METHODS

Fault Prediction for Capacitor of Power Converters Based on CEEMDAN and GWO-RELM

LINGHUI MENG[1,2], **QUAN SUN**[3], **(Member, IEEE), ZHENWEI ZHOU**[1,2], **AND LICHEN YANG**[3]
[1]National Key Laboratory of Science and Technology on Reliability Physics and Application of Electronic Component, Guangzhou 510370, China
[2]China Electronic Product Reliability and Environmental Testing Research Institute, Guangzhou 510370, China
[3]School of Automation, Nanjing Institute of Technology, Nanjing 211167, China

Corresponding author: Linghui Meng (menglinghui@ceprei.com)

ABSTRACT Aluminum electrolytic capacitors (AECs) get multiple superior functions such as filtering, energy storage and decoupling, which have a great effect on the performance and lifetime for power converters. Therefore, analyzing and predicting the faults of Aluminum electrolytic capacitors (AECs) is conducive to improve the safety and reliability of the power converters. In order to establish the AECs' fault prediction model and improve the accuracy, an integrated model based on complete ensemble empirical mode decomposition with adaptive noise, grey wolf optimization algorithm and regularized extreme learning machine (CEEMDAN-GWO-RELM) is proposed. The CEEMDAN is used to decompose the time series of AEC degradation process into several sequences, which can decouple the feature of local fluctuations from global degradation in the AEC time series. Then, the RELM optimized by GWO is used to predict each sequence after decomposition. RELM has the advantages of fewer hyperparameters and less operation time, and GWO with strong astringency is used for its optimization to obtain better fault prediction. Eventually, the predicted values are reconstructed to obtain the predicted values of the integrated model. The results show that, based on the aging data of AEC, the integrated model based on CEEMDAN-GWO-RELM can provide better prediction progress than traditional models, and the maximum relative error of each prediction time point is lower than 1.6%.

INDEX TERMS Electrolytic capacitor, fault prediction, CEEMDAN, GWO-RELM.

ACRONYMS

AEC	Aluminum electrolytic capacitors
CEEMDAN	Complete ensemble empirical mode decomposition
GWO	grey wolf optimization algorithm
RELM	regularized extreme learning machine
VMD	variational mode decomposition
R	Residue
IMF	Intrinsic mode function
RMSE	Root mean square error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error

RE	Relative Error
TP	Time point

The associate editor coordinating the review of this manuscript and approving it for publication was Guillermo Valencia-Palomo.

I. INTRODUCTION

As the interface between source and load, power converter is widely used in rail transit, electric vehicle, aerospace and other fields [1]. The normal operation of the whole system is restricted by its security and reliability [2]. Therefore, in order to study the cause of its failure to improve reliability, many scholars have conducted researches on it [3], [4]. The circuit failure of power converter is mainly caused by the failure and degradation of its internal components. Moreover, as an important part of the power electronic converter, it is also one of the most vulnerable components. In Fig. 1, based on more than 200 products from 80 companies, 30% of capacitors failed [5]. Thus, fault prediction of AEC is conducive to
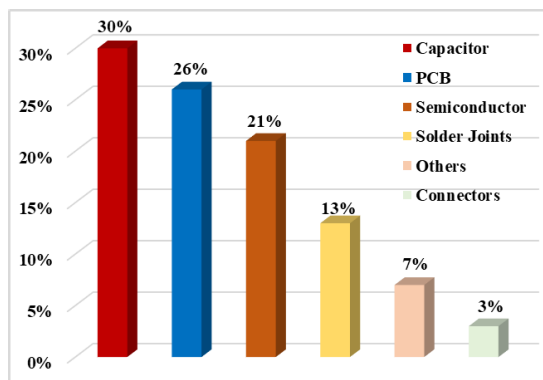
FIGURE 1. Percentage of power electronic systems failures [5].

timely replacement or maintenance of the failed capacitors in the circuit, to ensure the normal operation and improving the reliability of the circuits.

The mainstream methods used in fault prediction can be roughly divided into filter based and data-driven methods. However, most of filtering based methods need to realize the establishment of complex empirical models [7], [8]. However, in the degradation process of AEC, it is easy to be affected by the external environment, including temperature and voltage, as well as the internal capacitance regeneration, which increases the complexity of empirical model building.

