Hindawi
International Journal of Energy Research
Volume 2023, Article ID 8569161, 13 pages
https://doi.org/10.1155/2023/8569161

Research Article

State-of-Health Estimation for Lithium-Ion Battery Based on an Attention-Based CNN-GRU Model with Reconstructed Feature Series

Baolei Liu, Jinli Xu, and Wei Xia

School of Mechanical and Electrical Engineering, Wuhan University of Technology, Wuhan, China

Correspondence should be addressed to Baolei Liu; lbl9126@163.com

Received 30 November 2022; Revised 10 December 2022; Accepted 12 December 2022; Published 6 February 2023

Academic Editor: Suresh Kannan Balasingam

Health monitoring is an essential task for lithium battery systems. Recently, with the development of data-driven methods, deep learning has been successfully deployed for state-of-health (SOH) estimation. However, existing models trained using raw samples directly usually contain noise due to sensor errors. To enhance the performance of SOH prediction, short-term segments are extracted for SOH estimation based on reasonable SOC ranges. To address the measuring error that exists in the voltage and temperature samples, the reconstructed feature series (RFSs) is designed to restrain the signals' noise. Then, a CNN-GRU network with attention mechanism is proposed to achieve SOH estimation based on short-term RFSs' samples. To further enhance accuracy, a parallel structure is designed to fuse the feature information from both streams, raw samples, and RFSs in a reasonable manner. The performance of our proposed method is validated over a wide range of experiments on the Oxford battery degradation dataset, where the RMSE and MAE averaged 0.582% and 0.524%, respectively, demonstrating its forward estimation performance.

1. Introduction

New energy vehicles have received particular attention due to their superior environmental performance. Lithium-ion batteries have found applications in this field due to their flexible operation, fast response, and high-cost performance [1, 2]. Operation practice manifests that accuracy and dependable health monitoring technique has become an important part of lithium-ion battery application system [3, 4], and the principal mission is to predict the state of health (SOH) of lithium-ion batteries, giving important advice for battery management system (BMS) maintenance. Currently, various approaches to battery SOH prediction have been investigated, which fall into two categories: model-based and data-driven [5, 6].

For the model-based methods, electrochemical models (EMs) [7] or equivalent circuit models (ECMs) [8] are employed combined with the analysis of complicated electrochemical reaction inside the cell. In Reference [9], the summary and comparison of battery modeling techniques were given, providing ideas for the follow-up research. Liu et al. systematically returned EMs and discussed trends and challenges for further research [10]. Zhang et al. developed a ternary proportional-integral estimator for capacity estimation using EMs [11]. Yu et al. proposed a joint prediction method for state of health and state of charge (SOC) estimation based on the second-order ECM [12]. Wei et al. developed a joint estimation of SOC and SOH on multiple timescales using Kalman filter and recursive least squares, respectively, which is convenient for practical application on the premise of ensuring accuracy [13]. Kalman filter [14], particle filter [15], and other filtering algorithms are generally deployed to achieve SOH estimation. Model-based methods have achieved acceptable accuracy on battery health statue estimation. Nonetheless, the estimation accuracy of model-based approaches depends primarily on the model structure and prior knowledge. The limited representation of electrochemistry reaction or complexity of partial differential operations both restricted the model-based approaches' application [16].

Data-driven approaches are one type of artificial intelligence (AI) techniques that have been extensively investigated for battery state prediction and extended battery life strategy [17, 18]. The selection of health features (HFs) and the structure of the model are two key parts that affect the performance of SOH estimation [19]. For HFs' extraction, the commonly utilize trick is the incremental capacity (IC) analysis and thermal voltammetry method. Stroe and Schaltz deployed spikes, spike voltage, and spike acreage of IC to describe and estimate the aging process [20]. To address the noise in the differential operation, Li et al. smoothed the 3.8 V-4.1 V IC curve with a Gaussian filter [21]. While in some cases, valuable information may be eliminated by such filtering [22, 23]. To handle the condition that operating voltage outside the IC peak range, Wei et al. designed the morphological IC features and voltage entropy information combined with artificial neural network for SOH estimation [24]. In addition, voltage curves of charging phases have been widely used for feature extraction. Yang et al. selected the length, slope, and area of the charging phase as features for SOH estimation [25]. However, this causes time-consuming problems and is inflexible. From the present research, BP neural network [26], support vector regression (SVR) [27], and Gaussian process regression (GPR) [28] have been employed for battery capacity estimation combined with meaningful HFs.

Recently, deep learning techniques have attracted increasing attention to reinforce the adaptability of the aforementioned processes and reduce the influence of artificially designed features. Deep neural networks consist of multiple layers which can capture the complex nonlinear knowledge related to degradation data from initial data, such as terminal voltage, current, and surface temperature [29, 30]. Convolutional neural network (CNN), recurrent neural network (RNN), and stacked denoising autoencoder have been applied to battery SOH estimation [31–33]. As a variant of CNN, Hong et al. [34] established a novel HI to predict the state of health (SOH) of battery based on dilated CNN. Ren et al. [35] designed a parallel SOH prediction framework including CNN and long short-term memory (LSTM), and autoencoder is introduced to extract more valuable information from original signal. Li et al. [36] developed a deep LSTM network applying partial charging profiles, and estimation error fell to 0.76% for the online battery SOH prediction. To improve the model generalization, Che et al. joint the transfer learning technique to gated recurrent unit (GRU) to estimate the remaining useful life using partial voltage curves [37].

However, in the existing literature, it is common to leverage the entire charging curve for model training, which may not be available from battery operation. Some articles realized the SOH estimation using the short-term charging curves [38], but the extraction of the partial curve is usually based on the statistical characteristics of the curves [39], without considering SOC during the charging process, which inevitable loss some actual meaning. Meanwhile, these approaches make direct use of raw data, which often contains measurement noise. This stochastic noise may influence the accuracy of the SOH estimation for these approaches. To address the above issues, in this paper, we capture the charging curve based on the SOC of the charging phase. Then, the reconstructed feature series is designed to suppress the noise in the raw signal. In addition, an attention-based CNN-GRU model is proposed to enhance SOH estimation. The main contributions of this paper can be summarized as follows.

(i) A short-term segment is extracted based on the reasonable SOC operating range. RFSs are subsequently proposed to suppress the noise contained in the raw voltage and temperature signals

(ii) An attention-based CNN-GRU model is proposed to achieve battery SOH estimation. A parallel framework is further designed to fully exploit the information from raw samples and RFSs

(iii) Extensive experiments demonstrate that the proposed method can achieve SOH prediction with limited data, improving estimation robustness and accuracy

The reminder of this paper is arranged as follows: Section 2 elaborates the procedure of short-term curve extraction and RFSs' construction. Section 3 gives the structure of our attention-based CNN-GRU model. The comprehensive comparison results of the proposed approach are discussed in Section 4. Finally, the conclusions are provided in Section 5.

2. Degeneration Data Analysis

2.1. Data Description. In this paper, the Oxford battery degradation data [40] are utilized for battery SOH prediction, and the detailed test scheme is as follows.

Eight Kokam cells from Oxford battery degradation dataset are charged with constant current- (CC-) constant voltage (CV) charging scheme, and the urban Artemis driving condition test is then deployed for the cell every 100 cycles. The nominal capacity of the battery is 750 mAh, and the test temperature is set to 40°C. Test data including current, voltage, and surface temperature are sampled every second. The detailed test procedure and the degradation results are depicted in Figure 1.

2.2. Short-Term Profile Extraction. The choice of training samples inevitably affects the accuracy of the SOH prediction. Due to its relatively fixed charging protocol, the voltage and temperature series of the CC charging process will be analyzed and fed into the proposed model to implement the SOH estimates in this paper. In existing studies, the entire charging curve is generally trained by the model, which is not possible in real-world vehicle scenarios. Hence, we would like to estimate the SOH in combination with the short-term charging curve. The feasibility of deploying such an approach is clarified in detail below.

