

Lithium Battery Capacity Prediction Model Based on LSTM Optimized by Two-dimensional Circle Chaotic Map and Random Adjustment Factor Arithmetic Optimization Algorithm

Shi Li, Bing Liu*

Abstract—Lithium-ion batteries have problems such as capacity decline during long-term use, which affects the safety of the system. Achieving high-precision capacity prediction is a key challenge for battery management systems. A fusion prediction framework based on chaos improved arithmetic optimization algorithm (AOA) and long short-term memory (LSTM) network is proposed. The core innovation of the algorithm is that the key parameters of the algorithm are dynamically reconstructed through two-dimensional circle chaotic map, and the random adjustment factor is combined to improve the global exploration ability of the algorithm, which is named CRAOA. Breaking through the limitations of traditional gradient descent methods, CRAOA is used to directly optimize the LSTM weight and bias parameters, which is named CRAOA-LSTM. Experiments show that the CRAOA-LSTM model is superior to other neural networks in terms of prediction accuracy on the NASA battery data set, and the RMSE prediction index of the B0005 and B0006 is reduced by 2.5% ~ 15.2% and 0.87% ~ 21.2%, which provides an effective solution for real-time battery health monitoring.

Index Terms—arithmetic optimization algorithm, long short-term memory neural network, lithium-ion battery capacity

I. INTRODUCTION

With the gradual phase-out of internal combustion engine vehicles by new energy vehicles (NEVs), lithium-ion batteries have emerged as a pivotal research focus as the primary power source [1]. Battery degradation directly compromises both useful life and operational safety, where capacity fade serves as the critical indicator for evaluating state of health (SOH). Accurate SOH prediction not only ensures long-term reliable operation of battery systems by mitigating potential safety risks, but also enables optimized replacement strategies and reduced maintenance costs.

SOH prediction methods for lithium-ion batteries are primarily categorized into physics-based models and

data-driven approaches. Physics-based methodologies, including electrochemical models and equivalent circuit models, require precise physical parameters and theoretically achieve high accuracy, while suffering from substantial modeling complexity [2]. In contrast, data-driven approaches implement state prediction through analysis of historical operational data, circumventing intricate physical modeling processes. However, their predictive performance is significantly influenced by data quality and sample diversity [3]. Current research focuses predominantly on two aspects: SOH estimation and remaining useful life (RUL) prediction, where both methodological categories find extensive application in these domains.

Electrochemical models are constructed based on internal physicochemical mechanisms of batteries, systematically characterizing operational behaviors through fundamental processes including electrode reaction kinetics, mass transport phenomena, and thermodynamic properties [4]. These models integrate multi-physics characteristics (electrical, thermal, and mechanical properties) of materials, providing theoretical foundations for battery behavior analysis. Regarding electrochemical modeling, Li et al. [5] developed a simplified pseudo-two-dimensional (P2D) model using finite difference method for spatial discretization. Building upon degradation mechanisms induced by electrode side reactions, researchers have further proposed the Advanced Degradation Model for Electrochemical systems (ADME). Experimental validation demonstrates that this model significantly reduces maximum prediction error while maintaining computational efficiency, offering an effective approach for lithium-ion battery aging modeling.

Deng et al. [6] introduced a polynomial approximation-based P2D model simplification method, employing distinct polynomial orders to characterize electrolyte phase concentration distribution, solid-phase concentration profile, and heterogeneous reaction flux distribution. Comparative analysis reveals that higher-order polynomials (third-order) provide enhanced robustness and accuracy in reaction flux estimation, establishing an effective simplification strategy for lithium-ion battery modeling. Sadabadi et al. [7] developed a remaining useful life (RUL) prediction algorithm using a single-particle model framework. This approach innovatively utilizes

Manuscript received May 30, 2025; revised September 8, 2025.

Shi Li is a lecturer of the School of Mathematics, Anshan Normal University, Anshan, 114007, P. R. China (e-mail: aslishi@126.com).

Bing Liu is a professor of the School of Mathematics, Anshan Normal University, Anshan, 114007, P. R. China (Corresponding author, phone: 86-0412-2961036; e-mail: liubing529@126.com).

cyclable lithium mole quantity and internal resistance as dual SOH indicators. The study demonstrates that enhanced SOH estimation precision substantially improves RUL prediction accuracy.

Li et al. [8] established an electrochemical-thermal coupled (ECT) model for LiCoO₂ batteries, achieving reduced computational complexity through mechanistic parameter optimization and thermal dissipation analysis. The research implements parameter sensitivity analysis for rational model simplification, integrates comprehensive thermodynamic descriptions, and enables accurate identification of critical parameters based on stimulus response characteristics.

Equivalent circuit models (ECMs) characterize battery external characteristics through electrical component networks, establishing coupled relationships between state of charge (SOC) and multiple influencing factors, demonstrating significant application value in engineering practice. Hossain [9] developed a temperature-compensated parameter identification method enabling precise parameter extraction across an extended temperature range (-5° C to 45° C). Nikohan et al. [10] proposed a smoothed adaptive estimation algorithm that effectively enhances model prediction robustness under varying thermal and load conditions. Ding et al. [11] significantly improved Thevenin model terminal voltage calculation accuracy by introducing a temperature-dependent open-circuit voltage correction term, achieving this enhancement through mathematical characterization of SOC-OCV relationships without increasing model complexity. These methods collectively advance ECM performance in practical applications.

Yuyang et al. [12] introduced an improved 2RC-PNGV battery model featuring an organic light-emitting resistor parallel circuit replacing conventional internal resistance, while employing RC networks to characterize both dynamic and static battery properties. The study innovatively accounts for differences between charge/discharge processes, implementing online parameter identification through a forgetting factor recursive least squares (FFRLS) method.

Recent advancements in machine learning have propelled the widespread application of data-driven approaches in battery state prediction. These methods learn nonlinear mapping relationships directly from historical data, circumventing the reliance on mechanistic parameters inherent to traditional physical models, while demonstrating superior prediction accuracy and adaptability. Xiong et al. [13] innovatively combined correlation coefficient analysis, LASSO regression, Neighborhood Component Analysis (NCA) and ReliefF algorithm to construct a hybrid feature selection framework, and extracted the most representative health status indicators from the measured parameters and derived features of the battery system through the multi-dimensional statistical analysis method. Qian et al. [14] innovatively developed a dynamic operating condition prediction framework combining historical states and future load information. Their approach employs an attention-mechanism-based multi-source sequence-to-sequence (AM-seq2seq) model with a dual-encoder-single-decoder architecture, where

attention layers effectively capture complex dependencies between prediction targets and multi-source inputs.

Ma et al. [15] introduced a SOH estimation method combining enhanced LSTM networks with health indicators derived from charge/discharge processes. Health indicators are constructed using external features including voltage, current, and temperature profiles during cycling. Key capacity-correlated features are selected through Pearson correlation coefficient analysis, while redundant information is eliminated via neighborhood component analysis (NCA) to reduce computational complexity. The study further implements a differential evolution grey wolf optimizer (DEGWO) for LSTM hyperparameter optimization.

Ren et al. [16] developed an Auto-CNN-LSTM prediction framework integrating improved convolutional neural networks (CNN) and LSTM architectures. This method enhances feature dimensions of limited data through autoencoder techniques, combines deep CNN and LSTM networks to extract latent degradation patterns, and incorporates filtering algorithms to smooth prediction outputs.

Chen et al. [17] designed a hybrid neural network architecture fusing multi-channel CNN, bidirectional LSTM, and gated recurrent unit (GRU) with attention mechanisms for precise SOH and RUL assessment. The methodology first applies the Pauta criterion to eliminate abnormal data for capacity trajectory smoothing, subsequently extracts multidimensional health features from charging profiles, and selects critical degradation indicators through correlation analysis. Comparative experiments demonstrate exceptional prediction accuracy and generalization capabilities.

A novel optimization methodology for Long LSTM networks is proposed, utilizing a metaheuristic algorithm enhanced with two-dimensional circle chaotic mapping and random adjustment factors. Departing from conventional gradient descent-based weight and bias optimization, the AOA is fundamentally improved through chaotic initialization and adaptive parameter control mechanisms, the search space of the algorithm was increased. In this study, the proposed CRAOA-LSTM model is systematically verified based on NASA standard lithium-ion battery datasets (Battery #5 and Battery #6). The experimental results show that the framework shows excellent accuracy and robustness in battery capacity prediction, which fully verifies its effectiveness in practical applications.

The sections of the manuscript are structured as follows. The enhanced AOA is elaborated in Section II. The foundational architecture of LSTM networks is presented in Section III. The CRAOA-LSTM optimization methodology is systematically described in Section IV. Comprehensive experimental simulations and comparative analyses are provided in Section V. Finally, conclusive findings and future research directions are discussed in Section VI.

II. THE PROPOSED AOA

This subsection provides a more detailed description of standard equivalent circuit models and gives the second-order RC model used; it introduces the basic theory

of KF, with an emphasis on the extended Kalman filter and AEKF. This paper aims to establish a battery SOC estimation model through comparison of various estimation methods. The SOC is then estimated using the adaptive extended Kalman filter method and the SORC equivalent circuit model. Finally, the model's accuracy is verified through simulation using MATLAB.

A. AOA

Arithmetic is the foundation of number theory and also one of the core branches of modern mathematics. Its basic operators include: Addition (A “+”), Subtraction (S “−”), Multiplication (M “×”), and Division (D “÷”). These operators serve not only as fundamental tools for numerical analysis but also facilitate mathematical optimization by identifying optimal solutions from candidate sets. The AOA framework is constructed based on the dynamic interactions of arithmetic operators. This algorithm establishes a mathematical model to capture the dynamic behavior of arithmetic operators during optimization. It has extensive application potential in fields such as engineering and computer science. In the following sections, the specific influence of arithmetic operators on AOA will be specifically discussed [18]. Fig. 1 shows the search strategy of the AOA.

(1) Initialization phase

The optimization process of the AOA starts with a randomly generated initial set of candidate solutions. In the iterative process, the population is dynamically evaluated and updated, where the best solution of each generation will be used as the candidate of the current global optimal solution, and the optimal solution of the problem is gradually approximated by continuous optimization.

