An Adaptive Combined Method for Lithium-Ion Battery State of Charge Estimation Using Long Short-Term Memory Network and Unscented Kalman Filter Considering Battery Aging

Longchen Lyu,[a, b] Bo Jiang,*[a, b] Jiangong Zhu,[a, b] Xuezhe Wei,[a, b] and Haifeng Dai*[a, b]

The accurate estimation of battery state of charge (SOC) enables the reliable and safe operation of lithium-ion batteries. Data-driven SOC estimation is considered an emerging and effective solution. However, existing data-driven SOC estimation methods typically involve direct estimation and lack effective feedback correction. Moreover, battery degradation poses additional challenges to accurate SOC estimation. Therefore, this study proposes an adaptive combined method for battery SOC estimation based on a long short-term memory (LSTM) network and unscented Kalman filter (UKF) algorithm considering battery aging status. First, an LSTM model is constructed to characterize the battery's dynamic performance instead of traditional battery models. Then, the UKF algorithm is employed to perform SOC estimation through the feedback of terminal voltage prediction. To enhance estimation accuracy under different aging statuses, a proportional-integral-derivative controller is employed to correct the capacity fading during the SOC estimation process. Validation results indicate that the terminal voltage prediction model demonstrates exceptional robustness against interference from current and voltage noise. Compared to the traditional estimation method combining the deep learning model and Kalman filter algorithm, the proposed method demonstrates superior estimation accuracy under various complex operating conditions. Furthermore, the proposed method outperforms the traditional method in estimation performance during battery aging.

Introduction

The desire for carbon neutrality is driving the development of clean energy storage systems, with electric vehicles (EVs) playing a pivotal role. Lithium-ion batteries are known for their high energy/power density, long service life, and environmental friendliness, and they are one of the most promising technologies utilized in EVs.[1,2] Due to the widespread adoption of EVs, the essential functions of the battery management system such as state estimation, optimizing charging, and discharging control have become increasingly important.[3] State of charge (SOC), defined as the ratio of the residual capacity to the nominal capacity, indicates the remaining charge state of the battery and represents its safe operating range. However, SOC cannot be directly measured or even cannot be linearly related to external measurements such as voltage and temperature. Therefore, reliable estimation for battery SOC still faces significant challenges.

Currently, SOC estimation methods for lithium-ion batteries can be classified into three categories: principle-based methods, model-based methods, and data-driven methods.

Principle-based methods are relatively basic SOC estimation methods, including the Coulomb counting method, internal resistance method, and open-circuit voltage (OCV) method. Each of these methods has its limitations. The Coulomb counting method, also known as the ampere-hour integration method, is based on the initial SOC and current integration It is susceptible to influences from the initial SOC and current measurement errors.[4] For lithium-ion batteries, the impedance phase at specific frequencies and constant temperature has a strong linear relationship with SOC, making it a viable parameter for SOC estimation.[5] However, the impedance phase will change due to battery aging, and its measurement is easily affected by the accuracy of current and voltage sampling. The OCV method is based on the nonlinear mapping relationship between the battery SOC and the OCV.[6] However, the acquisition of battery OCV requires a long period of rest, making it difficult to achieve in practical applications. Furthermore, the SOC-OCV relationship can shift during the battery degradation, affecting the accuracy of SOC estimation. In conclusion, these SOC estimation methods are easily influenced by factors such as the initial state and current and voltage sampling. Currently, the Coulomb counting method and OCV method are used as components of more complex SOC estimation algorithms.[7]

Model-based methods are currently one of the widely used estimation techniques. Battery models primarily include electrochemical models[8,9] and equivalent circuit models (ECMs).[10,11] The electrochemical model contains complex reaction processes and numerous internal parameters.[9] Although it can better explain the internal performance of the battery theoret-

[a]   L. Lyu, B. Jiang, J. Zhu, X. Wei, H. Dai
      School of Automotive Studies, Tongji University, Shanghai 201804, China
      E-mail: jiangbo15@tongji.edu.cn
              tongjidai@tongji.edu.cn

[b]   L. Lyu, B. Jiang, J. Zhu, X. Wei, H. Dai
      Clean Energy Automotive Engineering Center, Tongji University, Shanghai 201804, China

Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202400441

Chemistry
Europe

European Chemical
Societies Publishing

ically, it involves substantial computational effort and poses challenges to parameter identification. In contrast, the ECM is the most commonly used model for SOC estimation due to its adaptability to various operating states, simple structure, and fewer required parameters. Model-based methods rely on the accuracy of the model and the precision of parameter identification, and they are typically combined with various filtering or observer algorithms, such as the Kalman filter (KF),[12] H-infinity filter,[13] particle filter,[14] and sliding mode observer,[15,16] to achieve the estimation of unknown states. Cui et al.[17] used an extended Kalman filter (EKF) to address estimation errors caused by Gaussian noise. Lin et al.[18] employed the exogenous Kalman filter for SOC estimation, achieving better results than the EKF. Maletić et al.[19] discussed the fast-scale SOC estimation using dual extended Kalman filters to ensure high SOC estimation accuracy. Shrivastava et al.[20] reviewed the role of the KF algorithm family in online SOC estimation and evaluated the performance of various KF algorithms. Battery aging is accompanied by changes in internal parameters.[21] Consequently, model-based methods that rely on these parameters lack adaptability to battery aging. Therefore, a model that does not depend on parameters is more desirable.

The data-driven method directly learns the nonlinear mapping relationship between measurement data and SOC, without relying on the principles and models of battery, avoiding errors in battery modeling. With the proliferation of machine learning, the data-driven method has become a new focus area.[22] Machine learning models such as Support Vector Machines (SVM),[23,24] Gaussian Process Regression (GPR),[25–28] and Deep Neural Networks (DNN)[29,30] learn the nonlinear mapping between battery measurement data and SOC. SVM and GPR, as traditional machine learning models, support the learning of nonlinear mappings with small samples but rely heavily on manually extracted features, limiting their ability to fully capture the complex nonlinear features of large sample data.[25–27] Deep learning, suitable for training on large sample data, has seen broader usage, such as Backpropagation Neural Networks[31] and Recurrent Neural Networks (RNNs).[32–34] RNN with multiple layers of artificial neurons adjust weights and biases for inter-layer data transfer, extracting information layer by layer and passing it downwards. Theoretically, RNN can map to any function, learning the complex mapping relationship between large sample input and output data. Currently, many scholars pay attention to RNN for SOC estimation. Zhao et al.[32] proposed a multi-level optimized training RNN to address RNN gradient issues. Dang et al.[35] improved the SOC estimation performance by combining differential equations with RNNs. Shahriar et al.[36] and Kannan et al.[30] improved model performance by introducing hybrid RNN models. Shahriar introduced a hybrid RNN model combining Convolutional Neural Network (CNN), Gate Recurrent Unit, and long short-term memory (LSTM), while Kannan proposed a new hybrid architecture combining DNN and Neural Basis Expansion Analysis for Time Series. Jiang et al.[37] proposed a mechanics-based battery SOC estimation approach using LSTM which can be adaptable to different training and operating conditions. Yuan et al.[38] and

