Energy Proceedings
Vol 37, 2023
ISSN 2004-2965

SOC Estimation Method Based on Fusion of Data-Driven Method and Model-Based Filtering Method: Arithmetically Optimized LSTM Network and Adaptive Unscented Kalman Filter

Chunsheng Wang[1], Mutian Li[1], Yuan Cao[1,*], Yukun Hu[2]

1 The School of Automation, Central South University, Changsha, Hunan, 410083, China

2 University College London, London, 999020, UK

(*Corresponding Author: caoyuan3116@csu.edu.cn)

ABSTRACT

This paper proposes a state of charge (SOC) estimation model that combines data-driven method with model-based filtering method. Firstly, an improved arithmetic optimization algorithm (AOA) is employed to optimize the initial values of the long short-term memory (LSTM) network, and the optimized LSTM network is utilized for the preliminary estimation of SOC. Then, an adaptive unscented Kalman filter (AUKF) is employed to correct the SOC estimation results. Experimental results demonstrate that the proposed model achieves accurate and smooth SOC estimation while being able to quickly respond to initial SOC errors.

Keywords: State of charge, Long short-term memory network, Arithmetic optimization algorithm, Adaptive unscented Kalman filter.

NONMENCLATURE

Abbreviations	
State of Charge	SOC
Arithmetic Optimization Algorithm	AOA
Long Short-Term Memory	LSTM
Adaptive Unscented Kalman Filter	AUKF
Dynamic Stress Test	DST
Federal Urban Driving Schedule	FUDS
Root Mean Square Error	RMSE
Maximum Error	MAX
Symbols	
Output variables at time t	o_t
Approximate covariance of output error	H_t
Process noise variance	Q_t
Measurement noise variance	R_t

1. INTRODUCTION

Batteries serves as crucial energy storage devices in various industrial sectors. SOC is a vital parameter for assessing the remaining capacity of a battery. Accurately estimating the battery SOC contributes to prolonging battery lifespan and maximizing its performance utilization [1].

Common methods for SOC estimation include Coulomb counting, open-circuit voltage (OCV) method, model-based filtering methods, and data-driven methods. The Coulomb counting calculates the battery's capacity by integrating the current over time [2]. However, this method heavily relies on the accuracy of current measurements and is prone to cumulative measurement errors, which can result in high measurement costs. The OCV method estimates the corresponding SOC based on the battery's OCV [3]. However, it can only be used when the battery is in a static state.

Model-based filtering methods estimate the SOC by establishing an equivalent model that represents its behavior, such as equivalent circuit model, electrochemical model, and data-driven model. These methods utilize techniques such as Kalman filtering or extend Kalman filtering to optimize the SOC estimation based on the model's dynamics and measurements [4]. However, Kalman Filtering is not applicable to nonlinear systems, and extended Kalman filtering may neglect higher-order terms, leading to a decrease in estimation accuracy. Therefore, the Unscented Kalman Filter is a better choice as it can be used for nonlinear systems and provides a more comprehensive description of system nonlinearity. Indeed, constructing an equivalent model for most batteries that accurately captures their internal reactions and chemical characteristics can be a complex task. Developing such models requires a deep understanding of battery chemistry and extensive experimental data for parameter estimation [5].

Data-driven methods can be used to establish battery data-driven models by leveraging a large volume of relevant data [6]. By training on a diverse dataset,

This is a paper for the international conference on energy, ecology and environment 2023 (ICEEE2023), Aug. 14-18, 2023, London, UK.

data-driven models can capture the complex relationships between input variables (e.g., current, voltage, temperature) and the corresponding SOC, enabling accurate estimation without explicit knowledge of the underlying chemistry. Common data-driven models used for battery SOC estimation include support vector machines, artificial neural networks, and other machine learning algorithms [7]. These models are capable of learning complex relationships from the input-output data and can be trained to estimate SOC accurately based on the provided dataset. Among them, LSTM network has found wide application in SOC estimation due to its ability to effectively model the dynamic behavior of batteries and handle long-term dependencies in the input data. Reference [8] proposed a method that utilized LSTM network to predict multiple parameters of a battery. This method was able to establish an equivalent model and accurately estimate the parameters based on the data collected from an electric vehicle over a period of one year. However, data-driven models have high requirements for the accuracy of the input data and are prone to getting trapped in local optima, leading to significant deviations in the results [9].

