

Article

An Improved Gated Recurrent Unit Network Model for State-of-Charge Estimation of Lithium-Ion Battery

Wenxian Duan ¹, Chuanxue Song ^{1,2}, Silun Peng ¹, Feng Xiao ^{1,2,*}, Yulong Shao ³ and Shixin Song ⁴

¹ State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China; dwx342977542@126.com (W.D.); songchx@126.com (C.S.); pengsilun@jlu.edu.cn (S.P.)

² Taizhou Automobile Power Transmission Research Institute, Jilin University, Taizhou 210008, China

³ Zhengzhou Yutong Bus Co., Ltd., Zhengzhou 450016, China; shaoyl13@mails.jlu.edu.cn

⁴ School of Mechanical and Aerospace Engineering, Jilin University, Changchun 130022, China; songshx202@126.com

* Correspondence: xiaofengjl@jlu.edu.cn; Tel.: +86-137-5647-3659

Received: 23 October 2020; Accepted: 29 November 2020; Published: 3 December 2020



Abstract: An accurate state-of-charge (SOC) can not only provide a safe and reliable guarantee for the entirety of equipment but also extend the service life of the battery pack. Given that the chemical reaction inside the lithium-ion battery is a highly nonlinear dynamic system, obtaining an accurate SOC for the battery management system is very challenging. This paper proposed a gated recurrent unit recurrent neural network model with activation function layers (GRU-ATL) to estimate battery SOC. The model used deep learning technology to establish the nonlinear relationship between current, voltage, and temperature measurement signals and battery SOC. Then the online SOC estimation was carried out on different testing sets using the trained model. The experiments in this paper showed that the GRU-ATL network model could realize online SOC estimation under different working conditions without relying on an accurate battery model. Compared with the gated recurrent unit recurrent neural (GRU) network model and long short-term memory (LSTM) network model, the GRU-ATL network model had more stable and accurate SOC prediction performance. When the measurement data contained noise, the experimental results showed that the SOC prediction accuracy of GRU-ATL model was 0.1–0.4% higher than the GRU model and 0.3–0.7% higher than the LSTM model. The mean absolute error (MAE) of SOC predicted by the GRU-ATL model was stable in the range of 0.7–1.4%, and root mean square error (RMSE) was stable between 1.2–1.9%. The model still had high prediction accuracy and robustness, which could meet the SOC estimation in complex vehicle working conditions.

Keywords: state-of-charge; lithium-ion battery; gated recurrent unit; non-Gaussian noises; robustness

1. Introduction

With the development of science and technology, the world's dependence on energy is increasing year by year. At present, oil and coal are the main non-renewable energy sources, and the emission of waste gas will inevitably cause some serious environmental consequences [1]. To solve these problems, governments of various countries strongly advocate that electric vehicles replace fuel vehicles, helping to reduce the emission of harmful gases. Lithium-ion batteries have made considerable progress in the past decade [2]. Compared with other chemical materials, lithium batteries have the advantages of high energy density, long life, and high power. They have been widely used in mobile phones, computers, electric vehicles, and satellites [3]. Because the chemical reaction inside the battery is a

highly nonlinear dynamic system, it is a very challenging task to obtain an accurate state-of-charge (SOC) for the battery management system [4].

1.1. Literature Review

Contemporary methods are mainly based on current, voltage, and temperature measurement signals to obtain battery SOC value; these methods include the ampere-hour integration [5], open-circuit voltage [6], model-based [7], and data-driven methods [8,9]. The ampere hour integration method is mainly used to obtain the SOC value by integrating the current. However, inaccuracy of SOC initial value accumulates errors, which eventually leads to the estimated SOC value deviating from the actual SOC value. The open-circuit voltage method mainly establishes the relationship between SOC and open-circuit voltage by looking up the table or polynomial fitting. However, in order to obtain an accurate SOC value, this method requires the battery to be at rest for more than two hours. Therefore, the ampere-hour integration method and open-circuit voltage method are not suitable for online SOC estimation under complex dynamic conditions.

The electrochemical model uses some partial differential equations to establish the relationship between SOC and battery impedance or lithium concentration in a battery to obtain accurate SOC values [10,11]. However, this electrochemical model equation is excessively complex and involves many model parameters; thus, the accuracy of online SOC estimation is low, and it can only be applied to the offline battery design and performance analysis under laboratory conditions. At present, most of the research focuses on the method based on the equivalent circuit model, which includes the extended Kalman filter algorithm [12], unscented Kalman filter algorithm (UKF) [13], square root unscented Kalman filter algorithm [14], particle filter algorithm (PF) [15], and the H infinite filter [16]. The least square method with the forgetting factor is used to identify the parameters online to obtain more accurate SOC values [17,18]. Xia et al. [19] used the Kalman filter algorithm as a density function to improve the PF algorithm, which improved the prediction accuracy of the model. Although these model-based methods could obtain accurate SOC values online, their prediction performance was highly dependent on the accuracy of battery models. This kind of method requires designers to spend considerable time to select model parameters and noise covariance in the filter algorithm. In addition, they are carried out under the condition that the battery system is Gaussian noise by default.

The data-driven method does not need an accurate battery model. It uses the strong learning ability of the model to train a large number of data to get a network model which can reflect the relationship between battery measurement signals (such as current, voltage, temperature, and internal resistance) and SOC. Then, the network model is used to estimate the SOC on the unknown data set. Fotouhi et al. [20] proposed an adaptive neuro-fuzzy inference system combined with coulomb counting method for SOC estimation of lithium sulfur battery. However, the internal resistance and open circuit voltage were used as input eigenvectors, which were not easy to measure directly. Guo et al. [21] proposed an improved back propagation neural network (BPNN), which used the measured voltage, current, and temperature as input characteristics to predict the actual SOC. However, the convergence speed of the BPNN algorithm was slow, and it was very sensitive to the initial weight. Further, it was easy to appear gradient diffusion or fall into local minimum problem. In order to solve this problem and obtain an accurate SOC value, Wang et al. [22] used the artificial fish swarm algorithm to optimize some parameters of the BP neural network. However, this swarm intelligence optimization algorithm greatly increased the computational cost in the case of a large amount of data. With the development of computer hardware, some complex deep learning networks have been applied to battery SOC estimation. Deep learning technology can easily capture the complex and changeable relationship between the measurement signal and SOC through the deep neural network of multilayer nonlinear transformation. Chemali et al. [23] proposed a long short-term memory (LSTM) network to estimate the battery SOC based on signal characteristics such as current, voltage, and temperature. This method can complete SOC estimation under different working conditions without using the Kalman filter algorithm. In order to further improve the prediction accuracy of battery SOC, Bian et al. [24] used a bidirectional

long short-term memory (LSTM) neural network to predict battery SOC. However, the disadvantage of the LSTM network is that the decoding accuracy is seriously affected when the decoding begins without enough input sequence information. Vidal et al. [25] proposed a deep feed-forward neural network (FNN) approach for battery SOC estimation. The method still had a good SOC estimation performance when the measurement signal was added with errors. But this method needed to calculate the average value of the measured signal in a certain time step, and the average value in different time periods had a certain influence on the prediction results. Yang et al. [26] used the gated recurrent unit recurrent neural network (GRU-RNN) to predict the battery SOC using the measured current, voltage, and temperature signals. Jiao et al. [27] used the momentum gradient to optimize the weight of the network to improve the prediction accuracy of the GRU network model. In addition, some methods have combined neural networks and filter algorithms that also obtained accurate SOC estimation. Dong et al. [28] proposed a hybrid model of the wavelet neural network and particle filter algorithm and estimated the energy state of a battery under different working conditions. Tian et al. [29] proposed a method combining LSTM with an adaptive cubature Kalman filter for battery SOC estimation, and realized the accurate estimation of battery SOC with less data. Although this type of method can reduce the SOC estimation error to some extent, the final SOC estimation value excessively depends on the SOC value estimated by deep learning or neural network. In order to achieve good SOC estimation results, it also needs a lot of experiments to adjust the process noise covariance and measurement noise covariance.

1.2. The Contributions

Inspired by reference [25], this paper proposed a GRU-RNN network model with an activation function layer (GRU-ATL) for battery SOC estimation. This model mainly added a Tanh layer, a leaky rectified linear unit (ReLU) layer, and a clipped rectified linear unit (ReLU) layer to GRU-RNN network model. The relationship between current, voltage, temperature, and SOC was established in the training set. Then the online SOC estimation was carried out on different testing sets using the trained model. The main contributions of this paper are as follows:

- (1) Deep learning technology can solve the problem of low prediction accuracy of battery SOC by a traditional neural network. This technology does not need to establish an accurate battery equivalent circuit model. It can simplify the tedious parameter adjustment process based on the model method, and greatly save the time needed in the whole process of SOC estimation;
- (2) Compared with an LSTM network, a GRU network structure has the advantages of fewer parameters and a simple structure. It can save a lot of training and prediction time on the premise of ensuring the prediction accuracy of the model. Compared with FNN, LSTM, and GRU network models, it was found that the proposed model can obtain more accurate and stable SOC estimation results in different operating conditions. The Tanh activation function is a saturated activation function, which can enhance the nonlinear learning ability of neural network. The leaky ReLU and the clipped ReLU are unsaturated activation functions, which can solve the problem of gradient disappearance encountered in the neural network. The output of the neural network was limited to a certain area, which improved the prediction performance of the model. Adding the above three activation function layers in the GRU-RNN network can improve the prediction accuracy and robustness of the model;
- (3) The SOC prediction performance of LSTM, GRU, and GRU-ATL network models was compared when the measurement signals contain Gaussian noise and non-Gaussian noise. The experimental results showed that the SOC prediction accuracy of GRU-ATL model was 0.1–0.4% higher than GRU model and 0.3–0.7% higher than LSTM model;
- (4) The battery model obtained by the deep learning network can be used for SOC online prediction in different temperature working conditions. The GRU-ATL model still had high prediction

accuracy and robustness when the measurement data contains noise. Therefore, it could meet the complex dynamic conditions of vehicles.

