

Received December 17, 2020, accepted December 28, 2020, date of publication January 12, 2021, date of current version January 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3049944

An Improved Bidirectional Gated Recurrent Unit Method for Accurate State-of-Charge Estimation

ZHAOWEI ZHANG^{1,2}, (Graduate Student Member, IEEE), ZHEKANG DONG^{1,2}, HUIPIN LIN^{1,2}, ZHIWEI HE^{1,2}, (Member, IEEE), MINGHAO WANG³, (Member, IEEE), YUFEI HE³, XIANG GAO³, AND MINGYU GAO^{1,2}, (Member, IEEE)

¹School of Electronics Information, Hangzhou Dianzi University, Hangzhou 310018, China

²Zhejiang Provincial Key Lab of Equipment Electronics, Hangzhou 310018, China

³Department of Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong

Corresponding author: Mingyu Gao (mackgao@hdu.edu.cn)

This work was supported in part by the Key Research and Development Plan of the Ministry of Science and Technology under Grant 2020YFB1710600, in part by the Key Research and Development Program of Zhejiang Province under Grant 2021C01111, and in part by the National Natural Science Foundation of China under Grant 61671194 and Grant U1609216.

ABSTRACT State-of-Charge (SOC) estimation of lithium-ion batteries have a great significance for ensuring the safety and reliability of battery management systems in electrical vehicle. Deep learning method can hierarchically extract complex feature information from input data by building deep neural networks (DNNs) with multi-layer nonlinear transformations. With the development of graphic processing unit, the training speed of the network is faster than before, and it has been proved to be an effective data-driven method to estimate SOC. In order to further explore the potential of DNNs in SOC estimation, take battery measurements like voltage, current and temperature directly as input and SOC as output, an improved method using the Nesterov Accelerated Gradient (NAG) algorithm based Bidirectional Gated Recurrent Unit (Bi-GRU) network is put forward in this paper. Notably, to address the oscillation problem existing in the traditional gradient descent algorithm, NAG is used to optimize the Bi-GRU. The gradient update direction is corrected by considering the gradient influence of the historical and the current moment, combined with the estimated location of the parameters at the next moment. Compared to state-of-the-art estimation methods, the proposed method enables to capture battery temporal information in both forward and backward directions and get independent context information. Finally, two well-recognized lithium-ion batteries datasets from University of Maryland and McMaster University are applied to verify the validity of the research. Compared with the previous methods, the experimental results demonstrate that the proposed NAG based Bi-GRU method for SOC estimation can improve the precision of the prediction at various ambient temperature.

INDEX TERMS State-of-Charge estimation, lithium-ion batteries, Bidirectional Gated Recurrent Unit, Nesterov Accelerated Gradient.

I. INTRODUCTION

With global warming and the emergence of various extreme climates, the greenhouse gas emissions caused by diesel and gasoline vehicles have been paid more and more attention. The development of electric vehicles and hybrid vehicles has triggered extensive research to reduce fossil fuel consumption and greenhouse gas emissions. Battery management system (BMS) is widely used, to guarantee the safety, durability,

reliability and efficiency of electric vehicles, then perform the necessary management and diagnostic functions [1]. The State-of-Charge (SOC) provides reliable information about the potential charging and discharge strategies, and it is one of the essential states that BMS needs to monitor [2], [3]. Therefore, the accurate estimation of SOC plays an essential role in BMS [4]. Actually, SOC is not an inherent parameter of the battery, and it cannot be measured directly by any sensor. The parameters such as current, voltage, temperature, internal resistance, battery ageing etc. often affects the estimation of SOC. Furthermore, SOC estimation in electric vehicles

The associate editor coordinating the review of this manuscript and approving it for publication was Jiayong Li.

requires a good performance of accuracy and real-time, different battery materials and working conditions make it challenging to obtain the battery SOC accurately in real-time [5].

At present, the commonly used SOC estimation methods can be divided into four categories: look-up table method, ampere-hour method, model-based method and data-driven method. The look-up table method is a simple method to estimate the SOC. By using the mapping relationship between battery characteristic and SOC, the battery behavior is characterized by a large number of experiments. The conventional look-up table methods mainly includes open-circuit voltage method [6] and AC impedance method [7]. Still, the estimation accuracy depends on the accuracy of the look-up table established by experiments. The main idea of the ampere-hour method is: by measuring the current that though the battery and calculate the current integration value at this period of time, and then add it to the initial value of the battery, finally get the SOC [8]. The ampere-hour method is an open-loop estimation method. Due to the influence of current sensor sampling accuracy and sampling frequency, there will be accumulated error in the integration for a long time [9]–[11]. The battery model is combined with the nonlinear observer to form a model-based method. Yan *et al.* [12] developed a Lebesgue-sampling-based extended Kalman filter to estimate the SOC. Shehab El Din *et al.* [13] proposed an unscented Kalman filter with the autocovariance least-squares technique to estimate SOC, and evaluated using a novel battery cell model. Zhang *et al.* [14] presented an adaptive weighting cubature particle filter to estimate SOC. Considering that the cubature particle filter is sensitive to noise, the statistics of system noise, prediction state, measurement vector and covariance are estimated by an adaptive weighting method. Also, some nonlinear observers are applied to model-based method for SOC estimation, such as H infinity filter [15], proportional integral observer [16], sliding mode observer [17], Luenberger observer [18], etc. The model-based method uses a closed-loop structure, which can estimate the SOC with an unknown initial state. However, the process of establishing the model is complex and time-consuming, and it is difficult to identify the parameters online, which requires a large amount of calculation. Besides, the estimation accuracy depends on the accuracy of the battery model and parameter identification. In recent years, data-driven method has drawn more attention from scholars because it only requires battery measured signals for SOC estimation, instead of relying on any complicated battery model. With the rapid development of graphic processing units, building and training neural networks have been more and much easier and faster than before, the neural network-based deep learning method has drawn more attention from the research world [19]. Therefore, data-driven method for SOC estimation often use machine learning platform, automatically learn network parameters by intelligent algorithms [20]. Lipu *et al.* [21] using the back-propagation neural network to estimate the SOC for a lithium-ion battery, principal component analysis and particle swarm optimization algorithms were used to enhance

the accuracy and robustness of the SOC estimation model. Song *et al.* [22] proposed a combined convolutional neural network–long short-term memory network to infer battery SOC from measurable data, such as current, voltage, and temperature. Li *et al.* proposed a method for SOC prediction based on particle swarm support vector regression [23]. The data-driven method relies on the analysis of battery measured signals, which avoids the difficulties in parameter identification and high computational complexity in the model-based method, and the data-driven method have the advantages of estimation speed and generalization. Therefore, in recent years, the use of the data-driven method to estimate SOC has received more and more attention [19].

It is essential to choose the network structure and the optimization function of network weight in the data-driven model. Appropriate network structure and optimization function can minimize the network loss function and improve the training speed. Lipu *et al.* [24] proposed a long short-term memory model with an attention mechanism to estimate the charging status of lithium-ion batteries. Mamo *et al.* [25] proposed a state-of-charge estimation model for a lithium-ion batteries using an improved extreme learning machine (ELM) algorithm. A battery states estimation method by combining recurrent neural network modeling and particle-filtering-based errors redress was investigated [26]. Hannan *et al.* [27] developed an SOC estimation approach for a lithium-ion batteries by improving BPNN capability using backtrack-search algorithm. However, the classical unidirectional structure cannot capture battery temporal information in both forward and backward directions. Furthermore, the classical gradient descent algorithm are easy to fall into the local optimum, and have a slow convergence speed. To solve the above problems, this paper further investigates the network structure and optimization algorithm suitable for SOC estimation. The main contributions are summarized as follows:

(1) A GRU with bidirectional structure for SOC estimation is presented, which is able to capture the long-term time dependencies of battery sequence in forwarding and backward directions, and improve the accuracy of the model.