To address the limitations of above methods, this paper proposes a fusion model that combines model-based filtering methods with data-driven methods. This model eliminates the need for constructing complex equivalent circuit models or chemical models while effectively mitigating estimation errors caused by the data. The innovations of this paper are as follows:

(1) Integration of Model-Based Filtering and Data-Driven Methods: The paper proposes a novel approach that combines the strengths of model-based filtering methods and data-driven techniques. By leveraging the advantages of both approaches, the proposed model aims to achieve more accurate and robust SOC estimation.

(2) Arithmetically Optimized LSTM Network: The paper introduces an improved AOA to optimize the initial values of the LSTM network used in the SOC estimation. This optimization enhances the performance of the LSTM network, leading to improved initial SOC estimation results.

(3) AUKF for SOC Correction: The paper utilizes the AUKF to correct the SOC estimation results obtained from the data-driven model. By incorporating the AUKF, the proposed model can effectively adjust and refine the SOC estimation, resulting in more accurate and reliable SOC values.

The structure of this paper is as follows. Section 2 describes the modeling approach used in the proposed hybrid model. Section 3 presents the experimental setup and methodology used to validate the effectiveness of the proposed model. Section 4 summarizes this paper.

2. METHODS

This section describes the methods used in the study that are necessary for reproducing the results. Specifically, it covers three aspects: the construction of the LSTM network, the optimization of initial parameters, and the development of the hybrid model.

2.1 The construction of the LSTM network

The structure of the LSTM network is illustrated in Fig. 1. LSTM network is a type of recurrent neural network (RNN) that is designed to capture long-term dependencies in sequential data. Unlike standard RNN, LSTM network has additional memory cells, or gates, that enable them to selectively remember or forget information over time. LSTM network can be divided into three parts: the forget gate, the input gate, and the output gate. These gates allow the LSTM to selectively filter long-term historical data.

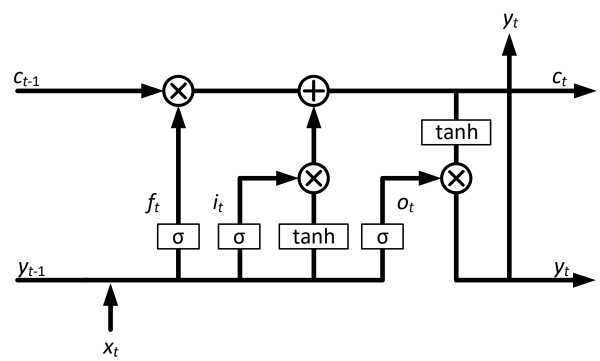

Fig. 1. Structure of an LSTM unit

In the context of battery SOC estimation, there are several parameters that influence it. This paper considers voltage, current, and temperature as the inputs to the LSTM network, and SOC as the outputs. Additionally, the number of hidden layers, the number of neurons, the maximum number of iterations, and the learning rate are parameters that need to be predefined for the LSTM network. To address this, the paper utilizes the AOA to initialize these parameters.

2.2 Arithmetic optimization algorithm

The AOA utilizes arithmetic optimization operators to search for the optimal individuals and obtain the optimal parameter values. These operators, such as addition, subtraction, multiplication, and division, are used to iteratively explore the solution space and optimize the parameters for the LSTM network. Through random selection, each individual has an equal probability of entering the exploration stage and development stage. This ensures that all individuals have an equal chance to explore the solution space and contribute to the development of the optimal parameter values.

During the exploration stage, each individual explores the solution space by using multiplication and division operations within a wide range of values. This allows for a broad search of potential solutions, enabling the algorithm to explore different parameter combinations and identify promising regions in the search space. During the development stage, each individual further explores the optimal solution by using addition and subtraction operations with a smaller range. This allows for a more focused search around promising solutions identified. The combination of exploration and development helps the algorithm converge towards the optimal parameter values for the LSTM network.