1.3. The Organization of the Paper

The remainder of the paper is organized as follows. The second part introduces the basic principle of the GRU-ATL network model in detail. The third part describes the battery data used in the experiment and the evaluation index of the model. The fourth part discusses and analyzes in detail the SOC prediction performance of the GRU-ATL network model under various operating conditions. The fifth part is the conclusion of this paper.

2. GRU-FC Network Model

2.1. GRU Structural Unit

A GRU network was formed by adding a gating mechanism to a simple RNN network, which was used to control the transmission of information in the neural network. A GRU network could effectively capture the dependence of large step in time series, and solve the problem of gradient attenuation or explosion in long-term memory and back-propagation [30]. Compared with an LSTM network, the GRU network had the advantages of fewer parameters, a simpler structure, and higher computational efficiency, which were suitable for building larger networks [31,32]. At present, GRU networks have been proven effective in some application scenarios. A GRU network was mainly composed of the reset gate and update gate units. Its structural unit is shown in Figure 1.

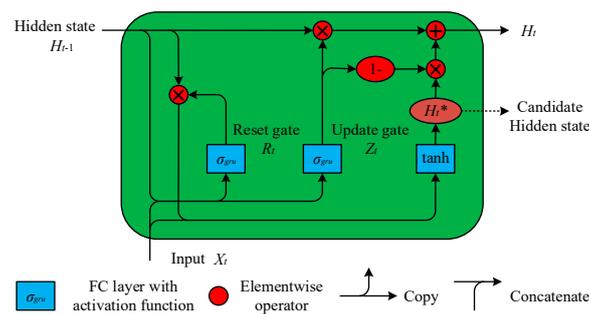


Figure 1. Gated recurrent unit (GRU) structural unit [32].

Figure 1 shows that the input of the reset and update gates of the gating unit were the hidden state H_{t-1} and the input X_t , and the output was calculated by the Sigmoid activation function. Suppose that the number of hidden cells is h , the small-batch input of time step t is $X_t \in \mathbb{R}^{n \times d}$ (the length of sample time is n , the number of eigenvectors is d) and the hidden state of the previous time step is $H_{t-1} \in \mathbb{R}^{n \times h}$. The calculation of reset gate $R_t \in \mathbb{R}^{n \times h}$ and update gate $Z_t \in \mathbb{R}^{n \times h}$ is as follows:

$$\begin{aligned} R_t &= \sigma_{gru}(X_t W_{xr} + H_{t-1} W_{hr} + b_r), \\ Z_t &= \sigma_{gru}(X_t W_{xz} + H_{t-1} W_{hz} + b_z), \end{aligned} \quad (1)$$

where $W_{xr}, W_{xz} \in \mathbb{R}^{d \times h}$, and $W_{hr}, W_{hz} \in \mathbb{R}^{h \times h}$ are network weight parameters. b_r and b_z are network bias parameters. σ_{gru} is a Sigmoid activation function. Its main function is to convert the values in both gates to 0–1.

The candidate hidden state $H_t^* \in \mathbb{R}^{n \times h}$ was obtained by a series of operations between the output of the reset gate and the hidden state H_{t-1} of the previous time step. When the element value in the reset gate was close to 0, then the hidden state H_{t-1} was discarded. When the element value was close to 1, the hidden state H_{t-1} was retained. The calculation method is as follows:

$$H_t^* = \tanh(X_t W_{xh} + (R_t \otimes H_{t-1}) W_{hh} + b_h) \quad (2)$$

where $W_{xh} \in \mathbb{R}^{d \times h}$ and $W_{hh} \in \mathbb{R}^{h \times h}$ are the network weight parameters and b_h is the network bias parameter. The function of the Tanh activation function was to convert all element ranges into between $[-1, 1]$.

In an LSTM network, the input and forget gates are complementary and have certain redundancy. By contrast, a GRU network directly uses an update gate to control the balance between the input and forgetting. The hidden state $H_t \in \mathbb{R}^{n \times h}$ of the current time moment is obtained by combining the hidden state H_{t-1} of the previous time moment and the candidate hidden state H_t^* of the current time through the update gate Z_t of the current time step.

$$H_t = Z_t \otimes H_{t-1} + (1 - Z_t) \otimes H_t^* \tag{3}$$

Equations (1)–(3) show that when $Z_t = 0$ and $R_t = 1$, a GRU network will degenerate into a simple RNN network. When $Z_t = 0$ and $R_t = 0$, the current state H_t is only related to the current input X_t , and is not involved with the historical state H_{t-1} . When $Z_t = 1$, the current state H_t is equal to the previous hidden state H_{t-1} .

2.2. SOC Estimation Based on GRU-ATL Network

As shown in Figure 2, a GRU-RNN [8,26] network model consists of a sequence input layer, a GRU network layer, a fully connected layer with neurons, a fully connected layer without neurons, and a regression output layer. The input of the GRU network model is composed of current, voltage, and temperature measurement signals, and the output is the battery SOC at the current time, namely $x_t = [I_t, V_t, T_t]$, $y_t = SOC_k$. In the training set, the nonlinear relationship between current, voltage, and temperature and SOC was established. The prediction accuracy of the model was verified in the test set. Inspired by reference [25], this paper proposed a GRU-RNN network model with activation function layer (GRU-ATL) for battery SOC estimation. As shown in Figure 2, this model mainly added a Tanh layer, a leaky ReLU layer, and a clipped ReLU layer to the GRU-RNN network model.

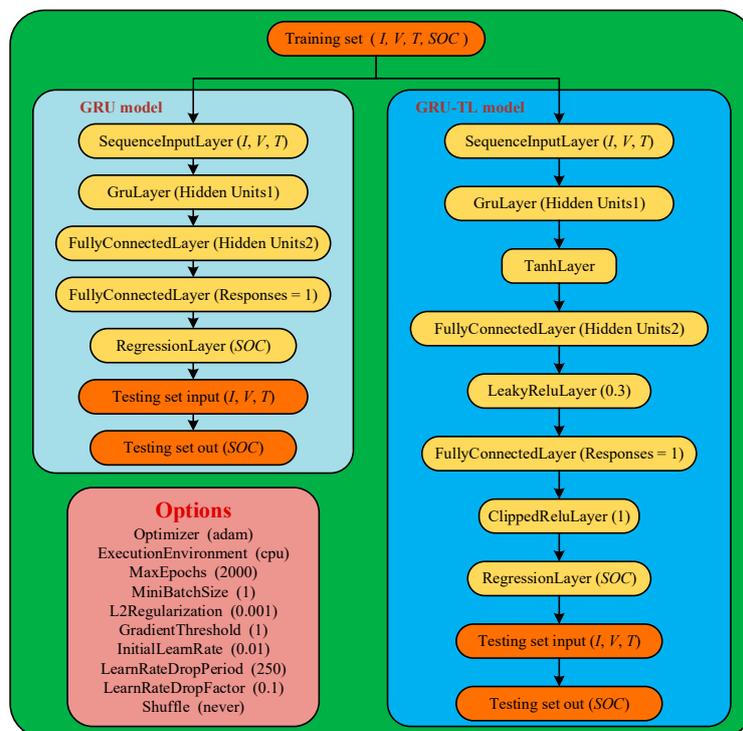


Figure 2. Schematic diagram of a GRU-RNN network model with an activation function layer (GRU-ATL) network structure.

The leaky ReLU layer played the role of execution threshold. It mainly multiplied any input value less than zero by a fixed factor coefficient [33], and its expression is as follows:

$$f(x) = \begin{cases} x & x \geq 0 \\ scale \cdot x & x < 0 \end{cases} \quad (4)$$

where $scale$ is the factor coefficient when the input x is negative.

The clipped ReLU layer also performed threshold operation. It mainly set any input value less than zero to zero, and any value higher than the clipping upper limit as the clipping upper limit [34]. The expression is as follows:

$$f(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < T_C \\ T_C & x \geq T_C \end{cases} \quad (5)$$

where T_C is the threshold of the clipped ReLU layer.

2.3. Selection of Other Parameters in Network

In deep learning networks, the common optimizers include the adaptive moment estimation (Adam) optimizer, stochastic gradient descent (SGD) optimizer, adaptive gradient (Adagrad) optimizer, root mean square prop (RMSProp) optimizer, and batch gradient descent (BGD) optimizer. The Adam optimizer was proposed by Kingma and Ba, and it combines the advantages of momentum and RMSProp [35]. The principle and implementation of the algorithm are simple, and the memory requirement is low. The updating of parameters is not affected by the scaling change of gradient. The super parameters have good interpretability, and the learning rate can be automatically adjusted without adjustment or little fine-tuning. The Adam optimizer has achieved good results in practical application, and it is suitable for gradient sparse or gradient noise problems.

The updating equation of the bias parameter θ_{adam} is as follows:

$$\theta_{adam,t} = \theta_{adam,t-1} - \alpha_{adma} \cdot m_t^* / (\sqrt{v_t^*} + \varepsilon) \quad (6)$$

where m_t^* is the bias-corrected first raw moment estimate, v_t^* is the bias-corrected second raw moment estimate, α_{adma} is the learning rate, and ε is a constant value.

The above formula that Adam algorithm can adjust adaptively from two aspects of gradient mean and gradient square. To prevent calculation errors, the constant value ε is set to 1×10^{-8} . The convergence speed of the network becomes very small when the α_{adma} is too small. The loss function of the network oscillates and even deviates from the minimum when the α_{adma} is too large. Therefore, the α_{adma}^0 is set as 0.01. The α_{adma} is attenuated once every 250 iterations with a magnification factor of 0.1.