(2) Unlike the other optimization algorithms, NAG considers the gradient influence of historical and current moment as the factor, and modifies the direction of gradient update by predicting the position of parameters in the next moment. Because of these, NAG can effectively reduce the vibration and improve the stability in the process of gradient descent.

(3) The influence of various parameters on the accuracy of the model was analyzed through experiments, and two public lithium-ion battery datasets is used to verify the validity of the proposed Bi-GRU model.

The rest of the paper is organized as follows. Section 2 briefly describes the Bi-GRU and NAG algorithm. A NAG algorithm based Bi-GRU method is proposed for battery SOC estimation in Section 3. The experimental simulation and results analysis are discussed in section 4. Section 5 concludes this paper.

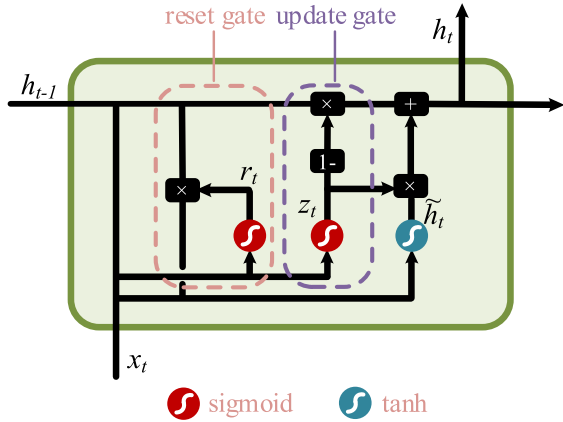


FIGURE 1. The structure of the GRU.

II. A NAG ALGORITHM BASED Bi-GRU

A. Bi-GRU NETWORK

GRU [28] is a kind of recurrent neural network (RNN). Like LSTM (long short term memory) [29], it is also proposed to address the problems of long-term memory and gradient in back-propagation. We choose GRU in our work since it performs similarly to LSTM, but is computationally cheaper. The structure of the GRU is shown in Fig. 1. GRU is composed of the current input x_t , the previous hidden state h_{t-1} , the update gate z_t , the reset gate r_t , and the new hidden state h_t . In the following process, the input of GRU is x_t and h_{t-1} , and the output is h_t .

- Update gate

The update gate z_t is used to control the degree that the information of the previous moment is brought into the current state. The greater the value of the update gate, the more information is brought in from the previous moment. Update gate z_t can be calculated by:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (1)$$

where W_z and b_z are the weight matrix and bias of the update gate respectively. $\sigma(x) = 1/[1+\exp(-x)]$ means the Sigmoid function, by which the data is converted into values in the range of $0 \sim 1$ (acting as the gate-control signal).

- Reset gate

Reset gate r_t is used to control how much of the hidden layer information of the previous moment needs to be forgotten. It can be calculated by:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (2)$$

where W_r and b_r are the weight matrix and bias of the reset gate respectively. If there is no need to remember the hidden state of the previous moment, sigmoid will set the output of the reset gate to 0 to forget the information of the hidden state of the previous moment. On the contrary, sigmoid will set the output of the reset gate to 1, so that all the hidden states from the previous moment pass through, that is, the smaller the reset gate is, the less information is retained from the previous state.

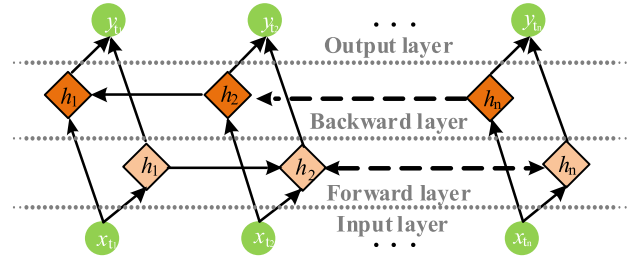


FIGURE 2. The structure of the Bi-GRU.

- Candidate output state

Candidate output state \tilde{h}_t can be calculated by:

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \quad (3)$$

where W_h and b_h are the weight matrix and bias of candidate output state \tilde{h}_t respectively. \tanh activates the function to scale the data to a range from -1 to 1 . \odot is Hadamard Product, namely the multiplication of the corresponding elements in the matrix. The output of the reset gate directly affects the candidate output state of the GRU unit. The product of r_t and the hidden layer state of the previous moment determines how much the output of the neuron is retained from the previous moment. The greater the value of r_t , the more the output of the neuron is retained from the previous moment, and vice versa.

- New hidden state

The output of GRU is the new hidden layer state h_t , it is determined by z_t , h_{t-1} and \tilde{h}_t . The mathematical expression is as follows:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

The larger z_t is, the greater dependence of h_t on \tilde{h}_t is, and the smaller determining role of h_{t-1} on the output is. $(1 - z_t) \odot h_{t-1}$ indicates the selective “forgetting” to previous hidden state h_{t-1} , $z_t \odot \tilde{h}_t$ indicates selective “memory” to \tilde{h}_t which contains the current node information.

Unlike traditional GRU which can only predicate the output of the next moment based on the temporal sequence information of the previous moment, Bi-GRU combines the information of the input sequence in both the forward and backward directions. The bidirectional structure is shown in Fig. 2. The bidirectional structure can be regarded as two hidden layers with opposite directions of information transfer and GRU units. At the moment t , the input x_t supplies the hidden layers in two opposite directions simultaneously. The output y_t at moment t is jointly determined by these two one-way hidden layers. The forward GRU layer has the information of moment t and the previous moment in the input sequence, while the backward GRU layer has the information of moment t and the subsequent moment in the input sequence. The hidden layer propagation process of Bi-GRU can be defined as :

$$\vec{h}_t = f(\vec{W}x_t + \vec{V}\vec{h}_{t-1} + \vec{b}) \quad (5)$$

$$\overleftarrow{h}_t = f(\overleftarrow{W}x_t + \overleftarrow{V}\overleftarrow{h}_{t+1} + \overleftarrow{b}) \quad (6)$$

rewritten as:

$$\theta'_t = \theta'_{t-1} - \gamma m_t - \eta \nabla L(\theta'_{t-1}) \quad (12)$$

$$m_t = \gamma m_{t-1} - \eta \nabla L(\theta'_{t-1}) \quad (13)$$

Define the loss function for the network:

$$L(\theta') = \frac{1}{2} (SOC_t - SOC'_t)^2 \quad (14)$$

The process of NAG to optimize network parameters is as follows.

Step 3 (SOC Estimation and Performance Evaluation): In this paper, mean absolute error (MAE) and root mean square error (RMSE) is used to evaluate the performance of well-established model. MAE represents the average of the absolute error between the predicted and the observed value. RMSE is an indication of the dispersion of the samples, and large RMSE indicates high degree of dispersion.

$$MAE = \frac{1}{T} \sum_{t=1}^T |SOC - SOC'_t| \quad (15)$$

$$RMSE = \frac{1}{T} \sqrt{\sum_{t=1}^T (SOC - SOC'_t)^2} \quad (16)$$

Step 4 (Model Optimization): Replace parameters and retrain the network, analyze the performance of different network parameters on SOC estimation problem.

IV. EXPERIMENTAL SIMULATION AND RESULTS

A. DATASET DESCRIPTION

In this work, two public lithium-ion battery datasets are used to evaluate the proposed method. The electrical characteristics of the batteries are listed in Table 2. One of the experimental dataset is collected from the Samsung 18650 LiNiMnCoO₂/Graphite lithium-ion batteries by the Center for Advanced Life Cycle Engineering (CALCE) at University of Maryland [30]. It includes the current and voltage data of BJDST, DST, FUDS and US06 drive cycles in varying ambient temperatures. Test process at three ambient temperatures of 0°C, 25°C and 45°C, and the sampling frequency of voltage and current parameters is 1s. Samsung INR 18650-20R dataset includes four stages: charging, discharging at a current rate of 1 C, standing and discharging according to drive cycles. Fig. 4 shows the current data of the four drive cycles.

