Improved multi-head bi-directional long and short-term memory temporal convolutional network for lithium-ion batteries state of charge estimation in energy storage systems.

LI, Y., WANG, S., LIU, D., CUI, Y., FERNANDEZ, C. and BLAABJERG, F.

2024

Improved Multi-head Bi-directional Long and Short-term Memory Temporal Convolutional Network for Lithium-ion Batteries State of Charge Estimation in Energy Storage Systems

Yang Li
Southwest University of Science and Technology
School of Information Engineering
Mianyang, China
liyang-1998@qq.com

Shunli Wang*
Southwest University of Science and Technology
School of Information Engineering
Mianyang, China
wangshunli@swust.edu.cn

Donglei Liu
Southwest University of Science and Technology
School of Information Engineering
Mianyang, China
liudonglei1994@qq.com

Yixiu Cui
Institute of Electronic Engineering
China Academy of Engineering Physics
Mianyang, China
cyxiu@caep.cn

Carlos Fernandez
School of Pharmacy and Life Sciences
Robert Gordon University
Aberdeen, United Kingdom
lledofernandez76@hotmail.com

Frede Blaabjerg
Department of Energy Technology
Aalborg University
Aalborg East, Denmark
fbl@energy.auu.dk

***Abstract*—Lithium-ion batteries with their high voltage, large capacity, high discharge rate, no memory effect, and green environmental protection advantages are widely used in communication, power stations, backup power, and other energy storage fields. Accurate estimation of the state of charge (SOC) of lithium-ion batteries is a key prerequisite to ensure the safe, reliable, and efficient operation of battery systems. To address this core challenge, this paper innovatively proposes a composite model combining the Kepler optimization algorithm, temporal convolutional network (TCN), bi-directional long and short-term memory network (BiLSTM), and multi-head attention mechanism (MHA). Kepler optimization algorithm was used to search the optimal hyperparameters in the TCN-BiLSTM structure so that the model could adjust the structural parameters and extract the input features accurately under different working conditions and temperatures. The multi-head self-attention mechanism is introduced to assign different weights to the feature outputs extracted by time convolution according to the different importance of information to improve the adaptability of the model. Finally, the proposed model is tested and compared with other models under different temperatures and working conditions.**

Keywords—soc, tcn, bilstm, multi-head attention mechanism, Kepler optimization algorithm

I. INTRODUCTION

As an important link between the user and the battery, the battery management system (BMS) controls the whole process of battery charging and discharging and timely intervenes in possible safety failures, which plays an important role in ensuring the safety and reliability of the battery during use. Therefore, accurate monitoring of SOC real-time changes is of great significance for improving user security [1]. Because lithium batteries show strong time-varying characteristics and nonlinear characteristics. It cannot be estimated directly by the instrument and can only be estimated indirectly by measuring other relevant parameters in the battery.

In recent years, the industry has developed diverse technical strategies for SOC estimation of lithium-ion batteries, which mainly include three categories: basic methods, observer methods, and data-driven methods [2]. Among them, basic methods provide a straightforward and easy-to-understand framework for SOC estimation as a starting point. Ampere-hour (AH) is a straightforward and widely adopted technique for estimating the SOC of batteries due to its intuitive nature and simplicity of implementation. Wang et al. [3] used a dynamic matrix control algorithm to update battery model parameters and then combined it with an extended Kalman filter algorithm to estimate SOC. Lv et al. [4] proposed a multi-model SOC estimation algorithm based on EKF that considered the low temperature and high current ratio of the battery. This kind of observer method usually requires the construction of the battery model, and the model accuracy directly affects the SOC estimation accuracy. Data-driven methods mainly include the neural network method, fuzzy inference method, and linear regression method. The data-driven method does not need to pay attention to the internal characteristics of the battery, directly input the external characteristic data of the battery into the model, train the nonlinear relationship between these data and SOC, and then output the SOC estimate. However, the data-driven method also has many shortcomings, such as the model hyperparameters are difficult to adjust, and the data demand is large. Due to battery aging and other effects, the mapping relationship changes, and the model cannot adapt to this change in time, thus affecting the estimation accuracy. In response to these problems, some scholars put forward improving methods. Etse Dablu Bobobee and other researchers [5] cleverly employed a particle swarm optimization algorithm to fine-tune the parameters of the LSTM to construct a novel network model that integrates a temperature compensation mechanism. This innovation significantly enhances the estimation accuracy and accelerates the convergence process of the algorithm. Paul Takyi-Aninakwa et al. [6] constructed a correlation LSTM for rough estimation of lithium battery SOC, and then input the rough estimate into the square gain EKF for filtering and noise reduction. This hybrid model has high robustness and stability. Based on LSTM, Chen et al. [7] introduced a self-attention mechanism to assign different weights to the output of the hidden layer to improve the efficiency of the model in

