time t_0 and the present time t , respectively, C is the capacity of the battery, CE is the Coulomb efficiency, and I represents the current that passes through the voltage source. Accurate prediction of SOC values is crucial for efficient energy utilization and optimization of battery performance, and the LSTM utilizes sequential patterns in battery data to achieve this goal. Each LSTM unit uses a memory C_t at time t . Here, h_t is an output or LSTM unit activation determined by

$$h_t = o_t \odot \tanh(C_t) \tag{3}$$

where o_t is the output gate that controls the amount of content that is provided through memory.

The output gate is calculated through the following:

$$o_t = \sigma(w_0 \cdot [h_{t-1} + x_t] + b_0) \tag{4}$$

where σ is the sigmoid activation function, w_0 is the weight matrix, x_t is the input at time t , b_0 is a model parameter, and h_{t-1} is the hidden state from the previous time step.

With the partial forgetting of the current memory and addition of new memory content as \hat{C}_t , memory cell C_t is updated to

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \tag{5}$$

where f_t is the activation vector of the forget gate, and i_t is the activation vector of the input/update gate. The amount of current memory that should be forgotten is controlled by the forgetting gate (f_t), and the amount of new memory content that should be added

to the memory cell is controlled by the update gate (i_t) which is sometimes known as the input gate. This is done with the following calculations:

$$f_t = \sigma(w_f \cdot [h_{t-1} + x_t]) + b_f \quad (6)$$

$$i_t = \sigma(w_u \cdot [h_{t-1} + x_t] + b_u) \quad (7)$$

The new memory content is expressed through the following:

$$\hat{C} = \tanh(w_c \cdot [h_{t-1} + x] + b_c) \quad (8)$$

At every time step, an LSTM is provided with details regarding the battery, such as its voltage, current, temperature, and so on. Additionally, the LSTM takes the previous hidden state h_{t-1} , as well as the previous cell state C_{t-1} from the previous time step $t - 1$, into account. Through a particular formulation, an LSTM makes predictions of the SOC. So, we can define SOC at time t as

$$SOC_t = \sigma(w \cdot [h_{t-1} + x_t] + b_0) \quad (9)$$

The problem of estimating the SOC was approached as a supervised learning problem in this study, wherein the model was provided with numerous input–output pairs from which to learn. An offline model was trained using a dataset with historical information obtained from the battery manufacturer. The dataset consisted of 45,122 samples of multivariate time-series data with eight characteristics, including the cycle ID, measured voltage, measured current, measured temperature, capacity, charging current, charging voltage, and time. The objective of the offline model was to utilize the known SOC values to compare and evaluate them against the digital twin's SOC predictions. During the offline model's training, five essential input features, namely, the measured voltage, measured current, cycle ID, measured temperature, and time, were selected to capture the battery's behavior over time. The output vector for the offline model was the known SOC values. In a given sample, the input vector X comprised concatenated values of five selected input features, and it was represented by $[x_1, x_2, \dots, x_n]$. The output vector Y , which was represented by $[y_1, y_2, \dots, y_n]$, corresponded to the SOC value ($SOC_0, SOC_1, \dots, SOC_n$) for the input sample. The goal was to train the LSTM model to accurately map the input vector X to the output vector Y . The LSTM neural network used in the offline model comprised four hidden layers, as shown in Figure 8. The LSTM's hidden layers sequentially processed the input and passed the hidden state to the following layer. This model received sequential input data at the input layer, where each time step's features were processed. Each LSTM unit in a layer comprised an input gate, a forget gate, and an output gate, which determined what information to store in the cell state, what information to discard, and what part of the cell state to output as the hidden state to the next layer or the final output. The activation functions utilized by the LSTM units were sigmoid and hyperbolic tangent (\tanh) functions. The input, forget, and output gates utilized the sigmoid function to regulate the flow of information based on relevance and importance, and the values were limited to a range of 0 to 1. On the other hand, \tanh was employed to compute new candidate values that could be included in the memory cell; values were compressed between -1 and 1 while taking the magnitude and significance of new data into account. The output layer processed the final hidden state or cell state to produce the desired output, that is, the prediction of the next value or the estimation of the SOC of a battery. During training, backpropagation through time (BPTT) was used to compute gradients and update the weights. To prevent overfitting, the L1/L2 regularization technique was included in the LSTM model by adding a penalty term to the loss function. Furthermore, two dense layers were used in the output layer to take the sequence of hidden states produced by the LSTM layers and transform them into a meaningful prediction of the state of charge.

Figure 8 illustrates that the input was a matrix with dimensions of $(n, 256)$ and a batch size of 256. After passing through the LSTM and dense layers, the output was transformed into a matrix of dimensions $(n, 1)$.

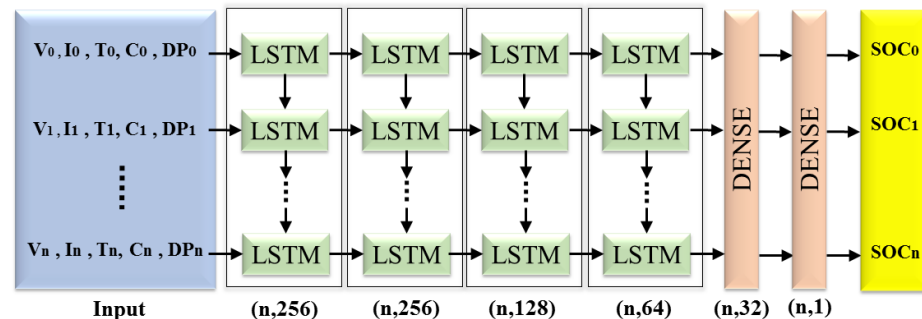


Figure 8. Architecture of the LSTM neural network in the offline model.

The LSTM model was trained using three different types of optimizers, namely, SWATS, Adam, and SGD, with different learning rates. This approach provided several benefits, such as the ability to choose the most effective optimizer, ensuring good generalization and convergence, adapting to various data characteristics and learning rates, being robust against local minima, and gaining insights for future optimization strategies. Despite the presence of L1/L2 regularization, overfitting could still occur in the LSTM model, particularly if the learning rate was high. This was because a high learning rate could lead to the model learning the training data too well, including the noise, and consequently, the model may not have performed well on unseen data. To mitigate this problem, one can use the SWATS optimizer. The SWATS optimizer is a new optimizer that prevents overfitting by adaptively adjusting the learning rate during training. It is based on the Adam optimizer but adds features such as a moving average of gradients and a decay factor to gradually decrease the learning rate over time. It can be used with any machine learning model but is particularly useful for LSTM models that are prone to overfitting. The Switching Adam to SGD optimizer is a variant that improves the convergence speed of the optimizer by switching to SGD after a certain number of epochs. The SWATS optimizer has shown promising results in preventing overfitting in LSTM models.

In This study, the SWATS optimizer was used to train the LSTM model with different learning rates. The results showed that the SWATS optimizer was able to achieve better performance than that of the other optimizers, including Adam and SGD. The SWATS optimizer prevented overfitting while still achieving good accuracy on the training data.

The LSTM model was trained using the battery manufacturer's past data to create an offline model that could be used as a benchmark to measure the reliability and precision of the digital twin's SOC forecasts. This enabled a quantitative assessment of the digital twin's effectiveness by comparing its predictions with the actual SOC values.

Additionally, we compared the LSTM models trained with different optimizers (SWATS, Adam, and SGD) with a GRU network to understand their strengths and weaknesses in LIB behavior forecasting. LSTM and GRU are both RNN variants but differ in their architecture, which impacts their ability to capture temporal dependencies and handle long-term sequences. GRU networks have a simplified architecture compared to that of LSTM [48]. GRU networks have a reset gate and an update gate, which control the flow of information within the network. The reset gate determines which parts of the past information should be forgotten, while the update gate decides which parts of the new information should be incorporated. By comparing the performance of both models, we evaluated their suitability for accurate battery behavior forecasting while considering factors such as prediction accuracy, convergence speed, generalization capabilities, and computational efficiency. Two different evaluation metrics, the root mean square error (RMSE), and mean

absolute error (MAE), were utilized to evaluate the effectiveness of the proposed model. The formulas for calculating these metrics are presented below.

$$MAE = \frac{1}{N} \sum_{i=1}^N |SOC_{actual} - SOC_{pred}| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (SOC_{actual} - SOC_{pred})^2} \tag{11}$$

Following every forward propagation, the model loss was computed as the mean square error (MSE), which involved assessing the deviation between the predicted SOC value and the actual SOC value.

$$Loss = \frac{1}{N} \sum_{i=1}^N |SOC_{actual} - SOC_{pred}|^2 \tag{12}$$

Table 1 presents a summary of the MAE and RMSE outcomes obtained from the LSTM and GRU models while utilizing three distinct optimizers with learning rates of 0.05, 0.03, 0.01, and 0.008. Compared to the other two optimization techniques, the SWATS optimizer exhibited superior performance, effectively mitigating issues related to overfitting and gradient instability. Furthermore, it was evident that the choice of learning rate significantly impacted the accuracy and quality of the predictions. To assess their impacts, we used a range of learning rate values in our simulations. Notably, for battery B0005, the employment of a learning rate of 0.03 in conjunction with the SWATS optimizer yielded the most optimal results in terms of SOC calculations. This particular configuration proved to be highly effective for estimating the SOC in this battery model. Moreover, we provide similar simulation results for battery B0006 in Table 2, battery B0007 in Table 3, and battery B0018 in Table 4.