Chung et al.[39] believe that the LSTM-RNN model has better SOC estimation accuracy compared to other estimation models. However, research by Tao et al. and Sun et al.,[40,41] still shows that current data-driven SOC estimation suffers from issues such as data overfitting, slow convergence, and local optimality, leading to significant estimation errors. Currently, data-driven methods directly obtain SOC from measurements, but unavoidable estimation errors still exist. Therefore, addressing this issue requires new approaches. The coupling of multiple algorithms for multi-estimation target error compensation has become a crucial method for improving SOC estimation and enhancing the accuracy of SOC estimation under battery aging. Many scholars have researched the coupling of multi SOC estimation algorithms: Yuan et al.[38] used the output of the KF algorithm as input for machine learning to estimate SOC, obtaining better performance; Yun et al.[42] optimized the SOC estimation algorithm by adjusting the Unscented Kalman Filter (UKF) parameters through a fuzzy Proportional-Integral-Derivative (PID) controller; Yang et al.[43] achieved stable and smooth SOC estimation results by CNN, a system error feedback mechanism, and a KF post-processor. These methods retain a direct structure that obtains output results directly through input conditions. Although the accuracy of the data-driven method can be enhanced through coupling with other algorithms, the inherent direct structure of the model cannot be altered, resulting in poor accuracy and lack of stability.

The above literature review shows that the researchers have investigated many effective SOC estimation methods and indicates that the data-driven method is considered an emerging technique. However, the existing data-driven SOC estimation methods still face the following issues: (1) The traditional data-driven SOC estimation is a typically direct estimation structure, using the basic form of external measurement as the model input and battery SOC as the model output. The lack of feedback in the direct structure can lead to algorithm instability and further constrain the accuracy of SOC estimation. (2) Battery degradation significantly influences the relationship between the external measurement and battery SOC. The traditional data-driven method usually adopts large datasets from different aging cycles to deal with this problem. How to reasonably correct the estimation model considering battery aging still needs to be investigated.

To address these issues, this paper proposes an adaptive combined method: (1) An indirect estimation structure is constructed in the proposed method. In this structure, the LSTM model is disposed to characterize the battery's dynamic performance. SOC is updated adaptively and fed into the LSTM model for voltage prediction. Meanwhile, the UKF algorithm is employed to perform SOC estimation through the feedback of terminal voltage prediction. This structure contains a feedback mechanism and increases the performance of SOC estimation. (2) Considering the capacity fading under battery degradation, a PID controller is employed to correct the capacity. Adaptive correction makes it more accurate to estimate the SOC under battery aging using the model trained by datasets before battery aging.

The remainder of this paper is organized as follows: Section 2 illustrates the basic idea and detailed estimation framework including voltage prediction, SOC estimation, and capacity correction. The experimental instruments and test procedures are introduced in Section 3. Section 4 presents and analyzes the performance of the proposed SOC estimation approach, and the research limitations are presented. Finally, Section 5 concludes this research.

Methodologies

The Basic Idea of the Proposed SOC Method

As shown in Figure 1(a), the KF algorithm ensures the stability of SOC estimation during dynamic battery processes through feedback regulation. The model applied in the model-based methods is a description of the dynamic performance of the battery, which requires the identification of numerous parameters. However, the model is susceptible to the interference of parameter changes under battery aging. As shown in Figure 1(b), the traditional data-driven SOC estimation uses the current and voltage as the input and battery SOC as the output. And it does not rely on a description of the battery operation mechanism. However, the direct structure in traditional data-driven methods lacks effective feedback correction. Based on the above methods, this paper puts forward a new idea combining the advantages of the two methods. In Figure 1(c), an LSTM model for terminal voltage prediction serves as the model for the UKF algorithm to estimate SOC, constructing an indirect structure. The accuracy of SOC estimation can be strengthened due to effective feedback. Moreover, adaptive capacity correction improves the dynamic performance characterized under battery aging.

Terminal Voltage Prediction Based on the LSTM-RNN Model

The RNN is a basic type of artificial neural network characterized by internal cyclic connections. RNN achieves the storage and processing of long-term dependencies of historical data using recurrent information flow. RNNs are structured into three layers: the input layer, which receives input sequences and forwards them; the hidden layer, which maintains network memory through recurrent connections; and the output layer, which produces output results. The operational principles of an RNN can be succinctly described in the equations governing the transformation from the input layer to the hidden layer and from the hidden layer to the output layer:

$$\begin{cases} h_t = tanh(W_{hh}h_{t-1} + W_{xh}in_t) \\ out_t = W_{hy}h_t \end{cases} \quad (1)$$

where, h_t is the current hidden state, h_{t-1} is the previous hidden state, in_t is the current input, y_t is the current output, and W is the weight matrix learned during training process. The $tanh(\cdot)$ is the hyperbolic tangent function.

Basic RNN struggles with long-term memory retention, particularly when trained over extended time steps, leading to gradient vanishing or explosion. Consequently, LSTM is developed to address these deficiencies. As illustrated in Figure 2(a), to maintain a stable gradient, LSTM incorporates a cell state in the hidden layer, which is modified by the hidden state and input. As shown in Table 1, modifications to the cell structure are achieved through gate structures, where, f_t is the forget gate, i_t is the input gate, o_t is the output gate, C_t is the new cell state, C_{t-1} is the previous state, $\tilde{C}_t$ is the candidate, an $\sigma(\cdot)$ is logistic sigmoid function.
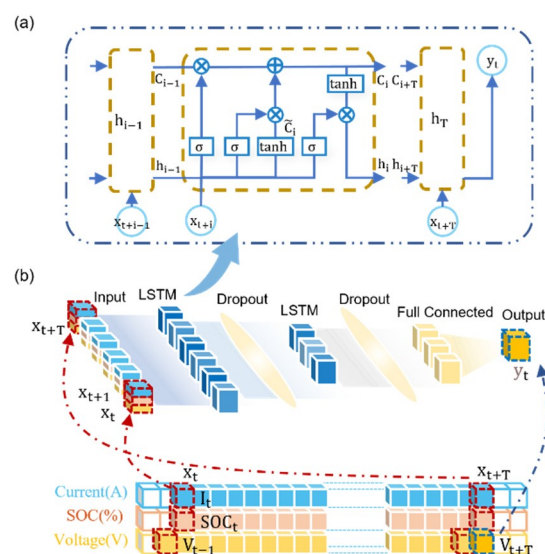

Figure 2. LSTM-RNN Model Structure. (a) The Basic Structure of an LSTM Unit. (b) LSTM Network Diagram.