*Corresponding Author

processing data. Chen et al. [8] proposed an improved long short-term memory neural network, which uses a time-varying average voltage sliding window at the input port to limit the data flow between network layers at the output port, reduce the large fluctuations of network output, and improve the adaptability of the network to unknown data. Qian et al. [9] introduced a unified estimation model that integrates convolutional neural networks, self-attention mechanisms, and long short-term memory neural networks. They utilized a joint loss function to enhance network optimization and bolster the model's resilience.

II. Theoretical Analysis

A. Bidirectional Long Short-Term Memory Network

LSTM is a type of recurrent neural network (RNN) that addresses several issues found in traditional RNNs. These issues include gradient vanishing or exploding, as well as the difficulty in capturing long-range dependencies. LSTM achieves this by incorporating three gates—input, forget, and output, depicted in Fig. 1 below. LSTM architectures are widely applied in tasks such as time series prediction and other domains requiring modeling of sequential data.

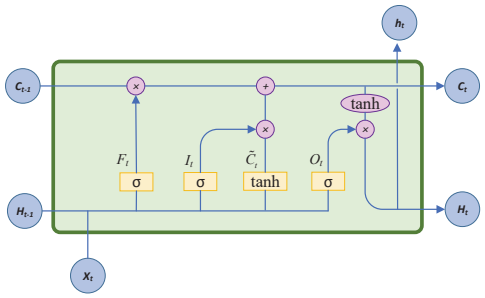

Fig. 1. Architecture of unidirectional long short-term memory network cells

The LSTM processes X_t as the current input and H_{t-1} from the previous time step as its hidden state input. It uses a sigmoid activation function to compute its output, which is expressed as:

$$\begin{cases} I_t = \sigma\left(X_t M_{xi} + H_t M_{hi} + b_i\right) \\ F_t = \sigma\left(X_t M_{xf} + H_{t-1} M_{hf} + b_f\right) \\ O_t = \sigma\left(X_t M_{xo} + H_{t-1} M_{ho} + b_o\right) \end{cases} \tag{1}$$

M_{xi}, M_{xf}, M_{xo} M_{hi}, M_{hf}, and M_{ho} are weight parameters, and b_i, b_f, and b_o are bias parameters.

The candidate memory cells use the tanh function as an activation function.

$$\tilde{C}_t = \tanh\left(X_t M_{xc} + H_{t-1} M_{hc} + b_c\right) \tag{2}$$

M_{xc} and M_{hc} are the weight parameters and b_c is the bias parameter.

Traditional LSTM can only process data in the positive direction, and the BiLSTM network can process data in both positive and negative directions at the same time. This kind of network can process information more efficiently by adding a neural network layer that transmits information in the reverse direction.

The framework structure of BiLSTM is illustrated in Fig. 2, which cleverly incorporates a bidirectional parallel LSTM layer design, one processing the sequence from beginning to end, while the other processes the sequence in the opposite direction.