Fortunately, with the rapid development of artificial intelligence in recent years, data driven methods have been widely used in the establishment of fault prediction models in various fields. A conditional deep neural network (DNN) with a dropout technique was proposed by Hao et al. for fault prediction of AEC, which greatly reduce the testing time of AEC in aging test [9]. Vanilla Long Short Time Memory (LSTM) was proposed by Jeong et al. to establish the fault prediction of AEC [10]. Its root mean square error and mean absolute error are decreased by 8.6% and 1.7%, respectively. Based on NASA data set, Mesquita applied different artificial neural networks to predict the remaining life time (RUL) [11], and Delanyo proposed a prediction method based on Bi-LSTM [12]. Broad learning system and LSTM are combined into a fusion network. It realizes the fault prediction and RUL prediction of lithium-ion batteries with a small proportion of training data based on two public data sets [13]. A hybrid model is proposed by Ma et al. based on k-fold cross validation, metaheuristic support vector regression (SVR) optimized by different meta-heuristics and the nonparametric fried-man test is proposed for the landslide displacement prediction [14]. Moreover, a comprehensive comparison of twenty meta-heuristics is beneficial for the researches of hyperparameters tuning in other fields [15].

A defect of the models based on data driven is that the local fluctuations and regeneration phenomenon of times series are not considered [31]. Furthermore, AEC generally presents typical nonlinear characteristics in the actual degradation process. In order to improve fault prediction and establish a more accurate prediction model for nonlinear time series. Feature based methods are used in time series processing. These methods can effectively mine the internal information of the time series, which is conducive to the establishment of the prediction model, so as to improve the prediction accuracy. Furthermore, in the process of establishment of model, although it will increase the complexity of the model and the calculation time, optimization algorithms are often used in adaptive optimization of model parameters to obtain better model effects.

Multiobjective grasshopper optimization algorithm is used by Feng to determine the gear transmission dynamic model [16]. A hybrid model is proposed by Wang et al. based on ARIMA-BO-Bi-LSTM, which makes full use of Autoregressive Integrated Moving Average model (ARIMA) to process linear features in time series and BILSTM's ability to predict the linear part [17]. the time series of lithium-ion batteries through ensemble empirical mode decomposition (EEMD), and then a RUL prediction model is built based on GWO-SVR is decomposed by Zhang et al. [18]. Compared with the GWO-SVR, the average relative error is reduced by 3.125%. CEEMDAN is used by Shi et al. to divide the lithium battery capacity into main degradation trends and several local degradation trends, and then LSTM is used to predict the RUL of the decomposed data respectively. Finally, the predicted data was effectively integrated to improve prediction accuracy [19].

Considering the limitations of the variational mode decomposition (VMD), a novel fault information-guided VMD is proposed to enhance the sensitivity of bearing fault signature. The mode number and bandwidth control parameter can be optimally determined and the diagnosis of bearing is realized based on the novel method. [20]. An integrated model based on VMD, particle filter and Gaussian process regression is proposed to predict the RUL of lithium-ion battery pack, which improves the accuracy of prediction and reduced errors [21]. However, before VMD, its modal number needs to be determined. The unreasonable mode number may lead to incomplete decomposition or large reconstruction error. Relative root mean square is used by Wu to divide the stage of bearing degradation, and Pearson correlation coefficient combined with Entropy Weight Method is used to select sensitive features as the input of RELM, thus realizing the RUL prediction of bearings. By injecting regularization factor, RELM inherits the advantages of ELM such as simple parameter selection and fast operation speed. Moreover, it further prevents the over fitting phenomenon in the operation process of ELM, obtaining better performance. On the other hand, compared with the algorithm based on neural network, its operation speed is faster and its stability is higher [22].

To sum up, considering the advantages and disadvantages of each algorithm, a fault prediction model based on feature decomposition and machine learning is used in this paper for AEC. The major innovations and contributions including:

(1) An AEC fault prediction model based on the CEEMDAN-GWO-RELM is proposed. The superiority and feasibility of the integrated model are verified by the data set obtained from the accelerated aging experiment of AEC.

(2) The original time series is decomposed by CEEMDAN, which overcome the fluctuation of AEC in the degradation process and accurately capture the local fluctuations in the degradation process.

The architecture of this paper is organized as follows. Failure mechanism of AEC is shown in section II. Algorithm principle used for fault prediction of AEC is shown in section III. Prediction results are evaluated and analyzed in section IV. The conclusion is given in section V.