LG 18650HG2 lithium-ion battery dataset, which is performed at McMaster University in Hamilton, Ontario, Canada by Dr. Phillip Kollmeyer. In this dataset, eight driving cycles are included: UDD, HWFET, LA92, US06 and their random mix. The tests are repeated for ambient temperatures of 40°C, 25°C, 10°C, 0°C, -10°C, and -20°C, in that order. In this paper, LA92 and UDDS at 25°C are selected to test the universality of the proposed method. Fig. 5 shows the current of the LA92 and UDDS cycles.

Process Description of NAG

- 1 Initialization:
- 2 - Initialization the weights and bias of z_t, r_t, h_t :
 $W_{zh,0}, W_{zx,0}, b_{z,0}; W_{rh,0}, W_{rx,0}, b_{r,0}; W_{h,0}, W_{x,0}, b_{h,0}$
- 3 - Initialization the weights and bias of the output layer:
 $W_{o,0}, b_{o,0}$
- 4 **for** $t = 1, 2, \dots$ **do**
- 5 - Calculate the loss function $L(\theta')$
- 6 - Calculate the gradient of the loss function to the hidden layer parameters:

$$\begin{cases} \nabla L_{W,zh} = \frac{\partial L(\theta')}{\partial W_{zh}} \\ \nabla L_{W,zx} = \frac{\partial L(\theta')}{\partial W_{zx}} \\ \nabla L_{b,z} = \frac{\partial L(\theta')}{\partial b_z} \end{cases}, \quad \begin{cases} \nabla L_{W,rh} = \frac{\partial L(\theta')}{\partial W_{rh}} \\ \nabla L_{W,rx} = \frac{\partial L(\theta')}{\partial W_{rx}} \\ \nabla L_{b,r} = \frac{\partial L(\theta')}{\partial b_r} \end{cases},$$

$$\begin{cases} \nabla L_{W,h} = \frac{\partial L(\theta')}{\partial W_h} \\ \nabla L_{W,x} = \frac{\partial L(\theta')}{\partial W_x} \\ \nabla L_{b,h} = \frac{\partial L(\theta')}{\partial b_h} \end{cases}$$
- 7 - Calculate the gradient of the loss function to the output layer parameters:

$$\nabla L_{W,o} = \frac{\partial L(\theta')}{\partial W_o}, \quad \nabla L_{b,o} = \frac{\partial L(\theta')}{\partial b_o}$$
- 8 - Calculate the momentum of the hidden layer weight and bias:



$$\begin{cases} m_{W,zh,t} = \gamma m_{W,zh,t-1} - \eta \nabla L_{W,zh,t-1} \\ m_{W,zx,t} = \gamma m_{W,zx,t-1} - \eta \nabla L_{W,zx,t-1}, \\ m_{b,z,t} = \gamma m_{b,z,t-1} - \eta \nabla L_{b,z,t-1} \\ \\ m_{W,rh,t} = \gamma m_{W,rh,t-1} - \eta \nabla L_{W,rh,t-1} \\ m_{W,rx,t} = \gamma m_{W,rx,t-1} - \eta \nabla L_{W,rx,t-1} \\ m_{b,r,t} = \gamma m_{b,r,t-1} - \eta \nabla L_{b,r,t-1} \\ \\ m_{W,h,t} = \gamma m_{W,h,t-1} - \eta \nabla L_{W,h,t-1} \\ m_{W,x,t} = \gamma m_{W,x,t-1} - \eta \nabla L_{W,x,t-1} \\ m_{b,h,t} = \gamma m_{b,h,t-1} - \eta \nabla L_{b,h,t-1} \end{cases}$$
- 9 - Calculate the momentum of the output layer weight and bias:

$$\begin{aligned} m_{W,o,t} &= \gamma m_{W,o,t-1} - \eta \nabla L_{W,o,t-1}, \\ m_{b,o,t} &= \gamma m_{b,o,t-1} - \eta \nabla L_{b,o,t-1} \end{aligned}$$
- 10 - Update the weight and bias of the hidden layer:

$$\begin{cases} W_{zh,t} = W_{zh,t-1} - \gamma m_{W,zh,t-1} - \eta \nabla L_{W,zh,t-1} \\ W_{zx,t} = W_{zx,t-1} - \gamma m_{W,zx,t-1} - \eta \nabla L_{W,zx,t-1} \\ b_{z,t} = b_{z,t-1} - \gamma m_{b,z,t-1} - \eta \nabla L_{b,z,t-1} \\ \\ W_{rh,t} = W_{rh,t-1} - \gamma m_{W,rh,t-1} - \eta \nabla L_{W,rh,t-1} \\ W_{rx,t} = W_{rx,t-1} - \gamma m_{W,rx,t-1} - \eta \nabla L_{W,rx,t-1} \\ b_{r,t} = b_{r,t-1} - \gamma m_{b,r,t-1} - \eta \nabla L_{b,r,t-1} \\ \\ W_{h,t} = W_{h,t-1} - \gamma m_{W,h,t-1} - \eta \nabla L_{W,h,t-1} \\ W_{x,t} = W_{x,t-1} - \gamma m_{W,x,t-1} - \eta \nabla L_{W,x,t-1} \\ b_{h,t} = b_{h,t-1} - \gamma m_{b,h,t-1} - \eta \nabla L_{b,h,t-1} \end{cases}$$
- 11 - Update the weight and bias of the output layer:

$$\begin{aligned} W_{o,t} &= W_{o,t-1} - \gamma m_{W,o,t-1} - \eta \nabla L_{W,o,t-1}, \\ b_{o,t} &= b_{o,t-1} - \gamma m_{b,o,t-1} - \eta \nabla L_{b,o,t-1} \end{aligned}$$
- 12 - $t = t + 1$
- 13 **end**

TABLE 2. Electrical characteristics of the batteries.

Parameter	Samsung INR 18650-20R	LG 18650HG2
Minimum discharge voltage/Maximum Charge voltage	2.5V/4.2V	2V/4.2V
Nominal voltage	3.6V	3.6V
Nominal capacity	2000mAh	3000mAh
Initial internal impedance	≤20mΩ	≤35mΩ
Max. continuous discharge current	20A	20A
Min/max temperature	-20°C +60 °C	-20 °C +60 °C
Appearance		

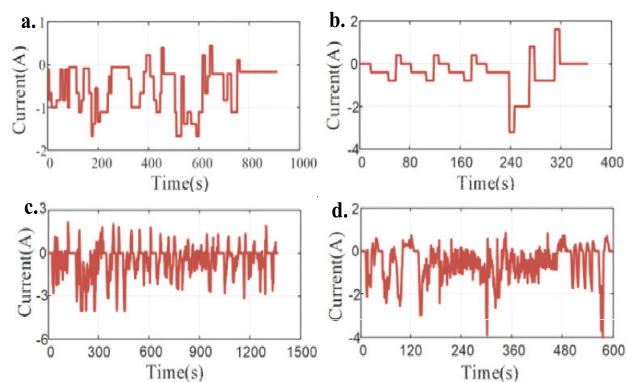


FIGURE 4. The current of four drive cycles: (a) BJDST, (b) DST, (c) FUDS, and (d) US06.