In the deep learning network, if many parameters are to be set and the network model is too complex, the overfitting phenomenon occurs. In other words, the model performs well in the training set but poorly in the actual test set, and the model does not have good generalization ability. Therefore, an L_2 regularization algorithm is selected to avoid the overfitting phenomenon. L_2 regularization is mainly based on the original loss function plus the sum of squares of weight parameters [36]. The expression is as follows:

$$E_{L_2}(\theta_{adam}) = E(\theta_{adam}) + \frac{C_{L_2}}{2} \|W\|^2 \quad (7)$$

where, $E_{L_2}(\theta_{adam})$ is the original loss function, C_{L_2} is the regularization coefficient, $C_{L_2} = 0.001$, and W is the weight vector.

3. Battery Data Description and Processing Process

3.1. Battery Data Description

To verify the feasibility of the deep learning network model proposed in this paper, we chose the open-source battery test data set of McMaster University's McMaster Institute for Automotive Research and Technology as the research target [25]. This data set conducted charging and discharging test experiments on LGHG2 batteries under multiple different working conditions, effectively simulating the real driving environment of electric vehicles. Each battery test data contained measurement signals such as time, current, voltage, temperature, and capacity. The various parameters of LGHG2 battery are shown in Table 1.

Table 1. The main specifications of LG 18650HG2.

Parameter	Value
Nominal capacity/voltage	3.0 Ah/3.6 V
Charge and discharge cut-off voltage	4.2 V/2.5 V
Normal end-of-charge current	50 mA
Max continuous discharge current	20 A
Standard charge current	1.5 A
Energy density	240 Wh/Kg

The four standard discharge test conditions mainly include Urban Dynamometer Driving Schedule (UDDS), Highway Fuel Economy Driving Schedule (HWFET), LA92 Dynamometer Driving Schedule (LA92), and Supplemental Federal Test Procedure Driving Schedule (US06). The charging test condition is a fast constant current and constant voltage (CC-CV) charging mode. In other words, the constant current of 3A is used for charging. When the cut-off voltage reaches 4.2 V, the battery is charged in the constant voltage mode until the cut-off charging current is 0.05 A. In addition, the data set includes eight mixed dynamic test conditions, each of which is a random combination of the above four standard dynamic conditions. The test data of -10 , 0 , 10 , 25 , and 40 °C were used to verify the feasibility of the model. Figure 3 shows the test data of LA92 dynamic test conditions. The figure shows a highly nonlinear relationship between the battery capacity and the three measurement signals of current, voltage, and temperature during the discharge process, and the discharge capacity at 25 °C is significantly greater than that at -10 and 0 °C. This finding shows that with the decrease of temperature, the chemical reaction inside the battery became slow, and the actual capacity value gradually decreased. It is worth noting that some aging phenomena appear in the battery after several charge–discharge cycle tests. In the original dynamic working condition data set, the time unit of recording data was 0.1 s, and in the charging working condition, the time unit of recording data was 60 s. To reduce calculation costs, the data set recorded every 1 s in the dynamic test conditions was selected as the research object of this article.

3.2. Data Processing Process and Evaluation Indicators

As shown in Figure 3, the measured signals (current, voltage, and temperature) in the data set fluctuated greatly, which affected the results of data analysis. Normalizing the data could not only speed up the learning speed of the network, but also eliminate the adverse effects caused by singular sample data. In this paper, the mapminmax function was used to normalize the three measurement signals to $[0, 1]$. The expression is as follows:

$$x_{normal} = (x_{signal} - x_{min}) / (x_{max} - x_{min}) \quad (8)$$

where x_{min} and x_{max} are the minimum and maximum values in real measurement data; x_{signal} is the real measurement data; and x_{normal} is the normalized data.

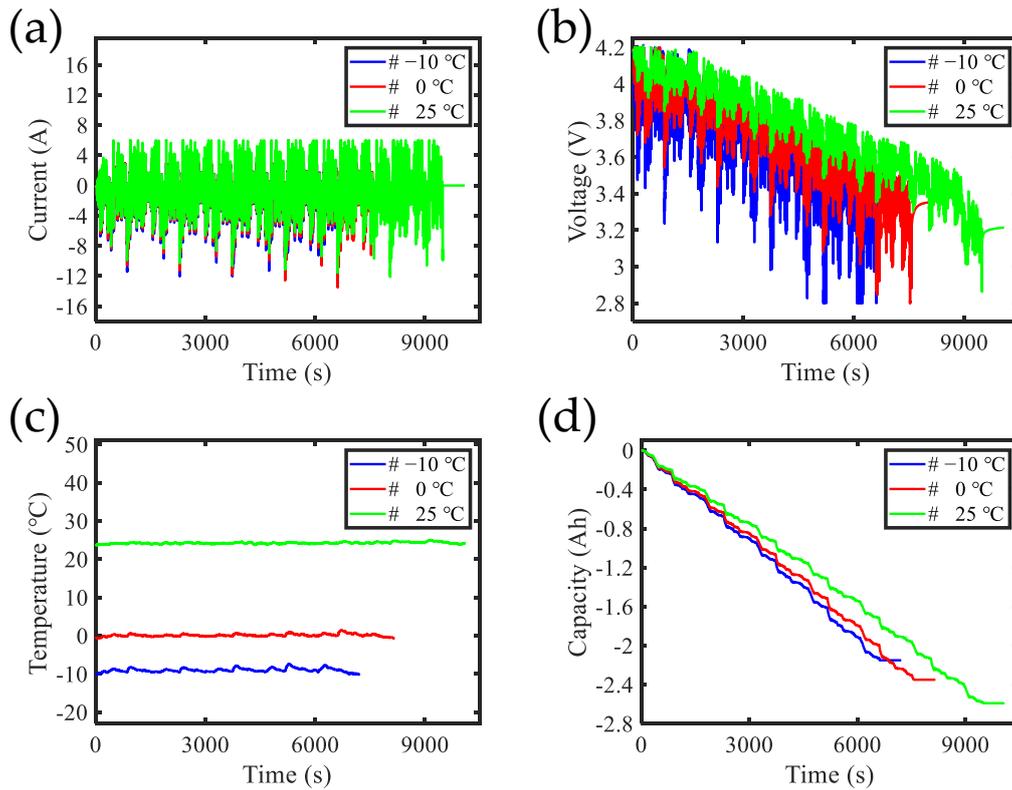


Figure 3. Test data of LA92 at three temperatures (−10, 0, and 25 °C); (a) current (b) voltage (c) temperature (d) discharge capacity.

The output of the GRU-ATL network model was the battery SOC, which was mainly obtained through the time integration of the current. The calculation process is as follows:

$$SOC(t) = SOC(t_0) - \int_{t_0}^t \eta I_t / (3600 \cdot C_N) d\tau \quad (9)$$

where $SOC(t_0)$ is the initial SOC value; $SOC(t)$ is the SOC value at current time t ; η is the coulomb rate, $\eta = 1$; I_t is the current flowing through the battery at time t ; and C_N is the actual capacity value of each working condition.

The data set was divided into a training set and five test sets. The GRU-ATL network model was used to obtain a training model on the training set. Then, the model was used to perform simulation verification on the test sets. It is worth noting that some conditions did not complete the charge–discharge experiment. For example, at 25 °C, the discharge time of HWFET was only 600 s. Therefore, this part of the data set was not added to the training set and test set. All the simulation experiments were carried out in the CPU simulation environment of MATLAB, and the computer processor was an Intel Core i5-7400 CPU@3.00 GHz with 8 GB memory. In each case, the experiment was carried out three times, and the average value of the three times was calculated as the final result. Finally, mean absolute error (MAE) and root mean square error (RMSE) were introduced to evaluate the prediction performance of the GRU-ATL network model:

$$MAE_{soc} = \frac{1}{n} \sum_{t=1}^n |SOC_{Acture} - SOC_{Estimated}| \quad (10)$$

$$RMSE_{soc} = \sqrt{\frac{1}{n} \sum_{t=1}^n (SOC_{Acture} - SOC_{Estimated})^2} \quad (11)$$

where MAE_{soc} is the mean absolute error of SOC; $RMSE_{soc}$ is the root mean square error of SOC; SOC_{Actual} is the actual SOC value; $SOC_{Estimated}$ is the predicted SOC value; and n is the sample size.

4. Experimental Results and Discussion

4.1. SOC Estimation Results of Four Network Models

In this experiment, three network models (FNN [25], LSTM, and GRU) were introduced to compare with the GRU-ATL network model. The training set was composed of eight mixed dynamic test conditions, CC-CV charging conditions, and four standard dynamic test conditions. The test set was composed of CC-CV charging conditions and four standard dynamic test conditions. Notably, the number of training sets was one and the number of test sets was five. The detailed division of the training set and test set is shown in Table 2. A large number of simulation experiments showed that the number of iterations had a certain impact on the SOC prediction accuracy of the four models. But this is beyond the scope of this article. According to reference [25], the maximum number of iterations in the FNN network model was 5500. In order to save computation while ensuring the accuracy, the maximum iteration times of the other three network models were set to 2000. The other parameters in the four network models were the same, and the number of neurons in each layer was 55.

Table 2. The detailed division of training set and test set.