Table 1. Evaluating the forecasting performance of the LSTM and GRU Models with varied optimizers and learning rates for battery B0005.

Type	MAE (LR = 0.008)	RMSE (LR = 0.008)	MAE (LR = 0.01)	RMSE (LR = 0.01)	MAE (LR = 0.03)	RMSE (LR = 0.03)	MAE (LR = 0.05)	RMSE (LR = 0.05)
LSTM (Adam)	1.432%	1.456%	1.447%	1.466%	2.525%	2.743%	7.182%	7.223%
LSTM (SGD)	9.432%	10.156%	2.943%	3.231%	11.112%	11.143%	11.109%	11.140%
GRU (Adam)	5.716%	5.762%	5.341%	5.373%	4.232%	4.336%	1.432%	1.456%
GRU (SGD)	3.286%	3.293%	3.777%	3.746%	3.214%	3.273%	1.714%	1.773%
LSTM (SWATS)	1.591%	1.598%	1.599%	1.603%	0.888%	0.912%	1.388%	1.452%

Table 2. Evaluating the forecasting performance of the LSTM and GRU Models with varied optimizers and learning rates for battery B0006.

Type	MAE (LR = 0.008)	RMSE (LR = 0.008)	MAE (LR = 0.01)	RMSE (LR = 0.01)	MAE (LR = 0.03)	RMSE (LR = 0.03)	MAE (LR = 0.05)	RMSE (LR = 0.05)
LSTM (Adam)	3.743%	3.559%	3.243%	3.319%	3.243%	3.319%	6.901%	7.103%
LSTM (SGD)	5.634%	6.751%	5.953%	6.761%	5.953%	6.761%	10.334%	10.440%
GRU (Adam)	10.023%	10.306%	9.841%	9.956%	9.841%	9.956%	1.257%	1.266%
GRU (SGD)	7.965%	7.944%	7.677%	7.416%	7.677%	7.416%	1.342%	1.553%
LSTM (SWATS)	3.265%	3.273%	3.117%	3.168%	3.211%	3.298%	1.005%	1.023%

Table 3. Evaluating the forecasting performance of the LSTM and GRU models with varied optimizers and learning rates for battery B0007.

Type	MAE (LR = 0.008)	RMSE (LR = 0.008)	MAE (LR = 0.01)	RMSE (LR = 0.01)	MAE (LR = 0.03)	RMSE (LR = 0.03)	MAE (LR = 0.05)	RMSE (LR = 0.05)
LSTM (Adam)	9.198%	9.911%	8.850%	9.080%	9.652%	9.701%	9.530%	9.670%
LSTM (SGD)	11.440%	11.990%	11.124%	11.320%	11.777%	11.889%	11.651%	11.872%
GRU (Adam)	2.690%	2.454%	1.607%	1.645%	2.021%	2.102%	1.153%	1.243%
GRU (SGD)	2.223%	2.303%	1.742%	1.943%	2.228%	2.501%	1.225%	1.382%
LSTM (SWATS)	1.889%	1.908%	1.365%	1.420%	1.301%	1.901%	1.023%	1.102%

Table 4. Evaluating the forecasting performance of LSTM and GRU models with varied optimizers and learning rates for battery B0018.

Type	MAE (LR = 0.008)	RMSE (LR = 0.008)	MAE (LR = 0.01)	RMSE (LR = 0.01)	MAE (LR = 0.03)	RMSE (LR = 0.03)	MAE (LR = 0.05)	RMSE (LR = 0.05)
LSTM (Adam)	1.877%	1.910%	1.996%	2.027%	9.332%	3.319%	6.901%	7.103%
LSTM (SGD)	10.672%	10.712%	10.856%	10.943%	11.602%	11.693%	11.777%	11.927%
GRU (Adam)	4.475%	4.631%	4.452%	4.651%	2.690%	2.721%	2.881%	2.911%
GRU (SGD)	2.372%	2.523%	2.697%	2.887%	2.438%	2.434%	2.6048%	2.674%
LSTM (SWATS)	1.355%	1.442%	1.612%	1.724%	1.923%	2.002%	2.281%	2.339%

This gives a thorough understanding of how our suggested method performs with and adjusts to different battery models. Finally, the loss function results for the initial 100 epochs of LSTM with the SWATS optimizer are shown in Table 5. Based on the results that were obtained, the optimal learning rate for battery B0006 with the SWATS optimizer was 0.05, while for battery B0007, it was 0.01, and for battery B0018, it was 0.008. The simulations using the SWATS optimizer yielded the most favorable outcomes and the lowest loss function values for SOC calculation in these batteries. The analysis of the results obtained from battery B0006 indicated that the SWATS optimizer worked best with a learning rate of 0.05. For battery B0007, after careful experimentation, it was found that the most favorable outcomes were achieved with a learning rate of 0.01 when used alongside the SWATS optimizer. On the other hand, a learning rate of 0.008 was the most effective value, highlighting the adaptability of the SWATS optimizer in B0018 across various battery models. It was noteworthy that the utilization of the SWATS optimizer consistently yielded superlative outcomes across these battery types, which were characterized by the lowest loss function values in the SOC calculation process. In Figure 9, we show in great detail the results of our thorough simulations of the B0005 battery. The findings are summarized in Figure 10, which showcases the use of the SWATS optimizer for batteries B0005, B0006, B0007, and B0018. In addition, we expanded our investigation to encompass the assessment of the effectiveness of the LSTM network in conjunction with the GRU network. The results of the SOC estimation for the B0005 battery are depicted in Figure 11, where we utilized different optimizers and learning rates to gain insights into the performance behavior of the network. We will specifically examine the outcomes of our LSTM modeling efforts that were conducted offline with a particular focus on forecasting the SOC using various learning rates in Section 5.

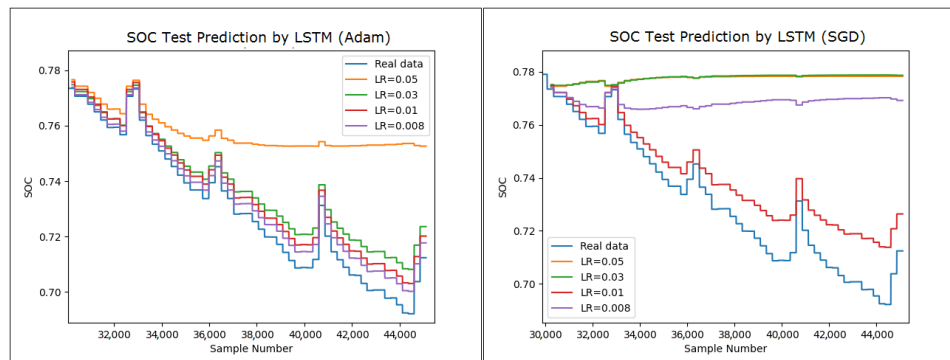


Figure 9. Comparative analysis of SOC prediction using offline LSTM modeling for 45,122 samples with learning rates of 0.05, 0.03, 0.01, and 0.008 with Adam and SGD in battery B0005.

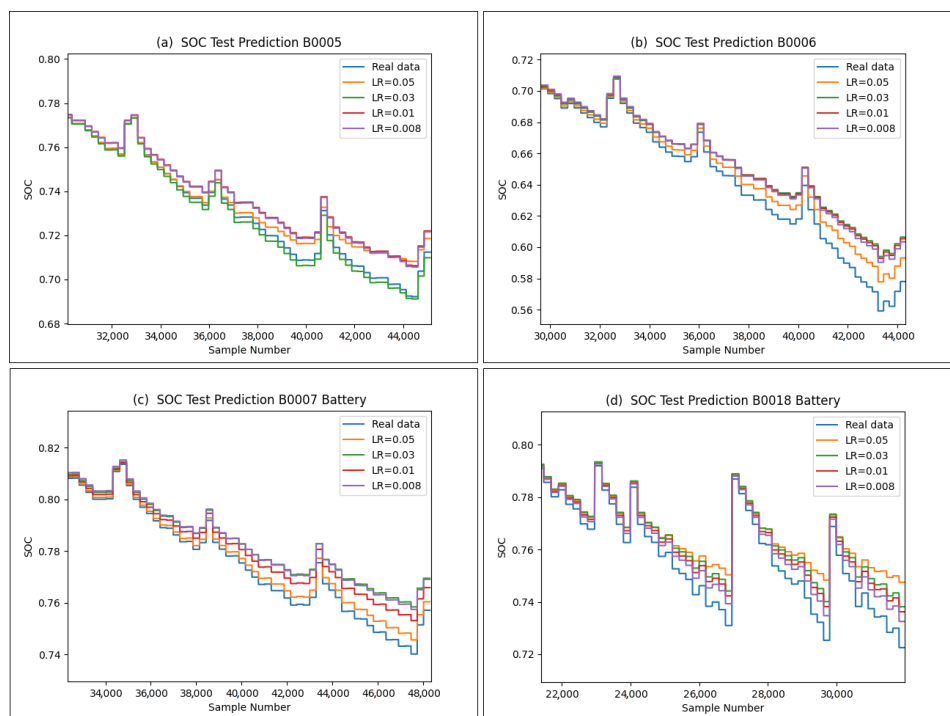


Figure 10. Comparative analysis of SOC prediction using offline LSTM modeling for 45,122 samples with learning rates of 0.05, 0.03, 0.01, and 0.008 with SWATS in batteries (a) B0005, (b) B0006, (c) B0007, and (d) B0018.