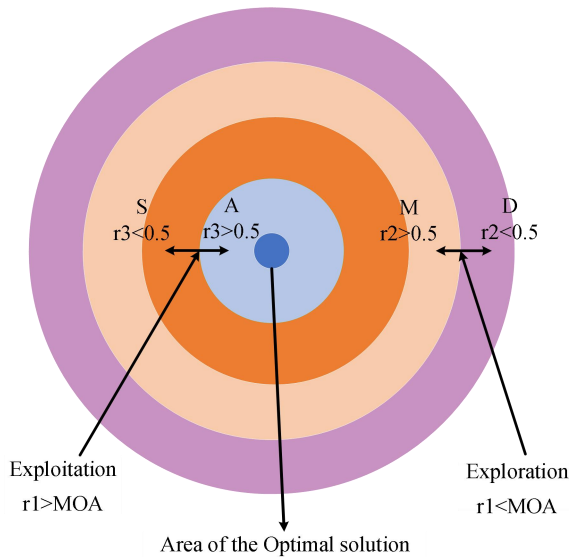


Fig. 1 Schematic diagram of the Rint model.

$$X(i, j) = rand \times (Ub - Lb) + Lb \quad (1)$$

Before initiating the AOA, the Mathematical Optimization Accelerator (MOA) selects the search phase (exploration or exploitation) based on Eq. (2).

$$MOA(C_Iter) = Min + C_Iter \times \left(\frac{Max - Min}{M_Iter} \right) \quad (2)$$

Where, the parameter $MOA(C_Iter)$ denotes the function value at the C_Iter iteration, while C_Iter indicates the current iteration count (bounded between 1 and the maximum iteration limit M_Iter). A random number $r_1 \in (0,1)$ is generated. The algorithm switches to exploration if $r_1 > MOA$, otherwise continues exploitation.

(2) Exploration phase

The exploration phase of AOA is performed using either the division operator (D) or the multiplication operator (M). This operator has been used to generate widely distributed candidate solutions due to its high dispersion. The search space can be efficiently explored and premature convergence avoided by this property, while gradually approaching the potential optimal solution through multiple iterations. Its position update formula is shown in Eq. (3).

$$X_{i,j} = \begin{cases} best(x_j) \div (MOP + \epsilon) \times \mu + LB_j, & r_2 < 0.5 \\ best(x_j) \times MOP \times ((UB_j - LB_j) \times \mu + LB_j) & \end{cases} \quad (3)$$

Where r_2 represents a random number between (0,1). $X_{i,j}(C_Iter)$ denotes the j -th position of the i -th solution in the current iteration. $best(x_j)$ denotes the j -th position of the best solution so far. ϵ is a small integer and μ is a control parameter that regulates the search process, $\mu = 0.5$.

$$MOP(C_Iter) = 1 - \frac{C_Iter^{1/\alpha}}{M_Iter^{1/\alpha}} \quad (4)$$

In Eq. (4), MOP is the mathematical optimization probability and α represents a fixed value, which is 5. Fig. 2 shows the dynamic evolution law of MOA and MOP with the iterative process.

(3) Exploitation phase

In the development stage of AOA algorithm, the local refinement search of solutions is realized by Subtraction operator (S) and Addition operator (A) with low dispersion characteristics. The synergy of these two operators can effectively guide the search individual to quickly converge to the optimal solution neighborhood, and the specific position update mechanism is given by Eq. (5).

Among them, r_3 represents a random number between (0,1). $X_{i,j}$ indicates the j -th position of the i -th solution in the current iteration. μ is the control parameter for adjusting the search process, and $\mu = 0.5$.

$$X_{i,j} = \begin{cases} best(x_j) - MOP \times ((UB_j - LB_j) \times \mu + LB_j) \\ best(x_j) + MOP \times ((UB_j - LB_j) \times \mu + LB_j) \end{cases} \quad (5)$$

B. Two-dimensional circle chaotic mapping AOA

As a random like method with deterministic characteristics, the essence of chaotic system is to generate pseudo-random sequences through nonlinear equations. Different from traditional random variables, chaotic variables can show complex random behaviors while maintaining the certainty of the system. Although one-dimensional chaotic maps have been widely used in function optimization and other fields, two-dimensional chaotic maps are currently mainly concentrated in security fields such as image encryption, and their application in optimization algorithms is still in the exploratory stage [19]. The one-dimensional circle chaotic map has the advantages

of uniform traversal, low correlation and low computational complexity, so it is constructed as a two-dimensional circle chaotic map and applied in AOA algorithm. The formula of two-dimensional circle chaotic map is as follows:

$$\begin{cases} x_{k+1} = \text{mod}\left(x_k + \Omega - \frac{K}{2\pi} \sin(2\pi y_k), 1\right) \\ y_{k+1} = \text{mod}\left(y_k + \Omega - \frac{K}{2\pi} \sin(2\pi(x_k)^2), 1\right) \end{cases} \quad (6)$$

Eq. (6) reveals the coupling characteristics of the X-axis and Y-axis chaotic sequences in the two-dimensional circle map, which is represented by the nonlinear interaction between two variables. As shown in Fig. 3, the chaotic trajectory generated by the mapping in the phase space shows typical ergodicity and pseudo-random characteristics, and this bidirectional dynamic coupling mechanism provides a rich chaotic source for parameter perturbation.

$$MOA = x_{k+1} \times \left(\text{Min} + C_Iter \times \left(\frac{\text{Max} - \text{Min}}{M_Iter} \right)^m \right) \quad (7)$$

$$MOP = y_{k+1} \times \left(1 - \frac{C_Iter^{1/\alpha}}{M_Iter^{1/\alpha}} \right) \quad (8)$$

Where $x_1 \in (0,1)$, $y_1 \in (0,1)$, k is the current iteration number. m denotes the order of the MOA. $m > 0$, and the value of m is 2. Fig. 4 shows the dynamic adjustment process of two key parameters of MOA and MOP by two-dimensional chaotic interference factors, in which the random sequence generated by chaotic mapping acts on the parameter update mechanism through nonlinear transformation, which intuitively shows the control effect of chaotic disturbance on the search behavior of the algorithm.

Where r_1 represents a random number between (0,1). When $r_1 > \text{MOA}$, the algorithm is in the exploration phase, otherwise the algorithm is in the exploitation phase. By observing the comparison between the improved MOA image and the original MOA image in Fig. 4, it can be found that the value of the improved MOA is smaller than the value of the original MOA in the early stage of the search. When $r_1 > \text{MOA}$, the algorithm is in the exploration stage, so this expands the ability and scope of the global search, and well balances the global search ability and local search ability of the algorithm. By accelerating the downward trend of MOP and adding random fluctuations, the convergence speed and accuracy of the algorithm can be improved.

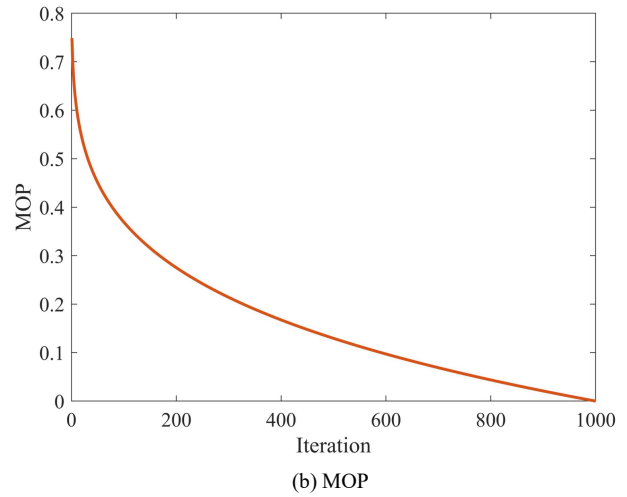
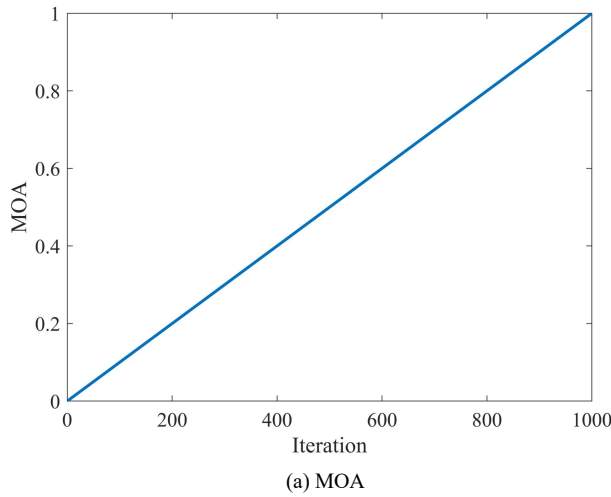


Fig. 2 Images of the MOA and the MOP.

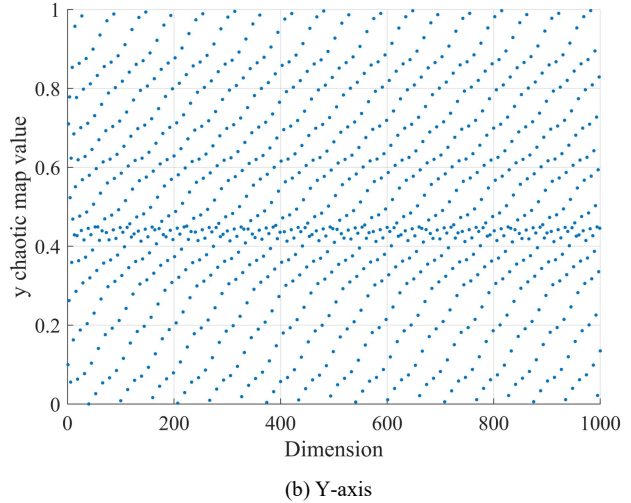
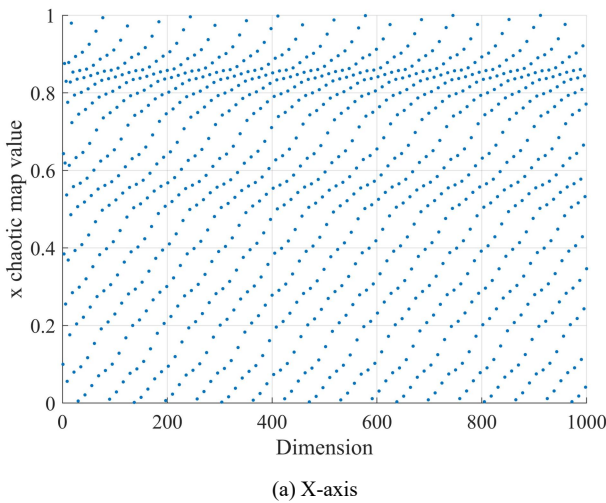


Fig. 3 Two-dimensional circle mapping X-axis and Y-axis chaotic mapping maps.