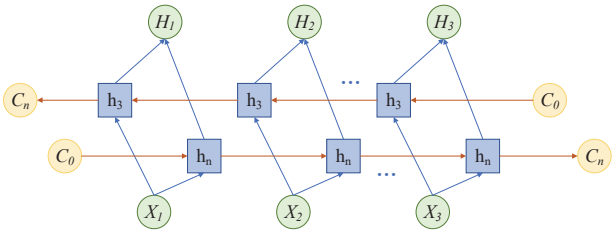

Fig. 2. Structure of bidirectional long short-term memory network

B. Temporal Convolutional Network Model

The TCN model utilizes a recurrent neural network architecture that integrates causal convolution with dilated convolution and residual modules. The causal convolution structure is depicted in Fig. 3. From the illustration, it is evident that the output at any given moment relies solely on the current input and the preceding inputs, reflecting a stringent temporal dependency model. The greater the use of historical data, the corresponding number of hidden layers increases, which increases network complexity. The result of expansive convolution is shown in Fig. 4. Adding a void into the convolutional kernel can effectively solve the problem of missing information on the pooling layer in the convolutional neural network without increasing the number of network parameters. Increasing the receptive field of the convolutional layer helps capture long-distance dependencies in the input.

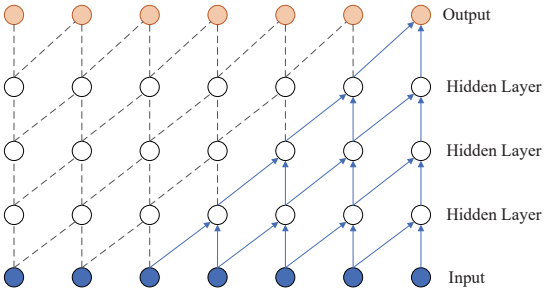

Fig. 3. Schematic diagram of causal convolution structure

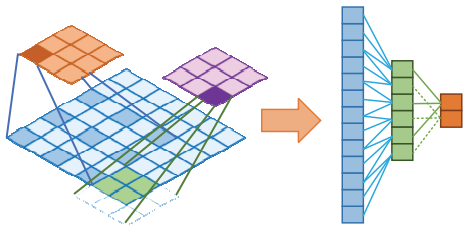

Fig. 4. Schematic diagram of expansion convolution structure

C. Multi Head-Attention Mechanism

The self-attention mechanism smartly evaluates the correlation weights between the elements of the input sequence, which in turn aggregates a weighted data representation. However, the self-attention mechanism tends to pay too much attention to its information and ignore the attention to other information. Therefore, multiple self-attentional heads are used for parallel computation, and the results are finally spliced together so that the model can capture different feature information at the same time and improve the performance of the model.

$$headi = attention\left(QW_i^q, KW_i^k, VW_i^v\right) \tag{3}$$

$$MultiHead(Q,K,V) = Attach(head_1,...,head_h)W^C \quad (4)$$

W_i^q, W_i^k, W_i^v are the parameter matrices for each head, W^C is the transformation matrix of the output.

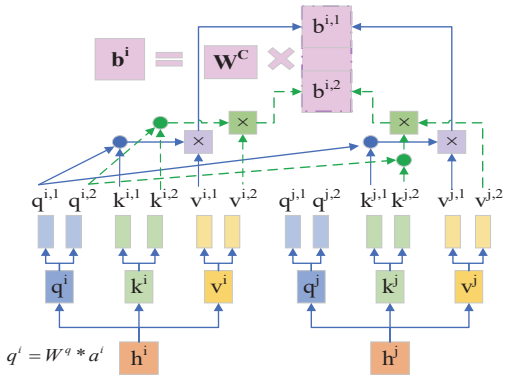

Fig. 5. The structural principle of multi-head self-attention mechanism

The structure of the multi-head self-attention mechanism (taking two heads as an example) is shown in Fig. 5. Firstly, the input sequence is calculated in parallel using multiple independent heads, then the output of all heads is spliced, and finally the spliced representation is linearly transformed, that is, the output of the multi-head attention mechanism. q stands for query, and k stands for key, q and k will match. The greater the correlation after matching, the greater the corresponding weight v will be.

D. Kepler Optimization Algorithm

The Kepler optimization algorithm finds the best solution in a complex search space by mimicking the mechanism by which planets orbit stars. A planet in an elliptical orbit, as a solution space, will be in different positions at different times. As the optimal solution, the gravitational attraction between the sun and the planet will affect the rotation speed of the planet, so it also determines how close the candidate solution in space is to the optimal solution.