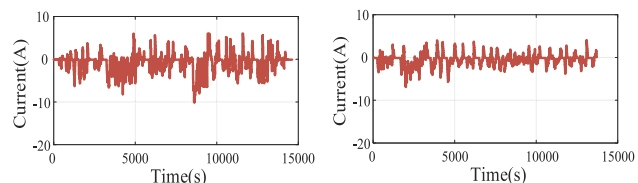


FIGURE 5. The current of LA92 and UDDS drive cycles.

B. SOFTWARE AND HARDWARE

In this study, the random mixture of BJDST, DST and US06 is used as a training set, the current and voltage profile at 25°C are displayed in Fig. 6. FUDS, LA92 and UDDS is used as a test set.

The experiments are conducted on the PC with CPU Intel(R) Core(TM) i7-8700 CPU @ 3.20 GHz, 16.0 GB RAM with PyCharm 2018.

C. THE INFLUENCE OF DIFFERENT PARAMETERS ON MODEL ACCURACY

1) SOC ESTIMATION AT MULTIPLE AMBIENT TEMPERATURES

The discharge process of the battery is always in a changing ambient temperature. And in the SOC estimation process, the influence of ambient temperature cannot be ignored [31]–[33]. It is relatively simple to estimation the SOC of the constant current charging and discharging process

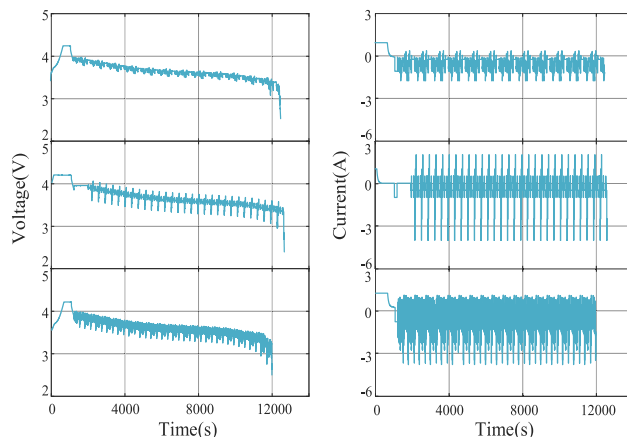


FIGURE 6. The INR 18650-20R measured data in the BJDST, DST and US06 test: (a) Current, (b) Voltage, (c) Capacity.

TABLE 3. Results on FUDS set T when the network is trained on multiple ambient temperature.

Temperature	MAE(%)	RMSE(%)	Error(%)
0	1.84	2.15	-0.70/6.07
25	1.02	2.15	-1.31/3.47
45	1.04	1.34	-1.05/3.72

of the battery, and the proposed model is mainly to solve the estimation problem of the battery discharge under different working conditions, so the discharge process data is used to test the model. Table 3 lists the test results on FUDS set at different ambient temperatures. Fig. 7 shows the performance on a FUDS discharge process on multiple ambient temperature with an initial SOC of 80%.

The results show that the proposed method have a good performance under different temperature conditions in predicting SOC. MAE is within 1.9% and RMSE is within 2.2%. Therefore, the model can accurately estimate battery SOC under the circumstance of temperature change. From the comparison results of the estimated performance of the models at different temperatures, it can be obtained that the model has poor estimation accuracy at low temperature. This is due to the more complex battery dynamics at low temperatures, which increases the difficulty of SOC estimation. In addition, due to battery capacity limitations, the drive cycle is shorter at low temperatures and less measurement data is collected from the battery. Without fully training the model using these low temperature data, the estimated performance of the model is lower.

2) SOC ESTIMATION AT DIFFERENT NETWORK PARAMETERS

In the process of network training, there are many network parameters that affect the accuracy of the model, of which the selection of number of hidden neurons and epoch has the greatest influence on the network [33]. When the network contains too many nodes, the limited information in the training set is not enough to train all the neurons in the

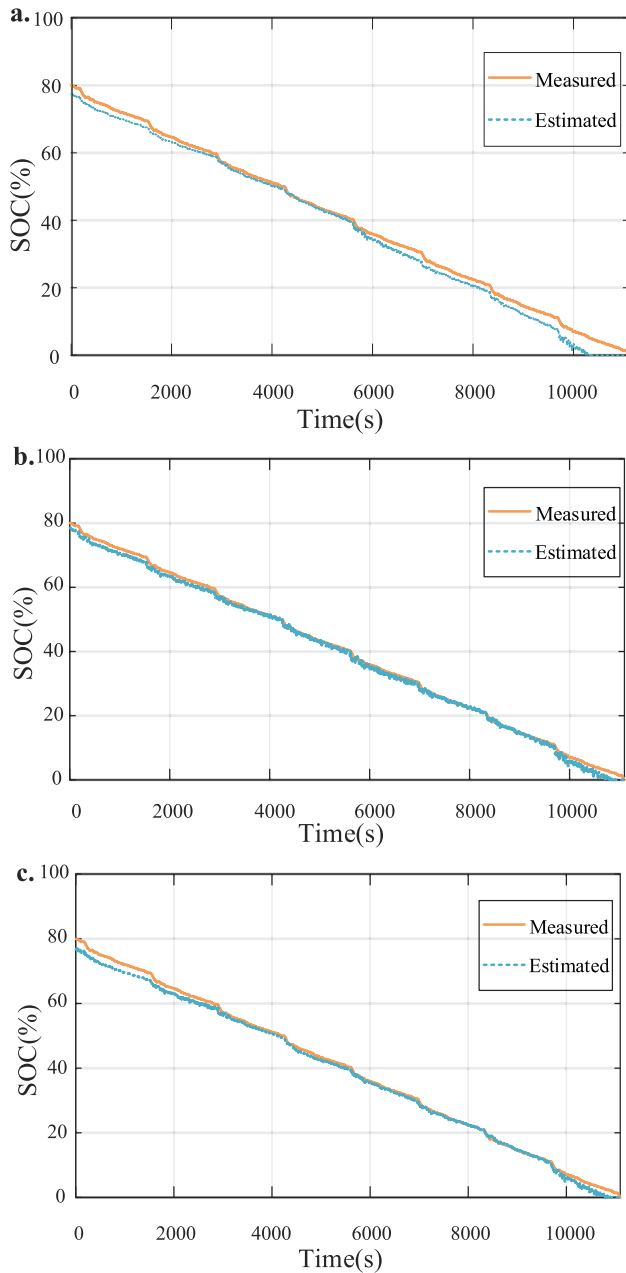


FIGURE 7. Estimation on the FUDS test set at ambient temperatures: (a) 0°C, (b) 25°C, and (c) 45°C.

hidden layer, which may lead to over fitting. Conversely, using too few neurons in the hidden layer results in under fitting. During the training, an epoch is a process in which all the data is sent into the network to complete a forward calculation and back transmission. The size of epoch is related to the degree of diversification of the data set. The greater the degree of diversification, the larger epoch should be selected. In addition, too many neurons in the hidden layer or too many times of training will increase the training time, making it difficult to achieve the desired effect.

It can be seen from Table 3 and Fig. 7 that the Bi-GRU model has the most stable performance of estimation at 25°C.

TABLE 4. RMSE of various hidden neurons and epochs tested in FUDS.