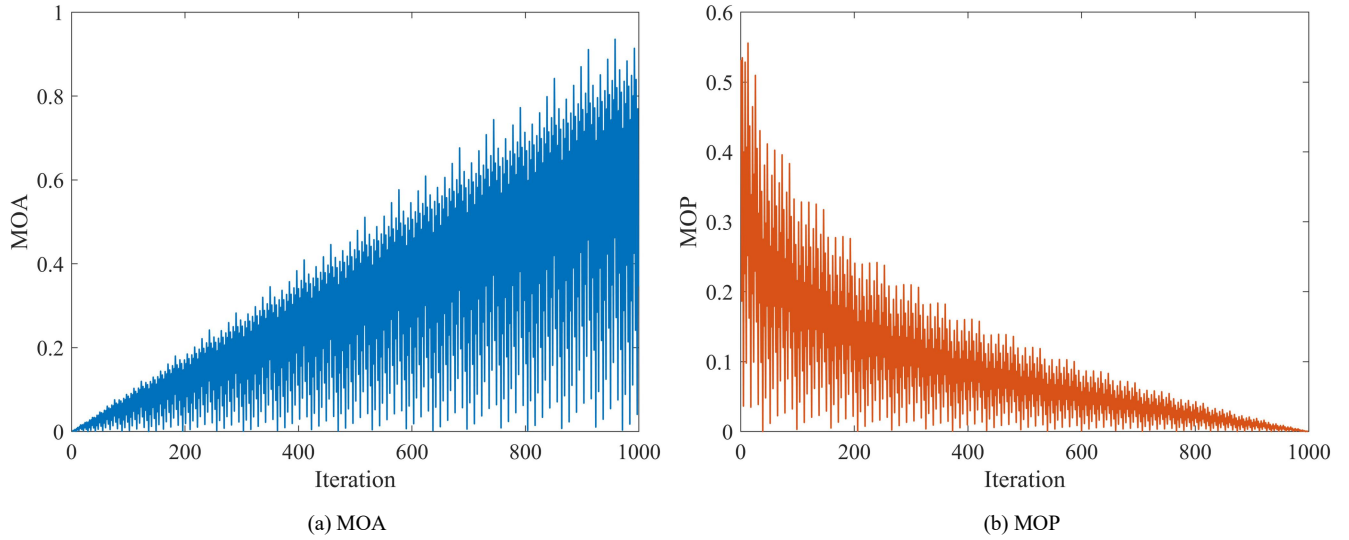
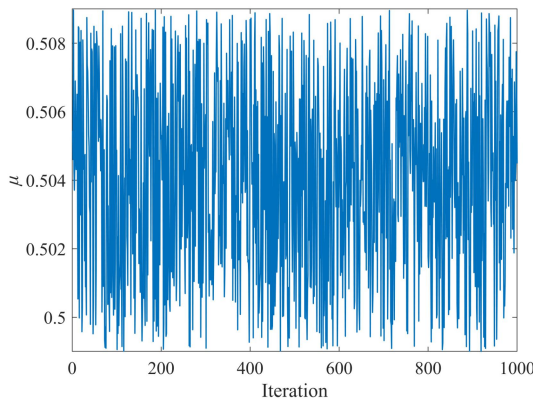


Fig. 4 Images of the MOA and the MOP.

C. AOA with random adjustment Factor

The search core strategy of AOA algorithm is reflected in the position update Eq. (3) and Eq. (5). In addition to the Addition, Subtraction and Multiplication and division operators, there is a most important adjustment factor μ . The original algorithm is set to a constant value of 0.5, which will affect the search space of the algorithm and lead to a local optimum. Therefore, a random adjustment factor is proposed, as shown in Eq. (9), where r is a random number between (0,1). The value fluctuates around 0.5, which enhances the search space of the original algorithm and improves the ability of global search and local search.

$$\mu = 0.499 \times 0.01 \times r \quad (9)$$


 Fig. 5 The random adjustment result of the μ value.

III. LONG SHORT-TERM MEMORY NEURAL NETWORK

As an improved architecture of Recurrent Neural Network (RNN), LSTM replaces the hidden neuron structure of traditional RNN by a unique memory cell unit, and its core innovation is to introduce a gating mechanism to solve the long-term dependence problem. In the cell

The core trainable parameters of LSTM include four types of weights and biases, corresponding to different gating mechanisms:

(1) Input gate parameters: $W_{i,x}$, $W_{i,h}$, b_i . It controls the selection and combination of new information, and adjusts the input retention rate and information intensity through the activation function.

structure shown in Fig. 6, the cooperative operation of input gate, forget gate and output gate enables the network to selectively store and update time series information, thus significantly improving the modeling ability of long sequence data [20]. Three gate controllers are placed in the LSTM cell, which are the input gate corresponding to the red dashed box, the forget gate corresponding to the blue dashed box, and the output gate corresponding to the orange dashed box. The corresponding parameters of each gate in Fig. 6 are calculated as follows.

$$f_t = \sigma(W_{f,x} \cdot x_t + W_{f,h} \cdot h_{t-1} + b_f) \quad (10)$$

$$i_t = \sigma(W_{i,x} \cdot x_t + W_{i,h} \cdot h_{t-1} + b_i) \quad (11)$$

$$\tilde{c}_t = \tanh(W_{\tilde{c},x} \cdot x_t + W_{\tilde{c},h} \cdot h_{t-1} + b_{\tilde{c}}) \quad (12)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (13)$$

$$o_t = \sigma(W_{o,x} \cdot x_t + W_{o,h} \cdot h_{t-1} + b_o) \quad (14)$$

$$h_t = o_t * \tanh(c_t) \quad (15)$$

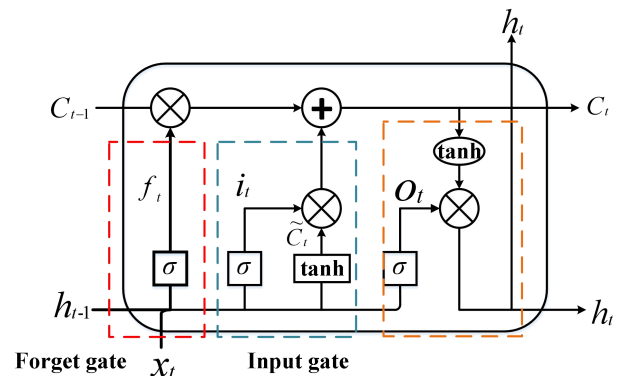


Fig. 6 LSTM cell structure diagram.

(2) Parameters of forgetting gate: $W_{f,x}$, $W_{f,h}$, b_f . Determine the proportion of retention of historical memory and regulates the degree of forgetting of preorder cell states.

(3) Cell state parameters: $W_{\tilde{c},x}$, $W_{\tilde{c},h}$, $b_{\tilde{c}}$. Update mechanisms that manage memory information and combine old and new information to form the current memory.

(4) Output gate parameters: $W_{o,x}$, $W_{o,h}$, b_o . Generate output gate that regulates the final output and calculates the output value based on the cell state and hidden state. The execution schematic of the LSTM is represented in Fig. 7.

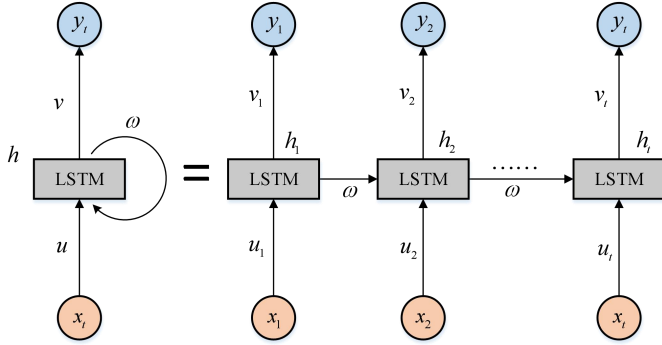


Fig. 7 LSTM execution schematic.

IV. OPTIMIZED LSTM BASED ON THE PROPOSED AOA

The problem of gradient disappearance in the parameter optimization process of deep neural networks can be traced back to the differential attenuation effect caused by the chain rule of the back propagation algorithm. The pioneering work of Hochreiter S. and his team in the field of recurrent neural networks was the first to theoretically prove that in time series modeling tasks, the learning framework based on gradient descent will cause the error gradient to decay exponentially with the time step, which seriously restricts the ability of the model to capture long-range dependencies [21].

Although advanced architectures such as LSTM alleviate the gradient decay problem in deep temporal modeling to some extent through the co-design of gating mechanism and cell state memory unit, their essential limitation still stems from the gradient backpropagation mechanism itself. Specifically, all deep learning models based on gradient descent are restricted by the topology of parameter space. Such optimization paradigms that rely on local gradient information are difficult to avoid the progressive weakening of gradient signals in deep networks or long sequences. Empirical studies show that although the joint application of regularization techniques such as residual connection and gradient clipping with adaptive optimizers can significantly improve the optimization effect, this problem has not been fundamentally solved at the theoretical level.

Based on this problem, a method of optimizing LSTM neural network by CRAOA is proposed, which is named CRAOA-LSTM. The gradient descent method used to update weights and biases is abandoned, and the swarm intelligence optimization algorithm that is popular and performs well in recent years is used instead.