KOA has the following steps:

1) Initialize the positions of the Sun and planets, as well as the orbital eccentricity e and period T

$$X_p^q = X_{p,l}^q + rand_{[0,1]}\left(X_{p,u}^q - X_{p,l}^q\right), \begin{cases} p = 1,2,\ldots,Y. \\ q = 1,2,\ldots,d. \end{cases} \quad (5)$$

p refers to the solution ranked p-th in the solution space, while Y comprehensively summarising the total count of solutions in that solution space. d is the dimensional attribute of the problem to be optimized. $X_{p,u}^q$ and $X_{p,l}^q$ define the specific ranges of upper and lower bounds, respectively, when considered for a J-dimensional problem.

$$e_p = rand_{[0,1]}, p = 1,\ldots,Y \quad (6)$$

$$T_p = |r|, p = 1,\ldots,Y \quad (7)$$

The rand function can generate a random value, which T_p is the period required for the object p to make one revolution around its orbit, and r is a randomly selected value based on a normal distribution.

2) Calculate the velocity of the celestial body

The rate of movement of a celestial body along its orbit is strongly regulated by the strength of the Sun's gravitational pull. As a celestial body approaches the Sun, the Sun's gravitational pull increases, prompting the body to move faster to maintain a stable orbit and avoid being attracted by the Sun's strong gravitational pull. Conversely, as the object moves away from the Sun, the gravitational force on it weakens, resulting in a corresponding slowing of its rate of motion.

$$V_p(t) = \begin{cases} \ell \times \left(2r_4\vec{X}_p - \vec{X}_B\right) + \ddot{\ell} \times \left(\vec{X}_A - \vec{X}_B\right) + \left(1 - R_{p-norm}(t)\right) \\ \quad \times \mathcal{F} \times \vec{U}_1 \times \vec{r}_5 \times \left(\vec{X}_{p,u} - \vec{X}_{p,l}\right) \quad if\ R_{p-norm}(t) \le 0.5 \\ r_4 \times \mathcal{L} \times \left(\vec{X}_A - \vec{X}_i\right) + \left(1 - R_{p-norm}(t)\right) \\ \times \mathcal{F} \times U_2 \times \vec{r}_5 \times \left(r_3\vec{X}_{p,u} - \vec{X}_{p,l}\right) \quad else \end{cases} \quad (8)$$

$$\ell = \vec{U} \times \mathcal{M} \times \mathcal{L} \quad (9)$$

$$\mathcal{L} = \left[\mu(t) \times \left(M_S + m_i\right)\left|\frac{2}{R_p(t) + \varepsilon} - \frac{1}{A_{p(t)} + \varepsilon}\right|\right]^{\frac{1}{2}} \quad (10)$$

$$\mathcal{M} = \left(r_3 \times \left(1 - r_4\right) + r_4\right) \quad (11)$$

$$\vec{U} = \begin{cases} 0 & \vec{r}_5 \le \vec{r}_6 \\ 1 & else \end{cases} \quad (12)$$

$$\mathcal{F} = \begin{cases} 1 & if\ r_4 \le 0.5 \\ -1 & else \end{cases} \quad (13)$$

$$\ddot{\ell} = (1 - \vec{U}) \times \vec{\mathcal{M}} \times \mathcal{L} \quad (14)$$

$$\vec{\mathcal{M}} = \left(r_3 \times \left(1 - \vec{r}_5\right) + \vec{r}_5\right) \quad (15)$$

$$\vec{U}_1 = \begin{cases} 0 & \vec{r}_5 \le r_4 \\ 1 & else \end{cases} \quad (16)$$

$$U_2 = \begin{cases} 0 & r_3 \le r_4 \\ 1 & else \end{cases} \quad (17)$$

where $V_p(t)$ represents the moving speed of the celestial body; r_3 and r_4 are randomly generated values [0,1]; $\vec{r}_5$, $\vec{r}_6$ is two random vectors; $\vec{X}_A$ and $\vec{X}_B$ represent schemes randomly selected from the population; M_s is the mass of X_s; m_P is the mass of X_P; $\mu(t)$ represents the universal gravitation constant; $R_p(t)$ measures the spatial distance between the optimal solution (the Sun) X_s and the candidate solution (a particular object), X_p; and $A_{p(t)}$ represents the length of the semi-long axis of the elliptical orbit of object p at time t.