Data Set	T (°C)	Condition
Training	−10, 0, 10, 25, 40	Mixed (1–8), CC-CV (3A), UDDS, HWFET, LA92, US06
Testing 1	−10	UDDS, HWFET, LA92, US06, CC-CV (3A)
Testing 2	0	UDDS, HWFET, LA92, US06, CC-CV (3A)
Testing 3	10	UDDS, HWFET, LA92, US06, CC-CV (3A)
Testing 4	25	UDDS, LA92, US06, CC-CV (3A)
Testing 5	40	UDDS, HWFET, LA92, US06, CC-CV (3A)

Figures 4 and 5 show the SOC prediction results of four network models at 0 and 25 °C, respectively. The figure shows that when the current, voltage, and temperature characteristics at time t were used to predict the SOC value at time t , the FNN network model had the worst prediction result, and the other three network models could effectively track the real SOC curve. Although the FNN network model needed the shortest time in the whole prediction process, the information in the network structure propagated in one direction, and there was no reverse information transmission. It led to the suggestion that the battery model could not reflect the relationship between the measurement signal and SOC very well. Therefore, the SOC error of the FNN network model was very large. From the SOC error curve of each working condition, we could see that the SOC error of the LSTM network model was larger than those of the GRU and GRU-ATL network models. The reason for this finding is that in the formula of network structure, the number of parameters of LSTM was four times that of simple RNN [37]. If the number of parameters was too much, the overfitting phenomenon would occur. However, GRU had only two gate switches, and the number of parameters was three times that of simple RNN [38]. Therefore, it could reduce the overfitting phenomenon, improve the prediction accuracy of the model, and save a considerable training time. The results in Table 3 show that the estimated SOC value of the GRU-ATL network model was more accurate than that of the GRU network structure under most dynamic conditions. Under the UDDS condition of 0 °C, the MAE and RMSE of the GRU network model were 1.2% and 2.3% respectively. In the UDDS and LA92 dynamic conditions, the GRU-ATL network model achieved good estimation results under different temperatures. Its MAE was less than 0.9%, and its RMSE was less than 1.2%. Owing to the large discharge current in US06 conditions, the chemical reaction inside the battery was severe, resulting in the larger error result of SOC compared with the previous two conditions. However, its MAE was also within 1.5%, and its RMSE was within 1.9%. The SOC prediction results of GRU model were more accurate than those of the GRU-ATL network model in the LA92 condition of 10 °C and US06 condition of 0 °C. It was

mainly because the difference of SOC prediction accuracy between GRU network model and GRU-ATL was not very obvious, and the SOC prediction results of GRU-ATL were inevitably lower than those of the GRU model in different time periods. However, the SOC prediction results of GRU-ATL were more accurate than those of GRU in the whole testing set at each temperature. In CC-CV conditions, the SOC prediction results of GRU-ATL model were more accurate than the other three models. But the SOC error results of the four network models were relatively large. The main reason is that the data acquisition unit was 60 s in CC-CV conditions, which did not provide enough data for network training. Through the analysis of the above results, the GRU-ATL network model had the highest accuracy and most stable SOC estimation among the four network models. This shows that the Tanh activation function, leaky ReLU activation function, and clipped ReLU activation function are helpful to improve the prediction accuracy and stability of GRU network model.

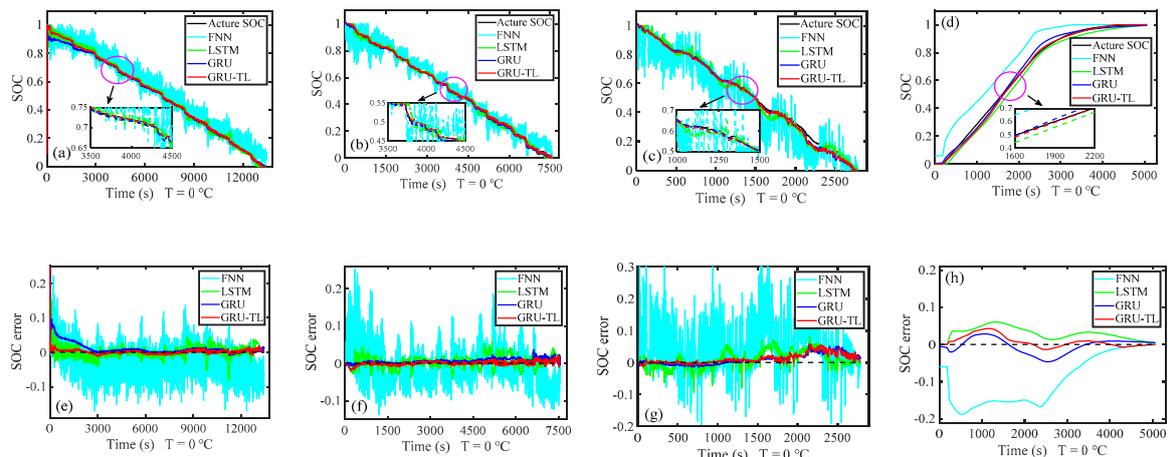


Figure 4. State-of-charge (SOC) and SOC error of four network models at 0 °C: (a,e) Urban Dynamometer Driving Schedule (UDDS), (b,f) LA92, (c,g) US06, (d,h) constant charge and constant voltage (CC-CV).

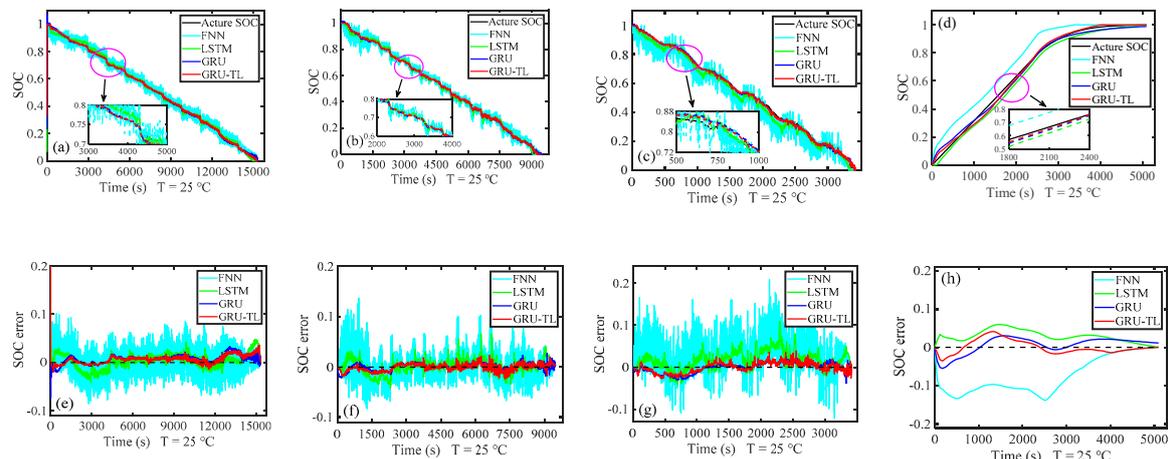


Figure 5. SOC and SOC error of four network models at 25 °C: (a,e) UDDS, (b,f) LA92, (c,g) US06, (d,h) CC-CV.

Table 3. SOC error results of four network models at five temperatures.

Network Model	T (°C)	UDDS		LA92		US06		CC-CV	
		MAE (%)	RMSE (%)						
FNN	−10	5.303	6.489	4.869	6.594	9.066	11.899	11.479	14.352
	0	4.041	4.821	3.437	4.507	7.519	9.632	9.475	11.597
	10	2.726	3.328	2.451	3.216	6.037	7.680	8.139	10.425
	25	1.612	2.078	1.784	2.404	4.305	5.490	7.427	8.827
	40	1.517	1.920	1.615	2.105	3.363	4.289	7.808	9.438
LSTM	−10	2.067	2.993	1.658	2.085	2.265	2.77	2.075	2.822
	0	1.256	2.471	1.271	1.613	2.216	2.834	2.868	3.313
	10	1.572	2.264	1.227	1.605	2.297	2.954	1.714	1.999
	25	1.261	2.017	1.080	1.416	2.225	2.787	2.869	3.249
	40	1.327	2.103	1.164	1.489	1.870	2.286	2.068	2.389
GRU	−10	0.558	0.899	0.454	0.551	1.587	1.724	3.165	4.387
	0	1.129	2.333	0.694	0.865	1.294	1.670	1.650	2.069
	10	0.702	1.314	0.461	0.554	1.474	1.842	1.409	1.901
	25	0.988	1.454	0.570	0.714	1.007	1.253	1.847	2.179
	40	0.612	1.046	0.620	0.800	0.989	1.321	1.604	2.072
GRU-ATL	−10	0.424	0.752	0.445	0.596	1.126	1.323	2.436	3.004
	0	0.406	1.161	0.361	0.500	1.324	1.898	1.142	1.685
	10	0.457	0.859	0.638	0.765	1.429	1.730	0.799	1.066
	25	0.814	1.165	0.526	0.652	0.978	1.179	1.407	1.763
	40	0.609	0.890	0.541	0.709	0.911	1.148	1.303	1.613

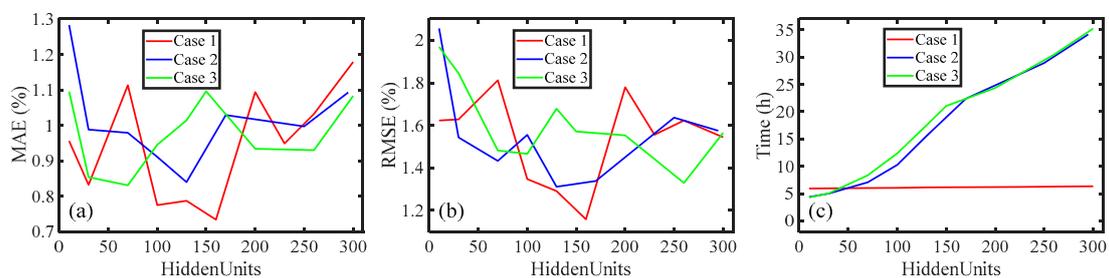
4.2. SOC Estimation Results under Unknown Conditions

In the process of driving, vehicles encounter a variety of complex working conditions. Therefore, to verify the prediction effect of the GRU-ATL network under unknown conditions, the training set used in this part was composed of eight mixed dynamic test conditions and CC-CV charging conditions. The testing sets consisted of CC-CV charging conditions and four standard dynamic test conditions. In other words, the testing sets were not added to the training set to train together. In addition, this part mainly discusses the influence of the number of neurons in the two network layers on the prediction results. The number of neurons in the two network layers had a certain influence on the final prediction results. However, no scientific and effective method to effectively select the number of neurons has been developed. The swarm intelligence optimization algorithm or evolutionary algorithm can find the best number of neurons in the case of minimal data to achieve the best prediction performance of the model. However, these algorithms need large amounts of calculation in some cases with large amounts of data; thus, they are not suitable for this paper. We then used three cases as examples to analyze the influence of the number of neurons on the accuracy of the model. In addition to the two evaluation indexes introduced in Section 3.2, the training time was also used as the evaluation index of the model. Case 1 indicated that the number of neurons in the GRU layer (N_{GRU}) was 55, and the number of neurons in the FC layer (N_{FC}) increased according to the number in Table 4. Case 2 indicated that the N_{FC} in the FC layer was 55, and the N_{GRU} in the GRU layer increased according to the number in Table 4. Case 3 indicated that the N_{GRU} and N_{FC} in both layers increased according to the number in Table 4.