Figure 1. Method Diagram. (a) The Model-Based Method. (b) The Data-Driven Method (c) The Proposed Method.

Table 1. LSTM Gate Structures.

Forget gate	$f_t = \sigma(W_f \cdot [h_{t-1}, in_t]^\top)$
Input gate	$i_t = \sigma(W_i \cdot [h_{t-1}, in_t]^\top)$
	$\tilde{C}_t = tanh(W_{Ci} \cdot [h_{t-1}, in_t]^\top)$
	$C_t = f_t C_{t-1} + i_t \tilde{C}_t$
Output gate	$o_t = \sigma(W_o \cdot [h_{t-1}, in_t]^\top)$

Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202400441

**Chemistry
Europe**
European Chemical
Societies Publishing

To train the LSTM-RNN model, battery data needs to be organized and processed: The input vector is defined as $IN = [in_1, in_2, ..., in_t, ...]$, where $in_t = [I_t, V_{t-1}, SOC_t]^T$ comprises the current I_t and SOC_t sampled in this instance, and the terminal voltage V_{t-1} sampled in the previous instance. The output vector is defined as $OUT = [out_1, out_2, ..., out_t, ...]$, where $out_t = V_{t+T}$, representing the terminal voltage of length T sampled after this instance. The SOC is computed according to the definition of battery SOC:

$$SOC_t = SOC_{t0} - \sum_{i=t0}^{t} \frac{I_i \Delta t}{Q} \qquad (2)$$

where, SOC_t is the SOC corresponding to the sequence number of battery data, SOC_{t0} is the initial SOC, Δt is the sampling time, and Q is the battery capacity.

The RNN dataset is constructed by progressively obtaining multiple sets of training data from IN and OUT, as illustrated in Figure 2(b). The training data consists of an input sequence of length T, denoted as $Input = [in_t, in_{t+1}, ..., in_{t+T}]$ and $Output = [out_t, out_{t+1}, ..., out_{t+T}]$. To eliminate the influence of dimensions and prevent gradient direction deviation during RNN training, the data needs to be normalized:

$$\widetilde{data} = \frac{data - data.mean}{data.std} \qquad (3)$$

where $\widetilde{data}$ represents the normalized data, $data$ represents the original data, and $data.mean$ and $data.std$ are the mean and standard deviation of the original data.

The LSTM model constructed in this paper is shown in Figure 2(b), which employs a two-layer LSTM architecture. First of all, the first layer of the model is the sequence input layer. Then, the first layer has 128 units and its output mode needs to be set to 'last'. The second layer has 64 units and its output mode needs to be set to 'last'. Each LSTM layer is followed by a dropout layer with a dropout rate of 0.2. The final layer is produced through a fully connected layer. The optimizer used is Adam, with the minimum batch size set to 64, an epoch number of 100, an initial learning rate of 0.01, and a learning rate decay strategy where the learning rate is reduced to 0.1 of its original value every 30 epochs. Other optimizer parameters are set to default values. During training, the data length T is set to 500. All parameters are randomly initialized, and the optimizer iteratively adjusts the weights and biases.

SOC Estimation Method Based on Voltage Deviation and UKF

The KF algorithm assumes that the current true state evolves from the previous state, utilizing measurements and predictions to obtain the optimal estimate of the current state, which is primarily used for analyzing linear systems. The method designed in this paper predicts terminal voltage through an LSTM-RNN model, calculates the deviation between the predicted and measured terminal voltage, and feeds back to

adjust the SOC. Since the LSTM-RNN model is a highly nonlinear system and cannot be linearized in an extended way, a UKF algorithm is required.[41] In the UKF algorithm, an unscented transform propagates the system's highly nonlinear state through a nonlinear function in the form of probability distribution while maintaining the accuracy of the state mean and covariance, enabling the KF algorithm to perform recursive operations.

As shown in Figure 3, based on the UKF algorithm and combined with the terminal voltage prediction model based on LSTM-RNN, a feedback estimation of SOC is implemented. SOC_k represents the SOC in the current state, which serves as the input for the UKF algorithm into the LSTM model and V_k represents the terminal voltage calculated by the LSTM model, which is used as the feedback in UKF. In the UKF, the state equation is provided by the definition of battery SOC, achieving an initial estimate of the SOC through state updates. The observation equation is provided by the voltage prediction model, where the SOC estimate is corrected by comparing the predicted and measured terminal voltage values. The system's state, input, and output are respectively as follows: $x_k = SOC_k$, $u_k = I_k$ and $z_k = V_k$, the state-space equation of the system is:

$$\begin{cases} x_k = f(x_{k-1}, u_k) = x_{k-1} - \frac{\Delta t}{Q_0} u_k \\ z_k = h(x_k, u_k) = LSTM(Input(x_k)) \end{cases} \qquad (4)$$

The input dataset $Input(x_k, u_k)$ is constructed from the state x_k and input u_k of the UKF and serves as the input data for the LSTM model. This dataset maintains the same structure as the training set described in the previous chapter. Each iteration updates the SOC sequentially according to the SOC definition formula. The current, SOC, and previous terminal voltage in the input sequence of the LSTM model are progressively updated.

The UKF algorithm process is shown in Table 2, where, x_k is the current state, x_{k-1} is the previous state, u_k is the input, ω is the system noise, and v is the observation noise, $f(\cdot)$ is the state function, $h(\cdot)$ is the observation function.
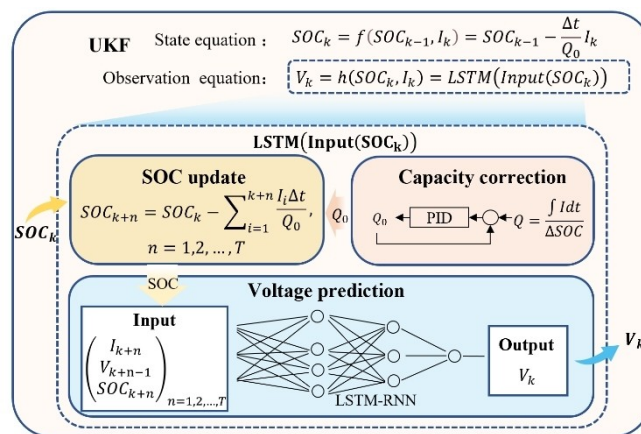

Figure 3. SOC Estimation Method Architecture.