$$A_p(t) = r_3 \times \left[T_p^2 \times \frac{\mu(t) \times \left(M_s + m_p\right)}{4\pi^2}\right]^{\frac{1}{3}} \quad (18)$$

If the semi-major axis of the object i in the elliptical orbit gradually decreases over time, then the candidate solution

progressively converges towards the region of the global optimal solution. $R_{i-norm}(t)$ means to normalize the Euclidean distance between X_s and X_i, which is normalized as follows:

$$R_{p-norm}(t)=\frac{R_p(t)-\min(R(t))}{\max(\mathrm{R(t)})-\min(R(t))} \tag{19}$$

When $R_{p-norm}(t)\le 0.5$ the celestial body is very close to the sun, it speeds up its movement due to the huge gravity generated by the sun, avoiding drifting toward the Sun.

3) Prevent jumping into local optimality

Most celestial bodies revolve around the Sun counterclockwise, but a few move in a clockwise direction. In KOA, $\mathcal{F}$ is used to change the search direction and ensure that the search scope covers the whole solution space, to avoid falling into local optimal.

4) Update position

In their journey around the Sun, celestial bodies undergo a cyclical cycle of moving towards and away from the Sun. Their speed in elliptical orbits is strongly regulated by the dynamics of the Sun's gravitational field. When a celestial body is far away from the Sun, this process provides a valuable opportunity for candidate solutions (celestial bodies) to deeply explore and effectively utilize the neighborhood of the optimal solution (the Sun); on the contrary, when a celestial body is close to the Sun, the speed of its motion is significantly accelerated, and this acceleration not only helps the celestial body to break away from the strong attraction of the Sun temporarily but also facilitates the exploration of the wider solution space and effectively avoids the trap of the local optimal solution that it may be trapped in.

$$\begin{aligned}\vec{X}_p(t+1)=\vec{X}_p(t)+\mathcal{F}\times\vec{V}_p(t)+\left(\boldsymbol{F}_{g_i}(t)+|r|\right)\\ \times\vec{U}\times\left(\vec{X}_S(t)-\vec{X}_p(t)\right)\end{aligned} \tag{20}$$

$\vec{X}_p(t+1)$ represents the position of object p at time $t+1$; $\vec{V}_p(t)$ represents the velocity required for object p to reach its new position; $X_s(t)$ represents the optimal solution position so far. $\boldsymbol{F}_{g_i}(t)$ represents gravitational effects of the Sun on celestial bodies, which is defined as follows:

$$\boldsymbol{F}_{g_i}(t)=e_p\times\mu(t)\times\frac{\bar{M}_s\times\bar{m}_p}{\bar{R}_p^2+\varepsilon}+r_1 \tag{21}$$

5) Update the distance from the Sun

To more accurately portray the "exploration" and "development" phases of the optimization process, an adaptive factor, H, is introduced as a key parameter to regulate the distance between the Sun and the celestial bodies at the time t. The h value is set at a high level to allow for a wider exploration of space through the action of the detection operator, whereas when the h value is set low, the celestial bodies tend to expand the distance between the Sun and the celestial bodies. Specifically, when the value of h is set high, the distance between the object and the Sun is enlarged by the detection operator, which promotes a wider exploration space; while when the value of H is adjusted low, the object tends to be closer to the Sun, which helps the object to develop and explore more carefully near the Sun (optimal solution).

$$\begin{aligned}\vec{X}_p(t+1)=\vec{X}_p(t)\times\vec{U}_1+\left(1-\vec{U}_1\right)\times\left(\frac{\vec{X}_p(t)+\vec{X}_S+\vec{X}_A(t)}{3.0}\right.\\ \left.+h\times\left(\frac{\vec{X}_p(t)+\vec{X}_S+\vec{X}_A(t)}{3.0}-\vec{X}_B(t)\right)\right)\end{aligned} \tag{22}$$

$$h=\frac{1}{e^{\eta r}} \tag{23}$$

η is a linear decreasing factor, determined as follows:

$$\eta=(a_2-1)\times r_4+1 \tag{24}$$

a_2 is the loop control parameter.