Under the 25 °C condition, the indexes under three conditions are shown in Figure 6, and the statistical results of each index are listed in Table 4. Figure 6a,b shows that with the increase of N_{GRU} and N_{FC} , the estimation accuracy of SOC had a certain degree of improvement. When the number of neurons increased to a certain value, the estimation accuracy tended to decline. However, no mathematical relationship was found between the number of neurons and the accuracy of the model. Table 4 shows that in cases 2 and 3, with the increase of N_{GRU} in the GRU layer, SOC estimation in some cases did not achieve ideal results. The reason for this finding was that too many neurons resulted in the underfitting phenomenon in the model, eventually leading to the failure of the GRU-ATL network model in SOC prediction. As can be seen from Figure 6c, under the conditions of cases 2 and 3, the training time doubled with the increase of N_{GRU} in the GRU layer. By contrast, the training time did not increase significantly with the increase of N_{FC} in the FC layer.

Table 4. Model evaluation indexes under three conditions at 25 °C.

Case 1	N_{FC}	10	30	70	100	130	160	200	230	260	300
	MAE (%)	0.956	0.832	1.113	0.775	0.787	0.734	1.094	0.949	1.032	1.179
	RMSE (%)	1.623	1.628	1.812	1.347	1.291	1.158	1.779	1.556	1.623	1.544
	Time (h)	5.92	5.95	6.01	6.03	6.13	6.15	6.18	6.21	6.26	6.33
Case 2	N_{GRU}	10	30	70	100	130	160	200	230	260	295
	MAE (%)	1.283	0.988	0.979	0.911	0.840	1.029	NA	0.997	NA	1.093
	RMSE (%)	2.056	1.543	1.433	1.555	1.311	1.338	NA	1.636	NA	1.575
	Time (h)	4.37	4.98	7.05	10.26	15.55	22.38	NA	28.93	NA	34.16
Case 3	N_{GRU} and N_{FC}	10	30	70	100	130	160	200	230	260	300
	MAE (%)	1.095	0.854	0.831	0.945	1.014	1.096	0.934	NA	0.930	1.083
	RMSE (%)	1.969	1.845	1.482	1.466	1.678	1.571	1.553	NA	1.329	1.565
	Time (h)	4.27	5.01	8.36	12.37	17.52	21.05	24.37	NA	30.45	35.25

**Figure 6.** The evaluation index values in three cases: (a) Mean absolute error (MAE), (b) root mean square error (RMSE), (c) training time.

A large number of simulation experiments showed that when the N_{GRU} in the GRU layer is 55 and the N_{FC} in the FC layer was 160, the proposed network model had higher accuracy for SOC estimation and saved considerable training time. Figure 7 shows the SOC estimation results at 25 °C. According to the enlarged layout of the three dynamic conditions in Figure 7c–e, the predicted SOC curve could effectively track the real SOC curve. Table 5 shows the statistical results of SOC error at five temperatures. At 0, 10, 25, and 40 °C, the MAE of SOC error was within 1.1%, and the RMSE was within 1.5%. At −10 °C, the error index of SOC was relatively high, the MAE was 1.271%, and the RMSE is 2.005%. The main reason for this finding is that the amount of test data at −10 °C was small, and the chemical reaction of the battery at low temperature also had a certain effect on the quality of the data set. From the analysis of the above experimental results, the GRU-ATL network model could obtain accurate SOC estimation results under unknown conditions.

4.3. SOC Estimation Results with Noise

In the working process of a battery management system, owing to the complex and changeable external environment and the poor accuracy of the signal acquisition sensor, the collected current, voltage, and temperature signals usually contain certain measurement errors. Gaussian and non-Gaussian noises were added to the measured signals to test the prediction performance of the GRU-ATL network model.

As shown in Figure 8a, Gaussian distribution noise (*Noise 1*) is white Gaussian noise with a mean value of 0 and a standard deviation of 0.02. Its calculation formula is as follows:

$$Noise_1 = \alpha_{Noise1} \cdot randn(1, n) \quad (12)$$

where α_{Noise1} is the coefficient of standard deviation, $\alpha_{Noise1} = 0.02$, and n is the length of prediction data.

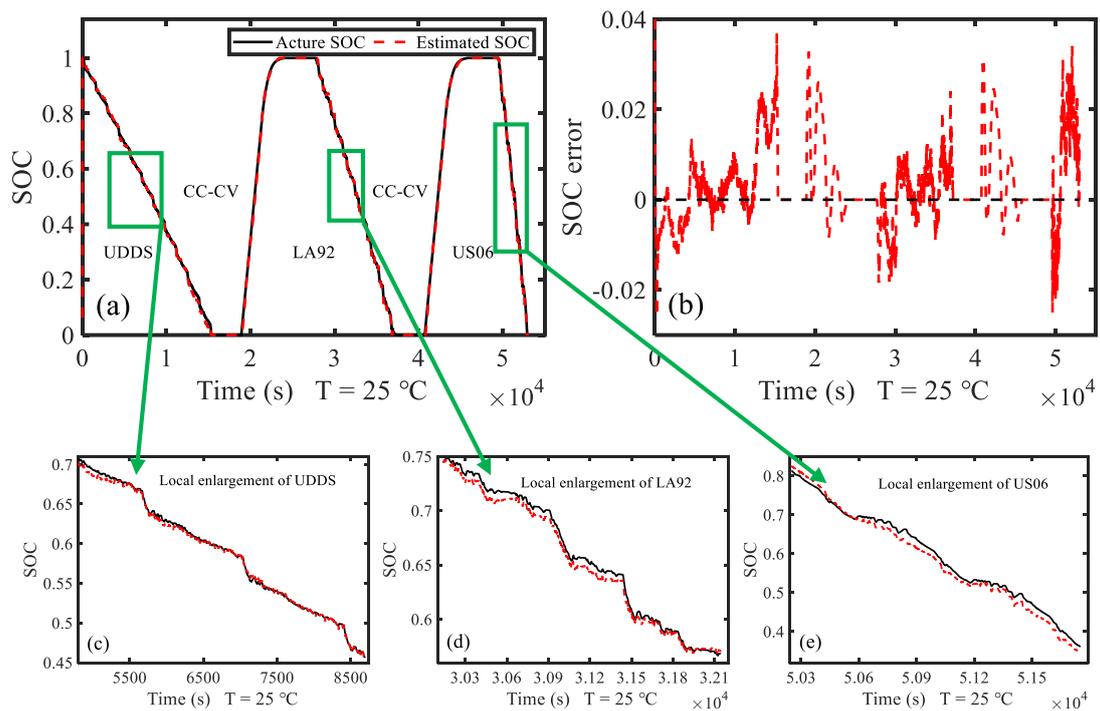


Figure 7. SOC estimation error of case 1 at 25 °C: (a) SOC, (b) SOC error, (c) local enlargement of UDDS, (d) local enlargement of LA92, (e) local enlargement of US06.

Table 5. SOC error results of case 1 ($N_{GRU} = 55, N_{FC} = 160$).

T (°C)	-10	0	10	25	40
MAE (%)	1.271	0.814	1.032	0.734	0.799
RMSE (%)	2.005	1.101	1.461	1.158	1.259

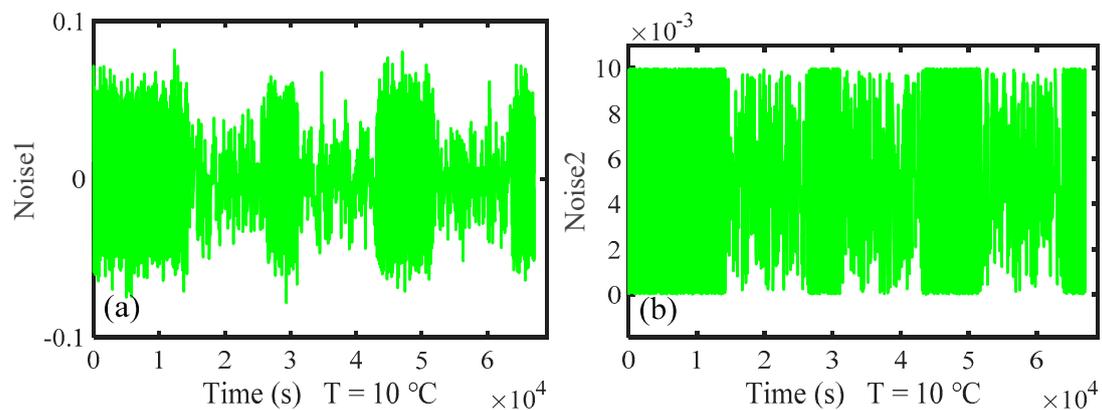


Figure 8. Two kinds of random noise: (a) Gaussian distributed noise, (b) uniformly distributed noise.