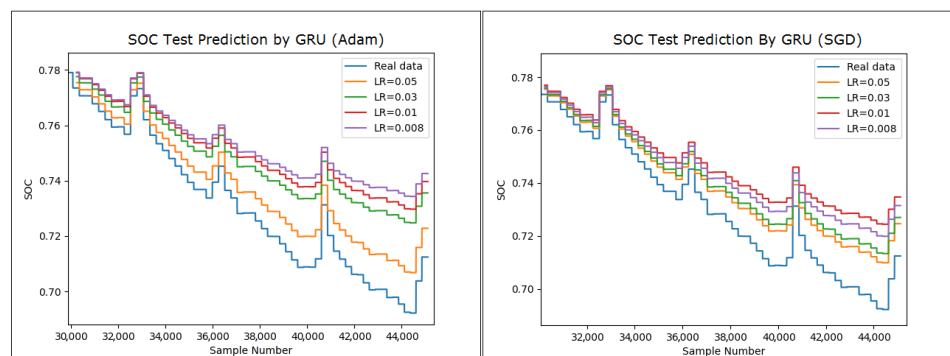


Figure 11. Comparative analysis of SOC prediction using offline GRU modeling for 45,122 samples with learning rates of 0.05, 0.03, 0.01, and 0.008 with SWATS, Adam, and SGD in battery B0005.

Table 5. Comparison of the loss function results for LSTM with the SWATS optimizer based on the number of epochs.

Epoch	B0005 (LR = 0.03)	B0006 (LR = 0.05)	B0007 (LR = 0.008)	B0018 (LR = 0.01)
10	1.765×10^{-1}	1.266×10^{-1}	2.443×10^{-1}	1.978×10^{-1}
20	9.219×10^{-2}	7.219×10^{-2}	3.053×10^{-2}	2.016×10^{-2}
30	8.031×10^{-3}	3.178×10^{-2}	1.724×10^{-2}	9.711×10^{-3}
40	5.045×10^{-3}	2.730×10^{-2}	7.203×10^{-3}	4.932×10^{-3}
50	3.724×10^{-3}	2.12×10^{-2}	5.125×10^{-3}	2.311×10^{-3}
60	9.589×10^{-4}	1.621×10^{-2}	2.489×10^{-3}	1.554×10^{-3}
70	7.956×10^{-4}	1.067×10^{-2}	1.223×10^{-3}	8.329×10^{-4}
80	9.271×10^{-5}	9.944×10^{-3}	8.175×10^{-4}	5.956×10^{-4}
90	8.331×10^{-5}	1.118×10^{-4}	2.105×10^{-4}	1.280×10^{-4}
100	8.317×10^{-5}	1.047×10^{-4}	2.079×10^{-4}	1.214×10^{-4}

4.1.4. Problem Formulation for the Generation of Real Data and Simulation Results

GANs were introduced in 2014 [49] and have proven effective in producing high-quality outputs via a mutual game-learning process between two modules: a generative model and a discriminative model. Generative models—or generators—are responsible for recovering real data distributions. The algorithm takes a random noise vector Z as an input and generates an output, $G(z)$, which can be an image or other data format. On the other hand, the discriminator is a type of discriminative model that has the job of differentiating between data samples that come from the training set and those that are created by the generator. It receives input in the form of x , which can either be actual training data or data that have been generated by the generator. The output score produced by the discriminator is either 1 or 0. If the score is 1, it means that the input is real data, while a score of 0 indicates that the input is false or generated data. The aim of the generative model G is to understand a distribution p_g of the data x . This is achieved by using a function $G(z; \theta_g)$ that maps a prior noise distribution $p_z(z)$ to the data space, where θ_g represents the model’s parameters. These parameters can be the weights of a multilayer perceptron that is used in G . On the other hand, the discriminative model $D(x; \theta_d)$ is created as a binary classifier that gives a scalar output that shows the probability of input x coming from the training data instead of p_g .

The training process involves a game between two models, which continues until they reach a stable balance point. This ensures that the generator can produce realistic data samples, while the discriminator can accurately distinguish between real and generated data. Both the generator and discriminator networks are trained simultaneously using min–max adversarial loss functions. The generation module’s objective function is stated as follows:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_x} [\log D(x)] - \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \tag{13}$$

where D represents the discriminator, G represents the generator, V is the adversarial loss function, the real data distribution is denoted by p_x , and the latent space distribution is represented by p_z . In order to transform a traditional GAN into a TS-GAN, some modifications in the architecture and training process have to be made. The following is a step-by-step guide for accomplishing this task. The TS-GAN architecture comprises two main components: an autoencoder and an adversarial network (Figure 12).

The autoencoder is responsible for learning a time-series embedding space that can capture the underlying patterns in the data, while the adversarial network generates artificial time-series data and distinguishes them from real data. The TS-GAN uses both supervised and unsupervised learning objectives during training, and it applies the adversarial loss to both real and synthetic sequences [43]. In addition, TS-GAN includes a stepwise supervised loss that rewards the model for accurately learning the distribution over transitions from one time point to the next, as observed in historical data.

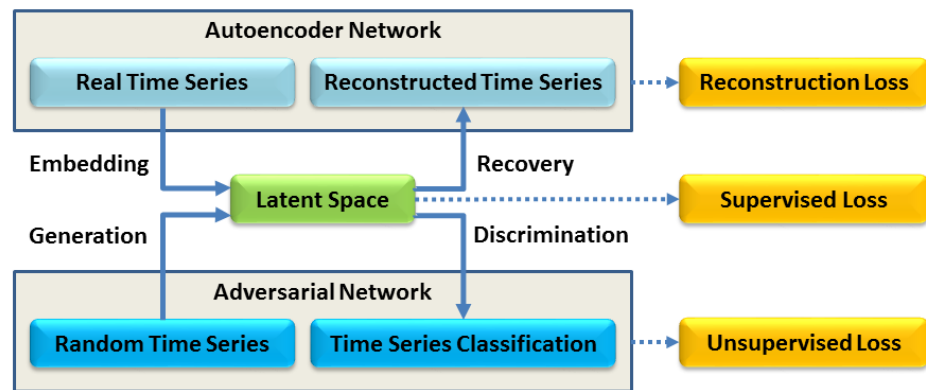


Figure 12. The hybrid TS-GAN architecture: combining a GAN with an autoencoder to generate lithium-ion battery data.

To implement the TS-GAN architecture, several steps are required; firstly, real historical time-series data, such as EV battery system data, are collected and made ready for training. Alongside this, random time-series data are generated to be used as a benchmark for comparison with the synthetic data that are produced. Next, the key components of the TS-GAN model, which include the autoencoder, sequence generator, and sequence discriminator, are established.

The TS-GAN architecture is made up of various components such as an Embedder, a Recovery, a Generator, a Discriminator, and a Supervisor. Its training involves 10,000 iterations. The Embedder is an RNN-based model that maps real data sequences x_t to a lower-dimensional space e_t and captures temporal dependencies.

$$e_t = E(x_t) \tag{14}$$

The Recovery (another RNN-based model) maps embeddings e_t back to the original data space to reconstruct the time-series data:

$$\hat{x}_t = R(e_t) \tag{15}$$

The Generator is an RNN-based model that generates synthetic data sequences $x_{(fake,t)}$ from random noise sequences z_t :

$$x_{fake,t} = G(z_t) \tag{16}$$

while the Discriminator distinguishes between real time-series data $x_{real,t}$ and generated data $x_{fake,t}$:

- $D(x_{real,t})$ represents the probability that $x_{real,t}$ is real;
- $D(x_{fake,t})$ represents the probability that $x_{fake,t}$ is fake.

The Supervisor acts as an intermediary between the Embedder and the Generator to enhance the quality of the generated sequences. There are two main objectives in the training process: adversarial loss and supervised loss. The adversarial loss can be defined as

$$\mathcal{L}_{adv} = \mathbb{E}_{x_{real} \sim p_{data}} [\log D(x_{real})] + \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \tag{17}$$

The supervised loss can be defined as

$$\mathcal{L}_{sup} = \mathbb{E}_{x_{real} \sim p_{data}} \left[\sum_t \text{loss}(x_{real,t+1}, x_{fake,t+1}) \right] \tag{18}$$

The combined adversarial and supervised losses make up the following overall objective:

$$\mathcal{L}_{\text{TS-GAN}} = \mathcal{L}_{\text{adv}} + \lambda \cdot \mathcal{L}_{\text{sup}} \quad (19)$$

where λ is an objective-balanced hyperparameter. During the initialization phase of the TS-GAN model, an autoencoder is employed to integrate the Generator and the Embedder. The main objective of this approach is to reconstruct genuine data sequences and obtain significant embeddings of the real data. During training, the Generator and Embedder are trained twice as often as the Discriminator to maintain balance. After training, the Generator generates synthetic data sequences, which are transformed back to the original data space using the Recovery model and inverse-scaled to obtain realistic-looking synthetic data. The results of the data generated by the TS-GAN algorithm are shown in Figure 13.