$$a_2=-1-1\times\left(\frac{t\%\frac{T_{\max}}{\bar{T}}}{\frac{T_{\max}}{\bar{T}}}\right) \tag{25}$$

6) Elite retention strategy

If the fitness value of the updated position is lower, the old position is replaced.

$$\vec{X}_{p,n}(t+1)=\begin{cases}\vec{X}_p(t+1) & if\ f\left(\vec{X}_p(t+1)\right)\le f\left(\vec{X}_p(t)\right)\\ \vec{X}_p(t) & else\end{cases} \tag{26}$$

III. Analysis of Experimental Results

To validate the algorithm's effectiveness, four temperatures (-5℃, 5℃, 15℃, 25℃) were employed to test and analyze its performance under BBDST working conditions. As depicted in Fig. 6, the algorithm's estimated values closely align with the actual SOC values of lithium-ion batteries across all four temperatures, whether at the top 50% SOC or the bottom 50% SOC.

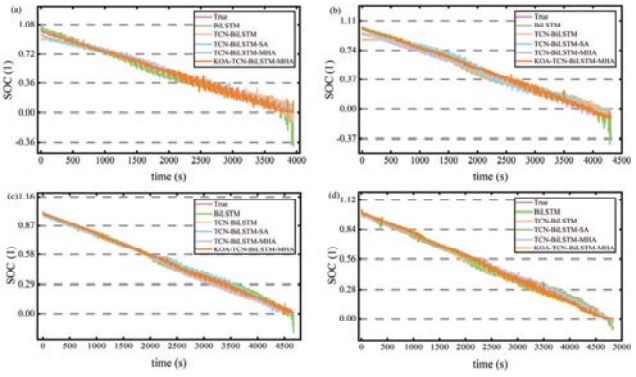

Fig. 6. SOC estimation results under BBDST conditions ((a) -5℃ (b) 5℃ (c) 15℃ (d) 25℃)

That can be concluded from Fig. 7, that the estimation error of the KOA-TCN-BiLSTM-MHA model is significantly reduced compared with other models at four temperatures under BBDST conditions. At the end of the estimation, due to the enhanced polarization of lithium batteries, the estimation errors of the BiLSTM model tend to be dispersed. However, although the estimation errors of the proposed model are increased, they always fluctuate within a small error band, which further proves the stability of the model.

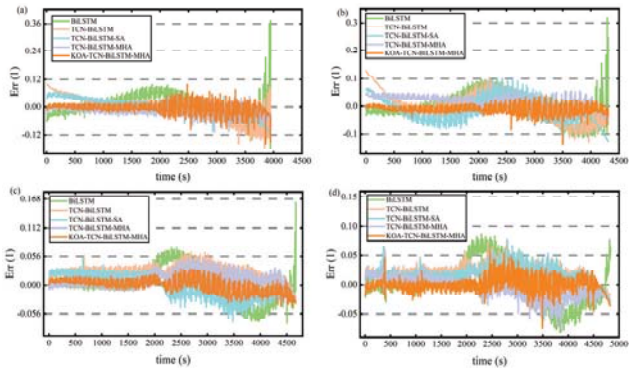

Fig. 7. SOC estimation error under BBDST conditions ((a) -5℃ (b) 5℃ (c) 15℃ (d) 25℃)

The experimental results were evaluated using three error indices: mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). Specifically, at 5 ℃, the optimized algorithm demonstrated an estimated accuracy that was 41.20% higher compared to its performance before optimization. Furthermore, the error margin was maintained within ±0.04%, as shown in Table 1.