As shown in Figure 8b, *Noise 2* was a uniformly distributed random noise between [0, 0.01]. *Noise 3* is non-Gaussian noise, which is the sum of *Noise 1* and *Noise 2*. The expression is as follows:

$$Noise_3 = Noise_1 + \alpha_{Noise2} \cdot rand(1, n) \tag{13}$$

where α_{Noise2} is the coefficient of uniformly distributed noise, $\alpha_{Noise2} = 0.01$.

To further study the influence of measurement signal noise on the prediction accuracy of the model, the current with *Noise 1* and *Noise 3* was labeled as case 1, the voltage with *Noise 1* and *Noise 3* was labeled as case 2, the temperature with *Noise 1* and *Noise 3* was labeled as case 3, and the

three measurement signals with *Noise 1* and *Noise 3* were labeled as case 4. The Gaussian noise and non-Gaussian noise of temperature are 25 times higher than those of *Noise 1* and *Noise 3*, respectively. For example, when *Noise 1* was 0.05, the Gaussian noise of current was 0.05 A, the Gaussian noise of voltage was 0.05 V, and the Gaussian noise of temperature was 1.25 °C. The training set used in the experiment was the same as that in Section 4.2. The test set was the data set with different noise added to the measured signals at each temperature. That is, the number of training set was 1 and the number of testing set was 40. In the GRU-FC network model, the N_{GRU} in the GRU layer was set to 55, the N_{FC} in the FC layer was set to 160, and the number of iterations was set to 2000. Other parameters were consistent with those in Section 2.3. At the same time, the LSTM network model and the GRU model were introduced for comparison.

Figure 9 shows the SOC prediction results of LSTM model with Gaussian noise at 10 °C. Figure 10 shows the SOC prediction results of GRU-ATL model with Gaussian noise at 10 °C. Figure 11 shows the SOC prediction results of a simple GRU model with non-Gaussian noise at 40 °C. Figure 12 shows the SOC prediction results of GRU-ATL model with non-Gaussian noise at 40 °C. As can be seen from Figures 9–12, the SOC prediction curves of three network models well tracked the actual SOC curve in four cases. Table 6 shows the statistical results of SOC error under each condition of the three models. According to the statistical results, the SOC prediction accuracy of GRU-ATL model was 0.1–0.4% higher than the GRU model and 0.3–0.7% higher than the LSTM model. The MAE of the SOC predicted by GRU-ATL model was stable in the range of 0.7–1.4%, and RMSE was stable between 1.2–1.9%. In some cases of -10°C , the MAE and RMSE of SOC predicted by the GRU model were 1.7% and 2.5%, respectively, and the MAE and RMSE of SOC predicted by LSTM model were 1.8% and 2.7%, respectively. This shows that the SOC prediction results of GRU-ATL model were more stable and accurate at a low temperature. Compared with LSTM and GRU network models, GRU-ATL network model had better prediction accuracy and stronger robustness in unknown conditions with noise.

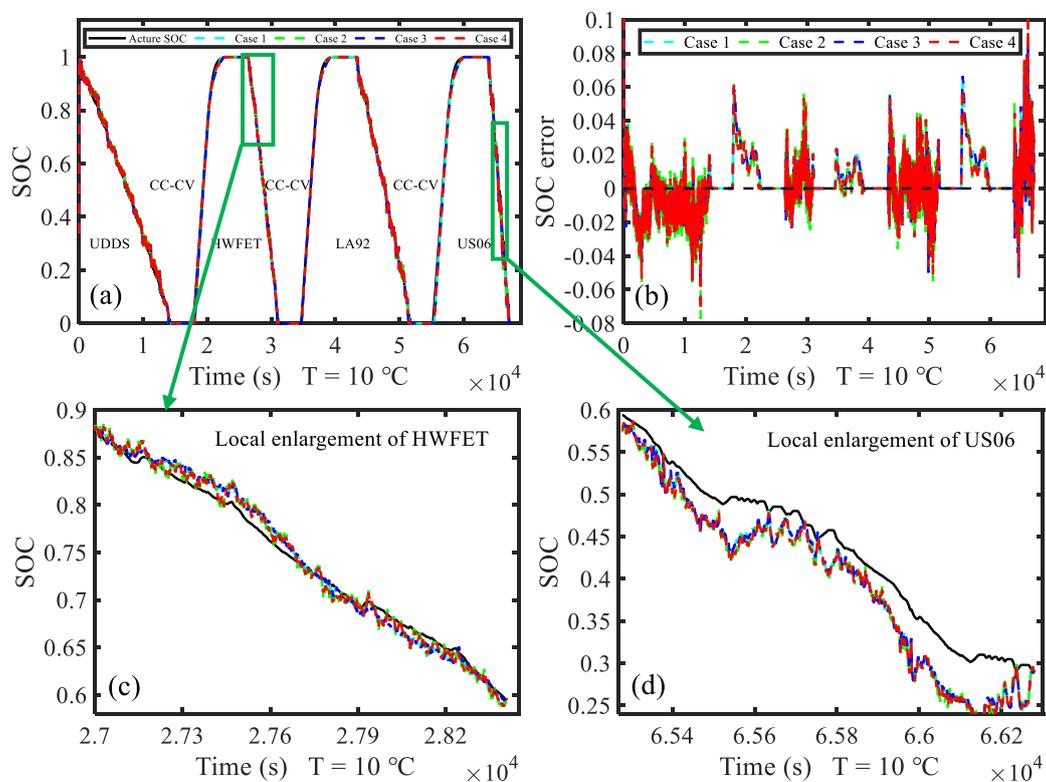


Figure 9. SOC and SOC error of long short-term memory (LSTM) model in four cases at 10 °C: (a) SOC, (b) SOC error, (c) local enlargement of HWFET, (d) local enlargement of US06.

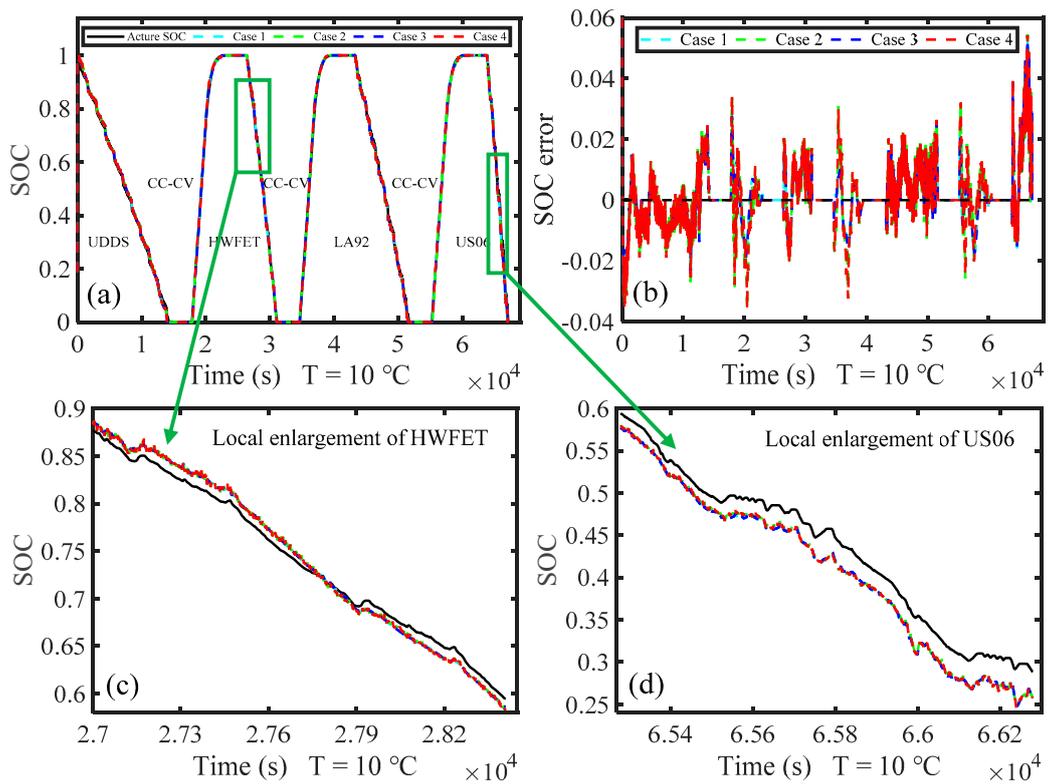


Figure 10. SOC and SOC error of GRU-ATL model in four cases at 10 °C: (a) SOC, (b) SOC error, (c) local enlargement of HWFET, (d) local enlargement of US06.