The charging operation in real scenarios usually depends on the current remaining SOC of the electric vehicles (EVs). Thus, it is easy to think of using SOC to segment the sample

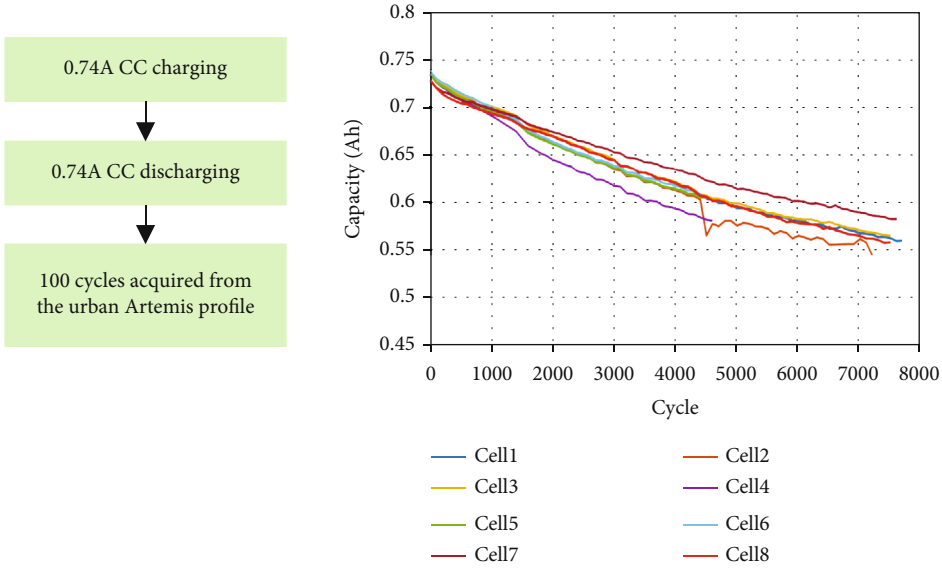


Figure 1: Testing procedure and degradation results for the battery.

during the charging phase. Reference [41] introduced that 20% to 80% SOC is the common use interval for battery charging in a research report on EVs. The dependence of the terminal voltage on SOC of cell1 can be seen in Figure 2. It is clear that the relationship between SOC and voltage changes periodically with the number of cycles. During each cycle, the voltage of the cell increases as the state of charge increases. This elaborates that the terminal voltage is an extrinsic feature of the charging state and can be considered as an alternative indicator of SOC.

Distribution statistics of SOC in degeneration process of eight cells are depicted in Figure 2(b). It can be concluded that the terminal voltage is typically around 3.75 V-4.01 V when the SOC ranges from 20% to 78%, in agreement with the real operation process of EVs. We want to use partial curves instead of the entire charging profile to predict the health status of the cell, which is more flexible and practical for applications in cloud battery management systems. With this in mind, from 3.75 V, six segments equidistantly spaced 0.05 V apart are selected. In addition, the corresponding temperature samples are also chosen as input to the model. To rapidly predict the battery SOH, 300 samples were selected from each partial voltage and temperature profile. The effect of sample size will be discussed in Section 4.1. The detailed process of extracting the six short-term profiles is shown in Figure 3.

2.3. Reconstructed Feature Series. As shown in Figure 3, due to the measuring error, the voltage and temperature are both fluctuations, especially for the temperature samples in Figure 3(b), which may have some negative impact on SOH estimation under direct utilization of raw noisy signals. In this section, a series of reconstructed feature series (RFSs) is designed to address the noise existing in the raw measured samples. Two time-domain statistical features, which have been proven to be feasible in [13], are calculated from the partial voltage and temperature curves: root mean square (RMS) and variance. Given a specified time interval T, the mean value of the sample flow from voltage and temperature is calculated as follows.

$$\begin{aligned} H_{\text{rms}}^{V} &= \sqrt{\frac{1}{T}\sum_{t=1}^{T}|V_t|^2}, \\ H_{\text{rms}}^{T} &= \sqrt{\frac{1}{T}\sum_{t=1}^{T}|T_t|^2}, \\ H_{\text{var}}^{V} &= \frac{1}{T-1}\sum_{t=1}^{T}\left|V_t - H_{\text{mean}}^{V}\right|, \\ H_{\text{var}}^{T} &= \frac{1}{T-1}\sum_{t=1}^{T}\left|T_t - H_{\text{mean}}^{T}\right|. \end{aligned} \tag{1}$$

The challenge is to apply these features to describe the sequential information of the measured signal. Thus, we construct RFSs by computing these features with different time intervals. As shown in Figure 4, given a sample of input data of length T, the variance and root mean square are computed at different time-lengths t, increasing from 1 to T, so that the length remains the same as the original one. Then, the computation is concatenated as follows.

$$\begin{aligned} H_{\text{RFS}}\left(X^{C}\right) = \text{concat}\big(&H_{\text{mse}}\left(X^{S},1\right),\left(H_{\text{var}}\left(X^{S},1\right),\right. \\ &\left(H_{\text{mse}}\left(X^{S},2\right),\left(H_{\text{var}}\left(X^{S},2\right),\cdots,\right.\right. \\ &\left.\left.\left(H_{\text{mse}}\left(X^{S},T\right),\left(H_{\text{var}}\left(X^{S},T\right)\right)\right)\right)\right)\big). \end{aligned} \tag{2}$$

Reconstructed samples from voltage and temperature sensors are concatenated as operations according to

$$\tilde{X} = \text{concat}\left(H_{\text{RFS}}\left(X^{V}\right), H_{\text{RFS}}\left(X^{T}\right)\right). \tag{3}$$

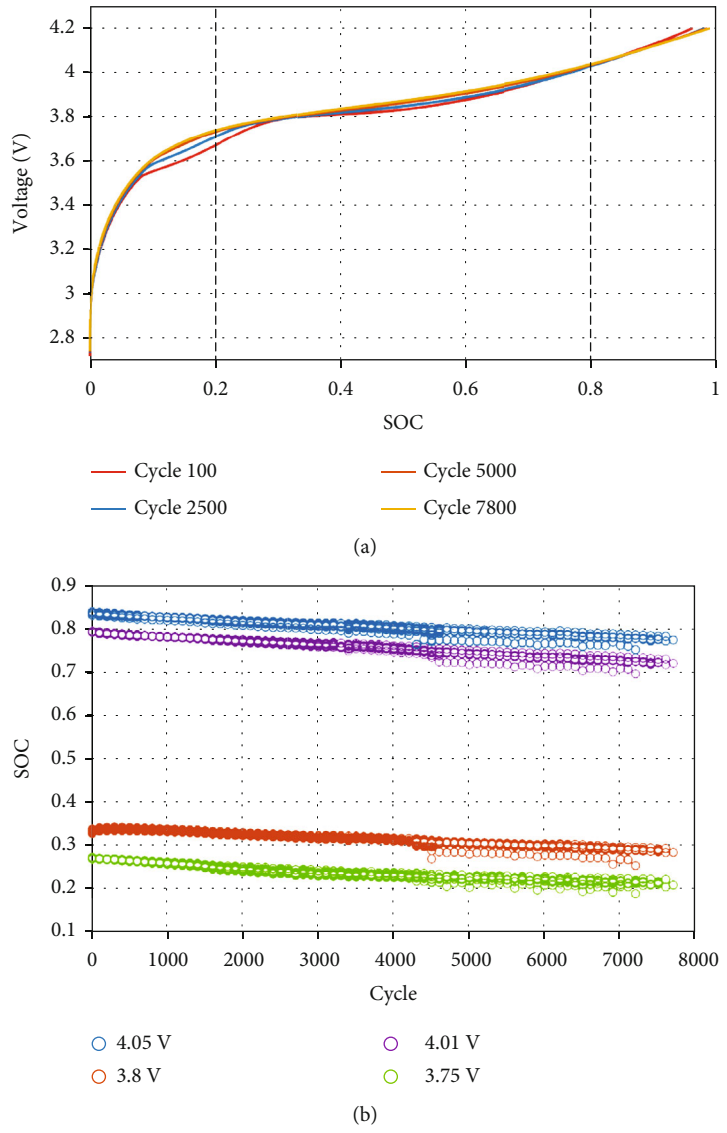


Figure 2: The relationship between terminal voltage and SOC. (a) Variation of voltage with SOC. (b) Distribution statistics of SOC in aging process.