2.3 Adaptive unscented Kalman filter

UKF is a variant of Kalman filter that addresses the limitations of linear models and Gaussian noise assumptions. It is designed to estimate the state of a nonlinear dynamic system in the presence of process and measurement noise. The key principle of the UKF is to approximate the probability distribution of the system's state using a set of representative sigma points. These sigma points are carefully chosen to capture the mean and covariance of the system's state distribution. The predicted state and sigma points are used to estimate the predicted measurement using the measurement model. The predicted measurement is compared with the actual measurement to compute the measurement residual. The covariance matrix is updated by incorporating the predicted state, predicted measurement, and the measurement residual. The corrected state estimate is computed by adjusting the predicted state using the updated covariance matrix and the measurement residual.

Indeed, the standard UKF requires predefining process noise and measurement noise, which may not accurately capture the variations in these noise sources that occur during battery operation. Such time-varying noise can introduce significant errors in SOC estimation results. To address this challenge, adaptive variants of the UKF have been developed. These adaptive UKF approaches dynamically estimate the process noise and measurement noise covariance matrices based on the available measurements. By continuously adapting the noise covariances, the adaptive UKF can effectively account for the varying noise characteristics and improve the accuracy of SOC estimation. The equation for noise adaptive update, as denoted by equation (1), is as follows:

$$\begin{cases} H_t = \dfrac{1}{N} \sum\limits_{i=t-N+1}^{t} (o_t - \bar{o}_t)(o_t - \bar{o}_t)^T \\ Q_t = K_t H_t K_t^T \\ R_t = H_t + \sum\limits_{i=1}^{2N} W_i (o_t - \bar{o}_t)(o_t - \bar{o}_t)^T \end{cases} \tag{1}$$

Where H_t means approximate covariance of output error at time t, N denotes dimension of variable, and Q_t and R_t represent process and measurement noise variance, respectively.

3. RESULTS AND DISCUSSION

In this section, a comparison is conducted between LSTM, LSTM-UKF, and the proposed model to validate the performance of the proposed AOA-LSTM-AUKF approach.

3.1 Dataset and indicators

In this study, the NASA open-source battery dataset is utilized, which includes various parameters of lithium batteries under dynamic stress test (DST), US06, and federal urban driving schedule (FUDS) driving cycles at different temperatures. To validate the generalization capability of the proposed model, the DST dataset is used as the training set, while the US06 and FUDS datasets are used as the testing sets.

The performance of the proposed model is evaluated using root mean square error (RMSE) and maximum error (MAX). These metrics provide insights into the accuracy and maximum deviation of the SOC estimation results compared to the ground truth values. The equations are as given by (2) and (3).

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{i=1}^{K} (SOC_{actual} - SOC_{predict})^2} \tag{2}$$

$$\text{MAX} = \max(|SOC_{actual} - SOC_{predict}|) \tag{3}$$

3.2 SOC estimation results

Using DST as training dataset, US06 and FUDS as test dataset, so as to obtain SOC estimation results at 0°C, 20°C, 30°C, and 50°C. Fig. 2 and Fig. 3 illustrate the results obtained from the LSTM network. Fig. 4 and Fig. 5 depict the results obtained from the LSTM-UKF model. Lastly, Fig. 6 and Fig. 7 showcase the results obtained from the proposed model. From the figures, it is evident that the proposed model yields more accurate and smoother SOC

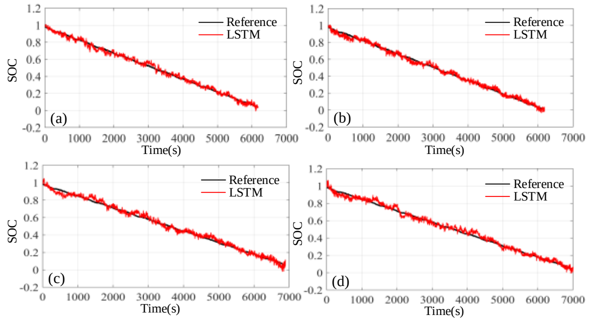

Fig. 2. Results of LSTM network for US06 cycles at: (a) 0°C; (b) 20°C; (c) 30°C; (d) 50°C.