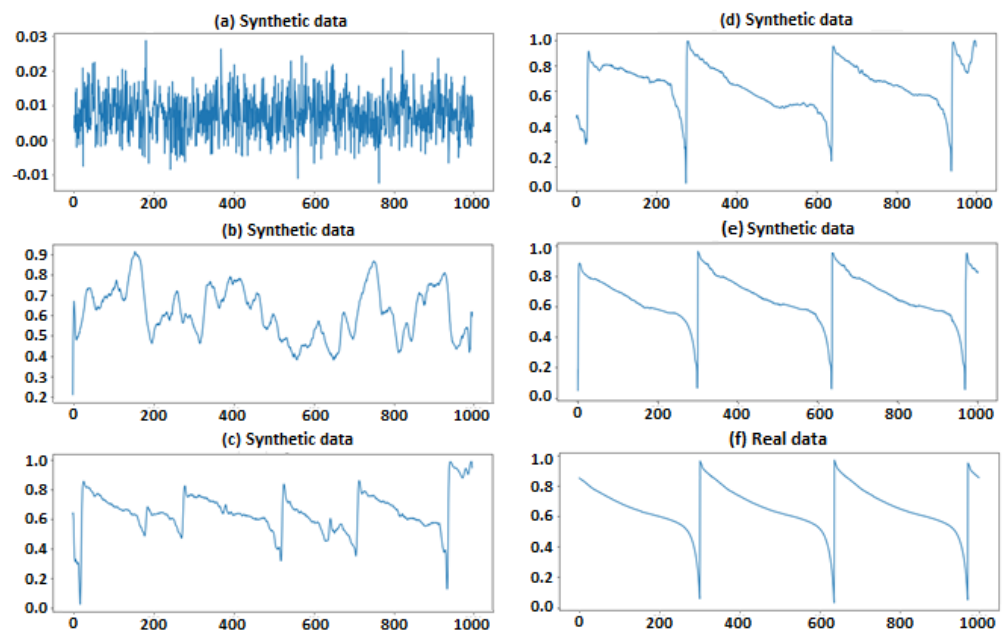


Figure 13. Synthetic voltage generated using the TS-GAN. (a) Epoch = 5, (b) epoch = 25, (c) epoch = 50, (d) epoch = 75, (e) epoch = 100, and (f) real data.

Furthermore, the battery time-series data for B0018 and B0005 are compared in Figures 14 and 15 to show the differences between the actual and generated data.

4.1.5. Evaluation of the Synthetic Data

The next step after synthesizing our data was to verify that the new data accurately reproduced the initial battery data. Using evaluation metrics is one of the best ways to compare real and synthetic data. In order to ensure the TS-GAN's reliability and applicability in real-world scenarios, it was important to accurately evaluate the data generated by it. It is important for the evaluation metric to be carefully selected when dealing with multivariate time-series data, such as those obtained from LIBs. A model's performance or predictive capabilities may not be significantly affected by small changes in time-series data. Hence, it is crucial to strike a balance between capturing meaningful differences in the generated data and being robust enough to tolerate minor variations. These requirements can be met by the Fréchet inception distance (FID). An objective measure of similarity between a generated data distribution and a real distribution is provided without being excessively influenced by minor fluctuations in the data distribution.

When evaluating data from a TS-GAN—especially historical data instead of real-time data—it is imperative to use an accurate evaluation metric, such as the FID. The FID approach provides objective and quantitative insights into the performance and general-

ization capabilities of a TS-GAN model, despite the fact that historical data are not ideal for evaluation. As a result, potential overfitting can be detected, generalization can be validated, and the model can be iteratively improved. Through the use of the FID, one can obtain significant insights into the TS-GAN model’s ability to learn from past data and the similarity between the generated data and the actual data distribution, even without real-time data.

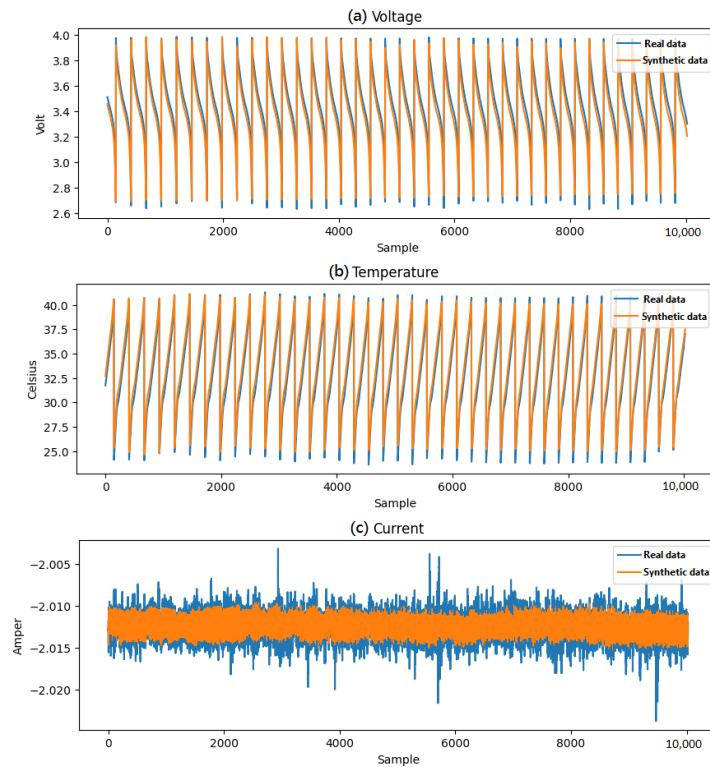


Figure 14. Analysis of actual (a) voltages, (b) temperatures, (c) currents, and their generated counterparts.

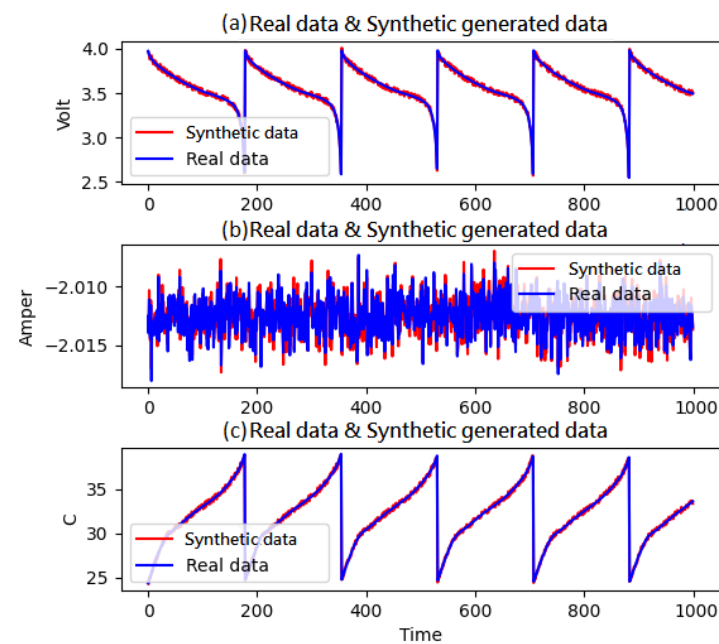


Figure 15. Analysis of actual (a) voltages, (b) currents, (c) temperatures, and their generated counterparts for B0005.

Formulation of the Fréchet inception distance: The Fréchet inception distance (FID) is a popular metric used in generative modeling, including for time-series data. The authors of [50] employed the FID metric to evaluate the performance of a new update rule called the two-time-scale update rule (TTUR) on various datasets. The TTUR is a strategy used in training generative adversarial networks (GANs) with stochastic gradient descent. It is designed to address convergence issues and enhance learning for GANs, leading to improved results in tasks such as image generation. The evaluation involves deep learning and feature extraction to gauge the dissimilarity between two probability distributions. In order to evaluate the accuracy of the data produced by TS-GAN, we utilized the FID to measure the disparity between the feature representations of the generated time-series data and the original time-series data. This was achieved by utilizing a pre-trained neural network, which is usually an Inception-v3 network (a deep convolutional neural network architecture), to extract high-level features from both types of time-series data. The FID is calculated as follows:

$$FID^2 = \|\mu_{real} - \mu_{fake}\|^2 + \text{Tr}(\Sigma_{real} + \Sigma_{fake} - 2(\Sigma_{real} * \Sigma_{fake})^{\frac{1}{2}}) \tag{20}$$

- μ_{real} is the mean of the feature representations of the real data samples;
- μ_{fake} is the mean of the feature representations of the generated data samples;
- Σ_{real} is the covariance matrix of the feature representations of the real data samples;
- Σ_{fake} is the covariance matrix of the feature representations of the generated data samples;
- $T_r(\cdot)$ denotes the matrix transposition operation.

By taking the multivariate nature of time-series data into account, the FID metric captures the relevant relationships and patterns for precise evaluation. It offers a reliable and informative measure of the similarity between the two distributions. A lower FID score suggests a higher similarity, indicating that the generated data closely resemble the actual data. This demonstrates the effectiveness of the TS-GAN model in replicating the underlying dynamics of the target system. In conclusion, the FID metric is a valuable tool for impartial and precise evaluation of TS-GAN-generated data, contributing to the advancement and application of this technology in real-world scenarios involving LIBs and beyond. Based on the FID metric, Figure 16 illustrates the results of the evaluation of the TS-GAN data.