The RFSs, taken at 100, 1000, and 2000 cycles of the data, are illustrated in Figure 5. The RMS values are smoother than those of the original sample and maintain a similar trend to the original signal as the number of cycle increases. The same can be found for the variance values, where the trend changes associated with aging are more pronounced compared to the original signal, especially for the temperature signal.

With our proposed RFSs, the noise in the minima data is constrained. At the same time, sequence acquisition is convenient for online operation, and aging information can be easily extracted. This enhances the accuracy of the SOH estimate.

3. Model Structure

3.1. CNN-GRU. The CNN-GRU network, which is one of the main units of the proposed model, can extract aging information from spatio-temporal dimensionality. The CNN-GRU module consists of a one-dimensional convolutional neural network (1D-CNN) and two GRU layers, respectively, which can be seen in Figure 6. The input data are short-term RFSs' samples with the shape of $M \times 300 \times 4$, where M is the number of total cycles. After that, the spatio-temporal features of the measured data can be extracted to better describe the aging process of the battery. The output of the CNN-GRU network can be calculated as follows.

$$\begin{aligned} C_t &= \sigma(W_s * X_t + b_s), \\ z_t &= \sigma(W_z i_t + U_z h_{t-1}), \\ r_t &= \sigma(W_r i_t + U_r h_{t-1}), \\ \tilde{h}_t &= \tanh\left(W_h i_t + r_t \circ U_h h_{t-1} + b_h\right), \\ h_t &= (1 - z_t) \circ h_{t-1} + z_t \circ \tilde{h}_t, \end{aligned} \tag{4}$$

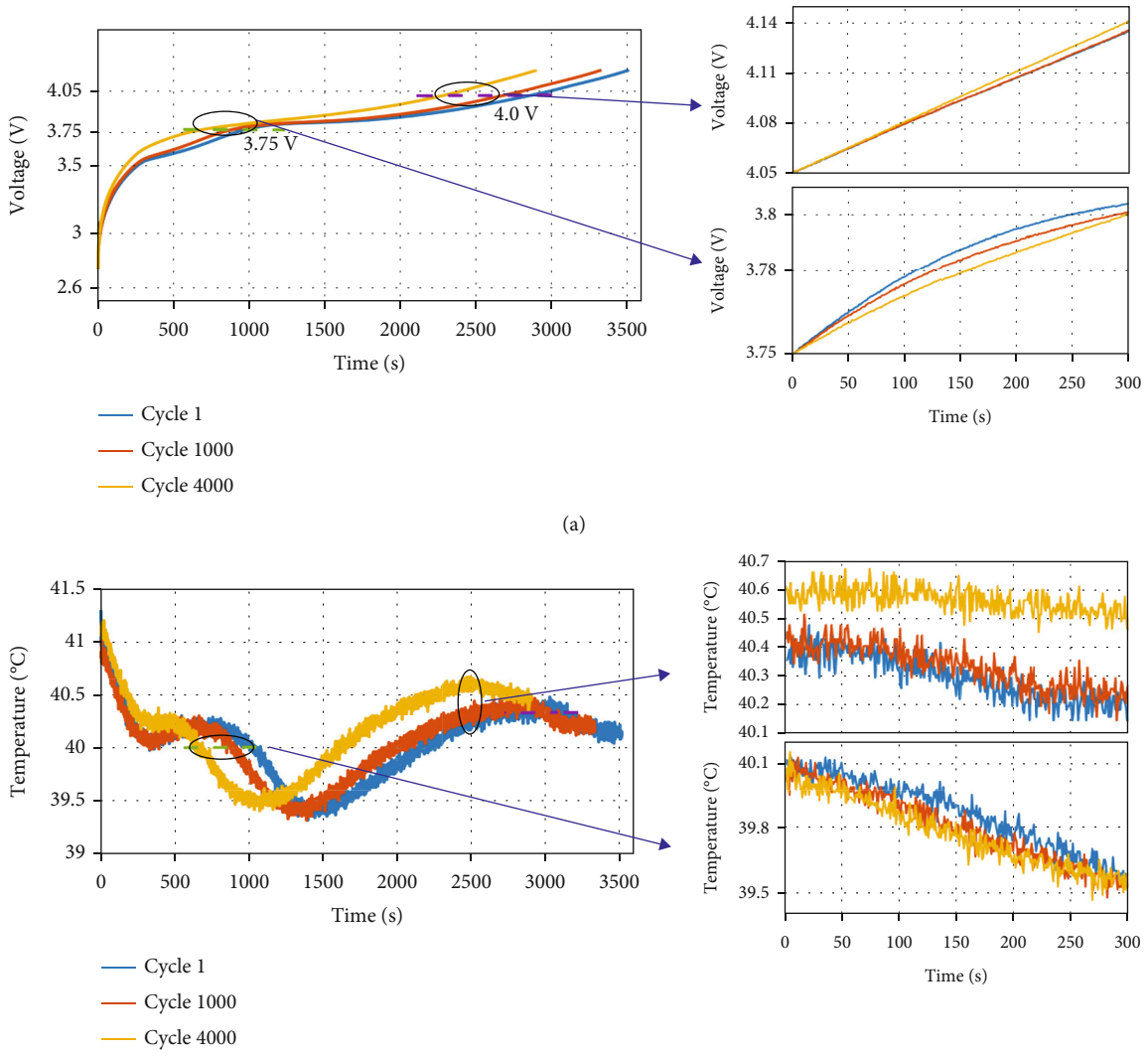


Figure 3: Procedure for extraction of the six short-term curves. (a) Voltage curves. (b) Temperature curves.

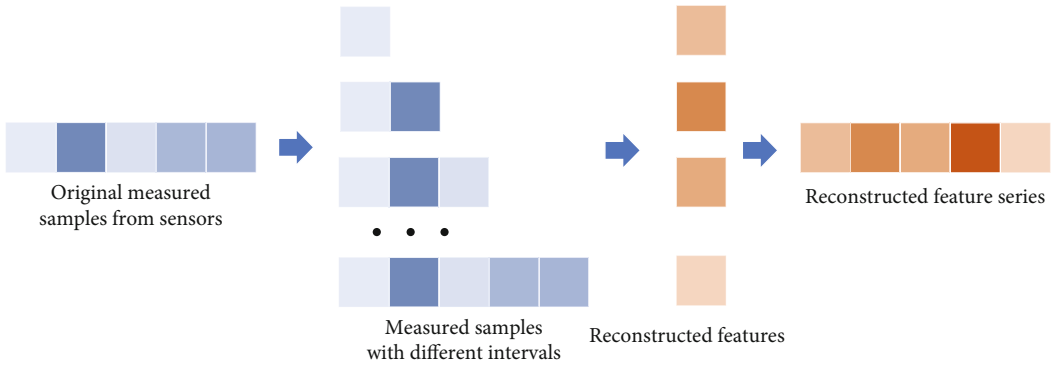


Figure 4: The construction process of RFSs.

where X_t is the RFSs designed above, i_t is the output of the CNN layer, W_s and b_s are the weight and bias of the CNN model, $*$ is the convolution operator, z_t and r_t are the output of update gate and reset gate at the previous moment t, σ denotes the sigmoid activate function, W and U represent the weight and parameter matrix of GRU unit, respectively, h_t denotes the state at time t, and $\tilde{h}_t$ is the candidate state, and $\circ$ is the Hadamard product.