II. FAILURE MECHANISM OF AEC

Accelerated degradation test of AEC is conducted by over-electric stress and over-heat stress in the NASA Ames Research Center [23]. In [3] and [24], Chen and Bhargava analyzed and summarized the failure mechanism of capacitors. Combined with the above studies, the failure of AEC is related to many factors. Environmental factors include temperature, humidity, air pressure, vibration, etc. Electrical factors include voltage, ripple current, charge and discharge times, etc. Physical factors include incomplete sealing, electrolyte leakage and evaporation, etc. Chemical factors include invasion of halogen ions and deterioration of sealing materials.

The failure of AEC can be divided into two types: structural failure (hard failure) and parametric failure (soft failure) [4]. Structural fault is mainly caused by short circuit or break in AEC equivalent circuit. The main causes of structural failure are short circuit between electrodes, damage of oxide film insulation, disconnection of terminal or poor contact. The degradation form of parametric failure is usually the change of circuit function index caused by capacitor parameter drift. It will not have the same intuitive impact on the system as a structural failure. However, with the pro-longed use, the impact of the environment and the usage of the system, parametric failure becomes more and more serious. Meanwhile, the reliability and security of the system decrease and related performance indicators change significantly. If not handled in time, it may further evolve into a more serious structural failure, resulting in irreversible consequences.

From the physical analysis, evaporation of electrolyte and degradation of dielectric are the main reasons for the deterioration of AEC aging performance. Evaporation of electrolyte is mainly due to the high temperature of external working conditions and the heating of internal equivalent series resistance. The operating temperature of AECs is generally lower than 85 °C. But in the actual working environment, the temperature will may be higher than 85 °C, accelerating the degradation of AEC. In normal operating circuits, when ripple current flows through an electrolytic capacitor, loss and heat are generated at the equivalent series resistance of the capacitor. The increase of external temperature and core

temperature of AEC are two main reasons for the accelerated evaporation of electrolyte. Evaporation of the electrolyte increases the equivalent series resistance of the capacitor and reduces the contact area between the electrolyte and the oxide layer leading to the deteriorating of AEC. The result is a decrease in capacitance and an increase in equivalent series resistance [25].

For electronic systems with capacitors, once one of the capacitors has degraded beyond its normal range of use. Other components are subjected to increasing electrical stress, accelerating their degradation. Finally, the degradation of the components in the power electronic system will inevitably lead to the degradation of the whole system, reducing the service life and safety reliability. On the other hand, the loss factor is the real impedance (equivalent series resistance) of the capacitor, which is also an important index to evaluate the quality of the AECs.

In brief, the changing trend of capacitance value and equivalent series resistance value are as well as the cause of AEC degradation.

III. AEC FAULT PREDICTION BASED ON THE CEEMDAN-GWO-RELM

A. CEEMDAN DECOMPOSITION

CEEMDAN is improved from EMD and EEMD [18], [19], [26]. Because these signal decomposition methods do not require prior analysis and research, the time series can be adaptively decomposed into multiple inherent mode functions (IMFs) and a residue (R) with different frequencies and scales. These methods have been widely used in the prediction models of nonlinear and non-stationary signals and time series. However, modal aliasing is actually easy to occur in the process of EMD [19]. EEMD cannot eliminate the added white noise, leading to incomplete decomposition and large reconstruction error. To solve these problems, Torres et al. improved the decomposition process and proposed CEEMDAN method [27]. It introduces additional white noise signal-to-noise ratio (SNR) in each decomposition process to control the noise level. The decomposition is more complete and the reconstruction error is smaller. The decomposition steps of AEC degradation time series are as follows:

(1) For the original data $f(t)$ $(t = 0, 1, \ldots n)$ of the AEC degradation time series, add the white noise capacity series, and the formula is as follows:

$$IMF2 = \frac{1}{n} \sum_{i=1}^{n} EMD_1(R_1(t) + \beta_1 EMD_1(w_i(t))) \quad (1)$$

where, t represents the number of cycles of AEC, and i represents that white noise is added for the i_{th} time, β_k represents the k_{th} SNR, $w_i(t)$ $(t = 0,1, \ldots n)$ represents the white noise subject to standard normal distribution added for the i_{th} time and data $f_i(t)$ is obtained.