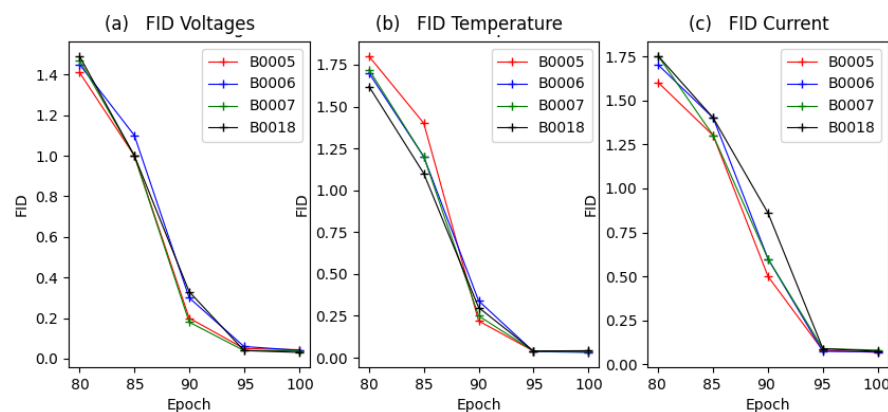


Figure 16. FID score results for the (a) voltage, (b) temperature, and (c) current generated by the TS-GAN.

To apply the FID on two distinct datasets, namely, a historical dataset and a real-time dataset of lithium-ion batteries produced by the TS-GAN, we utilized TensorFlow. The process involved loading and preprocessing the pre-trained Inception-v3 model, extracting features from both datasets using the same pre-trained Inception-v3 model, computing the

mean and covariance matrix for each dataset, and, finally, computing the FID score based on the mean and covariance matrix of the two datasets. The results are shown in Table 6.

Table 6. FID metric results for the evaluation of data generated by the TS-GAN.

Battery	Voltage	Temperature	Current
B0005	0.044	0.034	0.069
B0006	0.041	0.032	0.072
B0007	0.038	0.036	0.082
B0018	0.031	0.042	0.075

4.2. Training and Evaluation of the LSTM-Based SOC Estimator for the Digital Twin

To train the LSTM in the digital twin, these steps were followed:

- **Data Preparation:** The real-time data obtained from the TS-GAN model were pre-processed to make them compatible with the LSTM model’s input format. The pre-processed data were split into training and validation sets. In the training set, 67% of the 45,122 samples were included (30,080 samples), while in the validation set, 33% of the 45,122 samples were included (14,890 samples). Additionally, five important features were selected: the measured voltage, measured current, cycle ID, measured temperature, and time.
- **Model Architecture:** The LSTM model architecture within the digital twin should be the same or similar to that of the offline LSTM model used for historical data training. LSTM models with four layers, 25 hidden units in each layer, sigmoid activation functions, and L1/L2 regularization layers were used.
- **Training Process:** The LSTM model was compiled with an SWATS optimizer. To assess the accuracy and generalization capabilities of the LSTM model, a separate test dataset was used after training.

A comparison between the SOC predicted using the digital twin and the anticipated SOC obtained from historical or real data is shown in Figure 17. The gradient of the LSTM model for battery B0005 with learning rates of 0.05, 0.03, 0.01, and 0.008 is depicted in Figure 18. Additionally, Figure 19 illustrates the LSTM model’s loss during state-of-charge estimation for battery B0005, showcasing the impact of the learning rate (0.05, 0.03, 0.01, and 0.008) on the training convergence. Finally, Figures 20 and 21 present the distribution of prediction errors in SOC estimation for lithium-ion batteries using the LSTM network in the digital twin, thus providing insights into the accuracy and reliability of the model’s predictions during both the training and testing phases, respectively.

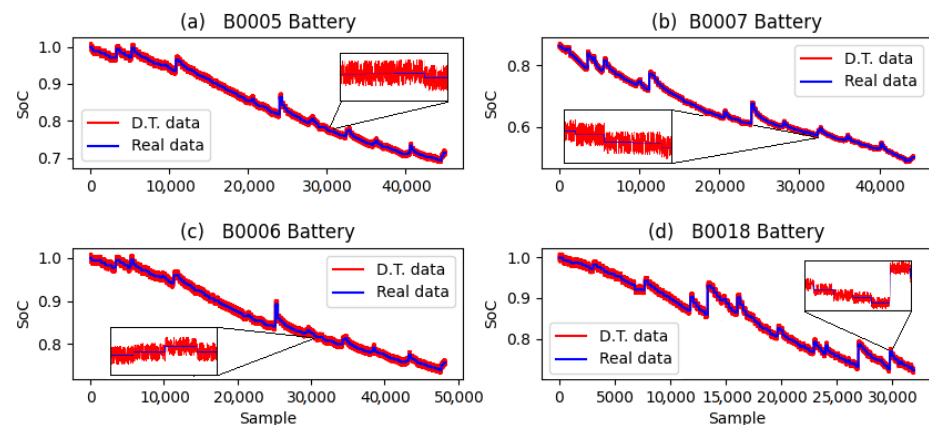


Figure 17. A comparison of the predicted SOC from the digital twin and the expected SOC from historical data.

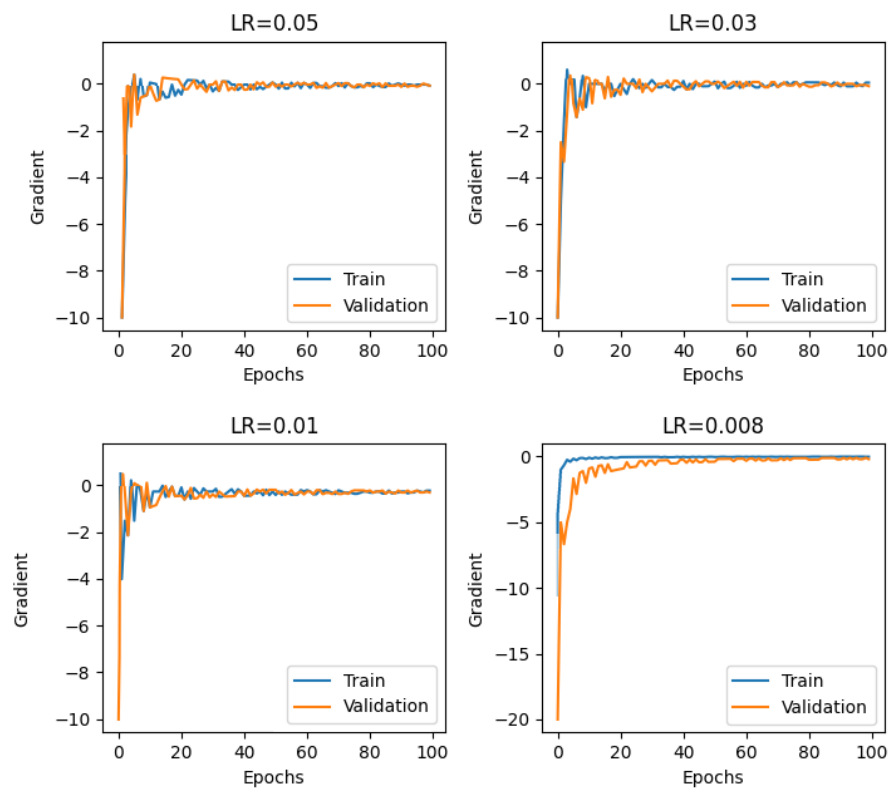


Figure 18. LSTM model gradient for battery B0005 in SOC estimation by the digital twin with learning rates of 0.05, 0.03, 0.01, and 0.008.

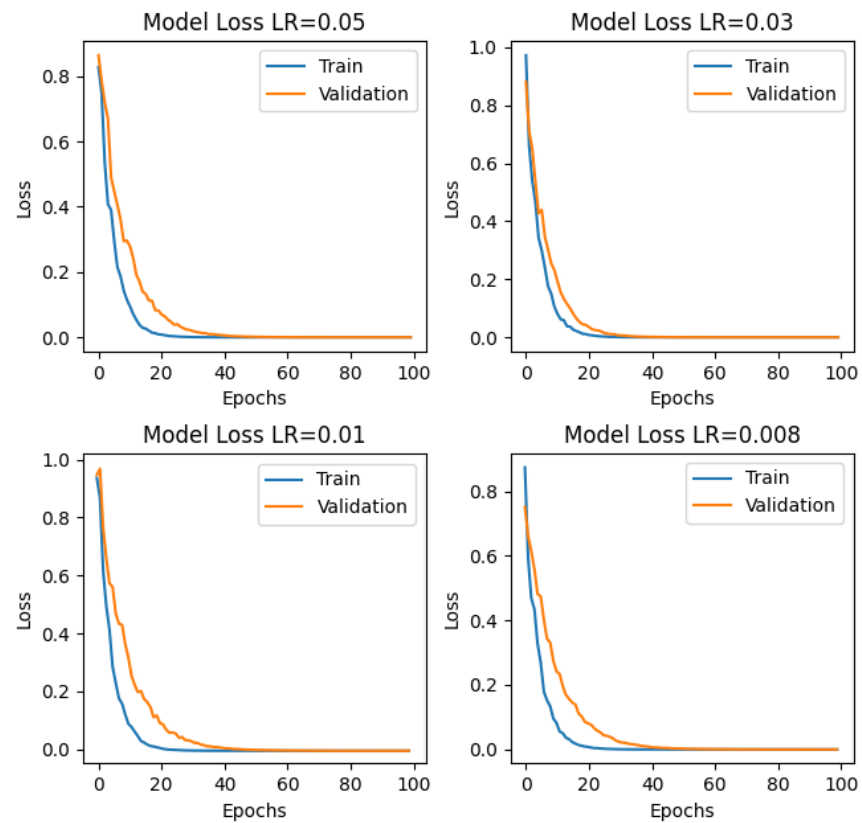


Figure 19. Training convergence analysis of the LSTM model loss in SOC estimation by the digital twin for battery B0005 with learning rates of 0.05, 0.03, 0.01, and 0.008.