Table 2. UKF Algorithm.

The state space expression is as follows,

$$\begin{cases} x_k = f(x_{k-1}, u_k) + \omega \\ z_k = h(x_k, u_k) + v \end{cases}$$

Perform an unscented transformation on the previous state to compute the prior covariance of the state.

$$X_{k-1} = \left[x_{k-1}^{[0]}, x_{k-1}^{[1]}, \ldots, x_{k-1}^{[2n]} \right]^{\mathrm{T}}$$

$$\begin{cases} x_{k-1}^{[i]} = x_{k-1}, i = 0 \\ x_{k-1}^{[i]} = x_{k-1} + \left(\sqrt{(n+\lambda)P_{k-1}} \right)_i, i = 1 \sim n \\ x_{k-1}^{[i]} = x_{k-1} - \left(\sqrt{(n+\lambda)P_{k-1}} \right)_i, i = n+1 \sim 2n \end{cases}$$

$$P_{k,k-1} = \sum_{i=0}^{2n} \omega_c^{[i]} \left(x_{k,k-1} - x_{k,k-1}^{[i]} \right) \left(x_{k,k-1} - x_{k,k-1}^{[i]} \right)^{\mathrm{T}} + Q$$

Compute the prior of the current state.

$$X_{k,k-1} = \left[f\left(x_{k-1}^{[0]}, u_k \right), f\left(x_{k-1}^{[1]}, u_k \right), \ldots, f\left(x_{k-1}^{[2n]}, u_k \right) \right]^{\mathrm{T}}, x_{k,k-1} = \sum_{i=0}^{2n} \omega_m^{[i]} x_{k,k-1}^{[i]}$$

Perform an unscented transformation on the prior of the current state to compute the state observation estimate.

$$X_k = \left[x_k^{[0]}, x_k^{[1]}, \ldots, x_k^{[2n]} \right]^{\mathrm{T}}, \begin{cases} x_k^{[i]} = x_{k,k-1}, i = 0 \\ x_k^{[i]} = x_{k,k-1} + \left(\sqrt{(n+\lambda)P_{k,k-1}} \right)_i, i = 1 \sim n \\ x_k^{[i]} = x_{k,k-1} - \left(\sqrt{(n+\lambda)P_{k,k-1}} \right)_i, i = n+1 \sim 2n \end{cases}$$

$$Z_k = \left[h\left(x_k^{[0]}, u_k \right), h\left(x_k^{[1]}, u_k \right), \ldots, h\left(x_k^{[2n]}, u_k \right) \right]^{\mathrm{T}}, \hat{z}_k = \sum_{i=0}^{2n} \omega_m^{[i]} z_k^{[i]}$$

Compute the covariance to obtain the Kalman gain.

$$P_{Z_k} = \sum_{i=0}^{2n} \omega_c^{[i]} \left(\hat{z}_k - z_k^{[i]} \right) \left(\hat{z}_k - z_k^{[i]} \right)^{\mathrm{T}} + R$$

$$P_{X_k, Z_k} = \sum_{i=0}^{2n} \omega_c^{[i]} \left(x_{k,k-1} - x_k^{[i]} \right) \left(\hat{z}_k - z_k^{[i]} \right)^{\mathrm{T}} K_k = P_{X_k, Z_k} P_{Z_k}^{-1}$$

Obtain the current state and the covariance of the current state

$$x_k = x_{k,k-1} + K_k(z_k - \hat{z}_k) P_k = P_{k,k-1} - K_k P_{k-1} K_k^{\mathrm{T}}$$

In the SOC indirect estimation combining LSTM and UKF, the system noise variance is set to 1e–6, the observation noise variance is set to 1e–3, and the initial covariance is set to 1e–3.

Capacity Correction Based on Incremental PID

Battery systems are prone to self-induced aging, which results in changes in electrochemical parameters under different battery aging states. The changes in capacity interfere with the input data for the LSTM model, affecting the terminal voltage output and consequently impacting the SOC estimation. To ensure the accuracy of SOC estimation results, the initial capacity parameter needs to undergo self-feedback correction on a variable time scale. During the self-correction process of the initial capacity, the capacity is calculated based on the SOC variation under a specific Ampere-hour integral, serving as the feedback quantity for the self-feedback correction system.

$$Q_C = \frac{\int_{\tau_0}^{\tau} Idt}{SOC_\tau - SOC_{\tau_0}} \tag{5}$$

where, τ_0 represents the moment when the Ampere-hour integral begins, and τ represents the moment when the Ampere-hour integral ends. The value of τ must ensure that $\int_{\tau_0}^{\tau} Idt$ equals a specific cumulative capacity amount. Upon completing this cumulative operation, the moment τ becomes the starting point for a new round of Ampere-hour integral, and the process restarts.

The incremental PID control algorithm will perform the primary function of capacity self-feedback correction. The incremental PID control algorithm is a recursive method that calculates the difference between the control quantities at the current moment and the previous moment, using the increment of the control quantity as the new control quantity. The formula is as follows:

$$\Delta u_t = K_P(e_t - e_{t-1}) + K_I e_t + K_D(e_t - 2e_{t-1} + e_{t-2}) \tag{6}$$

where $e_t = Q_C - Q_0$ represents the deviation between the process capacity Q_C and the initial capacity Q_0, in which initial capacity is set to 2.89 Ah. K_P, K_I, and K_D are the proportional, integral, and derivative coefficients, which are set to 0.1, 0.4 and 1e—3.

Evaluation of the Proposed Method

To evaluate the proposed model and methods in this paper, three metrics were established: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Maximum Absolute Error (MAXAE).

$$MAE = \frac{1}{n} \sum_{k=1}^{n} |\hat{y}_k - y_k| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (\hat{y}_k - y_k)^2} \tag{8}$$

$$MAXAE = \max_{k=1,2,\ldots,n} |\hat{y}_k - y_k| \tag{9}$$

where $\hat{y}_k$ is estimation value, and y_k is actual value.

Experimental Design and Data Description

This paper used a commercial 18650-type battery, along with a Li(NiCoMn)O$_2$ cathode and graphite anode. The battery has a nominal capacity of 2.9 Ah, a charging cut-off voltage of 4.2 V, and a discharging cut-off voltage of 2.5 V. The battery is tested in different aging cycles and selected cycle conditions, the whole test is carried out in the temperature chamber, the charge and discharge test is carried out by the Chroma 17011 system, and the battery information is transmitted to the upper computer. The experimental environment in this study was maintained at 25 °C, and the battery was tested under the conditions of New European Driving Cycle (NEDC), World Light Vehicle Test Procedure (WLTP), and Urban Dynamometer Driving Schedule (UDDS), as shown in Figure 4.