(2) $f_i(t)$ is decomposed n times by EEMD, and then the first modal component $IMF_1(t)$ is obtained by

averaging, $R_1(t)$ is obtained by (2):

$$R_1(t) = f(t) - IMF_1(t) \quad (2)$$

(3) For time series $R_1(t) + \beta_1 EMD_1(n_i(t))$ was decomposed n times repeatedly and averaged. Calculate $IMF_2(t)$ and $R_2(t)$ as follows:

$$IMF2 = \frac{1}{n}\sum_{i=1}^{n} EMD_1(R_1(t) + \beta_1 EMD_1(w_i(t))) \quad (3)$$

$$R_2(t) = R_1(t) - IMF_2(t) \quad (4)$$

where, EMD_k represents the k_{th} component after decomposition by EMD.

(4) Repeat the above steps until the number of extreme points of the margin sequence is less than or equal to two. That is, the margin signal cannot be decomposed again, then terminate the algorithm. The final time series after decomposition can be expressed as:

$$f(t) = R(t) + \sum_{k=1}^{K} IMF_i(t) \quad (5)$$

B. PRINCIPLE OF RELM

ELM is proposed on the basis of single hidden layer neural network [25], [28]. ELM has the advantages of less parameters, fast time for operation and small generalization error. However, only the principle of minimizing empirical risk is considered in the calculation process. Thus, over fitting may occur with the number of hidden layers increases, resulting in poor generalization ability [29]. In addition, the ELM directly calculates the least square solution of the weight value of the output layer. In the process of modeling the time series. it is difficult to adjust the information contained in the time window with more time series, resulting in insufficient controllability. Empirical risk ($\|\varepsilon\|^2$) and structural risk ($\|\beta\|^2$) can be balanced to obtain better generalization ability by adjusting the regularization parameters (λ). The purpose of RELM is to find the minimum value of the objective function (E) with the total risk as the objective function:

$$\min E = \min(\frac{\|\beta\|^2 + \lambda \|\varepsilon\|^2}{2}) \quad (6)$$

$$s.t.\ h(x_i)\beta = t_i - \varepsilon_i \quad (7)$$

where, the function $h(x)$ represents the activation function of hidden layer neurons. $\varepsilon_i\ (i = 1, 2, 3, \dots N)$ is the sum of training errors, and N is the number of samples. The Lagrange equation thus constructed is:

$$L(\alpha, \varepsilon, \beta) = \frac{\|\beta\|^2 + \lambda \|\varepsilon\|^2}{2} - \alpha(H\beta - t_i - \varepsilon) \quad (8)$$

where, $\alpha_i \in R\ (i = 1, 2, 3, \dots N)$ is the Lagrange multiplier. H is the output matrix of hidden layer. Its partial derivative is

shown in (9):

$$\begin{cases} \frac{\partial L}{\partial \beta} = 0 \rightarrow \beta^T = \alpha H \\ \frac{\partial L}{\partial \varepsilon} = 0 \rightarrow \lambda e^T + \alpha = 0 \\ \frac{\partial L}{\partial \alpha} = 0 \rightarrow H\beta = t_i + \varepsilon \end{cases} \quad (9)$$

From (9), the out weight matrix can be obtained:

$$(\frac{I}{\lambda} + H^T H)\beta = H^T H \quad (10)$$

In (10), I is identity matrix. The output matrix is:

$$Y = H\beta = H(\frac{I}{\lambda} + H^T H)\beta^{-1} H^T H \quad (11)$$

C. GREY WOLF OPTIMIZATION ALGORITHM

GWO has the advantages of simple principle, strong search ability and few parameters that is widely used in model optimization based on machine learning and deep learning [18, 0]. In GWO, according to the objective function in algorithms to be optimized, the four level definitions that govern the gray wolf are: the top level is α wolf with leadership, it is also the optimal solution of individual fitness. The second level is β wolf, which is the subordination of α wolf. The third layer is δ wolf. The fourth layer is ω wolves, are the basis of all wolves. Through the continuous change of the position of the wolves, the distance between the wolves and the prey is gradually narrowed, and finally the hunting is realized. The specific steps of GWO is shown as follows:

Step 1: Calculate the distance (D) between gray wolves and prey. Then with D decreases, the gray wolves are updated.

$$D = |C \cdot X_p(t) - X(t)| \quad (12)$$

$$X(t + 1) = X_p(t) - A \cdot D \quad (13)$$

$$A = 2a \cdot r_1 - a \quad (14)$$

$$C = 2 \cdot r_2 \quad (15)$$

where, $X_p(t)$ and $X(t)$ represent the prey and individual gray wolf. A and C are coefficient vectors, t is the number of iterations, r_1 and r_2 are random vectors between [0,1].