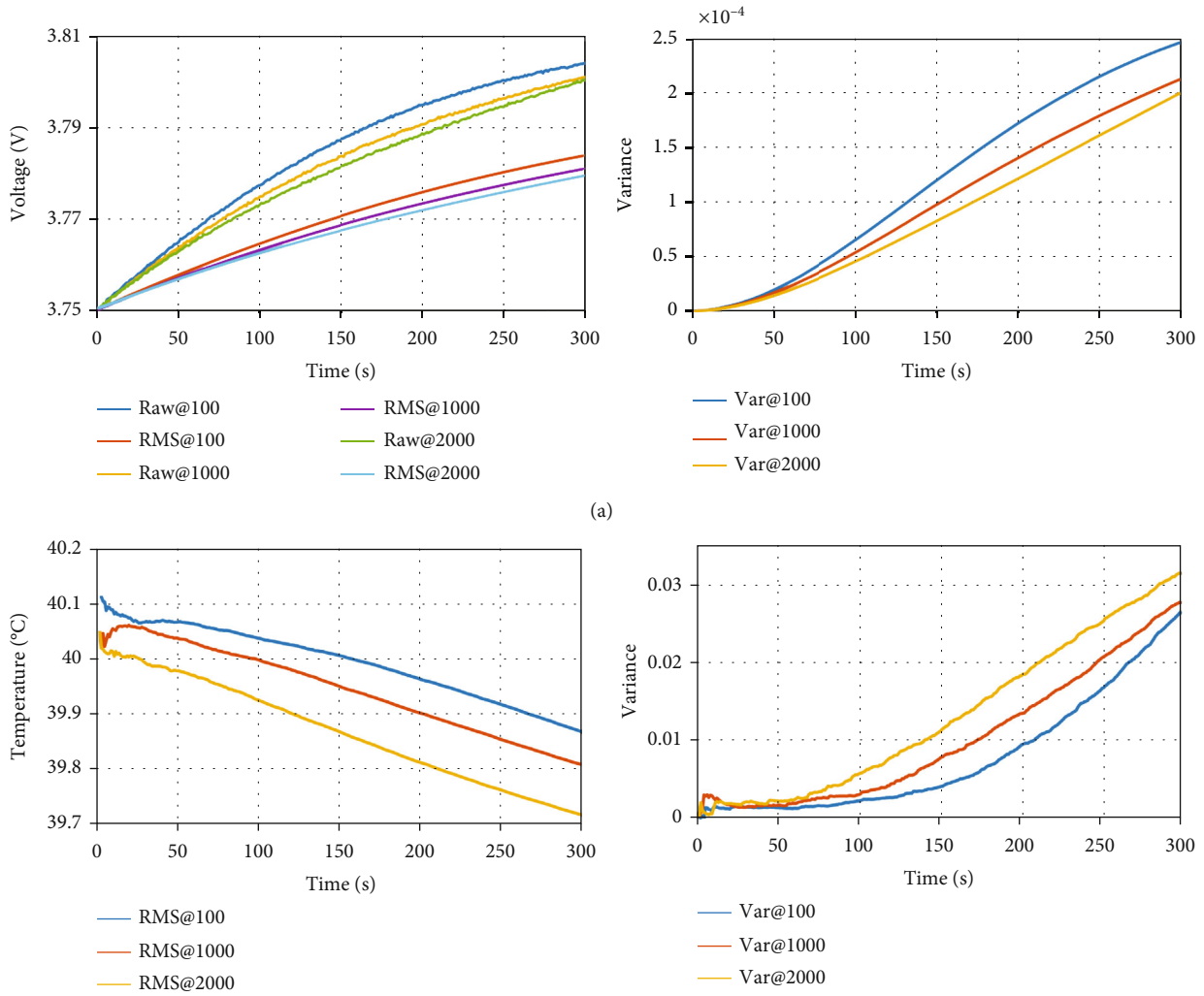


Figure 5: Reconstructed feature series of voltage and temperature: (a) voltage; (b) temperature.

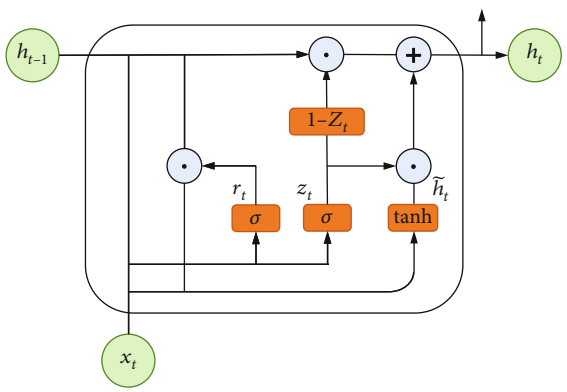


Figure 6: Structure of GRU unit.

3.2. Attention Mechanism. Attention models have been designed to extract the intrinsic characteristics of samples and enhance the speed and performance of feature procedures. It is well known that the information provided by the voltage and temperature at different times may have different effects on the SOH prediction. To overcome shortcomings of the RNN model in highlighting important information, the self-attention (SA) process is deployed to focus on the degraded features that are profitable to health monitoring and decreasing less beneficial ones. The integration process of the attention mechanism and the CNN-GRU model is illustrated in Figure 7.

As shown in Figure 7, after obtaining the attention weight of each time step, the final output H_t^o of the attention-based CNN-GRU is calculated as weighted sum of the output of CNN-GRU h_t as follows:

$$H_t^o = \sum_{k=1}^{n+1} \alpha_k h_{t-n+1}, \tag{5}$$

where $n+1$ is the size of flows; α_k is the attention value at time t-(k-1). The attention values α_k are expressed as

$$\alpha_k = \frac{\exp(e_k)}{\sum_{k=1}^{m+1} \exp(e_k)}. \tag{6}$$

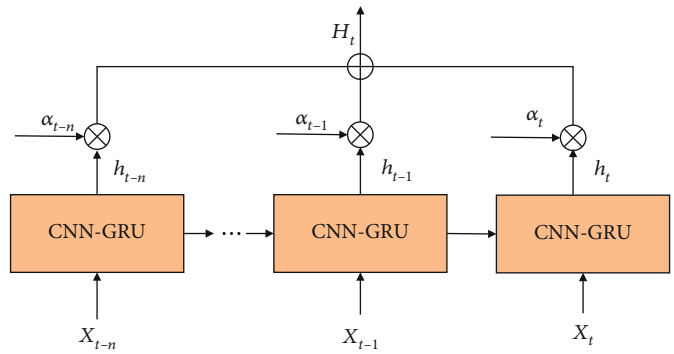


Figure 7: CNN-GRU networks with attention mechanism.

The scores $e = (e_1, e_2, e_3 \cdots e_{n+1})$ denote the importance of each position in the time series, and it can be computed as

$$e_t = v_s \tanh\left(W_s C_t + U_s h_t\right), \tag{7}$$

where e_t is the score of each time period; v_s, W_s, and U_s are the hyperparameters to be learned.