Additionally, to obtain operating data under battery aging, aging cycle experiments were conducted, with NEDC condition tests performed every 100 cycles. As shown in Table 3, capacity fading is more pronounced at 400, 500 and 600 cycles.
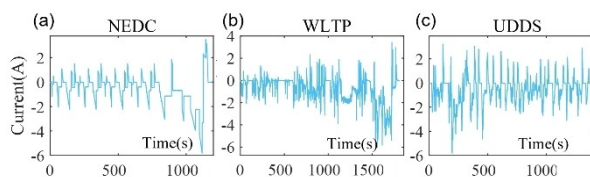
Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202400441

Chemistry
Europe
European Chemical
Societies Publishing

Figure 4. Battery Operating Condition. (a) NEDC Condition. (b) WLTP Condition (c) UDDS Condition.

Table 3. Battery Capacity Under Aging Cycles.

Cycles	0	100	200	300	400	500	600
Capacity (Ah)	2.89	2.87	2.85	2.81	2.76	2.68	2.58

Results and Discussion

Battery Terminal Voltage Prediction Results and Analysis

The setting of the data length significantly affects the learning performance of RNN, determining their memory capacity and the range of information transmission. An appropriate data length can enhance the ability to uncover hidden logical relationships within the data, thereby improving the estimation performance. Stacking LSTM layers, as opposed to merely increasing the number of units, further deepens the RNN, facilitating the exploration of intricate patterns and features within the data, and capturing long-term dependencies in sequences. However, this approach is prone to overfitting. As critical parameters in training the RNN-LSTM model, both the data length and the number of LSTM layers are crucial. This study validates the reasonableness of the paramete" set'Ings and the performance of the LSTM voltage prediction model.

This paper sets data lengths of 300 and 500 to validate the reasonableness of the training set data length. A single-layer LSTM structure (retaining only 128 LSTM units) is employed to validate the reasonableness of the model layer structure. To verify the model's performance, Gaussian noise with a variance of $\sigma^2 = 1e—2$ was respectively added to the current sampling data and voltage sampling data, while other data remains unchanged. The validation results are shown in Figure 5(a). The double-layer LSTM structure with a data length of 500s exhibits smaller RMSE and MAE, in which the voltage prediction RMSE is within 0.002 V. The impact of changing data length on estimation accuracy is greater than that of changing the number of LSTM layers. Under the condition of applied current noise, the deviation across all models showed no significant change, demonstrating the LSTM model's robustness to current noise. Under the condition of applied voltage noise, the increase in terminal voltage prediction error is consistent with the error introduced by the applied Gaussian noise, indicating that the terminal voltage prediction model inherits the noise error from the original data without amplifying it. The above analysis confirms the robustness of the model.

To verify whether the model can provide correct voltage feedback under conditions of initial SOC deviation, tests were
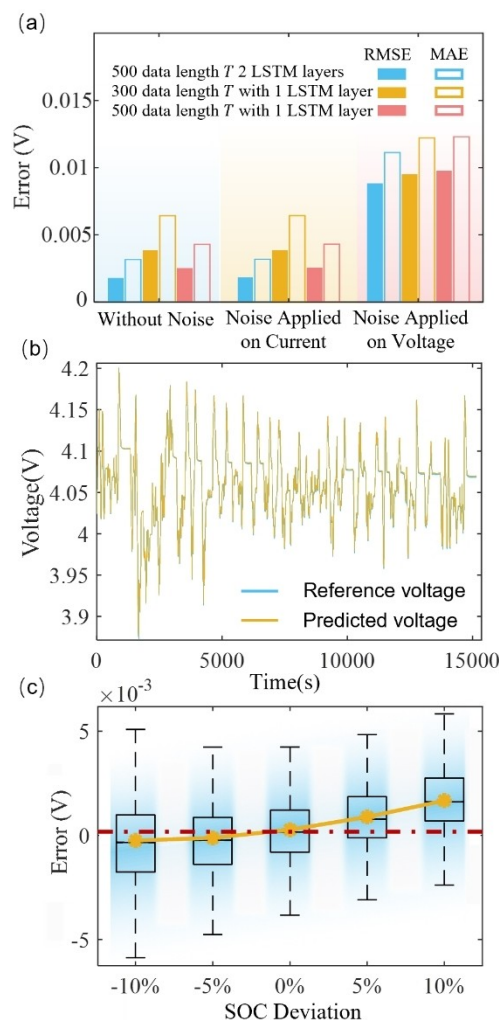

Figure 5. The Performance of Terminal Voltage Prediction Model. (a) The Error of Terminal Voltage Prediction Under Different Model Parameters and Noise Interference. (b) The Terminal Voltage Prediction Process Under a Double-Layer LSTM Structure With a Data Length of 500. (c) The Error of Terminal Voltage Prediction Under Different SOC Deviations.

conducted with SOC deviations of $\pm 5\%$ and $\pm 10\%$. As shown in Figure 5l, the mean error of the nominal SOC is close to zero, and the SOC deviation gradient can map onto the voltage deviation. The larger the SOC deviation, the larger the terminal voltage deviation. This indicates that voltage prediction deviation can serve as feedback for SOC estimation deviation.

Battery SOC Estimation Results and Analysis

SOC Estimation under Different Operating Conditions

To validate the performance of the proposed SOC estimation method, this section utilizes the NEDC, UDDS, and WLTP battery datasets, in which the battery has undergone 100 cycles of aging. Any one of these datasets is selected as the training set, while the others are used as the test set. The six cases for SOC estimation algorithm validation are shown in Table 4. Addition-

Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202400441

Chemistry
Europe

European Chemical
Societies Publishing

Table 4. The Cases for SOC Estimation Algorithm Validation.

Case	(1)	(2)	(3)	(4)	(5)	(6)
Training set	NEDC	WLTP	NEDC	UDDS	UDDS	WLTP
Test set	UDDS	UDDS	WLTP	WLTP	NEDC	NEDC

ally, in reference,[44,45] a traditional direct estimation model combining LSTM and KF is set as a comparison, in which the LSTM model takes current and voltage as inputs, and SOC as the output.

As shown in Figure 6, the SOC estimates and deviations in both proposed and traditional methods vary with the dynamic conditions. Notably, the dynamic SOC estimation error in the traditional method is larger and more pronounced, while the dynamic SOC estimation error in the proposed method is smaller. The proposed SOC estimation method exhibits a more stable estimation process than the traditional method, effec-

tively avoiding problems such as data overfitting, slow convergence, and local optimality. During the traditional direct estimation process, the SOC dynamic estimation process under different training set models for each test set exhibits relatively consistent error curve patterns.