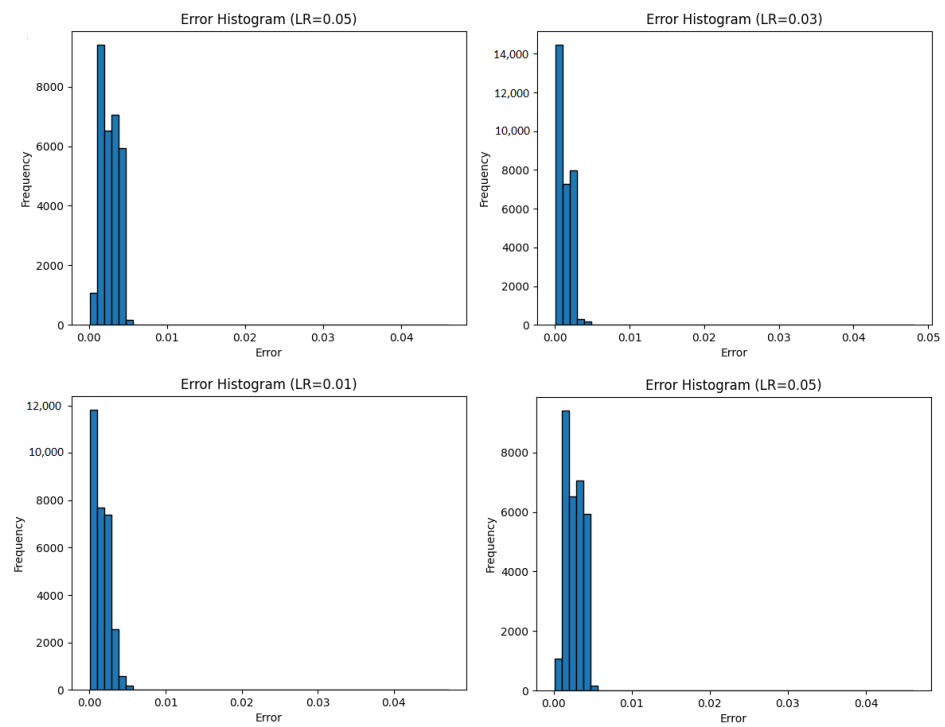


Figure 20. Distribution of prediction errors in the LSTM model’s training in SOC estimation for lithium-ion batteries in the digital twin.

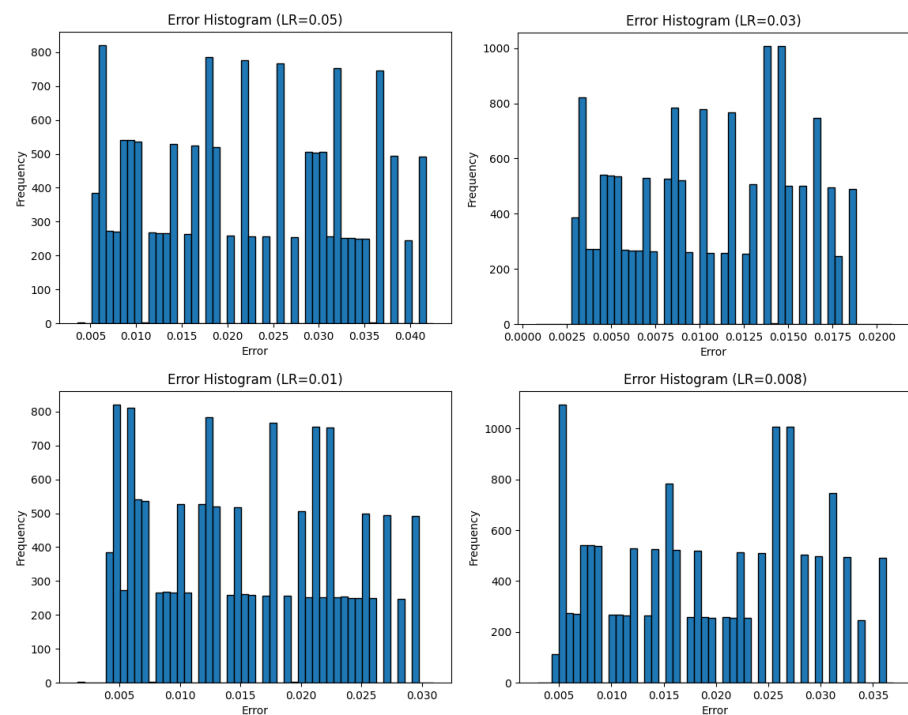


Figure 21. Distribution of prediction errors in the testing of the LSTM model in SOC estimation for lithium-ion batteries in the digital twin.

4.3. Evaluation of Digital-Twin-Based SOC Estimation

In this section, we compare the accuracy of the SOC estimation results obtained from two distinct models: the LSTM model within the digital twin and the LSTM model trained offline using historical data. Since both models are regression-based, we will be using error as the primary metric for evaluating their accuracy. The MSE serves as a key measure,

with lower values indicating superior performance. Simply comparing means between two models has limitations due to the effects of sample variation and inherent randomness. It is not practical to expect identical means between the two models. To effectively determine which model is superior, we followed standard statistical procedures that are more robust. In statistical testing, we established a null hypothesis and an alternative hypothesis that differed from the null. We then analyzed the data to demonstrate that the null hypothesis could not stand, thereby accepting the alternative hypothesis. This process ensured robust and credible claims. To evaluate the relative accuracy of the digital twin and offline LSTM model, we employed parametric tests—specifically, the *t*-test and analysis of variance (ANOVA).

These parametric tests were chosen for several important reasons. Firstly, parametric tests assume that the data follow a specific distribution (usually normal) and are designed for comparing means and variances. Secondly, they offer greater statistical power, which means that they are more likely to detect true differences if they exist.

Additionally, parametric tests provide precise estimates of the effect size, thus aiding in understanding the practical significance of the results. The *t*-test was used to analyze the variations in average values of metrics such as the MSE and MAE. ANOVA was used to compare the variance in the LIB datasets. The test results showed that there were no significant differences among the independent variables in the different groups ($p > 0.05$). Tables 7 and 8 display the results of the two statistical tests that were utilized to assess the performance of the two models. The usage of these parametric tests led to accurate statistical conclusions regarding the model’s precision. This method offered a thorough evaluation of the model’s performance that took the differences in both mean and variance into account.

Table 7. *t*-test for the evaluation of the digital twin.

Battery	t	df	cv	p	Result
B0005	0.012	90242	1.645	0.990	Fail ¹ to reject null hypothesis
B0006	0.039	88704	1.645	0.969	Fail to reject null hypothesis
B0007	0.043	96592	1.645	0.966	Fail to reject null hypothesis
B0018	0.086	311980	1.645	0.931	Fail to reject null hypothesis

¹ A low *p*-value (<0.05) indicates rejection of the null hypothesis, while a high *p*-value (>0.05) suggests support for the null hypothesis.

Table 8. Analysis of variance (ANOVA) for the evaluation of the digital twin.

Battery	State	p	Result
B0005	0.015	0.902	Probably ¹ the same distribution
B0006	0.025	0.874	Probably the same distribution
B0007	0.009	0.941	Probably the same distribution
B0018	0.024	0.876	Probably the same distribution

¹ “probability” indicates the likelihood of having similar distributions between two predicted values of SOC.

5. Discussion

In this section, we discuss the findings and significance of various simulation studies that centered around the utilization of digital twins for enhanced accuracy in forecasting the charge levels of batteries.

5.1. Performance of LSTM-Based Offline Modeling

This section discusses the performance of the LSTM-based offline modeling stage of the digital twin framework for SOC prediction. We trained the LSTM model with four

layers, three optimizers (Adam, SGD, and SWATS), and four learning rates (0.05, 0.03, 0.01, and 0.008) for four LIBs (B0005, B0006, B0007, and B0018).

Figure 7 shows the results for the Adam and SGD optimizers for battery B0005. As we can see, the plot for LR = 0.008 was the most similar to the real data plot. Additionally, Table 1 shows that the LSTM model with the Adam optimizer and LR = 0.008 had the lowest loss value (MAE = 1.432 and RMSE = 1.456). The LSTM model with the SGD optimizer and LR = 0.01 also had a low loss value (MAE = 2.945 and RMSE = 3.231).