Epochs	hidden neurons							
	32	64	96	128	160	192	224	256
50	6.01	4.38	2.48	2.28	3.22	1.80	2.81	3.59
100	5.23	2.80	1.35	1.52	2.09	1.29	2.41	2.76
150	4.88	3.03	1.48	1.73	1.14	2.18	2.77	2.22
200	5.02	2.78	1.33	1.54	0.94	1.75	2.66	2.31
250	4.57	2.63	2.39	1.57	2.02	2.20	2.81	2.23
300	5.67	3.45	2.48	1.02	1.96	2.76	3.09	2.76
350	5.59	3.74	2.51	2.04	1.87	3.64	2.51	2.85
400	4.97	3.36	1.99	2.06	1.61	1.93	2.99	2.71
450	5.40	2.18	2.73	2.15	1.51	2.19	2.52	2.22
500	5.44	3.29	2.35	1.73	1.70	2.29	2.71	2.99

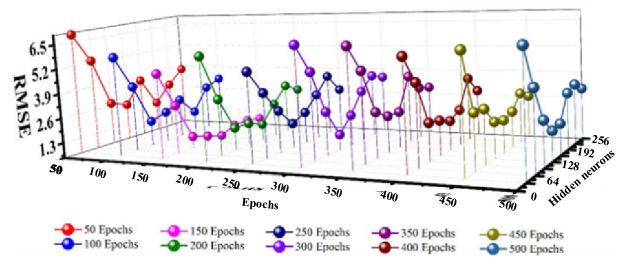


FIGURE 8. Model estimation performance with different number of hidden neurons and epochs.

In order to reduce the influence of temperature factors on the model error, the following tests are conducted under the data set of a temperature of 25°C. We constructed Bi-GRU models with different numbers of hidden neurons and different epochs to evaluate the influence of different numbers of neurons and epochs on the estimation performance of the model.

Table 4 and Fig. 8 show the comparison results of estimation performance of the model with different number of hidden neurons and iterations. The results show that when hidden neurons is too little, RMSE of the test results of the model fluctuates continuously with the increase of epoch, and the RMSE is at a higher level, and the network is under fitting. When hidden neurons are excessive, network accuracy is improved, but RMSE does not significantly decrease with the increase of epoch and fluctuates constantly. Select appropriate hidden neurons, such as around 128, with the increase of epochs, the RMSE of the model continuously decreases; however, when different hidden neurons are used, RMSE takes different times to reach its lowest point. When RMSE exceeds the lowest point, continuous increasing epoch will lead to the network over fitting and decreased accuracy of the model. When epoch is 300, Fig. 9 shows the performance of the Bi-GRU using different hidden neurons on a FUDS dataset with an initial SOC of 80%.

3) DIFFERENT γ ON NETWORK

Hyperparameter γ of NAG algorithm represents the influence of history gradient on current gradient update, and is generally valued between 0 and 1. With the increase of the number of iterations, the power of γ is more and more high, the coefficient before gradient matrix $\nabla L(\theta_t)$ continuously decreases,

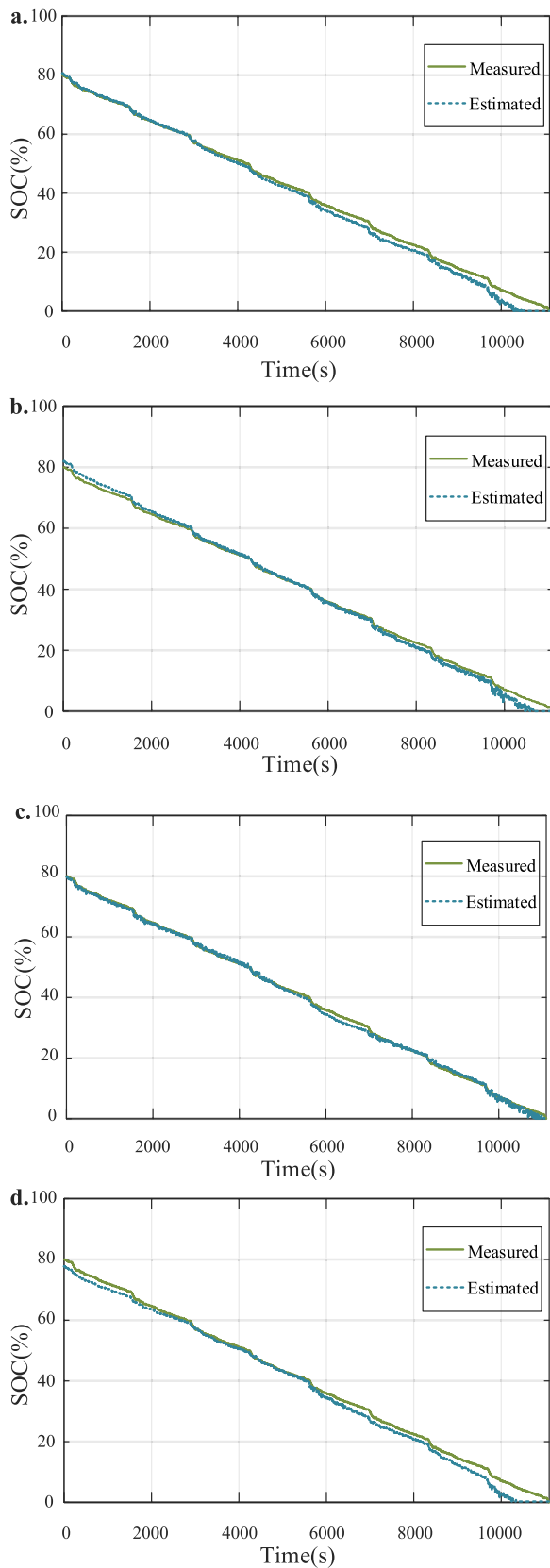


FIGURE 9. Estimates and errors of Bi-GRU using different hidden neurons on the FUDS test dataset: (a) 32, (b) 64, (c) 128, and (d) 256.

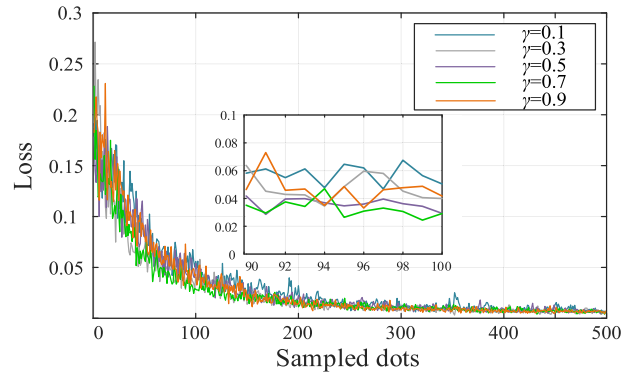


FIGURE 10. Loss changes of different γ .

showing that the influence of history gradient on the current gradient update is less and less, reaching the purpose that the position closer to the current moment has greater influence on the direction of gradient update. Fig.10 shows the changes of Loss during different training in NAG algorithm. As can be seen from Figure 10, when γ is 0.1, Loss appears obvious oscillation in the process of descent, and the descent speed is slow. When γ increased from 0.1 to 0.7, the descent rate of Loss was gradually accelerated; however, when γ continues to increase and is set to 0.9, the descent rate was somewhat reduced. However, different γ could eventually stabilize Loss. Selecting appropriate hyperparameter can speed up the training and improve the prediction accuracy of the model.

D. UNIVERSALITY OF THE MODEL

In order to evaluate the universality of the proposed model, we use the LG 18650HG2 dataset for SOC estimation. Fig. 11 shows the estimated results of LA92 and UDDS test cases at 25°C. It can be seen from it that the proposed Bi-GRU model achieves satisfactory estimation accuracy for other types of datasets. The RMSE and MAE of LA92 test case are 1.402% and 1.119% respectively, and the RMSE and MAE of UDDS test case are 1.370% and 1.132% respectively, and the maximum estimation error is within 5%. Due to its bidirectional recursive structure, the proposed model performs well in most test cases, and can perform better in SOC estimation.