The statistical data from the bar chart In Figure 7 shows that the proposed SOC estimation method exhibits smaller errors than the traditional method. The SOC estimation RMSE in the proposed method is within 1.2% and is reduced by 0.507% at least. Specifically, in Case (2) and (6), when the WLTP condition is used as the training set, the error is significantly smaller. While, in Case (5) and (6), when the NEDC condition is used as the test set, the error Is also notably smaller."For'he traditional SOC estimation model, the error is smaller when the WLTP condition is used as the test set model. In the traditional method of Case (1) and (5), when the NEDC and UDDS conditions are used as each other's training and test sets, the error increases significantly. This discrepancy may be attributed to the substantial differences in the dynamic characteristics between these conditions, but the proposed method addresses this issue and results in a smaller estimation error. In summary, the proposed SOC estimation method that combines the UKF and LSTM-RNN model demonstrates advantages.

Figure 6. SOC Estimation Error: (a–f) Present the Process Errors for Case (1), (2), (3), (4), (5) and (6).

SOC Estimation under Noise Interference

To validate the robustness of the proposed SOC estimation method, Gaussian noise with a variance of $\sigma^2 = 1e{-}2$ was added to the current sampling data and voltage sampling data, respectively, using Case (1), (2), and (3).

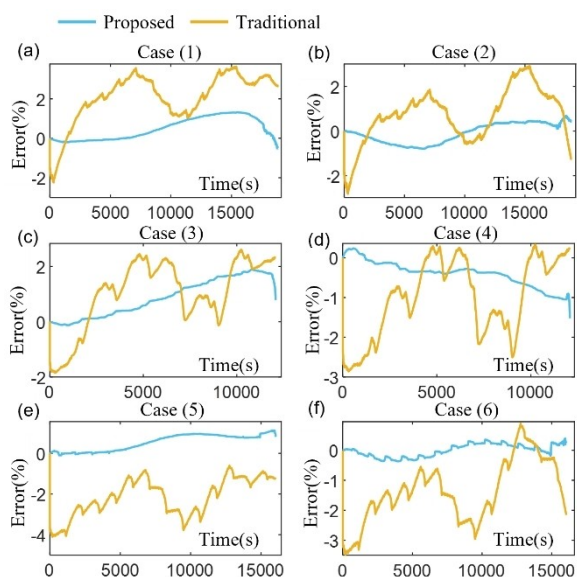As shown in Figure 8, different types of sampling noise have varying impacts on SOC estimation results under different
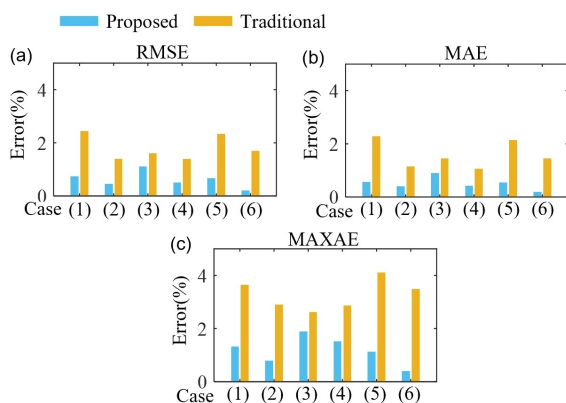

Figure 7. Statistics on SOC Estimation Errors: (a–c) Present the RMSE, MAE, and MAXAE for Case (1), (2), (3), (4), (5), and (6).
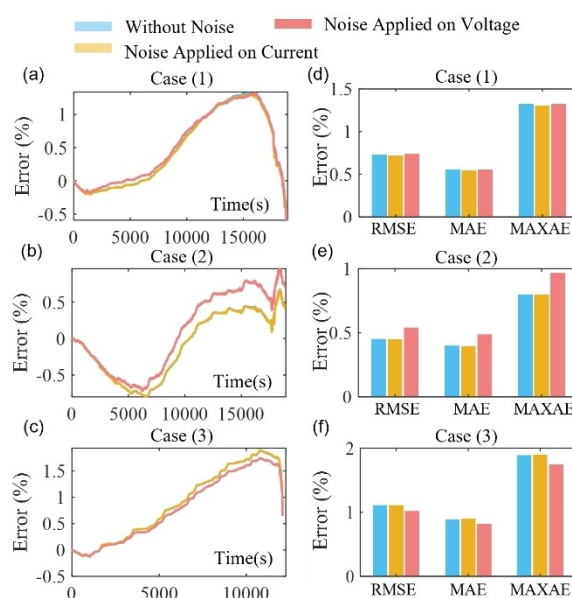

Figure 8. The Robustness of SOC Estimation Method: Figure (a–c) Present the Process Errors Under Noise Interference for Case (1), (2), and (3); (d–f) Present the RMSE, MAE, and MAXAE Under Noise Interference for Case (1), (2) and (3).

Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202400441

Chemistry
Europe
European Chemical
Societies Publishing

conditions. However, overall, noise does not significantly increase estimation errors, maintaining the same estimation accuracy, in which the maximum increase of RMSE is 0.0876 %. In Case (1) and (3), where the same training set model is applied to different training sets, the application of current noise and voltage noise does not increase estimation errors. In Case (1) and (2), where different training set models are applied to the same test set, only when NEDC is used as the training set and voltage noise is applied, there is a slight increase in estimation error. In all other cases, no increase in error is observed. From the bar chart in Figure 8, it is evident that the impact of applying current noise and voltage noise on estimation errors is minimal, further validating the robustness of the algorithm.

SOC Estimation under Initial SOC Deviation

To validate the correction capability of the proposed SOC estimation method under conditions of significant initial SOC deviations, tests were conducted with initial SOC deviations of 20 % and 50 %, using Case (1), (2), and (3). As shown in Figure 9, the SOC correction process involves quickly reducing the deviation to within 10 % and then further minimizing it. In this process, the correction time for a 20 % SOC deviation is shorter, and its impact on the SOC estimation accuracy is smaller compared to a 50 % SOC deviation. In Case (1) and (2), there is a slight overshoot in the SOC dynamic estimation process, and the larger the SOC deviation, the greater the resulting overshoot, with the maximum overshoot being less than 5 %. The overshoot is primarily caused by excessive SOC correction and will be gradually adjusted during the dynamic estimation process. During the SOC steady-state estimation phase, the SOC estimation is free from SOC deviation, and the estimation results are stable. The above analysis indicates that the SOC joint estimation model possesses significant SOC deviation correction capability: the smaller the deviation, the faster the correction, and the higher the subsequent SOC estimation accuracy.
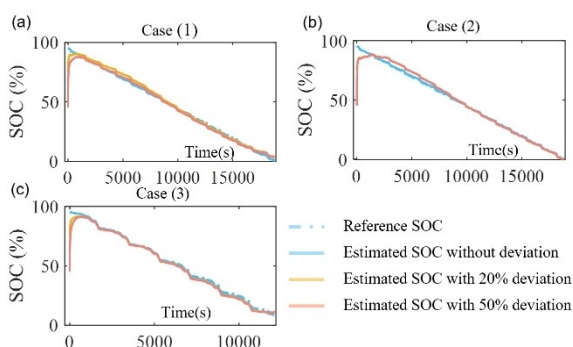
SOC Estimation under Battery Aging