TABLE I. SOC ESTIMATION ERROR EVALUATION INDEX UNDER BBDST WORKING CONDITION

	BBDST					
	MAE(%)	MSE(%)	RMSE(%)	MAE(%)	MSE(%)	RMSE(%)
	-5℃			5℃		
BiLSTM	0.0424	0.0029	0.0535	0.0412	0.0026	0.0514
TB	0.0373	0.0025	0.0497	0.0399	0.0025	0.0496
TBSA	0.0299	0.0014	0.0372	0.0322	0.0019	0.0438
TBMHA	0.0193	0.0006	0.0254	0.0233	0.0007	0.0265
KTBMHA	0.0148	0.0004	0.0204	0.0137	0.0003	0.0176
	15℃			25℃		
BiLSTM	0.0241	0.0010	0.0323	0.0272	0.0012	0.0353
TB	0.0240	0.0007	0.0265	0.0201	0.0005	0.0230
TBSA	0.0192	0.0005	0.0219	0.0175	0.0004	0.0208
TBMHA	0.0127	0.0003	0.0170	0.0140	0.0003	0.0174
KTBMHA	0.0103	0.0002	0.0128	0.0083	0.0001	0.0115

[a] K-KOA, T-TCN, B-BiLSTM, MHA-Multi head attention

To demonstrate the robustness of the algorithm, the model is well-trained and replaced with DST working data, and the estimated results of the proposed algorithm are analyzed. As shown in Fig. 8, the KOA-TCN-BiLSTM-MHA algorithm model can well capture the real-time change of SOC value of lithium-ion batteries at four different temperatures.

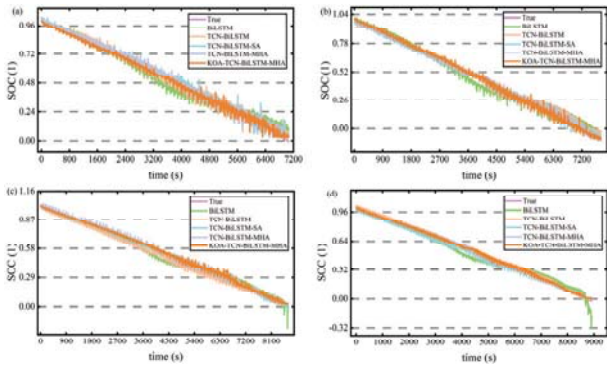

Fig. 8. SOC estimation results under DST conditions ((a) -5℃ (b) 5℃ (c) 15℃ (d) 25℃)

It can be seen from Fig. 9, that compared with the other four algorithms, the proposed algorithm has the smallest fluctuation of estimation error, which further verifies the stability of the algorithm.

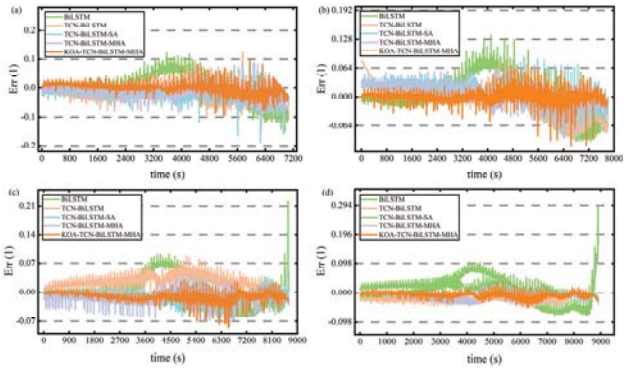

Fig. 9. SOC estimation error under DST conditions ((a) -5℃ (b) 5℃ (c) 15℃ (d) 25℃)

As shown in Table 2, the estimation accuracy of the KOA-TCN-BiLSTM-MHA algorithm is the highest at four different temperatures. Taking 25℃ as an example, the estimation accuracy of this algorithm increases by 70.64%, 56.56%, 42.08%, and 33.75%, respectively, compared with the other four algorithms.

TABLE II. SOC ESTIMATION ERROR EVALUATION INDEX UNDER DST WORKING CONDITION

	BBDST					
	MAE(%)	MSE(%)	RMSE(%)	MAE(%)	MSE(%)	RMSE(%)
	-5℃			5℃		
BiLSTM	0.0342	0.0020	0.0450	0.0361	0.0021	0.0457
TB	0.0267	0.0010	0.0311	0.0247	0.0012	0.0350
TBSA	0.0243	0.0007	0.0272	0.0213	0.0006	0.0240
TBMHA	0.0195	0.0005	0.0231	0.0156	0.0004	0.0188
KTBMHA	0.0103	0.0002	0.0141	0.0070	0.0001	0.0103
	15℃			25℃		
BiLSTM	0.0248	0.0011	0.0338	0.0361	0.0021	0.0453
TB	0.0259	0.0008	0.0287	0.0244	0.0007	0.0256
TBSA	0.0169	0.0005	0.0217	0.0183	0.0004	0.0211
TBMHA	0.0134	0.0002	0.0158	0.0160	0.0004	0.0190
KTBMHA	0.0116	0.0002	0.0141	0.0106	0.0002	0.0129