In CRAOA-LSTM, each individual vector position represents a set of weight bias combinations. In the LSTM architecture, the dimension of each optimization vector is determined by the total number of network parameters, which is expressed as the sum of the weight matrix and the bias term.

Let the input variable dimension be n and the number of hidden layer neurons be m , then the trainable

parameters of LSTM can be expressed as follows: (1) Input weights: $\bar{U}, \bar{U} \in R^{8m \times n}$; (2) Cycle weight: $\bar{W}, \bar{W} \in R^{8m \times m}$; (3) Full connection weight: $V \in R^{1 \times 2m}$; (4) Deviation: $\bar{b}, \bar{b}, c \in R^{(8m+1) \times 1}$. The total number of trainable parameters of the LSTM network can be calculated by the following formula, which is determined by the weight matrix from the input layer to the hidden layer, the recurrent weight, and the bias term.

$$x = 8mn + 8m^2 + 8m + 2m + 1 \quad (16)$$

To select the appropriate fitness value to test each CRAOA search agent, $RMSE$ is used as the fitness function, which represents the dissimilarity between the measured and predicted values, and a smaller value proves that the network predicts better. $RMSE$ is shown in Eq. (17), where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of training data set.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|} \quad (17)$$

Fig. 8 represents the assignment of the search agents of the CRAOA algorithm to the LSTM neural network, and the figure can also represent the process of extracting and reconstructing the weight and bias in the LSTM network into a one-dimensional column vector. After obtaining the optimal search agent position, it needs to assign the position corresponding to the weight and deviation to the network for prediction. The flow chart of the CRAOA-LSTM algorithm is shown in Fig. 9.

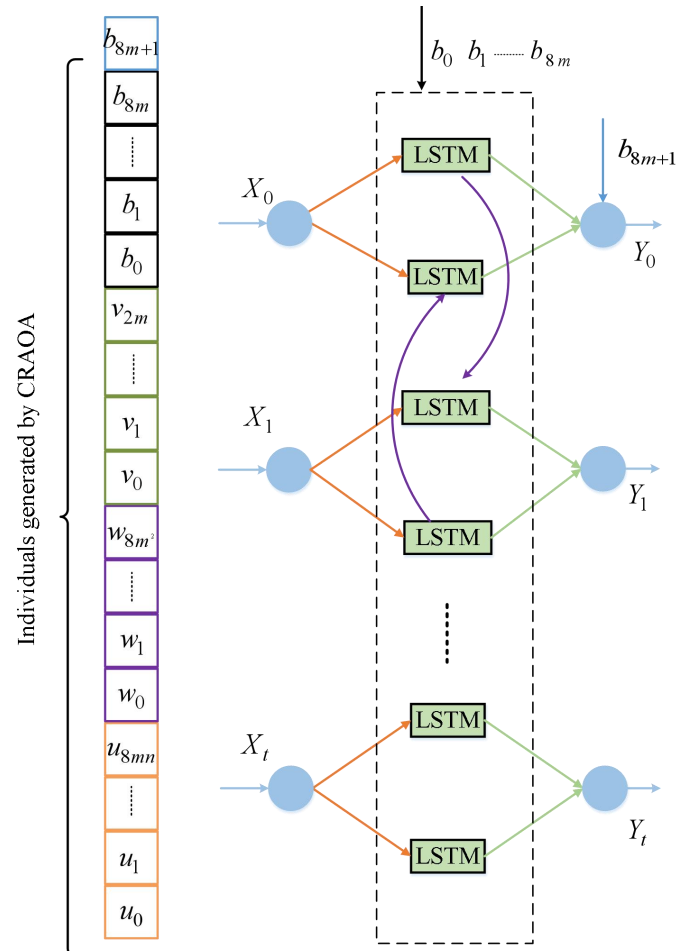


Fig.8 Assigning CRAOAsearch agent vectors to LSTM.

V. SIMULATION EXPERIMENT AND RESULT ANALYSIS

A. Solving function optimization problems based on CRAOA

In this section, the CEC 2022 benchmark function is adopted to verify the performance of CRAOA. The comparison algorithms and parameter settings are shown in Table 1, and the parameter settings are derived from the relevant literature.

To make the comparison experiments more fair, the experimental parameters were uniformly set as follows: population size was set to 20, maximum number of iterations of the evaluation function was set to 1000, dimension was set to 10 and each algorithm was run 20 times independently of each test function.

To verify the improved performance of the algorithm based on the original AOA algorithm, two improved strategies of two-dimensional circle chaotic map AOA (CAOA) and random adjustment factor AOA (RAOA) were introduced, and the ablation experiments of function optimization were carried out. Four sets of four different types of test function optimization effects are shown here, as shown in Fig. 10.

Fig. 10 shows that the red diamond curve is the CRAOA fused by the two improvements, the green quincunx curve is RAOA, the black five-pointed star curve is CAO, and the blue circle curve is the original AOA. Among them, the improved schemes of CRAOA, CAO and RAOA are better than the original algorithm, and CRAOA is the most influential algorithm among the three schemes under the same experimental conditions, and it is also the best to find the extreme point of the function.

The analysis results of the convergence characteristics of each test function are shown in Fig. 11, and the quantitative

evaluation data are detailed in Table 2. In the table, the convergence performance indicators of 30 independent experiments are calculated, including the average solution accuracy (Ave), the standard deviation of stability representation (Std) and the Best solution quality (Best). The performance Rank of each algorithm is given according to the comprehensive performance.

This multi-dimensional evaluation system not only intuitively shows the optimization process of the algorithm through the convergence curve, but also objectively compares the solution accuracy and robustness of different methods by means of statistical indicators.

From Table II, we can see that the overall performance of CRAOA is better than other advanced algorithms, such as function $f_1 \sim f_3$, $f_6 \sim f_{12}$, and the performance only in function $f_4 \sim f_5$ is not the best, but it is always ranked in the top three. However, CRAOA is still the best performer when ranking according to the average of all functions. The comparative analysis of the convergence curve in Fig. 11 shows that CRAOA algorithm (red diamond line) shows significant performance advantages.

TABLE I. COMPARISON OF ALGORITHM PARAMETERS

Algorithm	Parameters setting
BAT	$\alpha = 0.5, \lambda = 0.5, ro = 0.001$
AOA	$\alpha = 0.5, \mu = 0.5$
COA	$N = 30$
RSA	$\alpha = 0.5, \beta = 0.005$
CFOA	$\rho \in [0, 1]$
CRAOA	$\alpha = 0.5$

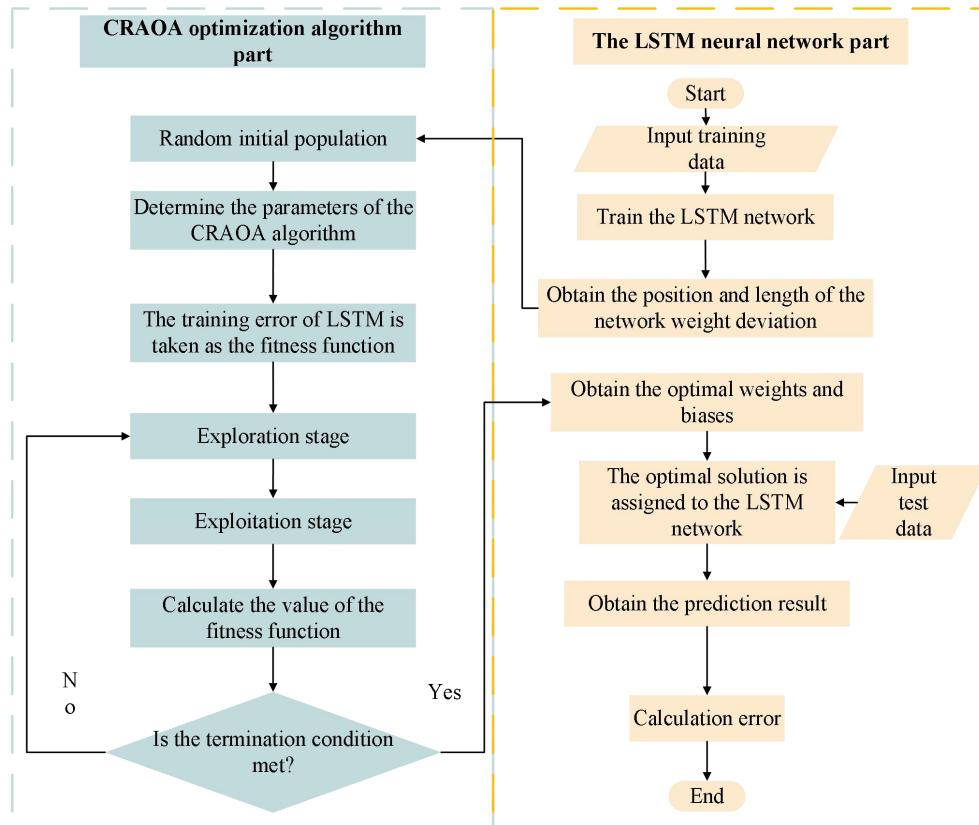


Fig. 9 CRAOA-LSTM flow chart.