4. Experiments

The full procedure of the developed method on SOH estimation is illustrated in Figure 8. First, six short-term data are explored based on a reasonable range of SOC operations. The RFSs are obtained by computing the root mean square and variance in each voltage range. To train the model, the RFSs and the corresponding real SOH of the first four datasets, namely, cell 1 to cell 4, are deployed for model training. Cross validation is applied to this process. Of the four subsamples, three are applied as the training set, while the single remaining one is used as the validation set for hyperparameter tuning with a Bayesian algorithm. Then, the data from cells 5-cell 6 are chosen as the test set to evaluate the performance of the proposed model. The effects of the starting voltage and sample size of RFSs are first discussed in the validation procedure. Four other classical models are compared with our proposed method. Finally, different inputs involving raw signals and RFSs are applied to manifest the effectiveness of RFSs on accuracy improvement. The proposed method is performed in the TensorFlow platform. The experiments are implemented using a workstation with an Intel Core i7-8700 CPU and an Nvidia GeForce. The root mean square error (RSME) and mean absolute error (MAE) are deployed to verify the precision and robustness of the model, and the two metrics can be calculated as follows.

$$\mathrm{RMSE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(\hat{y}_i - y_i)}, \tag{8}$$

$$\mathrm{MAE} = \frac{1}{L}\sum_{i=1}^{L}|\hat{y}_i - y_i|, \tag{9}$$

where L is the length of forecast data and $\hat{y}_i$ and y_i represent the estimated SOH and real SOH, respectively.

4.1. The Influence of Starting Point and Sample Size. From the above data acquisition details, the starting point of the voltage has a huge impact on the data characteristics of the selected segment, which will further affect the estimation results. The average RMSE and MAE of the proposed model at each initial voltage of 3.75 V, 3.8 V, …, and 4.0 V are shown in Table 1.

When the starting point of the segment is set as 3.9 V, both RMSE and MAE obtain a maximum, and the reason for this result is the flatten tendency of the voltage and temperature curves in (3.9 V-3.95 V), which can be clearly observed in Figure 5. For 4.0 V, the estimation error is smaller than for the segment starting at 3.9 V but higher than for the sample taken from 3.75 to 3.85 V. The smallest values are obtained at 3.8 V with 0.782% for the RMSE and 0.670% for the MAE, respectively. Consequently, the proposed AB CNN-CRU method is able to accurately predict the SOH over six different short-term segments of reasonable SOC range with superior convenience and speed, requiring only 300 samples.

As shown in Section 2.3, the segment length should first be chosen to structure the RFSs for model training and testing. The more samples are employed to estimate the cell SOH, the more accurate values are likely to be obtained in practice. However, the huge amount of data fed into the model tends to make it more computationally burdensome. Considering the charging scheme for the Oxford dataset, six segments of different lengths are obtained at intervals of 80, which implies an SOC of approximately 2% for adjacent segments, to evaluate their impact on health status prediction.

From Table 2, the mean RMSE of the prediction results using the proposed method is given for six different lengths. The value of RMSE decreases as the segment length increases. The RMSE is 1.751% for a length of 60 and gradually decreases to 0.726% for a length of 460. The AB CNN-GRU model can achieve RMSE of 0.782% when the sample size is chosen to be 300. On top of this, as the sample length continues to increase, the improvement in model prediction accuracy is not significantly enhanced, and the computation rate gradually decreases. Considering the trade-off between

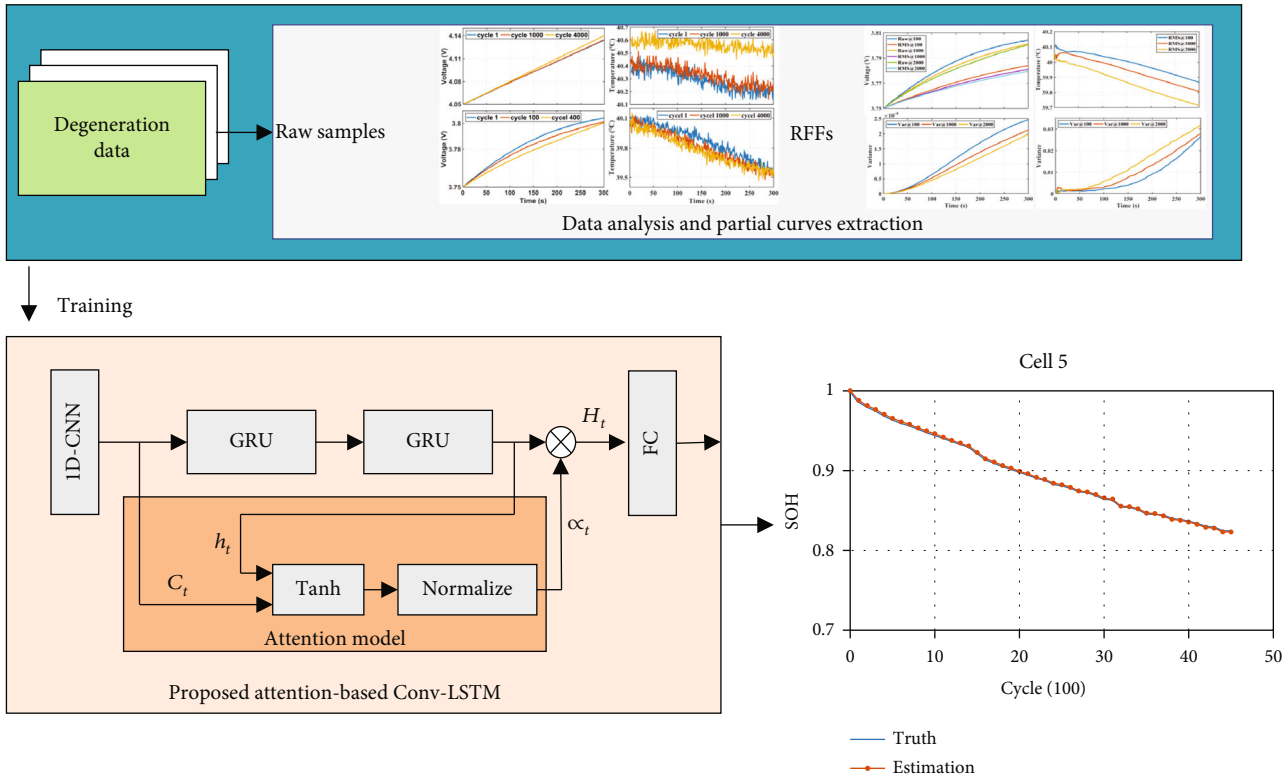


Figure 8: The procedure of the proposed model for cell SOH estimation.

Table 1: Average RMSE and MAE based on different starting point.

Initial voltage	RMSE (%)	MAE (%)
3.75 V	0.808	0.732
3.8 V	0.782	0.669
3.85 V	0.890	0.843
3.9 V	1.285	1.181
3.95 V	1.205	1.057
4.0 V	1.061	0.935

Table 2: Average RMSE and MAE based on different segment length.

Segment length	60	140	220	300	380	460
RMSE (%)	1.751	1.315	0.946	0.782	0.729	0.726

accuracy and computation, 300 samples are already sufficiently informative for cell SOH estimation and will be set as the final selection for follow-up validation.

4.2. Comparison with Other Approaches. In this section, the proposed attention-based CNN-GRU (AB CNN-GRU) approach will be compared with other advanced models. The base models include SVR, GPR, GRU, and CNN-GRU. The GRU and CNN-GRU models have the same structure as the proposed model, with one CNN layer and two GRU units. The proposed RFSs at a starting voltage of 3.8 V are chosen to estimate the SOH. Then, the predicted values combined with the true SOH are injected into Equations (8) and (9) to evaluate the model performance. The RMSE and MAE values for the four validated cells based on the different approaches are shown in Figure 9 and Table 3.

From Figure 9, the RMSE of cell5 is 1.481% based on SVR model, 1.504% based on GPR, 1.165% based on GRU, and 1.172% based on CNN-GRU. While, the SOH estimation result of our proposed method is only 0.776%, which is the smallest error compared to other approaches. The same result occurs for cells 6-8, where the RMSE is all limited to 0.9%. As known in Table 3, the mean MAE of the four cells achieves a minimum value of 0.815%, which is a reduction of 50.21% and 49.32% compared to the two commonly deployed methods, SVR and GPR. The CNN-GRU method has the second-best SOH prediction, while our proposed method obtains better estimation performance, benefiting from the attention mechanism that focuses on features with respect to the degradation process. Consequently, it can be confirmed that our proposed AB CNN-GRU approach has the better prediction precision for battery health monitoring.