To validate the adaptability of the proposed SOC estimation method under battery aging, the NEDC battery dataset, aged through 100 cycles, is used as the training set. Validation sets consisted of NEDC battery datasets aged through 400, 500, and 600 cycles. As shown in Figure 10, the proposed estimation algorithm exhibits significantly smaller estimation deviations and a more stable estimation process compared to the traditional estimation method. The SOC estimation RMSE in the proposed method is within 1.2 % and is reduced by 0.69 % at least. In all three aging states, the traditional estimation algorithm tends to underestimate SOC at high SOC levels, while it overestimates SOC at medium and low SOC levels. Conversely, in the proposed estimation algorithm, overestimation occurs at medium and low SOC levels, and as SOC decreases further, the estimation deviation diminishes, even showing negative deviations in the 500 and 600-cycle cases. The SOC error curves exhibit consistent shapes. It is likely due to the effect of battery cycle aging Introducing small systematic deviations In the electrochemical system of the battery.

Limitations and Prospects

This paper proposes an adaptive combined method for lithium-ion battery SOC estimation. The proposed method combines the advantages of LSTM and UKF and it shows preferable performance of SOC estimation under battery aging. However, there are still some limitations in this study that need to be addressed:

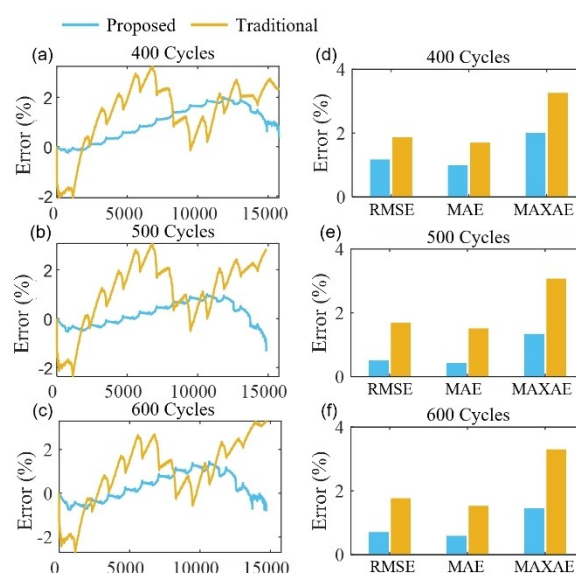(1) The experiments in this paper were conducted based on a battery operating environment at 25 °C. However, in actual

Figure 10. SOC Estimation Error Under Different Cyclic Aging Conditions: (a–c) Present the Process Errors Under Battery Aging of 400, 500, and 600 Cycles; (d–f) Present the RMSE, MAE, and MAXAE Under Battery Aging of 400, 500, and 600 Cycles.

Figure 9. SOC Estimation Process at Initial SOC Deviation: (a–c) Present the Process Errors Under SOC Deviation for Case (1), (2), and (3).

Batteries & Supercaps

Research Article
doi.org/10.1002/batt.202400441

Chemistry Europe
European Chemical
Societies Publishing

applications, the temperature under actual battery conditions will be constantly changing. Therefore, further exploration is needed for battery SOC estimation under varying temperature conditions.

(2) This paper employs the LSTM model for terminal voltage prediction, which requires a large time consumption. To meet the demands of actual applications, future research should aim to reduce the training time.

(3) In this paper, the verification object is the Ni–Co–Mn (NCM) battery, while the verification of the lithium iron phosphate (LFP) battery is lacking. Some scholars have made good progress in using data-driven methods for SOC estimation of LFP batteries,[46,47] while others show that the obstacle encountered in model methods can be overcome.[48] Therefore, the proposed method has the potential to be applied to the SOC estimation of LFP batteries.

Conclusions

Battery SOC is critical to the safe and reliable operation of lithium-ion batteries. This study proposes an adaptive combined method for battery SOC estimation considering battery aging status. First, an LSTM model is constructed to characterize the battery's dynamic performance and predict the terminal voltage. Then, the UKF algorithm is employed to perform SOC estimation through the feedback of terminal voltage prediction. To improve the estimation performance under different aging statuses, a PID controller is employed to correct the influence of battery capacity fading. The voltage prediction results show that the two-layer LSTM model with a data length of 500 achieves preferable accuracy and robustness to the measurement uncertainties, in which the voltage prediction RMSE is within 0.002 V. The verification results of the proposed SOC estimation show that the combined method conducts higher accuracy under dynamic conditions and ascendant adaptability to noise and initial SOC deviation. Moreover, during battery aging, the proposed adaptive method shows overwhelming superiority over the traditional direct method. At different battery aging cycles, the SOC estimation error is reduced by 0.69% compared to the traditional direct method at least. These results prove the efficacy of the proposed method and depict a promising prospect in the future.