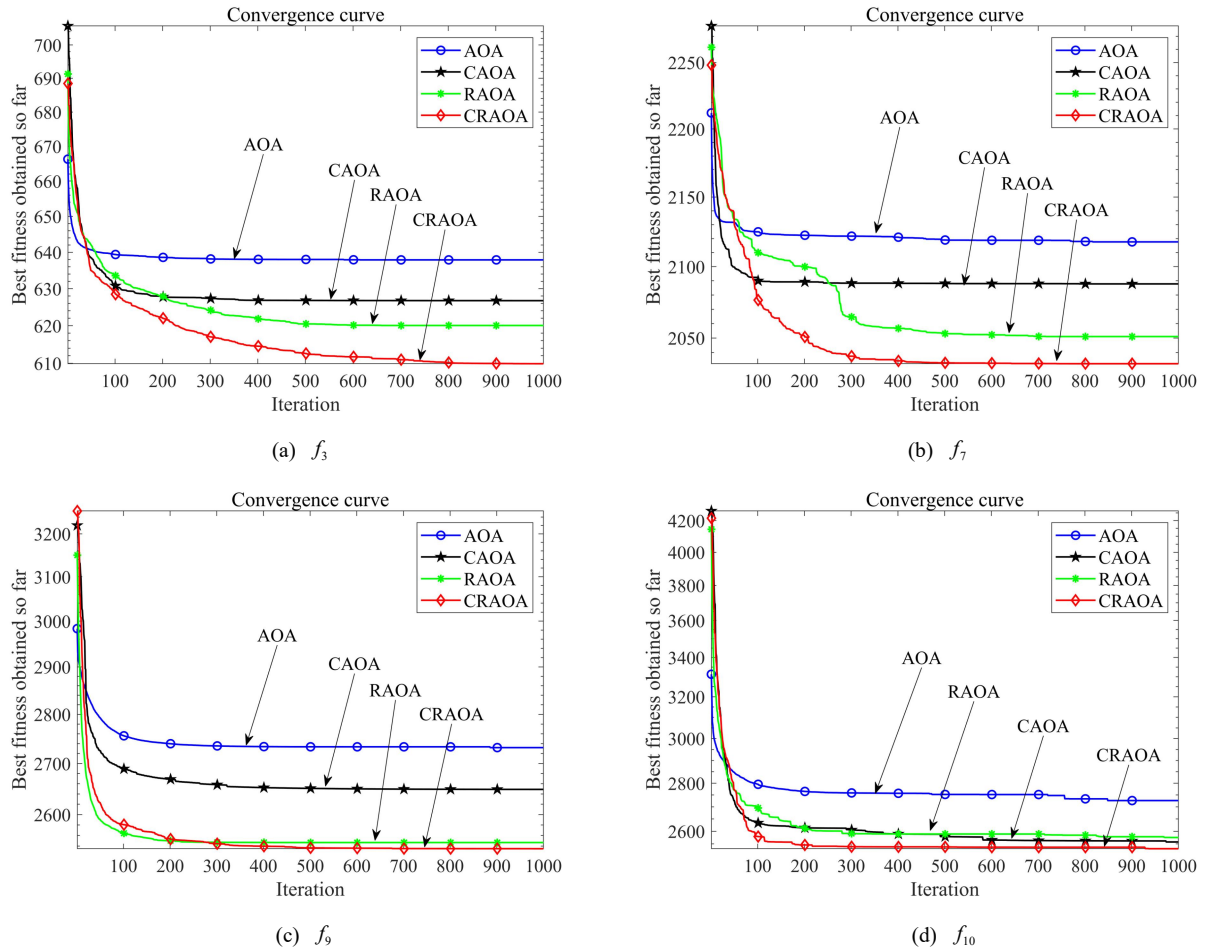
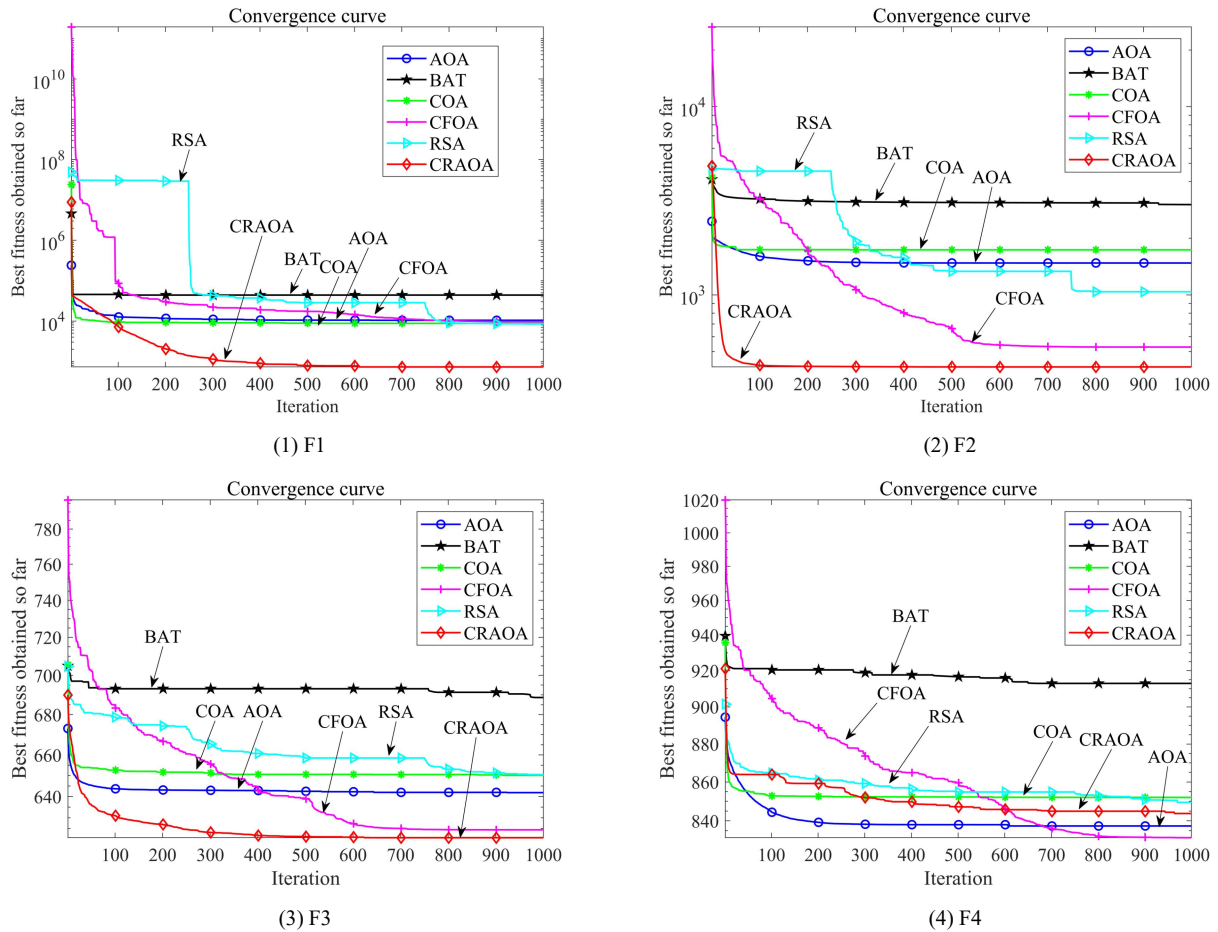
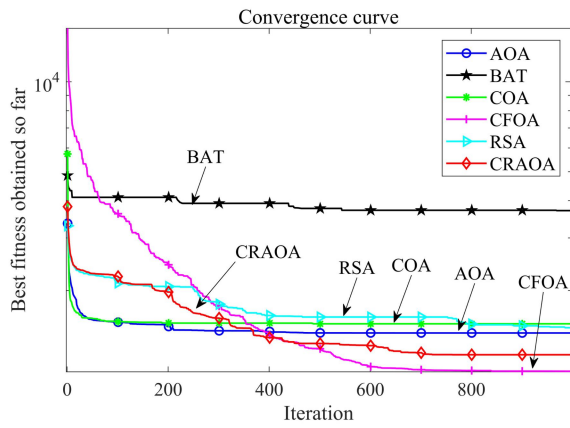
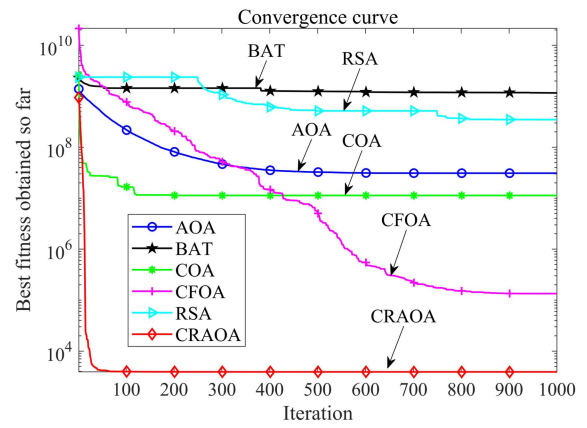


Fig. 10 Ablation experiments on four different types of test functions f_3, f_7, f_9, f_{10} .