E. COMPARISON OF DIFFERENT MODELS

In order to evaluate the estimation performance of the proposed model, we compared the proposed model with some existing methods, including the methods based on data driven such as LSTM, SVM, ELM, BPNN, RBFNN, GNN, and their improved methods [24]–[27]. In addition, there are filtering algorithms such as EKF, UKF, PF, CKF, AWCPF and their improved algorithms [13], [33]. Table 5 shows the RMSE of various models on FUDS, including ambient temperatures of 0°C, 25°C and 45°C. It can be divided into data-driven methods and filtering methods.

It can be seen from Table 4 that in the data-driven methods, temperature can influence SOC estimation accuracy, and the

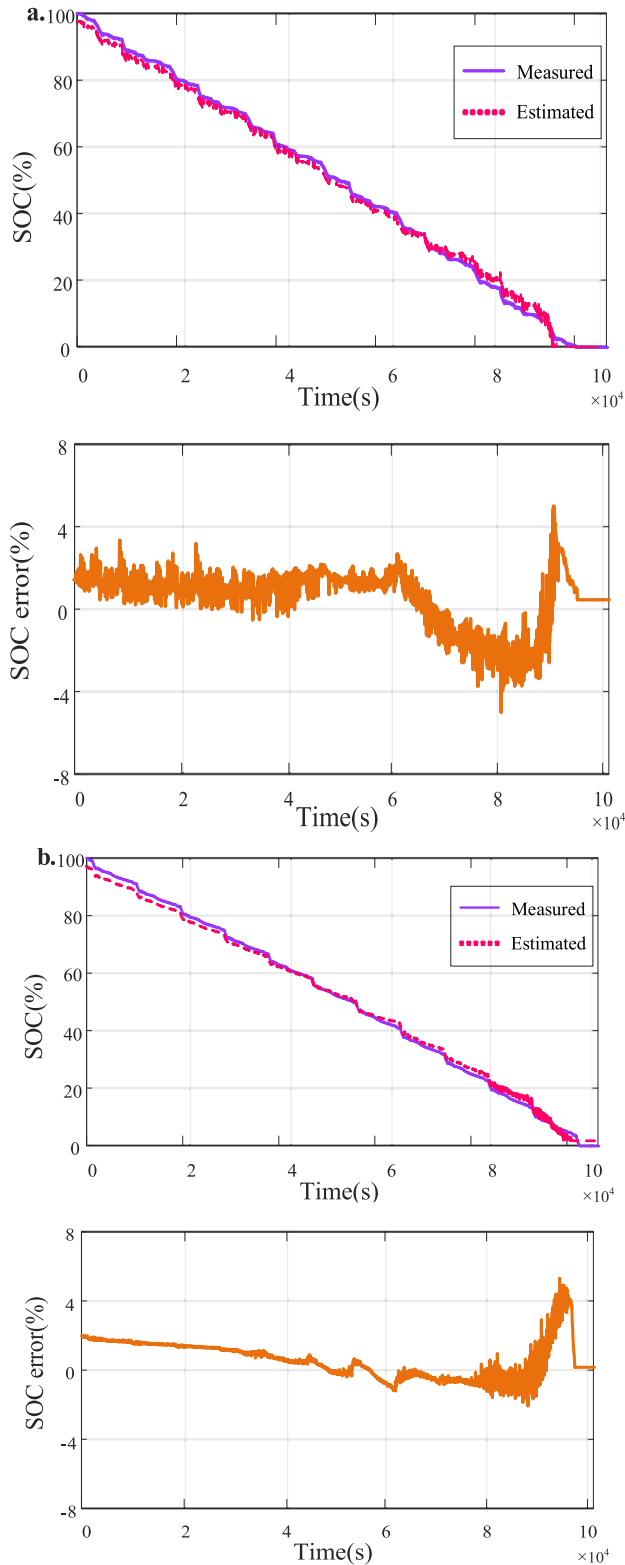


FIGURE 11. Estimates and errors of Bi-GRU on the different test cases at 25°C: (a) LA92, (b) UDDS.

RMSE of various models changes greatly with the change of temperature. In general, the models have more estimation error in low temperature state, but in high temperature state

TABLE 5. RMSE of various models on FUDS dataset.

Methods	Temperature		
	0°C	25°C	45°C
LSTM using optimal parameters	1.12	1.12	1.11
LSTM model with an attention mechanism.	0.96	0.87	0.92
SVM	5.96	3.28	2.50
LSTM	3.30	2.49	1.73
Standard RNN	5.72	2.73	1.96
LSTM-PF	1.42	1.50	1.68
SVM-PF	2.97	1.65	2.24
Standard RNN-PF	2.96	1.26	2.05
ELM based GSA		1.40	
BPNN based BSA	1.74	0.91	0.57
RBFNN based BSA	3.58	2.23	1.75
GRNN based BSA	3.22	2.17	1.81
ELM based GSA	3.72	2.34	2.27
CKF		5.21	
PF		4.27	
AWCPF	2.05		
Unconstrained-EKF		5.40	5.83
EKF		4.31	5.70
Unconstrained-UKF		4.61	3.70
UKF		3.97	3.08
Unconstrained-PF		2.04	3.13
proposed method	2.15	1.30	1.34

the error is little and have a satisfactory accuracy. In the proposed model, RMSE can be below 2.2% at different temperatures, which is less sensitive to temperature compared with other models. In contrast, there is less discussion on the influence of temperature on SOC estimation by filtering method.

It can be observed that under FUDS cycle conditions, the proposed models have a high estimation accuracy. Within the estimation models listed below, the LSTM model with an attention mechanism model [25] has the best estimation performance, that is, at 0°C, 25°C and 45°C, RMSE is 0.96%, 0.87% and 0.92% respectively. Secondly, the LSTM-PF model [26] has relatively better estimation accuracy and stability under the three temperatures, and its RMSE is 1.42%, 1.50% and 1.68% respectively. This is because PF is good at processing data under non-Gaussian noise. In addition, the BPNN based BSA model [27] also has a lower RMSE in the FUDS at the three temperatures. However, it is sensitive to temperature changes and has relatively poor stability. Compared with the above three models, the Bi-GRU based on NAG model which proposed in this paper has the advantages of simple network structure, reaching ideal accuracy of SOC estimation results and improved training speed. For filtering methods, such as the AWCPF method [14], the unconstrained-PF method [34], and others have the advantage of less time-consuming for modeling, relatively low

requirements for data set. It can be seen that the influence of temperature change is seldom discussed when SOC estimation is used in this kind of method, their SOC estimation performance under temperature variation is up for discussion. In addition, compared with the filtering methods, the NAG based Bi-GRU method proposed in this paper has a smaller RMSE on the FUDS data set in most cases, proving the effectiveness of this method for SOC estimation.

V. CONCLUSION

In this paper, we offers three unique contributions. The contribution on the battery modeling front is use the Bi-GRU maps battery measurements like voltage, current, and temperature directly to SOC of lithium-ion batteries. The contribution of parameterization front is that a NAG algorithm is used for ensure the stability and training speed of the network. In NAG algorithm, the current parameters are corrected by predicting the direction of gradient update at the next time, which can effectively reduce the oscillation and improve the network training speed in the process of optimizing the network weight. Finally, in order to verify the effectiveness of the proposed method, two types of lithium-ion battery sets are tested. The performance of SOC estimation at different ambient temperatures are discussed by experiments. The results show that the proposed method can learn the effect of ambient temperature very well. The RMSE of INR battery and NCR battery SOC estimation are less than 2.5% and 3.5% respectively at different temperatures. In addition, the effects of the number of neurons in the network, epoch and the super parameter γ in NAG on the accuracy of the model are also discussed. Selecting the appropriate number of neurons, epoch and γ can ensure the training speed and improve the prediction accuracy of the model. In addition, by comparing the RMSE of various methods in the same set, the proposed method has the advantages of simple network structure and high accuracy of SOC estimation results. When the trained model is used to test the test set, the test time is very short and can be controlled within 1s under CPU environment. Therefore, this method can realize real-time estimation of battery SOC, which is suitable for on-board real-time estimation.