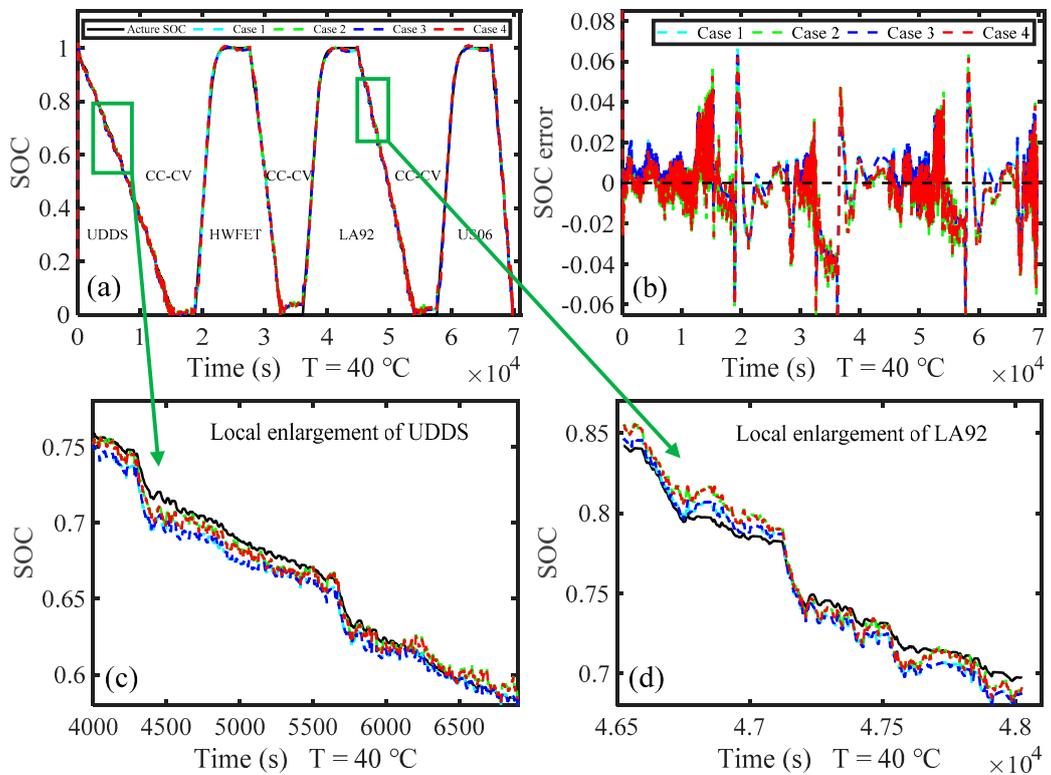


Figure 11. SOC and SOC error of GRU model in four cases at 40 °C: (a) SOC, (b) SOC error, (c) local enlargement of UDDS, (d) local enlargement of LA92.

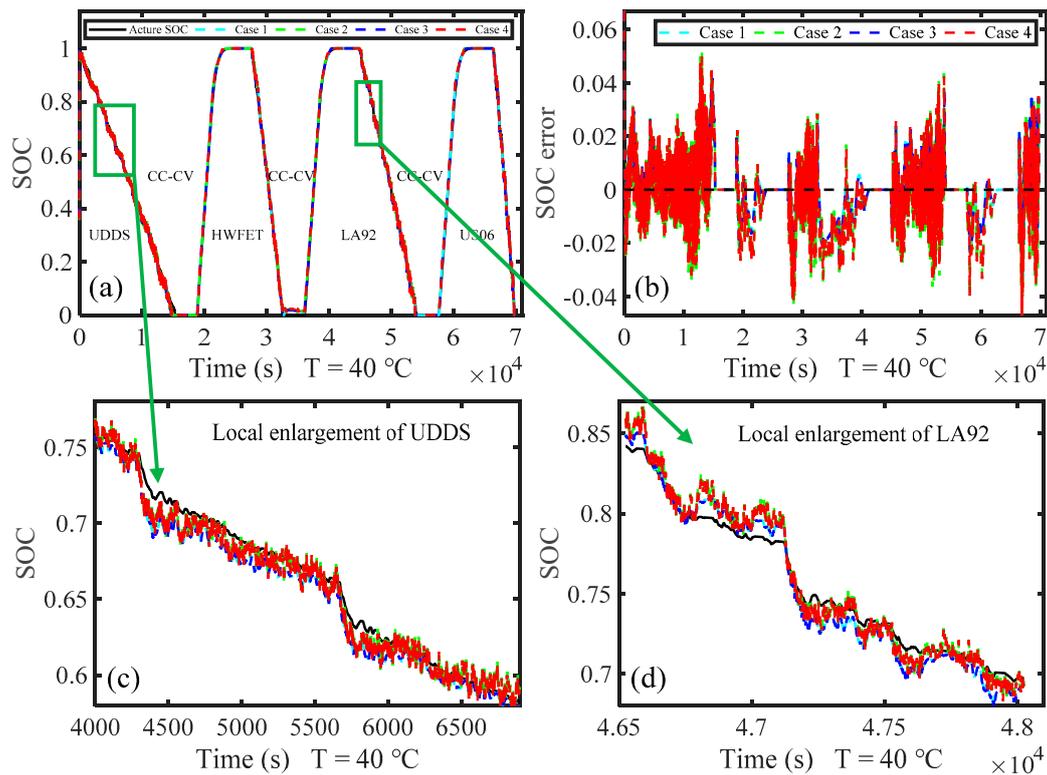


Figure 12. SOC and SOC error of GRU-ATL model in four cases at 40 °C: (a) SOC, (b) SOC error, (c) local enlargement of UDDS, (d) local enlargement of LA92.

Table 6. The statistical results of SOC error in four cases.

Network Model		T (°C)	Noise 1					Noise 3				
			-10	0	10	25	40	-10	0	10	25	40
LSTM	Case 1	MAE (%)	1.755	1.293	1.202	1.195	1.157	1.738	1.306	1.197	1.201	1.160
		RMSE (%)	2.369	1.879	1.695	1.751	1.689	2.358	1.894	1.693	1.754	1.691
	Case 2	MAE (%)	1.820	1.364	1.295	1.288	1.258	1.738	1.357	1.457	1.284	1.284
		RMSE (%)	2.449	1.943	1.815	1.195	1.782	2.669	1.878	1.962	1.853	1.813
	Case 3	MAE (%)	1.756	1.301	1.206	1.198	1.161	1.750	1.317	1.210	1.201	1.168
		RMSE (%)	2.370	1.886	1.699	1.753	1.692	2.372	1.911	1.704	1.755	1.670
	Case 4	MAE (%)	1.788	1.335	1.265	1.266	1.234	1.981	1.317	1.414	1.251	1.255
		RMSE (%)	2.410	1.915	1.780	1.815	1.759	2.572	1.850	1.913	1.823	1.786
GRU	Case 1	MAE (%)	1.687	1.277	1.169	1.272	1.010	1.687	1.295	1.165	1.281	1.015
		RMSE (%)	2.440	1.808	1.610	1.826	1.411	2.441	1.828	1.608	1.835	1.417
	Case 2	MAE (%)	1.697	1.299	1.180	1.305	1.113	1.687	1.160	1.410	1.250	1.062
		RMSE (%)	2.469	1.828	1.638	1.860	1.506	2.494	1.661	1.832	1.805	1.439
	Case 3	MAE (%)	1.692	1.278	1.169	1.272	1.011	1.700	1.297	1.168	1.273	1.019
		RMSE (%)	2.449	1.808	1.611	1.827	1.411	2.450	1.829	1.609	1.825	1.421
	Case 4	MAE (%)	1.697	1.298	1.176	1.299	1.089	1.700	1.176	1.388	1.238	1.035
		RMSE (%)	2.473	1.826	1.634	1.853	1.479	2.473	1.679	1.810	1.795	1.411
GRU-ATL	Case 1	MAE (%)	1.249	0.697	0.998	0.891	0.937	1.247	0.696	1.009	0.901	0.947
		RMSE (%)	1.859	1.155	1.409	1.248	1.273	1.863	1.159	1.422	1.257	1.282
	Case 2	MAE (%)	1.265	0.718	1.038	0.924	1.055	1.247	0.887	0.909	0.828	0.906
		RMSE (%)	1.888	1.171	1.45	1.288	1.413	1.865	1.235	1.278	1.197	1.275
	Case 3	MAE (%)	1.251	0.698	1.002	0.893	0.940	1.264	0.705	1.010	0.890	0.960
		RMSE (%)	1.864	1.158	1.413	1.250	1.278	1.878	1.175	1.422	1.247	1.294
	Case 4	MAE (%)	1.268	0.722	1.040	0.928	1.042	1.359	0.855	0.910	0.828	0.893
		RMSE (%)	1.892	1.176	1.450	1.292	1.395	1.865	1.218	1.281	1.195	1.259

5. Conclusions

This paper proposed a gated recurrent unit recurrent neural network model (GRU-RNN) with an activation function layer (GRU-ATL) for battery SOC estimation. This model mainly added a Tanh layer, a leaky rectified linear unit (ReLU) layer, and a clipped rectified linear unit (ReLU) layer to GRU-RNN network model. The relationship between current, voltage, temperature, and SOC was established in the training set. Then the online SOC estimation was carried out on different testing sets using the trained model. The main work of this paper are as follows:

- (1) Compared with an LSTM network, a GRU network structure has the advantages of fewer parameters and simple structure. It can save a lot of training and prediction time on the premise of ensuring the prediction accuracy of the model. Compared with FNN, LSTM, and GRU network models, it was found that the proposed model could obtain more accurate and stable SOC estimation results in different operating conditions. Adding the above three activation function layers in the GRU-RNN network could improve the prediction accuracy and robustness of the model;
- (2) The prediction accuracy of the model could be improved by appropriately increasing the number of neurons in the GRU layer and FC layer. But the excessive number of neurons in the two layers caused an over fitting phenomenon, which affected the SOC estimation accuracy. In order to save computation, the number of neurons in the GRU layer was 2–3 times less than that in the FC layer. A large number of experiments showed that the SOC was more accurate when the number of neurons in the GRU layer was 55 and that of the FC layer was 160. The MAE was less than 1.3% and RMSE was less than 2%;
- (3) The SOC prediction performance of LSTM, GRU, and GRU-ATL network models was compared when the measurement signals contained Gaussian noise and non-Gaussian noise. The experimental results showed that the SOC prediction accuracy of GRU-ATL model was 0.1–0.4% higher than the GRU model and 0.3–0.7% higher than the LSTM model. The MAE of SOC predicted by the GRU-ATL model was stable in the range of 0.7–1.4%, and the RMSE was stable between 1.2–1.9%. The SOC prediction results of the GRU-ATL network model were still more accurate and stable at a low temperature;
- (4) The experiments in this paper showed that the GRU-ATL network model could realize online SOC estimation under different working conditions without relying on an accurate battery model. When the measurement data contained noise, the model still had high prediction accuracy and robustness, which could meet the SOC estimation in complex vehicle working conditions.