Step 2: When the individual wolf recognizes the position of the prey, the prey is hunted down under the leadership of α, β and δ:

$$\begin{aligned} D_\alpha &= |C_\alpha \cdot X_\alpha(t) - X(t)| \\ D_\beta &= |C_\beta \cdot X_\beta(t) - X(t)| \\ D_\delta &= |C_\delta \cdot X_\delta(t) - X(t)| \\ X_1 &= X_\alpha(t) - A_\alpha \cdot D_\alpha \\ X_2 &= X_\beta(t) - A_\beta \cdot D_\beta \\ X_3 &= X_\delta(t) - A_\delta \cdot D_\delta \end{aligned} \quad (16)$$

The positions of ω wolves are updated on the basis of α, β and δ.

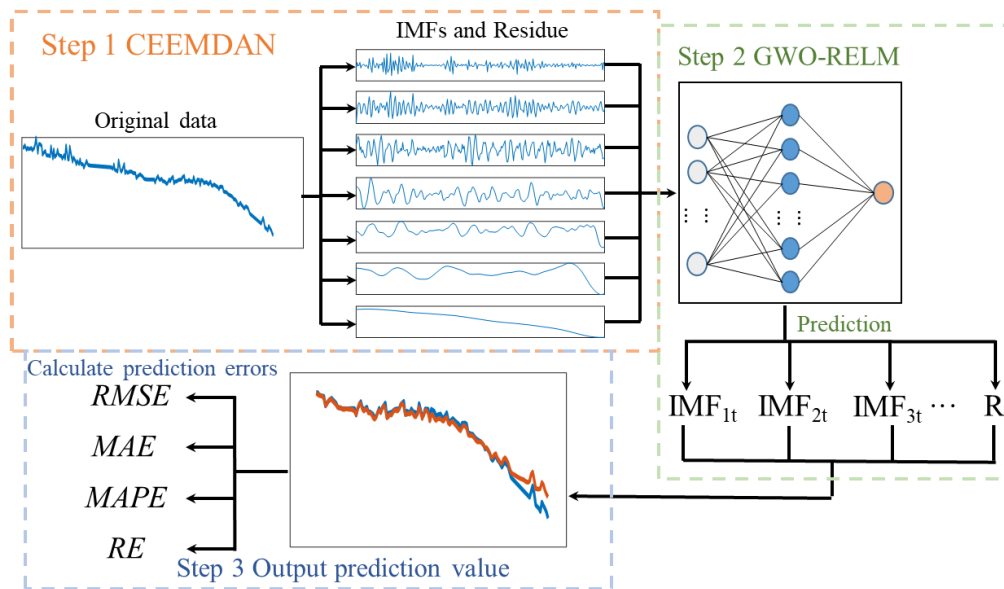$$X(t + 1) = \frac{(X_1 + X_2 + X_3)}{3} \quad (17)$$

FIGURE 2. Fault prediction of AEC based on CEEMDAN-GWO-RELM.

Step 3: Finally, after the convergence of GWO, the wolves attack the prey and find the optimal solution of objective function.

D. DATA SOURCE

In order to analyze the characteristic of AEC, aging experiment has been conducted under a high temperature on AEC and the specification of AEC is 330 μF/35V. AEC degraded in the chamber with the temperature is at 150°C. Each AEC has two wires extending to the outside of the drier that connect its positive and negative poles. So that the AECs in the process of degradation can be measured by LCR meter every 6 hours. At the initial stage of the experiment, the open circuit and short circuit of LCR meter are corrected to reduce the error of the measurement result caused by the wires. Finally, considering the factors of time and cost, the experiment lasted for 1608 hours and 268 data points were collected in the AEC degradation experiment.

E. FAULT PREDICTION BASED ON CEEMDAN-GWO-RELM

On the basis of the data set, the fault prediction model of AEC based on CEEMDAN-GWO-RELM is shown in Fig. 2. The process of the prediction model is as follows:

Step 1: AEC time series is decomposed into IMFs and *R* by CEEMDAN.

Step 2.1: Data set is divided into training set and testing set according to three different proportions by sliding time window. The size of window is 9 that is the last data is predicted by the past 8 data. RELM is used to make single-step and short-term prediction for each IMF and R, respectively. The short-term prediction shows the multi-steps prediction of AEC. It does fault prediction from earlier starting time point than single-step prediction.