4.3. Comparison with the Raw Samples. In this paper, RFSs are designed to replace the raw data in selected segments to address the noise in voltage and temperature. In this section, the performance of the proposed RFSs on the SOH prediction is compared with data obtained from the raw

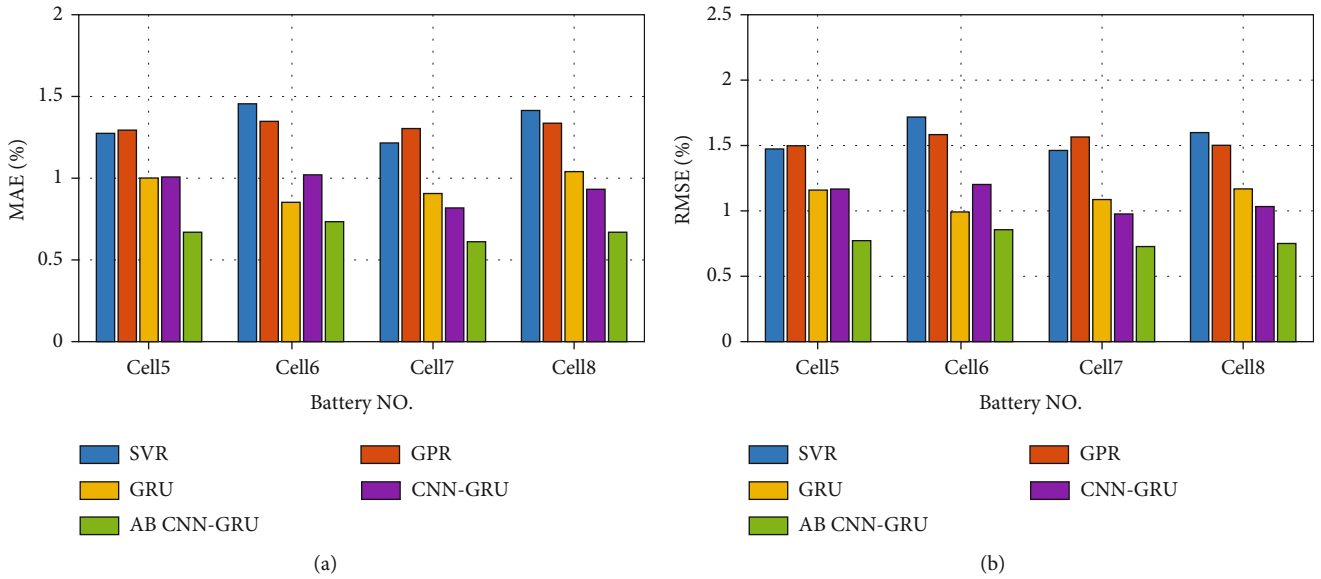


Figure 9: MAE and RMSE based on different approaches: (a) MAE results; (b) RMSE results.

Table 3: Average RMSE and MAE based on different methods.

Approach	SVR	GPR	GRU	CNN-GRU	AB CNN-GRU
RMSE (%)	1.567	1.541	1.107	1.103	0.782
MAE (%)	1.344	1.321	0.950	0.943	0.669

Table 4: RMSE and MAE based on different inputs.

Metrics	Data	cell5	cell6	cell7	cell8	Average
RMSE (%)	Raw samples	0.876	0.946	0.859	0.830	0.878
	RFSs	0.776	0.865	0.753	0.732	0.782
	Fusion	0.529	0.630	0.629	0.541	0.582
MAE (%)	Raw samples	0.753	0.880	0.704	0.730	0.767
	RFSs	0.665	0.735	0.67	0.608	0.669
	Fusion	0.446	0.590	0.586	0.475	0.524

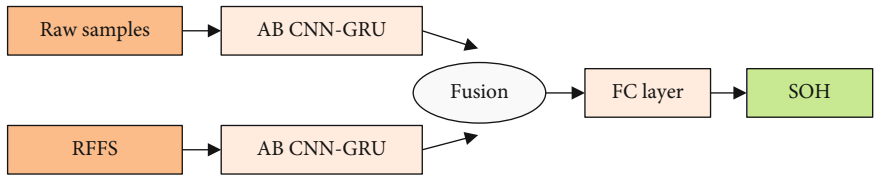
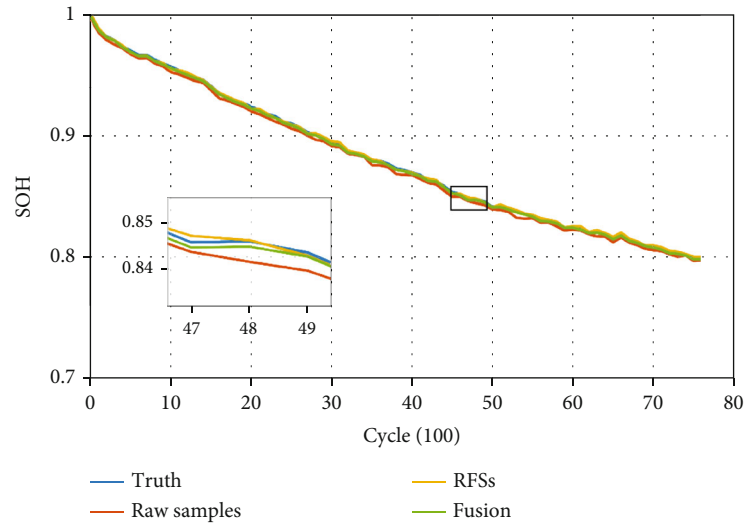


Figure 10: The parallel structure for the two stream integration.

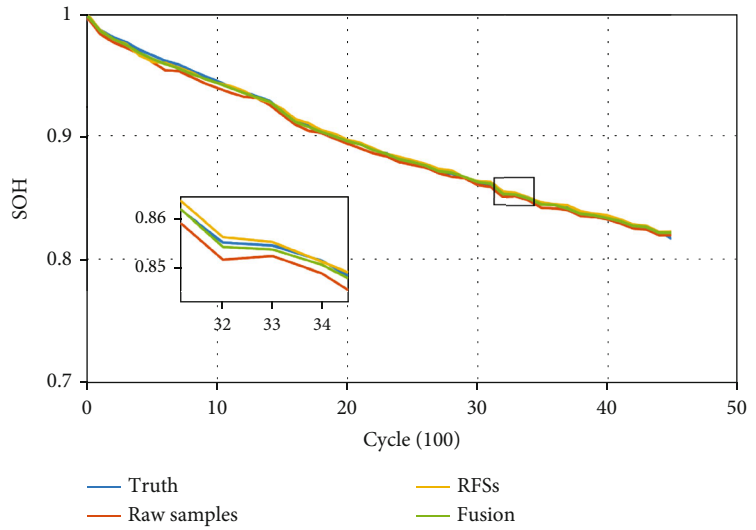
profiles. The results of the SOH estimation at 3.8 V for these two inputs are shown in Table 4.

As can be seen in Table 4, both RMSE and MAE with RFSs as input are improved compared to the original data for these four cells. Based on the proposed RFSs, the average RMSE and average MAE are 0.635% and 0.527%, respectively. It can be concluded that RFSs enhance the precision of the SOH estimation due to the noise reduction of the sources.