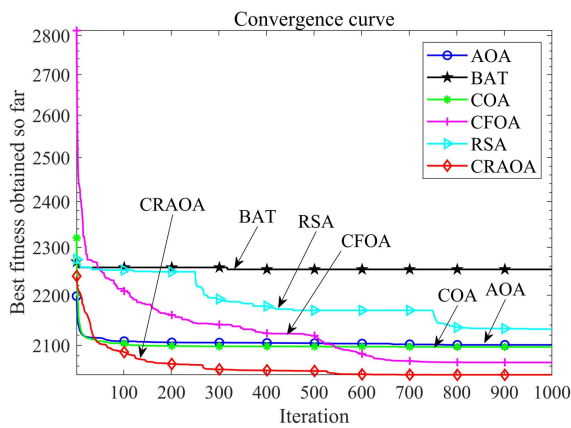




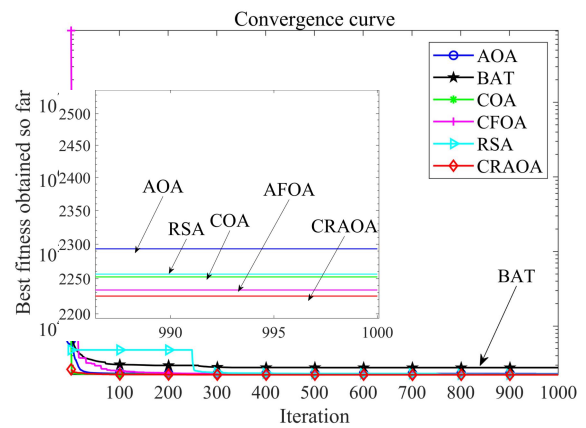
(5) F5



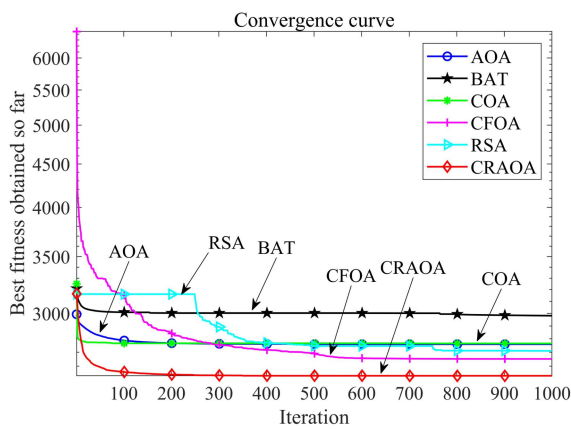
(6) F6



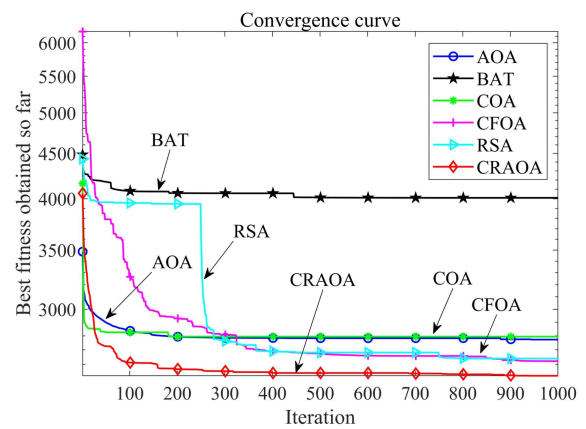
(7) F7



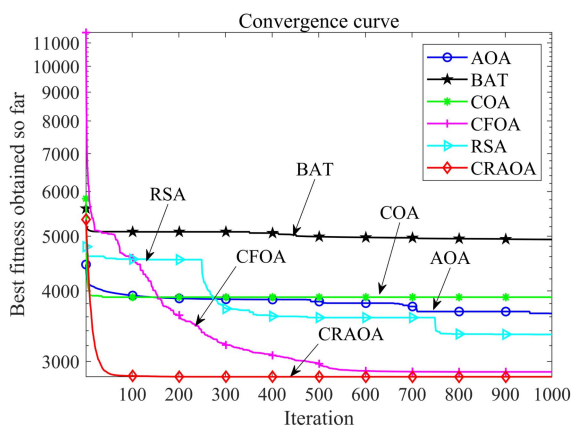
(8) F8



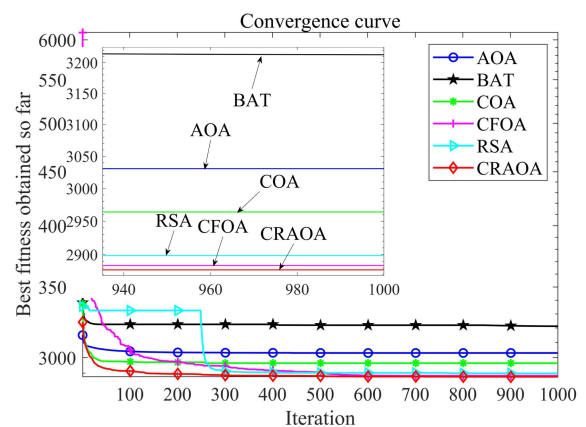
(9) F9



(10) F10



(11) F11



(12) F12

Fig.11 Comparison curves of CRAOA with other algorithms in function optimization.

TABLE II. COMPARISON DATA BETWEEN CRAOA AND OTHER ALGORITHMS ON TEST FUNCTION OPTIMIZATION

Function		AOA	BAT	COA	CFOA	RSA	CRAOA
F1	Best	5.164E+03	1.440E+04	4.796E+03	3.830E+03	3.644E+03	3.178E+02
	Ave	1.044E+04	4.406E+04	8.702E+03	9.615E+03	8.471E+03	7.291E+02
	Std	5.402E+03	2.252E+04	1.770E+03	4.169E+03	2.299E+03	4.396E+02
	Rank	5	6	3	4	2	1
F2	Best	5.955E+02	1.325E+03	8.861E+02	4.672E+02	5.916E+02	4.000E+02
	Ave	1.477E+03	3.017E+03	1.735E+03	5.281E+02	1.039E+03	4.142E+02
	Std	6.218E+02	1.426E+03	6.649E+02	5.349E+01	4.782E+02	2.346E+01
	Rank	4	6	5	2	3	1
F3	Best	6.303E+02	6.503E+02	6.338E+02	6.135E+02	6.361E+02	6.057E+02
	Ave	6.419E+02	6.887E+02	6.501E+02	6.244E+02	6.504E+02	6.208E+02
	Std	9.180E+00	2.141E+01	6.500E+00	6.540E+00	6.380E+00	9.010E+00
	Rank	3	6	4	2	5	1
F4	Best	8.187E+02	8.794E+02	8.344E+02	8.180E+02	8.387E+02	8.180E+02
	Ave	8.374E+02	9.129E+02	8.520E+02	8.315E+02	8.494E+02	8.438E+02
	Std	1.073E+01	2.047E+01	8.350E+00	1.010E+01	7.790E+00	1.571E+01
	Rank	2	6	5	1	4	3
F5	Best	1.121E+03	1.823E+03	1.233E+03	9.790E+02	1.201E+03	9.376E+02
	Ave	1.443E+03	3.706E+03	1.549E+03	1.075E+03	1.503E+03	1.221E+03
	Std	2.014E+02	9.506E+02	1.813E+02	8.472E+01	1.970E+02	1.719E+02
	Rank	3	6	5	1	4	2
F6	Best	1.970E+03	5.380E+07	4.280E+05	2.256E+03	2.852E+07	1.912E+03
	Ave	3.097E+07	1.163E+09	1.132E+07	1.343E+05	3.484E+08	3.902E+03
	Std	8.047E+07	1.122E+09	1.268E+07	2.726E+05	6.457E+08	2.243E+03
	Rank	4	6	3	2	5	1
F7	Best	2.045E+03	2.170E+03	2.053E+03	2.032E+03	2.073E+03	2.012E+03
	Ave	2.101E+03	2.253E+03	2.097E+03	2.067E+03	2.132E+03	2.043E+03
	Std	3.051E+01	5.450E+01	2.809E+01	2.314E+01	3.515E+01	1.440E+01
	Rank	4	6	3	2	5	1
F8	Best	2.224E+03	2.245E+03	2.228E+03	2.224E+03	2.239E+03	2.222E+03
	Ave	2.293E+03	2.775E+03	2.252E+03	2.234E+03	2.257E+03	2.225E+03
	Std	8.211E+01	6.182E+02	1.775E+01	7.940E+00	1.314E+01	2.260E+00
	Rank	5	6	3	2	4	1
F9	Best	2.653E+03	2.754E+03	2.670E+03	2.601E+03	2.620E+03	2.529E+03
	Ave	2.760E+03	2.982E+03	2.766E+03	2.652E+03	2.710E+03	2.533E+03
	Std	5.350E+01	1.913E+02	4.396E+01	2.756E+01	4.801E+01	5.340E+00
	Rank	4	6	5	2	3	1
F10	Best	2.531E+03	2.543E+03	2.527E+03	2.501E+03	2.515E+03	2.403E+03
	Ave	2.771E+03	4.009E+03	2.794E+03	2.622E+03	2.639E+03	2.524E+03
	Std	1.883E+02	9.055E+02	2.630E+02	2.406E+02	9.718E+01	5.467E+01
	Rank	4	6	5	2	3	1
F11	Best	2.964E+03	3.495E+03	3.037E+03	2.765E+03	2.879E+03	2.600E+03
	Ave	3.648E+03	4.934E+03	3.899E+03	2.875E+03	3.349E+03	2.818E+03
	Std	4.038E+02	8.314E+02	4.460E+02	1.547E+02	4.700E+02	1.609E+02
	Rank	4	6	5	2	3	1
F12	Best	2.928E+03	2.912E+03	2.881E+03	2.869E+03	2.878E+03	2.864E+03
	Ave	3.031E+03	3.213E+03	2.964E+03	2.884E+03	2.899E+03	2.877E+03
	Std	6.268E+01	1.426E+02	6.470E+01	1.481E+01	2.639E+01	2.343E+01
	Rank	5	6	4	2	3	1

B. Battery capacity prediction based on CRAOA-LSTM

(1) Battery dataset selection

It is proposed to estimate the capacity of lithium-ion batteries based on the voltage, current and temperature data collected by the Battery Management System (BMS). The experiment adopted data from eight 18650 lithium-ion batteries in NASA's database, including three working modes: charging, discharging and resting. The battery is charged by constant current and constant voltage (CCCV) method (1.5A constant current to 4.2V, then constant voltage to 20mA), and discharged at different currents (1A or 2A) to the cut-off voltage (2.0V - 2.7V) to accelerate the aging process [22].

Studies have shown that the capacity regeneration phenomenon will occur if lithium-ion batteries enter the resting period after many charging and discharging cycles, but the internal parameters change slowly, and the current technology is still difficult to achieve non-destructive monitoring of these microscopic parameters, and it is difficult to directly measure. In contrast, the external parameters of the charging process show significant changes with the number of cycles, making it more suitable for capacity prediction. Therefore, the characteristic parameters of the charging stage are selected for modeling. The experimental data have undergone preprocessing such as normalization. The specific batteries selected are Battery #5 and Battery #6, which are represented by the blue line and the red curve in Fig 12 respectively. Among them, the black dotted line represents the battery attenuation threshold. When the battery exceeds this range, it reaches the terminal life of the battery.