REFERENCES

- [1] C. Campestrini, M. F. Horsche, I. Zilberman, T. Heil, T. Zimmermann, and A. Jossen, "Validation and benchmark methods for battery management system functionalities: State of charge estimation algorithms," *J. Energy Storage*, vol. 7, pp. 38–51, Aug. 2016, doi: [10.1016/j.est.2016.05.007](https://doi.org/10.1016/j.est.2016.05.007).
- [2] Y. Tian, R. Lai, X. Li, L. Xiang, and J. Tian, "A combined method for state-of-charge estimation for lithium-ion batteries using a long short-term memory network and an adaptive cubature Kalman filter," *Appl. Energy*, vol. 265, May 2020, Art. no. 114789, doi: [10.1016/j.apenergy.2020.114789](https://doi.org/10.1016/j.apenergy.2020.114789).
- [3] J. Li, C. Zhang, Z. Xu, J. Wang, J. Zhao, and Y.-J.-A. Zhang, "Distributed transactive energy trading framework in distribution networks," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7215–7227, Nov. 2018.
- [4] E. Chemali, P. J. Kollmeyer, M. Preindl, and A. Emadi, "State-of-charge estimation of li-ion batteries using deep neural networks: A machine learning approach," *J. Power Sources*, vol. 400, pp. 242–255, Oct. 2018, doi: [10.1016/j.jpowsour.2018.06.104](https://doi.org/10.1016/j.jpowsour.2018.06.104).
- [5] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6730–6739, Aug. 2018, doi: [10.1109/tie.2017.2787586](https://doi.org/10.1109/tie.2017.2787586).
- [6] J. Li and M. Liu, "State-of-charge estimation of batteries based on open-circuit voltage and time series neural network," in *Proc. 6th Int. Conf. Syst. Informat. (ICSAI)*, Shanghai, China, Nov. 2019, pp. 257–262, doi: [10.1109/ICSAI48974.2019.9010535](https://doi.org/10.1109/ICSAI48974.2019.9010535).
- [7] S.-L. Wu, H.-C. Chen, M.-Y. Tsai, T.-C. Lin, and L.-R. Chen, "AC impedance based online state-of-charge estimation for li-ion battery," in *Proc. Int. Conf. Inf., Commun. Eng. (ICICE)*, Xiamen, China, Nov. 2017, pp. 53–56, doi: [10.1109/ICICE.2017.8479183](https://doi.org/10.1109/ICICE.2017.8479183).
- [8] D. Bing, T. Yantao, and Z. Changjiu, "One estimating method of the state of charge of power battery for electronic vehicle," in *Proc. 6th Int. Conf. Measuring Technol. Mechatronics Autom.*, Zhangjiajie, China, Jan. 2014, pp. 439–442, doi: [10.1109/ICMTMA.2014.107](https://doi.org/10.1109/ICMTMA.2014.107).
- [9] Y. Zheng, M. Ouyang, X. Han, L. Lu, and J. Li, "Investigating the error sources of the online state of charge estimation methods for lithium-ion batteries in electric vehicles," *J. Power Sources*, vol. 377, pp. 161–188, Feb. 2018, doi: [10.1016/j.jpowsour.2017.11.094](https://doi.org/10.1016/j.jpowsour.2017.11.094).
- [10] S. Sepasi, R. Ghorbani, and B. Y. Liaw, "A novel on-board state-of-charge estimation method for aged li-ion batteries based on model adaptive extended Kalman filter," *J. Power Sources*, vol. 245, pp. 337–344, Jan. 2014, doi: [10.1016/j.jpowsour.2013.06.108](https://doi.org/10.1016/j.jpowsour.2013.06.108).
- [11] H. Rahimi-Eichi, U. Ojha, F. Baronti, and M.-Y. Chow, "Battery management system: An overview of its application in the smart grid and electric vehicles," *IEEE Ind. Electron. Mag.*, vol. 7, no. 2, pp. 4–16, Jun. 2013, doi: [10.1109/MIE.2013.2250351](https://doi.org/10.1109/MIE.2013.2250351).
- [12] W. Yan, B. Zhang, G. Zhao, S. Tang, G. Niu, and X. Wang, "A battery management system with a Lebesgue-sampling-based extended Kalman filter," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3227–3236, Apr. 2019, doi: [10.1109/tie.2018.2842782](https://doi.org/10.1109/tie.2018.2842782).
- [13] M. S. El Din, A. A. Hussein, and M. F. Abdel-Hafez, "Improved battery SOC estimation accuracy using a modified UKF with an adaptive cell model under real EV operating conditions," *IEEE Trans. Transport. Electric.*, vol. 4, no. 2, pp. 408–417, Jun. 2018, doi: [10.1109/TTE.2018.2802043](https://doi.org/10.1109/TTE.2018.2802043).
- [14] K. Zhang, J. Ma, X. Zhao, D. Zhang, and Y. He, "State of charge estimation for lithium battery based on adaptively weighting cubature particle filter," *IEEE Access*, vol. 7, pp. 166657–166666, 2019, doi: [10.1109/ACCESS.2019.2953478](https://doi.org/10.1109/ACCESS.2019.2953478).
- [15] N. Chen, P. Zhang, J. Dai, and W. Gui, "Estimating the state-of-charge of lithium-ion battery using an H-Infinity observer based on electrochemical impedance model," *IEEE Access*, vol. 8, pp. 26872–26884, 2020, doi: [10.1109/ACCESS.2020.2971002](https://doi.org/10.1109/ACCESS.2020.2971002).
- [16] M. Cacciato, G. Nobile, G. Scarcella, and G. Scelba, "Real-time model-based estimation of SOC and SOH for energy storage systems," *IEEE Trans. Power Electron.*, vol. 32, no. 1, pp. 794–803, Jan. 2017, doi: [10.1109/TPEL.2016.2535321](https://doi.org/10.1109/TPEL.2016.2535321).
- [17] Z. Wei, S. Meng, B. Xiong, D. Ji, and K. J. Tseng, "Enhanced online model identification and state of charge estimation for lithium-ion battery with a FBCRLS based observer," *Appl. Energy*, vol. 181, pp. 332–341, Nov. 2016, doi: [10.1016/j.apenergy.2016.08.103](https://doi.org/10.1016/j.apenergy.2016.08.103).
- [18] Q. Zhong, F. Zhong, J. Cheng, H. Li, and S. Zhong, "State of charge estimation of lithium-ion batteries using fractional order sliding mode observer," *ISA Trans.*, vol. 66, pp. 448–459, Jan. 2017, doi: [10.1016/j.isatra.2016.09.017](https://doi.org/10.1016/j.isatra.2016.09.017).
- [19] D. N. T. How, M. A. Hannan, M. S. H. Lipu, and P. J. Ker, "State of charge estimation for lithium-ion batteries using model-based and data-driven methods: A review," *IEEE Access*, vol. 7, pp. 136116–136136, 2019, doi: [10.1109/ACCESS.2019.2942213](https://doi.org/10.1109/ACCESS.2019.2942213).
- [20] X. Hu, S. E. Li, and Y. Yang, "Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles," *IEEE Trans. Transport. Electric.*, vol. 2, no. 2, pp. 140–149, Jun. 2016, doi: [10.1109/TTE.2015.2512237](https://doi.org/10.1109/TTE.2015.2512237).
- [21] M. S. H. Lipu, M. A. Hannan, A. Hussain, and M. H. M. Saad, "Optimal BP neural network algorithm for state of charge estimation of lithium-ion battery using PSO with PCA feature selection," *J. Renew. Sustain. Energy*, vol. 9, no. 6, Nov. 2017, Art. no. 064102, doi: [10.1063/1.5008491](https://doi.org/10.1063/1.5008491).
- [22] X. Song, F. Yang, D. Wang, and K.-L. Tsui, "Combined CNN-LSTM network for state-of-charge estimation of lithium-ion batteries," *IEEE Access*, vol. 7, pp. 88894–88902, 2019, doi: [10.1109/ACCESS.2019.2926517](https://doi.org/10.1109/ACCESS.2019.2926517).