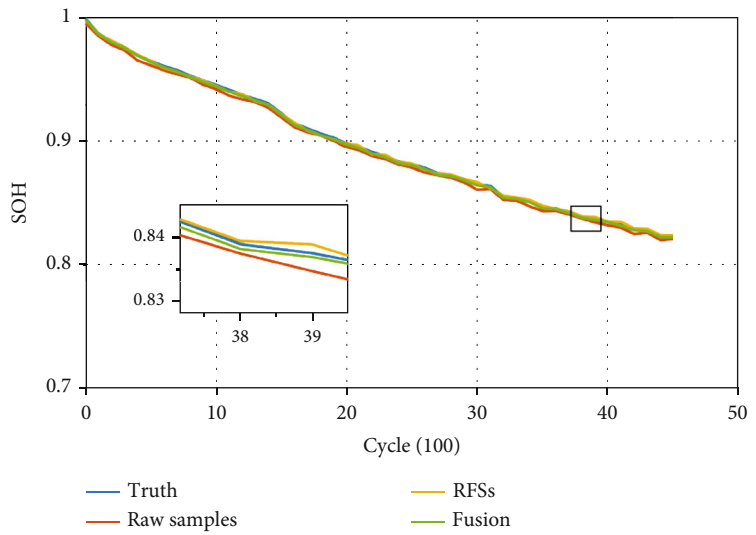
The accuracy of SOH estimation based on RFSs can only be improved to a limited extent in cases where the aging information is already obscured by noise. To fully exploit the features in the source data and RFSs, a parallel structure based on AB CNN-GRU is further designed to improve the battery SOH prediction results. The framework of the parallel network is presented in Figure 10. The uplink feeds the raw data to the proposed model, and the downlink deploys RFSs for information extraction. Equation (10) is used to



(a)

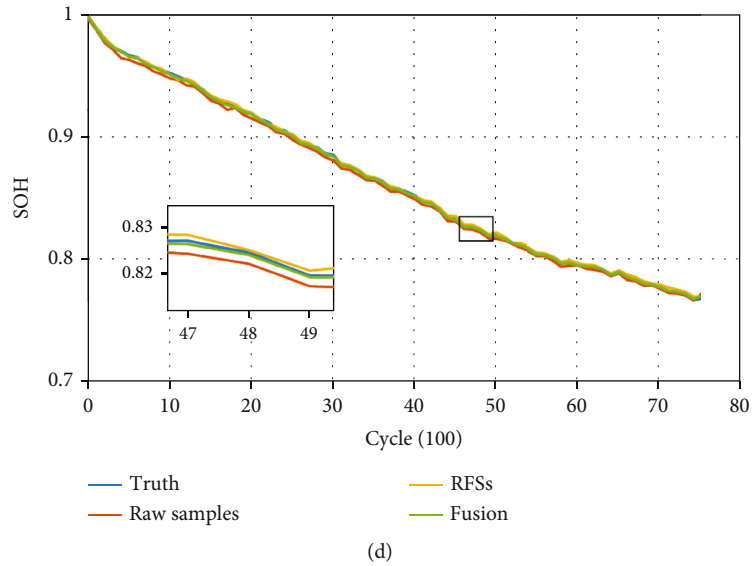


(b)



(c)

Figure 11: Continued.



(d)

Figure 11: The prediction results of four cells: (a) cell5; (b) cell6; (b) cell7; (d) cell8.

Table 5: RMSE and MAE based on different inputs.

Data	Training time (s)	Average prediction time (s)
Raw samples	42.86	0.38
RFSs	49.49	0.42
Fusion	59.56	0.46

obtain the fused features. The direct addition of two stream features does not improve the dimensionality of the data, which would be faster for SOH prediction. The parameter of the FC layer is 8.

$$F = O(X) + O(\tilde{X}), \tag{10}$$

where F is the fused feature of the two streams, O denotes the output of the proposed model, and X and $\tilde{X}$ are the input from the original curve and RFSs, respectively.

The results of the SOH estimation for the raw sample, the RFSs, and the fusions from cell 5 to cell 8 are shown in Figure 11. The estimation results of all three methods trace the trend of the true SOH, while the fusion method has the best performance for comparison. As shown in Table 4, RMSE and MAE achieve higher performance for SOH estimation based on the parallel ensemble approach. RMSE obtained reductions of 31.83%, 27.17%, 16.47%, and 26.09% for cell5-cell8 by fusing features from raw signals and RFSs. For MAE, the results for all four cells are constrained to within 0.6%. The RMSE and MAE average 0.582% and 0.524%, respectively, presenting a large improvement in the accuracy of SOH estimation. To quantitatively verify the computational burden, the training time and the average prediction time for the three types of samples described above are calculated. As shown in Table 5, the proposed model with raw samples achieves the shortest training time, and the fusion method has the longest training time of 59.56 s. While raw samples have a tenuous advantage in terms of computational burden, the model training process is usually performed offline using microcontrollers. For the prediction time, the difference between the three is negligible, which is acceptable considering the significant improvement in accuracy for the fusion method. As can be seen from the comparison results, the RFSs designed in this paper exhibit an advantage over the raw data due to the noisy filtering trick. Moreover, the fusion of raw samples and RFSs further enhances the performance of the model on SOH prediction, thus providing a novel high-precision and efficient means for battery system health monitoring in real-world scenarios.

5. Conclusion

In this paper, six short-term segments are exploited considering a reasonable range of SOC operation, and the RFSs are further designed to reduce the effect of noise in the voltage and temperature signals, reflecting the effective features associated with the degradation process. Meanwhile, to highlight important information, attention-based CNN-GRU models are proposed to focus on degenerate features that are profitable for health monitoring, while decreasing less beneficial features. Experimentally, our proposed approach is able to effectively enhance the SOH estimation performance compared to other classical models. For the Oxford battery dataset, a starting voltage of 3.8 V with 300 samples provides the best compromise between computational cost and accuracy in SOH estimation. The results of the SOH estimation using RFSs are more accurate than those of the raw samples, with both RMSE and MAE limited to 0.8%. To fully exploit the aging information in the raw signal and RFSs, a parallel structure based on AB CNN-GRU is developed to fuse the two streams, which achieves 0.582% for RMSE and 0.524% for MAE, which is at least 33.7% and 31.7% reduction, respectively, compared to a single stream. In this paper, the reconstructed characteristic series have verified the feasibility

on the SOH estimator. The validity of RFSs will be further investigated for other state prediction problems, including state of charge and state of energy for lithium-ion batteries, in our future work. For capacity regeneration phenomenon for batteries, the uncertainty quantification will also consider in our future work.

Data Availability

The data used to support the findings of this study have been deposited in the Oxford Battery Degradation Dataset 1 repository (doi:10.5287/bodleian:KO2kdmYGg).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Q. Lin, J. Wang, R. Xiong, W. Shen, and H. He, "Towards a smarter battery management system: a critical review on optimal charging methods of lithium ion batteries," *Energy*, vol. 183, pp. 220–234, 2019.

[2] Z. Wei, Z. Quan, J. Wu, Y. Li, J. Pou, and H. Zhong, "Deep deterministic policy gradient-DRL enabled multiphysics-constrained fast charging of lithium-ion battery," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 3, pp. 2588–2598, 2022.

[3] R. Xiong, L. Li, and J. Tian, "Towards a smarter battery management system: a critical review on battery state of health monitoring methods," *Journal of Power Sources*, vol. 405, pp. 18–29, 2018.

[4] X. Shu, G. Li, Y. Zhang, J. Shen, Z. Chen, and Y. Liu, "Online diagnosis of state of health for lithium-ion batteries based on short- term charging profiles," *Journal of Power Sources*, vol. 471, article 228478, 2020.

[5] M. F. Ge, Y. Liu, X. Jiang, and J. Liu, "A review on state of health estimations and remaining useful life prognostics of lithium-ion batteries," *Measurement*, vol. 174, article 109057, 2021.