To make a comparison, we trained a GRU model while using identical optimizers and learning rates. Figure 9 shows the results for battery B0005. The plot shows that the GRU model with the Adam optimizer and LR = 0.05 had the best results. Moreover, Table 1 shows that the GRU model with the Adam optimizer and LR = 0.05 had the lowest loss value (MAE = 1.432 and RMSE = 1.456). The GRU model with the SGD optimizer and LR = 0.05 also had a low loss value (MAE = 1.714 and RMSE = 1.773). The LSTM model was ultimately trained using the SWATS optimizer, as it demonstrated superior performance in previous experiments. Figure 8 and Table 1 show that the LSTM model with the SWATS optimizer and LR = 0.03 had the lowest loss value (MAE = 0.888 and RMSE = 0.912) for battery B0005. Table 2, Table 3, and Table 4 show the results for batteries B0006, B0007, and B0018, respectively. Table 5 shows the loss values for the LSTM model with the SWATS optimizer for various epochs. The results show that the loss value after 90 epochs no longer changed significantly, so we considered 100 epochs to be sufficient to reduce the calculation time. Based on the results of the offline modeling stage of the digital twin framework for SOC prediction, we can conclude that the LSTM model with the SWATS optimizer and LR = 0.03 achieved the best results. This model can be used to generate accurate SOC predictions for the online stage of the digital twin framework.

Extensive training of the LSTM and GRU models was conducted for various Li-ion batteries during the offline modeling stage of the digital twin framework for SOC prediction. The LSTM model that utilized the SWATS optimizer and had an LR of 0.03 consistently displayed superior performance to that of the other models. When applied to battery B0005, this particular LSTM setup resulted in an MAE of 0.888 and an RMSE of 0.912. These outcomes emphasized the ability of the LSTM model to produce accurate SOC forecasts for the digital twin's online phase.

5.2. Evaluating the Synthetic Data Generated by the TS-GAN

The discriminator and generator have contrasting roles in a GAN: Generators aim to create data that can be distinguished from real data, while discriminators determine the authenticity of data. Data generation is continuously improved as a result of this dynamic interplay. In Figure 11, we visualize the generator's output for battery B0005 at various epochs, ranging from 5 to 100. As the number of epochs increased, the synthetic data generated by the GAN progressively resembled real data more closely.

To evaluate the quality of the synthetic data produced by the TS-GAN, we employed the FID metric. According to Figure 15, there were no significant changes in the FID values after epoch 95, indicating that the generated data closely resembled the real data. Furthermore, Table 6 presents the FID results for the voltage, current, and temperature, demonstrating that at epoch 1000, we achieved the lowest FID scores. These findings demonstrate that the TS-GAN model generated synthetic data that closely resembled real-world observations of Li-ion batteries. As the FID values gradually became more stable after epoch 95, the TS-GAN model appeared to learn the underlying data distribution over time. Generating synthetic data is important in situations where obtaining real-time data is difficult, as it can be a useful replacement for real-time information.

Moreover, the lowest FID scores achieved at epoch 1000, as indicated in Table 6, signified the model's potential to consistently produce high-quality synthetic data. This has significant implications for scenarios where historical data must be leveraged for predictive modeling and decision making, as it offers a means of bridging the gap between historical and real-time data.

5.3. Performance of the LSTM-Based SOC Estimator in the Digital Twin

We used an LSTM in our digital twin application to estimate the real-time state of charge for lithium-ion batteries. It was crucial for the LSTM used in the online component to have the same architecture as that used offline. Figure 16 shows the forecasted SOC values for batteries B0005, B0006, B0007, and B0018, which closely resembled the anticipated SOC values from the manufacturer data. We used a *t*-test and F-test to compare these two SOC values. It appeared that all of the batteries (B0005, B0006, B0007, and B0018) exhibited similar error metrics (MSE and MAE) between the digital twin and offline LSTM model based on the provided table of statistical test results.

The values of *t* in Table 7 (test statistic) indicate how many standard errors the means of the two groups (digital twin and offline LSTM) were separated by. These *t*-values were very close to zero, indicating that the means of the two groups were very similar.

The data points used in the analysis were expressed in degrees of freedom (df). A large dataset suggests a robust analysis in this case, as there are many degrees of freedom. The critical value (cv) was used to compare the test statistic (*t*) against a threshold. The threshold in this case was set to 1.645, which is commonly used for a significance level of 0.05. The *p*-value represents the probability of obtaining results as extreme as those observed in the sample, assuming that the null hypothesis is true. In all cases, the *p*-values were very high (close to 1), suggesting that there was no statistically significant difference in the error metrics between the digital twin and the offline LSTM model. The interpretations via the critical value and *p*-value both led to the same conclusion: We accepted the null hypothesis. This meant that there was no statistically significant difference in the accuracy (as measured with the MSE and MAE) between the digital twin and the offline LSTM model for any of the batteries (B0005, B0006, B0007, and B0018).

When it came to estimating the SOC of these batteries, the digital twin model was just as effective as the offline LSTM model. Moreover, Table 8 displays the results of an ANOVA that was conducted to evaluate the digital twin's performance across different Li-ion batteries, namely, B0005, B0006, B0007, and B0018. The column labeled "State" represents the level of diversity present within each group of batteries. When considering this information, it can be inferred that values such as 0.015, 0.025, 0.009, and 0.024 indicate a relatively low degree of variability. The "*p*" column provides *p*-values associated with the ANOVA. The higher *p*-values, such as 0.902, 0.874, 0.941, and 0.876, indicate that there was no strong evidence to reject the null hypothesis, suggesting that the performance metrics across these batteries were probably from the same distribution. Put simply, the ANOVA suggested that the digital twin's accuracy remained consistent across all of the tested batteries, regardless of the specific type of Li-ion battery used, indicating that there was no significant difference in its performance.

6. Conclusions

We have demonstrated the immense potential of digital twin technology in simulating and predicting lithium-ion batteries' SOC. We found that our results were closely aligned with manufacturer data when we used a time-series generative adversarial network (TS-GAN) to generate synthetic data. These results have great potential for reducing energy usage in EVs and present interesting opportunities for future research efforts. We have provided valuable insight into SOC estimation for lithium-ion batteries, but we must acknowledge that there are certain limitations, particularly those related to data availability. Despite these challenges, our efforts are a significant addition to the rapidly developing area of energy for electric cars, and they highlight the importance of digital twin technology in overseeing batteries. The use of digital twin technology presents a significant change in the way that we oversee and enhance lithium-ion batteries, as demonstrated in our study. Moreover, by applying *t*-tests and ANOVAs, we demonstrated the consistency and accuracy of the digital twin across various Li-ion battery types. This comparison with previous work illustrates the robustness and versatility of our approach. Digital twins are used to generate a virtual replica of an actual battery, allowing for precise monitoring and

prediction in real time and ensuring optimal energy usage. This research will serve as a catalyst for further exploration and innovation in green energy, with the ultimate goal of fostering a sustainable and environmentally friendly future.