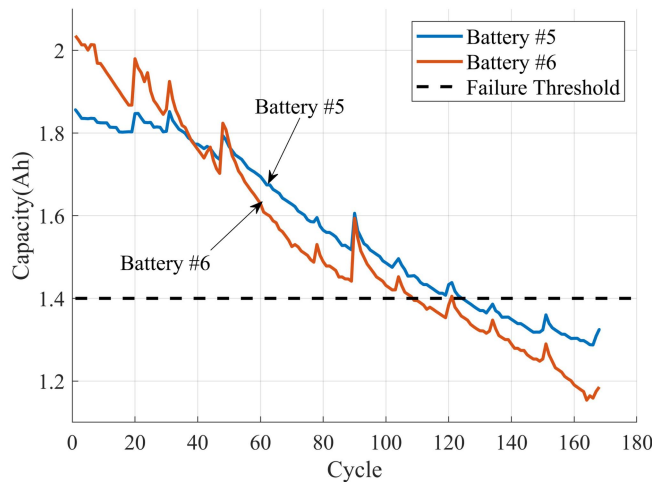


Fig. 12 Capacity attenuation curves of Battery #5 and #6 in NASA battery pack.

Through the visual analysis of the fitting curve and error curve, the prediction effect of each model on the battery capacity attenuation trend is intuitively presented. To further evaluate the performance of the model quantitatively, the three indicators of RMSE, MAE and MAPE are used for comprehensive evaluation, and the specific calculation formula is shown in Table III.

This qualitative and quantitative analysis method not only intuitively shows the overall fitting degree of CRAOA-LSTM and other classical models through image comparison, but also objectively compares the prediction

accuracy through multi-dimensional error indicators, which comprehensively verifies the performance differences of different algorithms.

TABLE III. QUANTITATIVE INDEX OF BATTERY CAPACITY PREDICTION MODEL

Index	Function
Mean absolute error	$MAE = \frac{1}{n} \sum_{i=1}^n \hat{y}_i - y_i $
Root mean square error	$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \right]^{\frac{1}{2}}$
Mean Absolute Percentage Error	$MAPE = \frac{1}{n} \sum_{i=1}^n \left \frac{\hat{y}_i - y_i}{y_i} \right $

(2) Neural network and optimizer parameter settings

To verify the superiority of CRAOA-LSTM, it is proposed to compare with classical deep learning models FNN and LSTM. For FNN, since there is not much battery data in the experiment, it does not need too many training times and layers, and the experiment is set to 500 times and 2 layers.

The optimizer of choice is Adam, which is widely used by many researchers. The parameter settings of LSTM are similar to FNN, the number of training is also 500 times, Adam optimizer is selected, the number of LSTM stacking layers is set to 5 layers, and the learning rate is set to 0.001.

Furthermore, different meta-heuristic algorithms for optimizing the LSTM model are also introduced. The AOA for optimizing LSTM (AOA-LSTM), the BAT Algorithm for optimizing LSTM (BAT-LSTM), and the Coati Optimization Algorithm for optimizing LSTM (COA-LSTM), Catch Fish Optimization Algorithm for optimizing LSTM (CFOA-LSTM), and Reptile Search Algorithm for optimizing LSTM (RSA-LSTM) to predict battery capacity. To ensure the fairness of the experiment, the number of populations is uniformly set to 20, and the maximum number of iterations is set to 500.

The parameter selection of each algorithm is consistent with the original paper. For the partition of the original dataset, 70% of the population was randomly selected as the training set, 15% as the validation set, and 15% as the test set.

(3) Capacity prediction of Battery #5 and Battery #6

In this part, the CRAOA-LSTM prediction model is compared with other optimization algorithms to optimize LSTM and other neural network prediction, as shown in Fig. 13 to 15. Fig. 13 shows the comparison test of LSTM optimized by CRAOA algorithm and other comparative optimization algorithms in solving the capacity prediction problem.

The fitness function is selected as shown in Eq. (16), the maximum number of iterations is set as 500 generations, and the number of populations is 20. The five-pointed star curve of the red line is the prediction results of CRAOA-LSTM in two groups of battery data, which has the lowest error compared with other optimization algorithms.

Secondly, CRAOA has a faster downward trend and convergence speed. And when other algorithms have fallen into local optimum at the early stage of iteration, CRASOA

still keeps a certain speed decline, which reflects the superiority of CRAOA in solving optimization problems. It can jump out of local optimum and has stronger ability in global search and local search.

Fig. 14 shows the capacity prediction curve of CRAOA-LSTM. Since too many curves will affect the reader's observation effect, only the comparison results of CRAOA-LSTM and AOA-LSTM and two kinds of deep learning networks are shown here.

The prediction errors of BAT-LSTM and other models are reflected in the data in Table IV and Table V. Among them, the yellow triangular curve corresponds to the predicted curve of CRAOA-LSTM, and the blue solid line is the true attenuation value in the test set. The proposed prediction model is more consistent with the real value as a whole, especially in the early stage of recurrent FNN, LSTM and other neural networks have higher deviation

from the real value, which is not as good as CRAOA-LSTM.

To show the prediction effect more directly, the prediction error curve is introduced, as shown in Figure 15, and the advantages of the CRAOA-LSTM model can be clearly seen. Table IV and Table V show the prediction errors of CRAOA-LSTM and other models, in the two sets of battery data, this model performs the best among the three errors, which are the minimum values, which directly reflects the advantages of the model. Among them, in Battery #5, CRAOA-LSTM is 2.5% ~ 15.2% lower than other prediction models in RMSE index, 0.67% ~ 13.9% lower in MAE, and 0.32% ~ 8.92% lower in MAPE. In Battery #6, compared with other prediction models, the RMSE index is reduced by 0.87% ~ 21.2%, MAE is reduced by 0.15% ~ 12.9%, and MAPE is reduced by 0.14% ~ 8.82%.

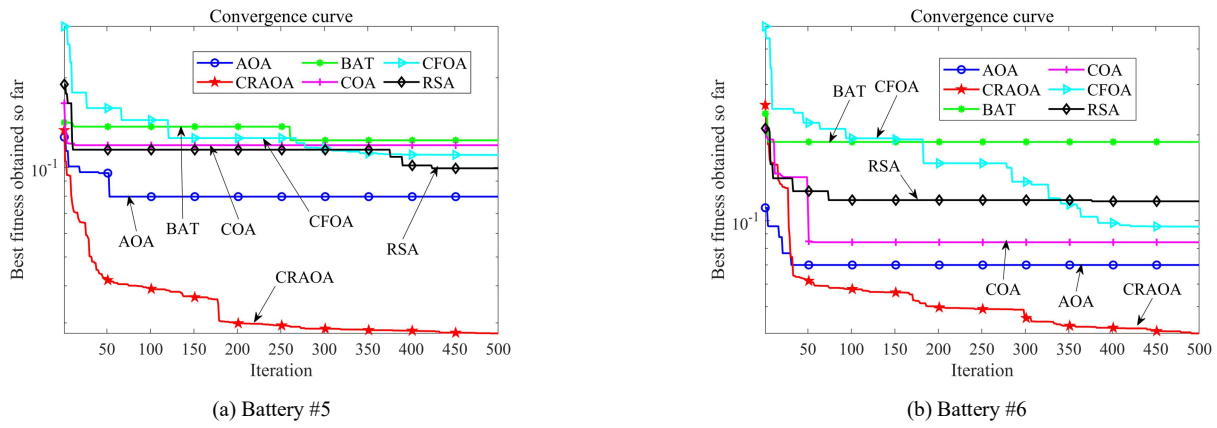


Fig. 13 CRAOA-LSTM and other optimization algorithms optimize the error decay curve of LSTM at Battery #5 and #6.

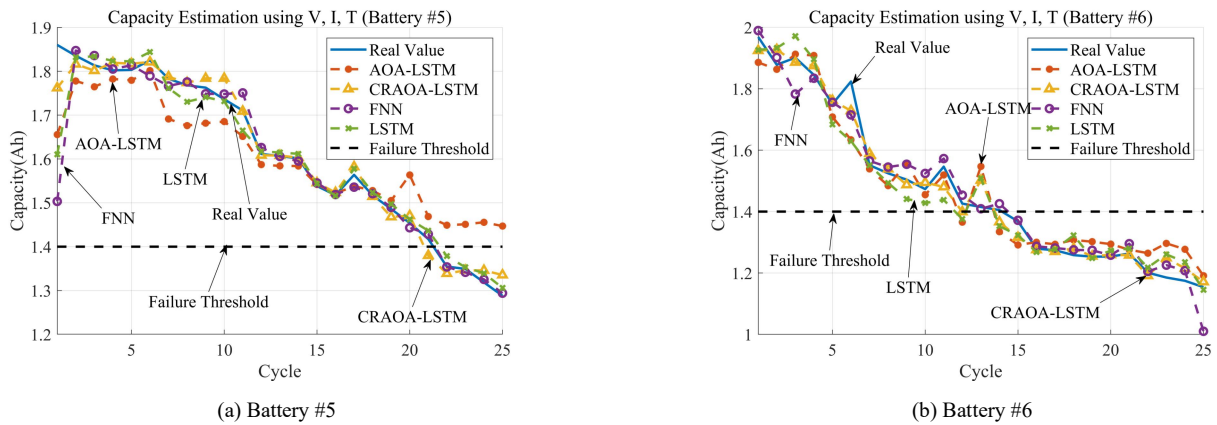


Fig. 14 Capacity attenuation curves of Battery #5 and #6 in NASA battery pack.

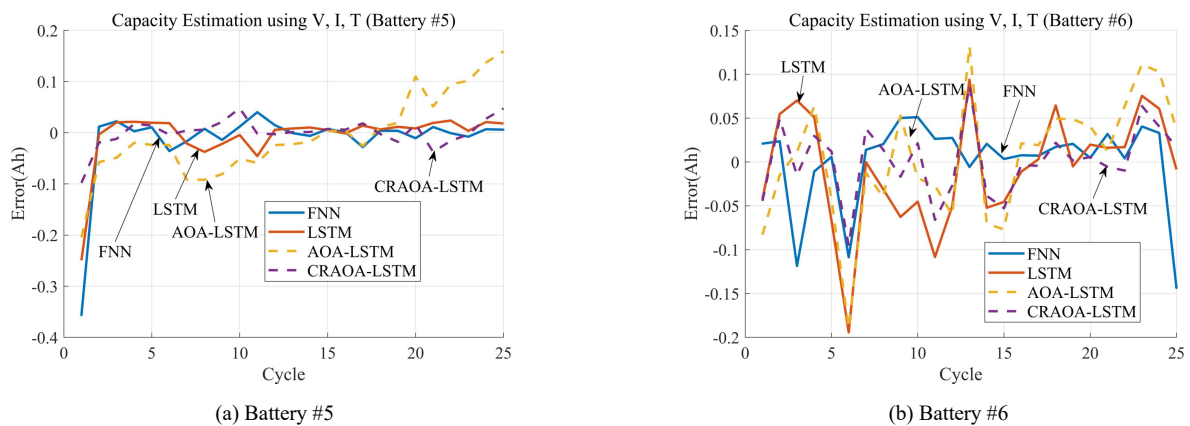


Fig. 15 Capacity prediction error curve of Battery #5 and #6 in NASA battery pack.