[6] K. Liu, Q. Peng, H. Sun, M. Fei, H. Ma, and T. Hu, "A transferred recurrent neural network for battery calendar health prognostics of energy-transportation systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 8172–8181, 2022.

[7] X. Shu, G. Li, J. Shen, W. Yan, Z. Chen, and Y. Liu, "An adaptive fusion estimation algorithm for state of charge of lithium-ion batteries considering wide operating temperature and degradation," *Journal of Power Sources*, vol. 462, article 228132, 2020.

[8] Y. Wang and Z. Chen, "A framework for state-of-charge and remaining discharge time prediction using unscented particle filter," *Applied Energy*, vol. 260, article 114324, 2020.

[9] C. Zhang, K. Li, S. Mcloone, and Z. Yang, "Battery modelling methods for electric vehicles-A review," in *2014 European Control Conference (ECC)*, pp. 2673–2678, Strasbourg, France, 2014.

[10] K. Liu, Y. Gao, C. Zhu et al., "Electrochemical modeling and parameterization towards control-oriented management of lithium-ion batteries," *Control Engineering Practice*, vol. 124, article 105176, 2022.

[11] H. Zhang, Q. Miao, X. Zhang, and Z. Liu, "An improved unscented particle filter approach for lithium-ion battery remaining useful life prediction," *Microelectronics and Reliability*, vol. 81, pp. 288–298, 2018.

[12] Q. Yu, R. Xiong, R. Yang, and M. G. Pecht, "Online capacity estimation for lithium-ion batteries through joint estimation method," *Applied Energy*, vol. 255, article 113817, 2019.

[13] Z. Wei, J. Zhao, D. Ji, and K. J. Tseng, "A multi-timescale estimator for battery state of charge and capacity dual estimation based on an online identified model," *Applied Energy*, vol. 204, pp. 1264–1274, 2017.

[14] N. Wassiliadis, J. Adermann, A. Frericks et al., "Revisiting the dual extended Kalman filter for battery state-of-charge and state-of-health estimation: a use-case life cycle analysis," *Journal of Energy Storage*, vol. 19, pp. 73–87, 2018.

[15] X. Qiu, W. Wu, and S. Wang, "Remaining useful life prediction of lithium-ion battery based on improved cuckoo search particle filter and a novel state of charge estimation method," *Journal of Power Sources*, vol. 450, article 227700, 2020.

[16] Z. Chen, H. Zhao, Y. Zhang, S. Shen, J. Shen, and Y. Liu, "State of health estimation for lithium-ion batteries based on temperature prediction and gated recurrent unit neural network," *Journal of Power Sources*, vol. 521, article 230892, 2022.

[17] K. Liu, Z. Wei, C. Zhang, Y. Shang, R. Teodorescu, and Q. L. Han, "Towards long lifetime battery: AI-based manufacturing and management," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 7, pp. 1139–1165, 2022.

[18] T. Hu, H. Ma, K. Liu, and H. Sun, "Lithium-ion battery calendar health prognostics based on knowledge-data-driven attention," *IEEE Transactions on Industrial Electronics*, vol. 70, pp. 407–417, 2022.

[19] G. Li, B. Li, C. Li, and S. Wang, "State-of-health rapid estimation for lithium-ion battery based on an interpretable stacking ensemble model with short-term voltage profiles," *Energy*, vol. 263, article 126064, 2023.

[20] D.-I. Stroe and E. Schaltz, "Lithium-ion battery state-of-health estimation using the incremental capacity analysis technique," *IEEE Transactions on Industry Applications*, vol. 56, no. 1, pp. 678–685, 2019.

[21] X. Li, C. Yuan, X. Li, and Z. Wang, "State of health estimation for Li-ion battery using incremental capacity analysis and Gaussian process regression," *Energy*, vol. 190, article 116467, 2020.

[22] R. R. Richardson, C. R. Birkl, M. A. Osborne, and D. A. Howey, "Gaussian process regression for in situ capacity estimation of lithium-ion batteries," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 127–138, 2018.

[23] Y. Li, C. Zou, M. Berecibar et al., "Random forest regression for online capacity estimation of lithium-ion batteries," *Applied Energy*, vol. 232, pp. 197–210, 2018.

[24] Z. Wei, H. Ruan, Y. Li, J. Li, C. Zhang, and H. He, "Multistage state of health estimation of lithium-ion battery with high tolerance to heavily partial charging," *IEEE Transactions on Power Electronics*, vol. 37, no. 6, pp. 7432–7442, 2022.

[25] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *Journal of Power Sources*, vol. 384, pp. 387–395, 2018.

[26] X. Li, C. Yuan, and Z. Wang, "State of health estimation for Li-ion battery via partial incremental capacity analysis based

on support vector regression," *Energy*, vol. 203, article 117852, 2020.

[27] J. Wen, X. Chen, X. Li, and Y. Li, "SOH prediction of lithium battery based on IC curve feature and BP neural network," *Energy*, vol. 261, article 125234, 2022.

[28] X. Li, C. Yuan, and Z. Wang, "Multi-time-scale framework for prognostic health condition of lithium battery using modified Gaussian process regression and nonlinear regression," *Journal of Power Sources*, vol. 467, article 228358, 2020.

[29] T. Tang and H. Yuan, "A hybrid approach based on decomposition algorithm and neural network for remaining useful life prediction of lithium-ion battery," *Reliability Engineering & System Safety*, vol. 217, article 108082, 2022.

[30] Y. Che, D. I. Stroe, X. Hu, and R. Teodorescu, "Semi-supervised self-learning-based lifetime prediction for batteries," *IEEE Transactions on Industrial Informatics*, pp. 1–10, 2022.

[31] M. Zhang, G. Kang, L. Wu, and Y. Guan, "A method for capacity prediction of lithium-ion batteries under small sample conditions," *Energy*, vol. 238, article 122094, 2022.

[32] Q. Zhang, X. Li, C. Zhou et al., "State-of-health estimation of batteries in an energy storage system based on the actual operating parameters," *Journal of Power Sources*, vol. 506, article 230162, 2021.

[33] J. Hong, Z. Wang, W. Chen, L. Wang, P. Lin, and C. Qu, "Online accurate state of health estimation for battery systems on real-world electric vehicles with variable driving conditions considered," *Journal of Cleaner Production*, vol. 294, article 125814, 2021.

[34] J. Hong, D. Lee, E. R. Jeong, and Y. Yi, "Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning," *Applied Energy*, vol. 278, article 115646, 2020.

[35] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, and M. J. Deen, "A data-driven auto-cnn-lstm prediction model for lithium-ion battery remaining useful life," *IEEE Transactions on Industrial Informatics*, vol. 17, pp. 3478–3487, 2020.

[36] Y. Li, H. Sheng, Y. Cheng, D. I. Stroe, and R. Teodorescu, "State-of-health estimation of lithium-ion batteries based on semi-supervised transfer component analysis," *Applied Energy*, vol. 277, article 115504, 2020.

[37] Y. Che, Z. Deng, X. Lin, L. Hu, and X. Hu, "Predictive battery health management with transfer learning and online model correction," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1269–1277, 2021.

[38] R. Li, J. Hong, H. Zhang, and X. Chen, "Data-driven battery state of health estimation based on interval capacity for real-world electric vehicles," *Energy*, vol. 257, article 124771, 2022.

[39] C. Zhang, Y. Kang, B. Duan et al., "An adaptive battery capacity estimation method suitable for random charging voltage range in electric vehicles," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 9, pp. 9121–9132, 2022.

[40] C. Birkl, *Diagnosis and Prognosis of Degradation in Lithium-Ion Batteries*, Department of Engineering Science, University of Oxford, 2017.

[41] Q. S. Jia and T. Long, "A review on charging behavior of electric vehicles: data, model, and control," *Control Theory and Technology*, vol. 18, no. 3, pp. 217–230, 2020.