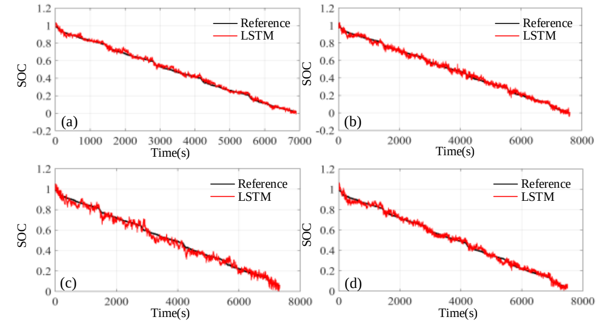

Fig. 3. Results of LSTM network for FUDS cycles at: (a) 0°C; (b) 20°C; (c) 30°C; (d) 50°C.

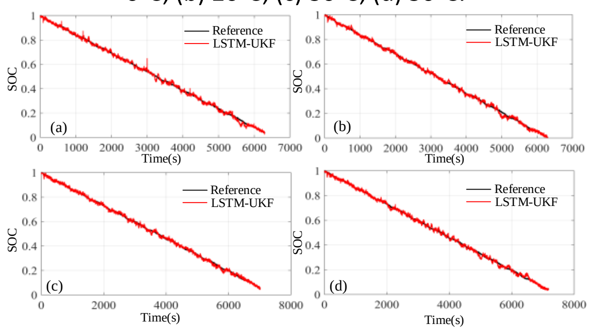

Fig. 4. Results of LSTM-UKF for US06 cycles at: (a) 0°C; (b) 20°C; (c) 30°C; (d) 50°C.

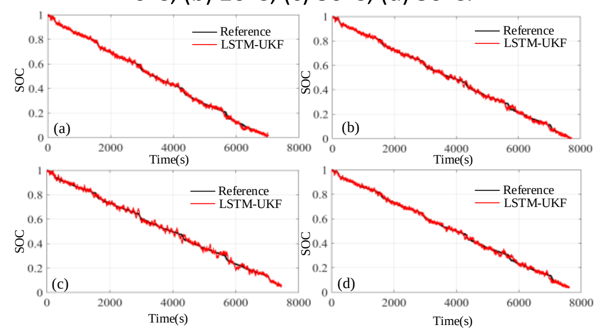

Fig. 5. Results of LSTM-UKF for FUDS cycles at: (a) 0°C; (b) 20°C; (c) 30°C; (d) 50°C.

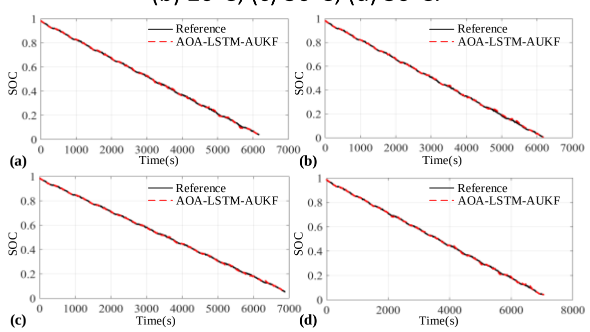

Fig. 6. Results of proposed model for US06 cycles at: (a) 0°C; (b) 20°C; (c) 30°C; (d) 50°C.

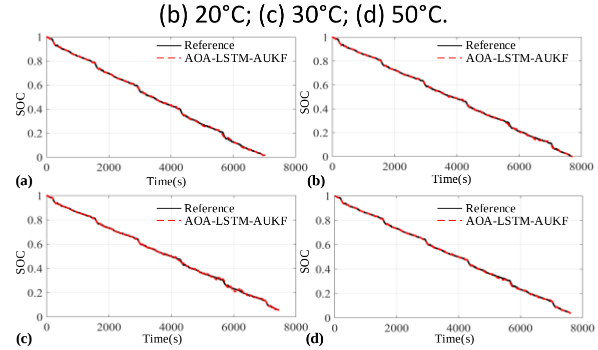

Fig. 7. Results of proposed model for FUDS cycles at: (a) 0°C; (b) 20°C; (c) 30°C; (d) 50°C.