TABLE IV. QUANTITATIVE METRICS FOR BATTERY #5 CAPACITY PREDICTION MODEL

Model	RMSE	MAE	MAPE (%)
FNN	0.0730	0.0256	1.45
LSTM	0.0531	0.0248	1.45
AOA-LSTM	0.0798	0.0613	3.94
CRAOA-LSTM	0.0277	0.0181	1.13
BAT-LSTM	0.1235	0.0978	6.19
COA-LSTM	0.1797	0.1573	10.05
CFOA-LSTM	0.1101	0.0845	5.46
RSA-LSTM	0.0993	0.0840	5.49

TABLE V. QUANTITATIVE METRICS FOR BATTERY #6 CAPACITY PREDICTION MODEL

Model	RMSE	MAE	MAPE (%)
FNN	0.0490	0.0327	2.2515
LSTM	0.0645	0.0500	3.3053
AOA-LSTM	0.0700	0.0561	3.9139
CRAOA-LSTM	0.0403	0.0312	2.1009
BAT-LSTM	0.1909	0.1598	10.9268
COA-LSTM	0.0841	0.0546	3.6481
CFOA-LSTM	0.0954	0.0788	5.4763
RSA-LSTM	0.1170	0.0961	6.5610

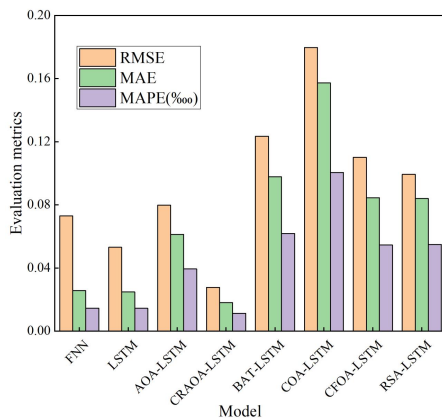


Fig. 16 Bar graph of CRAOA-LSTM versus other models in B0005 evaluation metrics.

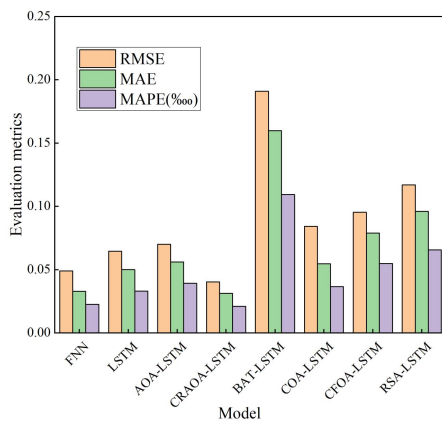


Fig. 17 Bar graph of CRAOA-LSTM versus other models in B0006 evaluation metrics.

To enhance the visualization effect of each model in B0005 and B0006 pack prediction, RMSE and other evaluation indicators are plotted into bar charts, as shown

in Figs. 16 and 17. Where the abscissa is each model and the ordinate is the evaluation index.

VI. CONCLUSION

LSTM is selected as the basic solution framework, and then a metaheuristic AOA algorithm is proposed to optimize the weight and bias of LSTM. At the same time, two-dimensional circle chaotic map and random adjustment factor AOA are proposed to increase the optimization ability of AOA. The simulation results show that the CRAOA-LSTM model has the best prediction effect and the lowest error value in NASA battery data. The specific conclusions are as follows:

(1) Aiming at the problem that the original AOA is difficult to find the optimal solution, the two-dimensional Circle chaotic map and random adjustment factor are introduced. The innovation of the two strategies increases the optimization ability of AOA, makes the algorithm achieve an effective balance in the global and local search stage, and avoids the situation that the optimization stops in the early iteration and the search cannot be refined at the end of the iteration.

(2) CRAOA is more competitive than the original AOA and recent excellent optimization algorithms in solving function optimization problems, and it is more likely to find the optimal solution of the function. CRAOA also plays a crucial role in lithium battery capacity prediction. CRAOA can find the weight and bias of the optimal LSTM, which effectively solves the problem of inaccurate prediction caused by inadequate LSTM training of traditional training methods.

In the larger context of future lithium battery capacity prediction, there is still much that can be innovated. We can consider using transfer learning, which is popular in recent years, by training a model on a large amount of historical data and transferring it to this problem. For the innovation of CRAOA, better variants of AOA and better optimization algorithms can be proposed to solve this problem.

REFERENCES

- [1] F. Degen, M. Winter, D. Bendig, et al. "Energy consumption of current and future production of lithium-ion and post lithium-ion battery cells", *J. Nature Energy*, vol. 8, no. 11, pp. 1284-1295, 2023.
- [2] J. Shi, A. Rivera, D. Wu. "Battery health management using physics-informed machine learning: Online degradation modeling and remaining useful life prediction", *J. Mechanical Systems and Signal Processing*, vol. 179, pp. 109347, 2022.
- [3] Y. Zhang, P. Gu, B. Duan, et al. "A hybrid data-driven method optimized by physical rules for online state collaborative estimation of lithium-ion batteries", *Energy*, vol. 301, pp. 131710, 2024.
- [4] M. Guo, G. H. Kim, R. E. White. "A three-dimensional multi-physics model for a Li-ion battery", *J. Journal of Power Sources*, vol. 240, pp. 80-94, 2013.
- [5] X. Li, Y. Yu, Z. Zhang, X. Dong. "Study on external characteristics of lithium-ion power battery based on ADME model", *J. Acta Physica Sinica*, vol. 71, pp. 345-353, 2022.
- [6] Z. Deng, L. Yang, H. Deng, et al. "Polynomial approximation pseudo-two-dimensional battery model for online application in embedded battery management system", *J. Energy*, vol. 142, pp. 838-850, 2018.
- [7] K. K. Sadabadi, X. Jin, G. Rizzoni. "Prediction of remaining useful life for a composite electrode lithium ion battery cell using an electrochemical model to estimate the state of health", *J. Journal of Power Sources*, vol. 481, pp. 228861, 2021.
- [8] J. Li, L. Wang, C. Lyu, et al. "New method for parameter estimation of an electrochemical-thermal coupling model for LiCoO2 battery", *J. Journal of Power Sources*, vol. 307, pp. 220-230, 2016.

- [9] M. Hossain, S. Saha, M. T. Arif, et al. "A parameter extraction method for the li-ion batteries with wide-range temperature compensation", *J. IEEE Transactions on Industry Applications*, vol. 56, no. 5, pp. 5625-5636, 2020.
- [10] A. Nikohan, J. Jaguemont, J. de Hoog, et al. "Complete cell-level lithium-ion electrical ECM model for different chemistries (NMC, LFP, LTO) and temperatures (-5 ° C to 45 ° C)-Optimized modelling techniques", *J. International Journal of Electrical Power and Energy Systems*, vol. 98, pp. 133-146, 2018.
- [11] X. Ding, D. Zhang, J. Cheng, et al. "An improved Thevenin model of lithium-ion battery with high accuracy for electric vehicles", *J. Applied Energy*, vol. 254, pp. 113615, 2019.
- [12] Y. Liu, S. Wang, Y. Xie, et al. "Research on Li-ion battery modeling and SOC estimation based on online parameter identification and improved 2RC-PNGV model", *J. Energy Storage Science and Technology*, vol. 10, no. 6, pp. 2312, 2021.
- [13] R. Xiong, Y. Sun, C. Wang, et al. "A data-driven method for extracting aging features to accurately predict the battery health", *J. Energy Storage Materials*, vol. 57, pp. 460-470, 2023.
- [14] C. Qian, B. Xu, Q. Xia, et al. "SOH prediction for Lithium-Ion batteries by using historical state and future load information with an AM-seq2seq model", *J. Applied Energy*, vol. 336, pp. 120793, 2023.
- [15] Y. Ma, C. Shan, J. Gao, et al. "A novel method for state of health estimation of lithium-ion batteries based on improved LSTM and health indicators extraction", *J. Energy*, vol. 251, pp. 123973, 2022.
- [16] L. Ren, J. Dong, X. Wang, et al. "A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life", *J. IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3478-3487, 2020.
- [17] B. Chen, Y. Liu, B. Xiao. "A novel hybrid neural network-based SOH and RUL estimation method for lithium-ion batteries", *J. Journal of Energy Storage*, vol. 98, pp. 113074, 2024.
- [18] L. Abualigah, A. Diabat, S. Mirjalili, et al. "The arithmetic optimization algorithm", *J. Computer Methods in Applied Mechanics and Engineering*, vol. 376, pp. 113609, 2021.
- [19] Y. X. Li, J. S. Wang, S. W. Zhang, et al. "Arithmetic optimization algorithm with cosine transform-based two-dimensional composite chaotic mapping", *J. Soft Computing*, vol. 29, pp. 1-41, 2025.
- [20] A. Sherstinsky. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network", *J. Physica D: Nonlinear Phenomena*, vol. 404, pp. 132306, 2020.
- [21] S. Hochreiter. "Recurrent neural net learning and vanishing gradient", *J. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, no. 2, pp. 107-116, 1998.
- [22] X. T. Wang, J. S. Wang, S. B. Zhang, et al. "Capacity prediction model for lithium-ion batteries based on bi-directional LSTM neural network optimized by adaptive convergence factor gold rush optimizer", *J. Evolutionary Intelligence*, vol. 18, no. 2, pp. 1-30, 2025.