Acknowledgements

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] F. Salek, S. Resalati, M. Babaie, P. Henshall, D. Morrey, L. Yao, *Batteries* **2024**, *10*, 79.
[2] J. Cai, X. Wei, X. Wang, J. Zhu, B. Jiang, Z. Tao, M. Tian, H. Dai, *Appl. Energy* **2025**, *377*, 124438.
[3] H. Wu, S. Chen, Y. Hong, C. Xu, Y. Zheng, C. Jin, K. Chen, Y. He, X. Feng, X. Wei, H. Dai, *J. Energy Chem.* **2024**, *91*, 59–72.
[4] K. S. Ng, C.-S. Moo, Y.-P. Chen, Y.-C. Hsieh, *Appl. Energy* **2009**, *86*, 1506–1511.
[5] X. Wang, J. Zhu, X. Wei, D. Wang, W. Xu, Y. Jin, H. Dai, *Energy Storage Mater.* **2024**, *65*, 103160.
[6] I. Snihir, W. Rey, E. Verbitskiy, A. Belfadhel-Ayeb, P. H. L. Notten, *J. Power Sources* **2006**, *159*, 1484–1487.
[7] K. Liu, K. Li, Q. Peng, C. Zhang, *Front. Mech. Eng.* **2019**, *14*, 47–64.
[8] L. X. Wu, K. Liu, H. Pang, J. M. Jin, *Energies* **2021**, *14*, 5265.
[9] S. Y. Wu, W. J. Pan, M. T. Zhu, *J. Electrochem. Soc.* **2022**, *169*, 090516.
[10] Z. Ren, C. Q. Du, *IFAC-PapersOnLine* **2022**, *55*, 197–202.
[11] L. P. Chen, X. B. Wu, A. M. Lopes, X. Li, P. H. Li, R. C. Wu, *Commun. Nonlinear Sci. Numer. Simul.* **2023**, *125*, 107365.
[12] Z. Cui, J. Dai, J. Sun, D. Li, L. Wang, K. Wang, A. M. B. Pereira, *Math. Probl. Eng.* **2022**, *2022*, 1–11.
[13] T. Ouyang, P. Xu, J. Chen, Z. Su, G. Huang, N. Chen, *Energy* **2021**, *226*, 120348.
[14] L. Zheng, J. Zhu, G. Wang, D. D.-C. Lu, T. He, *Energy* **2018**, *158*, 1028–1037.
[15] W. Qian, W. Li, X. Guo, H. Wang, *Energy* **2024**, *292*, 130585.
[16] X. Chen, W. Shen, M. Dai, Z. Cao, J. Jin, A. Kapoor, *IEEE Trans. Veh. Technol.* **2016**, *65*, 1936–1947.
[17] Z. Cui, W. Hu, G. Zhang, Z. Zhang, Z. Chen, *Energy Rep.* **2022**, *8*, 81–87.
[18] Q. Lin, X. Li, B. Tu, J. Cao, M. Zhang, J. Xiang, *Sensors* **2023**, *23*, 467.
[19] F. Maletić, J. Deur, I. Erceg, *IEEE Trans. Control Syst. Technol.* **2023**, *31*, 692–706.
[20] P. Shrivastava, T. K. Soon, M. Y. I. B. Idris, S. Mekhilef, *Renew. Sustain. Energy Rev.* **2019**, *113*, 109233.
[21] H. Z. You, B. Jiang, J. G. Zhu, X. Y. Wang, G. Y. Shi, G. S. Han, X. Z. Wei, H. F. Dai, *J. Power Sources* **2023**, *564*, 232892.
[22] D. D. Qiao, X. Z. Wei, B. Jiang, W. J. Fan, H. Gong, X. Lai, Y. J. Zheng, H. F. Dai, *IEEE Trans. Ind. Inf.* **2024**, *20*, 6751–6761.
[23] J. C. Á. Antón, P. J. G. Nieto, C. B. Viejo, J. A. V. Vilán, *IEEE Trans. Power Electron.* **2013**, *28*, 5919–5926.
[24] T. Hansen, C. J. Wang, *J. Power Sources* **2005**, *141*, 351–358.
[25] B. Jiang, Y. L. Zhu, J. G. Zhu, X. Z. Wei, H. F. Dai, *Energy* **2023**, *263*, 125802.
[26] Z. Deng, X. Hu, X. Lin, Y. Che, L. Xu, W. Guo, *Energy* **2020**, *205*, 118000.
[27] C. Qian, B. Xu, Q. Xia, Y. Ren, D. Yang, Z. Wang, *Materials (Basel)* **2022**, *15*, 5933.
[28] D. D. Qiao, X. Z. Wei, B. Jiang, W. J. Fan, X. Lai, Y. J. Zheng, H. F. Dai, *IEEE Trans. Ind. Electron.* **2024**, 13201–13210.
[29] S. El Fallah, J. Kharbach, Z. Hammouch, A. Rezzouk, M. Ouazzani Jamil, *J. Storage Mater.* **2023**, *62*, 106904.
[30] K. M. K. Sundareswaran, P. S. R. Nayak, S. P. Simon, *IEEE Trans. Veh. Technol.* **2023**, *72*, 7328–7337.
[31] S. Tong, J. H. Lacap, J. W. Park, *J. Storage Mater.* **2016**, *7*, 236–243.
[32] Y. Zhao, Y. Li, Y. Cao, L. Jiang, J. Wan, C. Rehtanz, *IEEE Trans. Veh. Technol.* **2023**, *72*, 11469–11481.
[33] Q. Wang, M. Ye, M. Wei, G. Lian, C. Wu, *Energy Rep.* **2021**, *7*, 7323–7332.
[34] J. Chen, Y. Zhang, W. Li, W. Cheng, Q. Zhu, *J. Storage Mater.* **2022**, *55*, 105396.
[35] L. Dang, J. Yang, M. Liu, B. Chen, *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–15.
[36] S. M. Shahriar, E. A. Bhuiyan, M. Nahiduzzaman, M. Ahsan, J. Haider, *Energies* **2022**, *15*, 8003.
[37] B. Jiang, S. Y. Tao, X. Y. Wang, J. G. Zhu, X. Z. Wei, H. F. Dai, *Energy* **2023**, *278*, 127890.
[38] H. Yuan, J. Liu, Y. Zhou, H. Pei, *Energies* **2023**, *16*, 2155.

[39] D.-W. Chung, J.-H. Ko, K.-Y. Yoon, *J. Electr. Eng. Technol.* **2022**, *17*, 1931–1945.

[40] S. Tao, B. Jiang, X. Wei, H. Dai, *Energies* **2023**, *16*, 2008.

[41] W. Sun, Y. C. Qiu, L. Sun, Q. S. Hua, *Int. J. Energy Res.* **2020**, *44*, 10307–10319.

[42] X. Yun, X. Zhang, X. Fan, *Electr. Eng.* **2023**, *105*, 3307–3318.

[43] Y. Yang, L. Zhao, Q. Yu, S. Liu, G. Zhou, W. Shen, *J. Energy Storage* **2023**, *70*, 108037.

[44] X. Shu, G. Li, Y. J. Zhang, S. Q. Shen, Z. Chen, Y. G. Liu, *IEEE Trans. Transp. Electrif.* **2021**, *7*, 1271–1284.

[45] J. Chen, Y. Zhang, J. Wu, W. Cheng, Q. Zhu, *Energy* **2023**, *262*, 125375.

[46] U. Primadusi, A. I. Cahyadi, O. Wahyunggoro, *Adv. Sci. Technol. Soc.* **2016**, *1755*, 090010.

[47] M. M. Liu, J. Xu, Y. H. Jiang, X. S. Mei, *Energy* **2023**, *274*, 127407.

[48] R. Xiong, Y. Duan, K. Zhang, D. Lin, J. Tian, C. Chen, *Appl. Energy* **2023**, *349*, 121581.