The future work was mainly to use the model proposed in this paper to verify SOC estimation on other battery data sets. The SOC value used in the experiment was calculated according to the actual capacity. It was difficult to obtain the actual capacity value due to the aging phenomenon of the battery. Therefore, how to use deep learning networks to predict the actual capacity is also an important task in the future.

Author Contributions: Conceptualization, W.D. and F.X.; methodology, W.D. and C.S.; software, S.P. and S.S.; validation, S.P., S.S. and Y.S.; formal analysis, S.P. and Y.S.; writing—original draft preparation, W.D. and F.X.; writing—review and editing, W.D. and C.S.; project administration, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Tackling Item in Science and Technology Department of Jilin Province, China Grant number [20150204017GX], Jilin Provincial Natural Science Foundation, China Grant number [20150101037JC].

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

References

1. Berecibar, M.; Gandiaga, I.; Villarreal, I.; Omar, N.; Van Mierlo, J.; Van den Bossche, P. Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renew. Sustain. Energy Rev.* **2016**, *56*, 572–587. [[CrossRef](#)]
2. Lipu, M.S.H.; Hannan, M.A.; Hussain, A.; Hoque, M.M.; Ker, P.J.; Saad, M.H.M.; Ayob, A. A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations. *J. Clean. Prod.* **2018**, *205*, 115–133. [[CrossRef](#)]
3. Zhang, C.; He, Y.; Yuan, L.; Xiang, S. Capacity prognostics of lithium-ion batteries using EMD denoising and multiple kernel RVM. *IEEE Access* **2017**, *5*, 12061–12070. [[CrossRef](#)]
4. Shrivastava, P.; Soon, T.K.; Idris, M.Y.I.B.; Mekhilef, S. Overview of model-based online state-of-charge estimation using Kalman filter family for lithium-ion batteries. *Renew. Sustain. Energy Rev.* **2019**, *113*, 109233. [[CrossRef](#)]
5. Ng, K.S.; Moo, C.; Chen, Y.; Hsieh, Y. Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Appl. Energy* **2009**, *86*, 1506–1511. [[CrossRef](#)]
6. Petzl, M.; Danzer, M.A. Advancements in OCV Measurement and Analysis for Lithium-Ion Batteries. *IEEE Trans. Energy Convers.* **2013**, *28*, 675–681. [[CrossRef](#)]
7. Wang, Q.; Wang, J.; Zhao, P.; Kang, J.; Yan, F.; Du, C. Correlation between the model accuracy and model-based SOC estimation. *Electrochim. Acta.* **2017**, *228*, 146–159. [[CrossRef](#)]
8. Li, C.; Xiao, F.; Fan, Y. An approach to state of charge estimation of lithium-ion batteries based on recurrent neural networks with gated recurrent unit. *Energies* **2019**, *12*, 1592. [[CrossRef](#)]
9. Yang, F.; Song, X.; Xu, F.; Tsui, K. State-of-charge estimation of lithium-ion batteries via long short-term memory network. *IEEE Access* **2019**, *7*, 53792–53799. [[CrossRef](#)]
10. Kim, M.; Chun, H.; Kim, J.; Kim, K.; Yu, J.; Kim, T.; Han, S. Data-efficient parameter identification of electrochemical lithium-ion battery model using deep Bayesian harmony search. *Appl. Energy* **2019**, *254*, 113644. [[CrossRef](#)]
11. Bartlett, A.; Marcicki, J.; Onori, S.; Rizzoni, G.; Yang, X.G.; Miller, T. Electrochemical model-based state of charge and capacity estimation for a composite electrode lithium-ion battery. *IEEE Trans. Contr. Syst. Technol.* **2015**, *24*, 384–399. [[CrossRef](#)]
12. Xiong, R.; Sun, F.; Chen, Z.; He, H. A data-driven multi-scale extended Kalman filtering based parameter and state estimation approach of lithium-ion polymer battery in electric vehicles. *Appl. Energy* **2014**, *113*, 463–476. [[CrossRef](#)]
13. Partovibakhsh, M.; Liu, G. An adaptive unscented kalman filtering approach for online estimation of model parameters and state-of-charge of lithium-ion batteries for autonomous mobile robots. *IEEE Trans. Contr. Syst. Technol.* **2015**, *23*, 357–363. [[CrossRef](#)]
14. Wang, S.; Fernandez, C.; Cao, W.; Zou, C.; Yu, C.; Li, X. An adaptive working state iterative calculation method of the power battery by using the improved Kalman filtering algorithm and considering the relaxation effect. *J. Power Sources* **2019**, *428*, 67–75. [[CrossRef](#)]
15. Ye, M.; Guo, H.; Cao, B. A model-based adaptive state of charge estimator for a lithium-ion battery using an improved adaptive particle filter. *Appl. Energy* **2017**, *190*, 740–748. [[CrossRef](#)]
16. Chen, C.; Xiong, R.; Shen, W. A lithium-ion battery-in-the-loop approach to test and validate multiscale dual h infinity filters for state-of-charge and capacity estimation. *IEEE Trans. Power Electron.* **2018**, *33*, 332–342. [[CrossRef](#)]
17. Dai, H.; Xu, T.; Zhu, L.; Wei, X.; Sun, Z. Adaptive model parameter identification for large capacity Li-ion batteries on separated time scales. *Appl. Energy* **2016**, *184*, 119–131. [[CrossRef](#)]
18. Yang, H.; Sun, X.; An, Y.; Zhang, X.; Wei, T.; Ma, Y. Online parameters identification and state of charge estimation for lithium-ion capacitor based on improved Cubature Kalman filter. *J. Energy Storage* **2019**, *24*, 100810. [[CrossRef](#)]
19. Xia, B.; Sun, Z.; Zhang, R.; Cui, D.; Lao, Z.; Wang, W.; Sun, W.; Lai, Y.; Wang, M. A comparative study of three improved algorithms based on particle filter algorithms in soc estimation of lithium ion batteries. *Energies* **2017**, *10*, 1149. [[CrossRef](#)]
20. Fotouhi, A.; Auger, D.J.; Propp, K.; Longo, S. Lithium–sulfur battery state-of-charge observability analysis and estimation. *IEEE Trans. Power Electr.* **2018**, *33*, 5847–5859. [[CrossRef](#)]

21. Guo, Y.; Zhao, Z.; Huang, L. Soc estimation of lithium battery based on improved bp neural network. *Energy Procedia* **2017**, *105*, 4153–4158. [CrossRef]
22. Wang, Q.; Wu, P.; Lian, J. SOC estimation algorithm of power lithium battery based on AFSA-BP neural network. *J. Eng.* **2020**, *13*, 535–539. [CrossRef]
23. Chemali, E.; Kollmeyer, P.J.; Preindl, M.; Ahmed, R.; Emadi, A. Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries. *IEEE Trans. Ind. Electron.* **2018**, *65*, 6730–6739. [CrossRef]
24. Bian, C.; He, H.; Yang, S. Stacked bidirectional long short-term memory networks for state-of-charge estimation of lithium-ion batteries. *Energy* **2020**, *191*, 116538. [CrossRef]
25. Vidal, C.; Kollmeyer, P.J.; Naguib, M.; Malysz, P.; Gross, O.; Emadi, A. Robust xev battery state-of-charge estimator design using a feedforward deep neural network. *SAE Int. J. Adv. Curr. Prac. Mobil.* **2020**, *2*, 2872–2880.
26. Yang, F.; Li, W.; Li, C.; Miao, Q. State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network. *Energy* **2019**, *175*, 66–75. [CrossRef]
27. Jiao, M.; Wang, D.; Qiu, J. A GRU-RNN based momentum optimized algorithm for SOC estimation. *J. Power Sources* **2020**, *459*, 228051. [CrossRef]
28. Dong, G.; Zhang, X.; Zhang, C.; Chen, Z. A method for state of energy estimation of lithium-ion batteries based on neural network model. *Energy* **2015**, *90*, 879–888. [CrossRef]
29. Tian, Y.; Lai, R.; Li, X.; Xiang, L.; Tian, J. 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* **2020**, *265*, 114789. [CrossRef]
30. Cho, K.; Courville, A.; Bengio, Y. Describing Multimedia Content Using Attention-Based Encoder-Decoder Networks. *IEEE Trans. Multimed.* **2015**, *17*, 1875–1886. [CrossRef]
31. Song, Y.; Li, L.; Peng, Y.; Liu, D. *Lithium-Ion. Battery Remaining Useful Life Prediction Based on GRU-RNN*; IEEE: Piscataway, NJ, USA, 2018; Volume 2018, pp. 317–322.
32. Zhang, A.; Lipton, Z.C.; Li, M.; Smola, A.J. *Dive into Deep Learning*. 2020. Available online: <https://github.com/d2l-ai/d2l-zh> (accessed on 29 November 2020).
33. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. In Proceedings of the 30th International Conference on Machine Learning, Stanford University, Stanford, CA, USA, 17–19 June 2013; p. 3.
34. Hannun, A.; Case, C.; Casper, J.; Catanzaro, B.; Diamos, G. Deep Speech: Scaling up end-to-end speech recognition. *arXiv* **2014**, arXiv:1412.5567.
35. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2015**, arXiv:1412.6980.
36. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, UK, 2012.
37. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural. Comput.* **1997**, *9*, 1735–1780. [CrossRef]
38. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In Proceedings of the EMNLP 2014—2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 1724–1734.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).