Author Contributions: Conceptualization, M.P.; investigation, M.P.; writing—original draft preparation, M.P.; writing—review and editing, I.S.; supervision, I.S.; funding acquisition, I.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (No. RS-2023-00252328).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, Q.; Wang, D.; Yang, B.; Dong, H.; Zhu, C.; Hao, Z. An electrochemical impedance model of lithium-ion battery for electric vehicle application. *J. Energy Storage* **2022**, *50*, 104182. [[CrossRef](#)]
2. Xiong, R.; Tian, J.; Shen, W.; Sun, F. A novel fractional order model for state of charge estimation in lithium ion batteries. *IEEE Trans. Veh. Technol.* **2018**, *68*, 4130–4139. [[CrossRef](#)]
3. Ma, L.; Xu, Y.; Zhang, H.; Yang, F.; Wang, X.; Li, C. Co-estimation of state of charge and state of health for lithium-ion batteries based on fractional-order model with multi-innovations unscented Kalman filter method. *J. Energy Storage* **2022**, *52*, 104904. [[CrossRef](#)]
4. Cao, M.; Zhang, T.; Wang, J.; Liu, Y. A deep belief network approach to remaining capacity estimation for lithium-ion batteries based on charging process features. *J. Energy Storage* **2022**, *48*, 103825. [[CrossRef](#)]
5. Wang, Y.; Huang, H.; Wang, H. A new method for fast state of charge estimation using retired battery parameters. *J. Energy Storage* **2022**, *55*, 105621. [[CrossRef](#)]
6. Li, Y.; Zou, C.; Berecibar, M.; Nanini-Maury, E.; Chan, J.; Van den Bossche, P.; Van Mierlo, J.; Omar, N. Random forest regression for online capacity estimation of lithium-ion batteries. *Appl. Energy* **2018**, *232*, 197–210. [[CrossRef](#)]
7. Nuhic, A.; Terzimehic, T.; Soczka-Guth, T.; Buchholz, M.; Dietmayer, K. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. *J. Power Source* **2013**, *239*, 680–688. [[CrossRef](#)]
8. Muh, K.; Caliwag, A.; Jeon, I.; Lim, W. Co-Estimation of SoC and SoP Using BiLSTM. *J. Korean Inst. Commun. Sci.* **2021**, *46*, 314–323.
9. Du, J.; Liu, Z.; Wang, Y.; Wen, C. A fuzzy logic-based model for Li-ion battery with SOC and temperature effect. In Proceedings of the 11th IEEE International Conference on Control & Automation (ICCA), Taichung, Taiwan, 18–20 June 2014; pp. 1333–1338. [[CrossRef](#)]
10. Cui, Z.; Wang, L.; Li, Q.; Wang, K. A comprehensive review on the state of charge estimation for lithium-ion battery based on neural network. *Int. J. Energy Res.* **2022**, *46*, 5423–5440. [[CrossRef](#)]
11. Du, Z.; Zuo, L.; Li, J.; Liu, Y.; Shen, H. Data-driven estimation of remaining useful lifetime and state of charge for lithium-ion battery. *IEEE Trans. Transp. Electrification* **2021**, *8*, 356–367. [[CrossRef](#)]
12. Tightiz, L.; Dang, L.; Yoo, J. Novel deep deterministic policy gradient technique for automated micro-grid energy management in rural and islanded areas. *Alex. Eng. J.* **2023**, *82*, 145–153. [[CrossRef](#)]
13. Pooyandeh, M.; Han, K.; Sohn, I. Cybersecurity in the AI-Based metaverse: A survey. *Appl. Sci.* **2022**, *12*, 12993. [[CrossRef](#)]
14. Alabugin, A.; Osintsev, K.; Aliukov, S.; Almetova, Z.; Bolkov, Y. Mathematical Foundations for Modeling a Zero-Carbon Electric Power System in Terms of Sustainability. *Mathematics* **2023**, *11*, 2180. [[CrossRef](#)]
15. Artetxe, E.; Uralde, J.; Barambones, O.; Calvo, I.; Martin, I. Maximum Power Point Tracker Controller for Solar Photovoltaic Based on Reinforcement Learning Agent with a Digital Twin. *Mathematics* **2023**, *11*, 2166. [[CrossRef](#)]
16. Tong, S.; Lacap, J.; Park, J. Battery state of charge estimation using a load-classifying neural network. *J. Energy Storage* **2016**, *7*, 236–243. [[CrossRef](#)]
17. Chaoui, H.; Ibe-Ekeocha, C. State of charge and state of health estimation for lithium batteries using recurrent neural networks. *IEEE Trans. Veh. Technol.* **2017**, *66*, 8773–8783. [[CrossRef](#)]
18. Wang, H.; Ou, S.; Dahlhaug, O.; Storli, P.; Skjelbred, H.; Vilberg, I. Adaptively Learned Modeling for a Digital Twin of Hydropower Turbines with Application to a Pilot Testing System. *Mathematics* **2023**, *11*, 4012. [[CrossRef](#)]
19. Ma, L.; Wang, Z.; Yang, F.; Cheng, Y.; Lu, C.; Tao, L.; Zhou, T. Robust state of charge estimation based on a sequence-to-sequence mapping model with process information. *J. Power Source* **2020**, *474*, 228691. [[CrossRef](#)]
20. Bian, C.; He, H.; Yang, S.; Huang, T. State-of-charge sequence estimation of lithium-ion battery based on bidirectional long short-term memory encoder-decoder architecture. *J. Power Source* **2020**, *449*, 227558. [[CrossRef](#)]
21. Savargaonkar, M.; Oyewole, I.; Chehade, A.; Hussein, A. Uncorrelated Sparse Autoencoder with Long Short-Term Memory for State-of-Charge Estimations in Lithium-Ion Battery Cells. *IEEE Trans. Autom. Sci. Eng.* **2022**, 1–12. [[CrossRef](#)]

22. Almaita, E.; Alshkoor, S.; Abdelsalam, E.; Almomani, F. State of charge estimation for a group of lithium-ion batteries using long short-term memory neural network. *J. Energy Storage* **2022**, *52*, 104761. [\[CrossRef\]](#)
23. Ren, X.; Liu, S.; Yu, X.; Dong, X. A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM. *Energy* **2021**, *234*, 121236. [\[CrossRef\]](#)
24. Jia, K.; Gao, Z.; Ma, R.; Chai, H.; Sun, S. An Adaptive Optimization Algorithm in LSTM for SOC Estimation Based on Improved Borges Derivative. *IEEE Trans. Ind. Inform.* **2023**, 1–12. [\[CrossRef\]](#)
25. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Ahmed, R.; Emadi, A. Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries. *IEEE Trans. Ind. Electron.* **2017**, *65*, 6730–6739. [\[CrossRef\]](#)
26. Pascual, A.; Caliwag, A.; Lim, W. Implementation of the Battery Monitoring and Control System Using Edge-Cloud Computing. *Korean J. Commun. Stud.* **2022**, *47*, 770–780. [\[CrossRef\]](#)
27. Wu, B.; Widanage, W.; Yang, S.; Liu, X. Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems. *Energy AI* **2020**, *1*, 100016. [\[CrossRef\]](#)
28. Ramachandran, R.; Subathra, B.; Srinivasan, S. Recursive estimation of battery pack parameters in electric vehicles. In Proceedings of the 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Madurai, India, 13–15 December 2018; pp. 1–7.
29. Peng, Y.; Zhang, X.; Song, Y.; Liu, D. A low cost flexible digital twin platform for spacecraft lithium-ion battery pack degradation assessment. In Proceedings of the 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Auckland, New Zealand, 20–23 May 2019; pp. 1–6.
30. Li, W.; Rentemeister, M.; Badeda, J.; Jöst, D.; Schulte, D.; Sauer, D. Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation. *J. Energy Storage* **2020**, *30*, 101557. [\[CrossRef\]](#)
31. Qu, X.; Song, Y.; Liu, D.; Cui, X.; Peng, Y. Lithium-ion battery performance degradation evaluation in dynamic operating conditions based on a digital twin model. *Microelectron. Reliab.* **2020**, *114*, 113857. [\[CrossRef\]](#)
32. Panwar, N.; Singh, S.; Garg, A.; Gupta, A.; Gao, L. Recent advancements in battery management system for Li-ion batteries of electric vehicles: Future role of digital twin, cyber-physical systems, battery swapping technology, and nondestructive testing. *Energy Technol.* **2021**, *9*, 2000984. [\[CrossRef\]](#)
33. Singh, S.; Weeber, M.; Birke, K. Implementation of battery digital twin: Approach, functionalities and benefits. *Batteries* **2021**, *7*, 78. [\[CrossRef\]](#)
34. Li, H.; Kaleem, M.; Chiu, I.; Gao, D.; Peng, J. A digital twin model for the battery management systems of electric vehicles. In Proceedings of the 2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, China, 20–22 December 2021; pp. 1100–1107. [\[CrossRef\]](#)
35. Qin, Y.; Arunan, A.; Yuen, C. Digital twin for real-time Li-ion battery state of health estimation with partially discharged cycling data. *IEEE Trans. Ind. Inform.* **2023**, *19*, 7247–7257. [\[CrossRef\]](#)
36. Li, A.; Weng, J.; Yuen, A.; Wang, W.; Liu, H.; Lee, E.; Wang, J.; Kook, S.; Yeoh, G. Machine learning assisted advanced battery thermal management system: A state-of-the-art review. *J. Energy Storage* **2023**, *60*, 106688. [\[CrossRef\]](#)
37. Chaudhari, P.; Agrawal, H.; Kotecha, K. Data augmentation using MG-GAN for improved cancer classification on gene expression data. *Soft Comput.* **2020**, *24*, 11381–11391. [\[CrossRef\]](#)
38. Yoon, J.; Drumright, L.; Van Der Schaar, M. Anonymization through data synthesis using generative adversarial networks (ads-gan). *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2378–2388. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Assefa, S.; Dervovic, D.; Mahfouz, M.; Tillman, R.; Reddy, P.; Veloso, M. Generating synthetic data in finance: Opportunities, challenges and pitfalls. In Proceedings of the First ACM International Conference on AI in Finance, New York, NY, USA, 15–16 October 2020; pp. 1–8. [\[CrossRef\]](#)
40. Pooyandeh, M.; Sohn, I. Edge network optimization based on ai techniques: A survey. *Electronics* **2021**, *10*, 2830. [\[CrossRef\]](#)
41. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Keskar, N.; Socher, R. Improving generalization performance by switching from adam to sgd. *arXiv* **2017**, arXiv:1712.07628. <https://doi.org/10.48550/arXiv.1712.07628>.
43. Yoon, J.; Jarrett, D.; Van der Schaar, M. Time-series generative adversarial networks. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 5509–5519.
44. Fei, C. Lithium-Ion Battery Data Set. 2022. Available online: <https://iee-dataport.org/documents/lithium-ion-battery-data-set> (accessed on 30 November 2023).
45. Sagheer, A.; Kotb, M. Unsupervised pre-training of a deep LSTM-based stacked autoencoder for multivariate time series forecasting problems. *Sci. Rep.* **2019**, *9*, 19038. [\[CrossRef\]](#)
46. Munir, M.; Siddiqui, S.; Dengel, A.; Ahmed, S. DeepAnT: A deep learning approach for unsupervised anomaly detection in time series. *IEEE Access* **2018**, *7*, 1991–2005. [\[CrossRef\]](#)
47. Fontes, C.; Pereira, O. Pattern recognition in multivariate time series—A case study applied to fault detection in a gas turbine. *Eng. Appl. Artif. Intell.* **2016**, *49*, 10–18. [\[CrossRef\]](#)
48. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555. <https://doi.org/10.48550/arXiv.1412.3555>.

49. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680.
50. Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; Hochreiter, S. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *arXiv* **2017**, arXiv:1706.08500.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.