Step 2.2: The parameters of GWO-RELM are set: the hidden layer and C of RELM are 50 and 0.01. The GWO is used to optimize the input layer weight and hidden layer biases of RELM. The objective function of GWO is RMSE. The population size and maximum iterations of GWO used in this paper is 5 and 500. After the optimization of GWO, the minimum RMSE is obtained and the parameters are used for testing set.

Step 3: Reconstruct the prediction data to realize the fault prediction of AEC and calculate the prediction error.

Furthermore, on the basis of the above single-step prediction, the multi-step prediction of AEC is realized by recursion. The operation mode is shown in Fig. 3. Firstly, the size of the time window is to 13 that is 8 historical data are used to predict the next 5 data. Secondly, supposed that, the obtained nonlinear model of the training set established by GWO-RELM is $f(x_t)$, where the t_{th} historical value ranked x_t in the time window. In the testing set, $[y_1, y_2, \ldots y_5] = f(x_t)$ is obtained. The, the $[y_1, y_2, \ldots y_5]$ will be re-divided into the training set to continue to predict the next 5 data. Repeat the above steps, multi-steps prediction is finally realized by recursion.

IV. RESULT AND ANALYSIS

In this study, the superior advantages of CEEMDAN-GWO-RELM integrated model are verified by the following three groups of contrast experiments:

(1) The time series sequence is decomposed by VMD, EEMD and CEEMDAN into IMFs and R. Then the decomposed sequences are reconstructed so that the effectiveness of feature decomposition is analyzed by comparing with the real value.
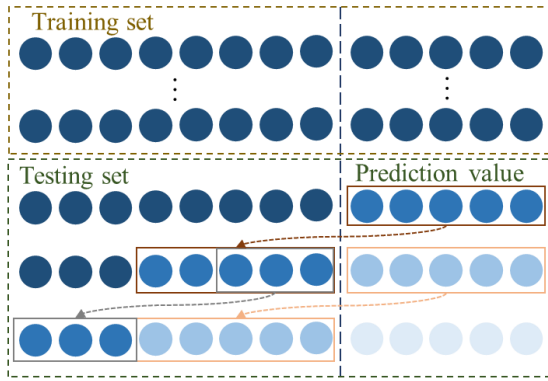
FIGURE 3. The operation mode of multi-steps prediction.

(2) The proportion between design training set and prediction set is 80%, 60% and 40%. The advantages of the model are confirmed by comparing with some commonly used data-driven methods.

(3) A contrast experiment of prediction models is designed with different prediction steps. The multi-steps prediction of AEC is realized by recursive five-steps method and the number of steps is 5, 10, 15, 20 and 25.

A. MODEL EVALUATION

For evaluating the accuracy of the model prediction, the following four error evaluation indicators are used in this paper.

(1) Root Mean Square Error

$$RMSE = \sqrt{\sum \frac{(P - Y)^2}{n}} \quad (18)$$

(2) Mean Absolute Error

$$MAE = \sum \frac{|P - Y|}{n} \quad (19)$$

(3) Mean Absolute Percentage Error

$$MAPE = \frac{\sum \left| \frac{P-Y}{Y} \right| \times 100\%}{n} \quad (20)$$

(4) Relative Error

$$RE = \frac{|Y - R|}{Y} \times 100\% \quad (21)$$

where, P represents the prediction values of different models and Y represents the real values. And the smaller $RMSE$, MAE, $MAPE$, and RE of prediction results are, the better predictions are.

The environment where the program runs is: Win10 x64, Intel(R) Core(TM) i5-10400F CPU @ 2.90GHz, GeForce GTX 1050Ti, MATLAB2021a.

B. ANALYSIS ON DECOMPOSITION RESULTS

CEEMDAN is used to decompose the degradation data set of AEC time series sequence into several IMFs and a R. The
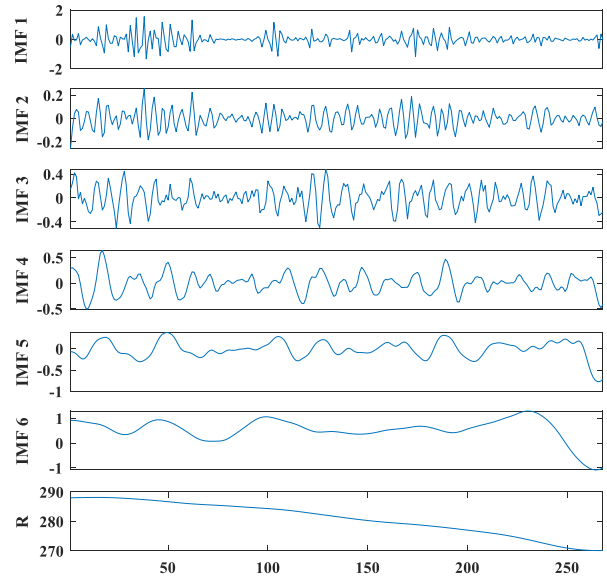