Table I
SOC estimation results for US06 and FUDS at different temperatures

Condition	Temperature	RMSE (%)			MAX (%)		
		LSTM	LSTM-UKF	Proposed model	LSTM	LSTM-UKF	Proposed model
US06	0 °C	2.31	1.32	0.58	9.78	8.73	1.83
	20 °C	2.36	1.21	0.54	8.33	5.40	1.89
	30 °C	2.77	1.09	0.52	9.06	4.22	2.19
	50 °C	2.73	1.26	0.61	9.74	5.05	2.27
FUDS	0 °C	1.87	1.17	0.45	7.07	5.02	1.99
	20 °C	2.19	1.29	0.43	8.27	6.52	2.02
	30 °C	3.14	1.60	0.48	11.37	7.31	2.21
	50 °C	2.26	1.24	0.49	8.50	5.32	1.52

estimation results compared to the others. The SOC trajectory in the proposed model exhibits less variation and closely follows the ground truth values, indicating its superior performance in estimating the battery's SOC.

Table I summarizes the RMSE and MAX values for each model. From the table, it is evident that the proposed model exhibits the lowest errors and highest accuracy across various driving conditions. The average RMSE is less than 0.6%, and the average MAX is less than 2.1%. In particular, under the FUDS driving condition, the RMSE is below 0.5%, and the lowest MAX value is 1.52%. These results highlight the superior performance of the proposed model in SOC estimation, even under challenging driving scenarios.

4. CONCULSTION

In conclusion, this study presents a novel hybrid model for SOC estimation in batteries, combining data-driven methods and model filtering techniques. By leveraging the strengths of LSTM networks and incorporating the AUKF for state correction, the proposed model achieves accurate and reliable SOC estimation without the need for complex equivalent circuit models. The proposed model demonstrates excellent generalization capability, as it achieves low RMSE and MAX values across different driving cycles. The average RMSE and MAX values remain below 0.6% and 2.1%, respectively, with even lower errors observed in the FUDS driving cycle.

Overall, the proposed hybrid model offers a promising solution for accurate and smooth SOC estimation in batteries. It addresses the limitations of existing methods and shows potential for applications in various industrial sectors that rely on battery energy storage systems. Future research directions may include further enhancing the model's robustness, investigating different adaptation strategies, and exploring its applicability in real-world battery systems.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation (NNSF) of China under Grant 62103443 and 61973322, and Hunan Provincial Natural Science Foundation under Grant 2022JJ40630.

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

REFERENCE

[1] Y. Cao and J. A. Qahouq, Small-Signal Modeling and Analysis for a Wirelessly Distributed and Enabled Battery Energy Storage System of Electric Vehicles. *Applied Sciences*, vol. 9, no. 20, p. 4249, Oct. 2019.

[2] A. Wadi, M. F. Abdel-Hafez, and A. A. Hussein, Mitigating the Effect of Noise Uncertainty on the Online State-of-Charge Estimation of Li-Ion Battery Cells. *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8593-8600, Sep. 2019.

[3] Y. Cao and J. A. Abu Qahouq, Hierarchical SOC Balancing Controller for Battery Energy Storage System. *IEEE Trans. Ind. Electron.*, vol. 68, no. 10, pp. 9386-9397, Oct. 2021.

[4] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu, State estimation for advanced battery management: Key challenges and future trends. *Renewable and Sustainable Energy Reviews*, vol. 114, p. 109334, Oct. 2019.

[5] Y. Cao and J. A. Abu Qahouq, Evaluation of bi-directional single-inductor multi-input battery system with state-of-charge balancing control. *IET power electron.*, vol. 11, no. 13, pp. 2140-2150, Nov. 2018.

[6] H. He, X. Zhang, R. Xiong, Y. Xu, and H. Guo, Online model-based estimation of state-of-charge and open-circuit voltage of lithium-ion batteries in electric vehicles. *Energy*, vol. 39, no. 1, pp. 310-318, Mar. 2012.

[7] R. Xiong, J. Tian, W. Shen, and F. Sun, A Novel Fractional Order Model for State of Charge Estimation in Lithium Ion Batteries. *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4130-4139, May 2019.

[8] J. Hong, Z. Wang, W. Chen, and Y. Yao, Synchronous multi-parameter prediction of battery systems on electric vehicles using long short-term memory networks. *Applied Energy*, vol. 254, p. 113648, Nov. 2019.

[9] Z. Cui, L. Kang, L. Li, L. Wang, and K. Wang, A combined state-of-charge estimation method for lithium-ion battery using an improved BGRU network and UKF. *Energy*, vol. 259, p. 124933, Nov. 2022.