IV. CONCLUSIONS

In this study, an improved multi-head bidirectional long short-term memory time convolution model is proposed for the estimation of lithium-ion battery SOC. The number of neurons in the network, the initial learning rate, and L2 regularization parameters were optimized by using KOA to find the optimal parameter values and improve the fitness of the network. The multi-head attention mechanism is introduced to give different inputs different weights so that the network can concentrate on processing important information and make rational use of computing power space. Finally, comprehensive experimental tests were implemented under diverse temperature and operating conditions, followed by meticulous analysis and comparison of the estimation results of four state-of-the-art network models, BiLSTM, TCN-BiLSTM, TCN-BiLSTM-SA, and TCN-BiLSTM-MHA.

The results show that under different working conditions and different temperatures, the MAE of the proposed algorithm is stable at about 0.01%, and the R2 determination coefficient exceeds 99%. Therefore, the KOA-TCN-BiLSTM-MHA model performs well in terms of estimation accuracy and stability.

V. REFERENCE

[1] L. Q. Zhang, H. Z. Liu, X. Y. Wang, and M. Li, "State-of-charge estimation for lithium primary batteries: Methods and verification," *Journal of Energy Storage,* vol. 86, May 2024, Art. no. 111189.

[2] L. Z. Xiao, X. N. Li, Q. Y. Jiang, and G. C. Geng, "Online state-of-charge estimation refining method for battery energy storage system

using historical operating data," *Journal of Energy Storage,* vol. 57, Jan 2023, Art. no. 106262.

[3] Y. Q. Wang, Y. Cheng, Y. Xiong, and Q. Z. Yan, "Estimation of battery open-circuit voltage and state of charge based on dynamic matrix control-extended Kalman filter algorithm," *Journal of Energy Storage,* vol. 52, Aug 2022, Art. no. 104860.

[4] H. Lv, Y. P. Liao, C. L. Zhao, X. H. Shang, and F. J. Zhang, "State of charge estimation of lithium-titanate battery based on multi-model extended Kalman filter considering temperature and current rate," *Journal of Energy Storage,* vol. 77, Jan 2024, Art. no. 109890.

[5] E. D. Bobobee, S. L. Wang, P. Takyi-Aninakwa, C. Y. Zou, E. Appiah, and N. Hai, "Improved particle swarm optimization-long short-term memory model with temperature compensation ability for the accurate state of charge estimation of lithium-ion batteries," *Journal of Energy Storage,* vol. 84, Apr 2024, Art. no. 110871.

[6] P. Takyi-Aninakwa, S. L. Wang, H. Y. Zhang, H. Li, W. H. Xu, and C. Fernandez, "An optimized relevant long short-term memory-squared gain extended Kalman filter for the state of charge estimation of lithium-ion batteries," *Energy,* vol. 260, Dec 2022, Art. no. 125093.

[7] G. X. Chen, W. W. Peng, and F. F. Yang, "An LSTM-SA model for SOC estimation of lithium-ion batteries under various temperatures and aging levels," *Journal of Energy Storage,* vol. 84, Apr 2024, Art. no. 110906.

[8] J. X. Chen, Y. Zhang, J. Wu, W. S. Cheng, and Q. Zhu, "SOC estimation for lithium-ion battery using the LSTM-RNN with extended input and constrained output," *Energy,* vol. 262, Jan 2023, Art. no. 125375.

[9] C. Qian, H. S. Guan, B. H. Xu, Q. Xia, B. Sun, and Y. Ren et.al, "A CNN-SAM-LSTM hybrid neural network for multi-state estimation of lithium-ion batteries under dynamical operating conditions," *Energy,* vol. 294, May 2024, Art. no. 130764.