- [23] R. Li, S. Xu, S. Li, Y. Zhou, K. Zhou, X. Liu, and J. Yao, "State of charge prediction algorithm of lithium-ion battery based on PSO-SVR cross validation," *IEEE Access*, vol. 8, pp. 10234–10242, 2020, doi: [10.1109/ACCESS.2020.2964852](https://doi.org/10.1109/ACCESS.2020.2964852).
- [24] M. S. H. Lipu, M. A. Hannan, A. Hussain, M. H. Saad, A. Ayob, and M. N. Uddin, "Extreme learning machine model for state-of-charge estimation of lithium-ion battery using gravitational search algorithm," *IEEE Trans. Ind. Appl.*, vol. 55, no. 4, pp. 4225–4234, Jul. 2019, doi: [10.1109/TIA.2019.2902532](https://doi.org/10.1109/TIA.2019.2902532).
- [25] T. Mamo and F.-K. Wang, "Long short-term memory with attention mechanism for state of charge estimation of lithium-ion batteries," *IEEE Access*, vol. 8, pp. 94140–94151, 2020, doi: [10.1109/ACCESS.2020.2995656](https://doi.org/10.1109/ACCESS.2020.2995656).
- [26] C. Zhang, Y. Zhu, G. Dong, and J. Wei, "Data-driven lithium-ion battery states estimation using neural networks and particle filtering," *Int. J. Energy Res.*, vol. 43, pp. 8230–8241, Aug. 2019, doi: [10.1002/er.4820](https://doi.org/10.1002/er.4820).
- [27] M. A. Hannan, M. S. H. Lipu, A. Hussain, M. H. Saad, and A. Ayob, "Neural network approach for estimating state of charge of lithium-ion battery using backtracking search algorithm," *IEEE Access*, vol. 6, pp. 10069–10079, 2018, doi: [10.1109/ACCESS.2018.2797976](https://doi.org/10.1109/ACCESS.2018.2797976).
- [28] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," 2014, *arXiv:1409.1259*. [Online]. Available: <http://arxiv.org/abs/1409.1259>
- [29] S. Hochreiter and J. Schmidhuber, "Neural Computation," *Long Short-Term Memory*, vol. 9, pp. 1735–1780, Nov. 1997.
- [30] F. Zheng, Y. Xing, J. Jiang, B. Sun, J. Kim, and M. Pecht, "Influence of different open circuit voltage tests on state of charge online estimation for lithium-ion batteries," *Appl. Energy*, vol. 183, pp. 513–525, Dec. 2016, doi: [10.1016/j.apenergy.2016.09.010](https://doi.org/10.1016/j.apenergy.2016.09.010).
- [31] Z. Huang, F. Yang, F. Xu, X. Song, and K.-L. Tsui, "Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries," *IEEE Access*, vol. 7, pp. 93139–93149, 2019, doi: [10.1109/ACCESS.2019.2928037](https://doi.org/10.1109/ACCESS.2019.2928037).
- [32] F. Yang, W. Li, C. Li, and Q. Miao, "State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network," *Energy*, vol. 175, pp. 66–75, May 2019, doi: [10.1016/j.energy.2019.03.059](https://doi.org/10.1016/j.energy.2019.03.059).
- [33] J. Li, Z. Xu, J. Zhao, and C. Zhang, "Distributed online voltage control in active distribution networks considering PV curtailment," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5519–5530, Oct. 2019.
- [34] G. Dong, J. Wei, and Z. Chen, "Constrained Bayesian dual-filtering for state of charge estimation of lithium-ion batteries," *Int. J. Electr. Power Energy Syst.*, vol. 99, pp. 516–524, Jul. 2018, doi: [10.1016/j.ijepes.2018.02.005](https://doi.org/10.1016/j.ijepes.2018.02.005).



ZHEKANG DONG received the B.E. and M.E. degrees in electronics and information engineering from Southwest University, Chongqing, China, in 2012 and 2015, respectively, and the Ph.D. degree from the School of Electrical Engineering, Zhejiang University, China, in 2019. He is currently an Associate Professor with Hangzhou Dianzi University, Hangzhou, China. He is also a Research Assistant (Joint-Supervision) with The Hong Kong Polytechnic University. His research interests include memristor and memristive systems, artificial neural networks, the design and analysis of nonlinear systems based on memristor, and computer simulation.



HUIPIN LIN received the B.E. degree in electrical engineering from China Jiliang University, Hangzhou, China, in 2009, and the Ph.D. degree from the School of Electrical Engineering, Zhejiang University, China, in 2019. He is currently a Lecturer with Hangzhou Dianzi University, Hangzhou. His research interests include power electronics, LLC, inverter, industrial robot control, power system optimization, energy system modeling, battery balancing topology, and battery energy management systems.



ZHIWEI HE (Member, IEEE) received the B.E. degrees in information engineering and information systems from Zhejiang University, in 2001 and 2006. He is also a Research Assistant with The Hong Kong Polytechnic University, in 2004, and went to Wayne State University, as a Senior Visiting Scholar, in 2012. He is currently a Professor with Hangzhou Dianzi University, Hangzhou, China. He is the author of more than 80 articles and more than 50 inventions. His research interests include deep learning, computer vision, and media signal processing.



MINGHAO WANG (Member, IEEE) received the B.Eng. degree (Hons.) in electrical and electronic engineering from the Huazhong University of Science and Technology, Wuhan, China, and the University of Birmingham, Birmingham, U.K., in 2012, and the M.Sc. and Ph.D. degrees in electrical and electronic engineering from The University of Hong Kong, Hong Kong, in 2013 and 2017, respectively. Since 2018, he has been with the Department of Electrical Engineering, The Hong Kong Polytechnic University, where he is currently a Research Assistant Professor. His research interests include smart grid technologies, ac-dc power conversion, power systems, electric springs, and power electronics.



ZHAOWEI ZHANG (Graduate Student Member, IEEE) received the B.E. degree in electronic information engineering and the M.E. degree in agricultural engineering from Hebei Agricultural University, China, in 2017 and 2019, respectively. She is currently pursuing the D.Phil. degree in electronic science and technology with Hangzhou Dianzi University. Her research interests include deep learning and uses data analytics to research battery management system for electric vehicles.



YUFEI HE received the B.Eng. degree in electronic and information engineering from Zhejiang University, Hangzhou, China, in 2016, and the Ph.D. degree in electrical engineering from The Hong Kong Polytechnic University, Hong Kong, in 2020. He is currently a Postdoctoral Research Fellow with the Department of Electrical Engineering, The Hong Kong Polytechnic University. His research interest includes power electronic control for grid-integration of renewables.



XIANG GAO is currently pursuing the Ph.D. degree in electrical engineering with The Hong Kong Polytechnic University, Hong Kong. Her research interests include renewable energy resources, stochastic optimization, and electricity market.



MINGYU GAO (Member, IEEE) received the M.S. degree in power electronics from Zhejiang University, China, in 1993, and the Ph.D. degree in information and communication engineering from the Wuhan University of Technology, China, in 2013. He joined Hangzhou Dianzi University, China, in 2001, where he is currently a Full Professor with the School of Electronic and Information. His current research interests include the areas of deep learning and computer vision.

...