FIGURE 4. Decomposition results by CEEMDAN.

decomposition results are shown in Fig. 4. It can be seen that R which after decomposition can be smoother and monotonic, which can better reflect the overall trend of AEC degradation. It can greatly improve the fault prediction accuracy of AEC.

The comparison between the decomposition and reconstruction of VMD, EEMD and CEEMDAN algorithms and the real value is shown in Fig. 5(a). The black one is realized as real value, the blue, yellow and orange ones are decomposition results based on EEMD, VMD and CEEMDAN, respectively. It can be seen intuitively that EEMD, compared with VMD and CEEMDAN is more deviated from the real value and has largest error.

Nevertheless, the data set reconstructed after VMD and CEEMDAN decomposition is close to the real value, which is difficult to see directly from the Fig. 5(a). For further comparing with the reconstruction errors of the two feature decomposition algorithms. REs at each TP are calculated after reconstruction compared with the real value.

The RE between the reconstructed sequence and the real value of CEEMDAN and VMD at each TP is shown in Fig. 5(b). The blue is based on VMD and the orange one is based on CEEMDAN. The minimum RE of these two algorithm are both lower than 0.006. In addition, it is obvious that CEEMDAN algorithm has the smallest error.

It can be seen that CEEMDAN algorithm is superior to EEMD and VMD for decomposition. It is more conducive to the establishment of AEC fault prediction model.
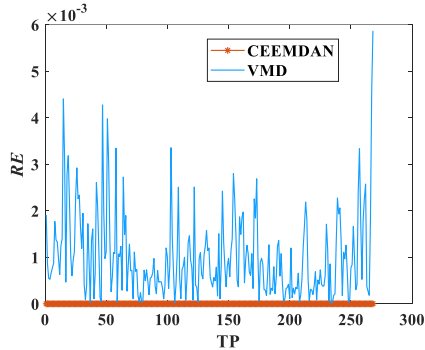
C. ANALYSIS OF FAULT PREDICTION RESULTS

1) RESULTS OF SINGLE-STEP PREDICTION

In this study, data set is divided by three ways: data set A: 80% data for training sets; data set B: 60% data for training sets; data set C, 40% data for training sets, as shown in TABLE 1.
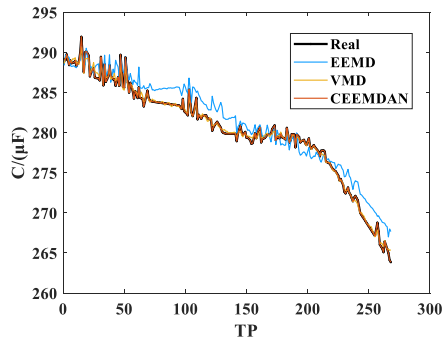
(a) Decomposition results by different algorithms

(b) REs of each TP by CEEMDAN and VMD

**FIGURE 5.** Decomposition results of degradation data set.

TABLE 1. Partition scale and starting point of training sets.

|	Data set A	Data set B	Data set C
Proportion of training sets	40%	60%	80%
Starting point (TP)	116	170	223

The purpose is to test whether the prediction effect of each model will change significantly due to the size of the training set. The starting point of the data set for A is TP_{223}, the starting point of the prediction set for data set B is TP_{170}, and the starting point of the prediction set for data set C is TP_{116}.

The prediction results of the three groups of experiments are shown in Fig. 6(a), where the blue one represents the real value, and the results of the three groups of experiments A, B and C are orange, yellow and purple, respectively. It can be seen from Fig. 6(a) that the deviation of the prediction results of data set A is greater, because the starting point of the prediction set of data set A is earlier, and the prediction error becomes larger and larger as time goes by. The RE of each TP in the three groups of experiments is shown in Fig. 6(b). The experimental RE_{Max} of data set C is 1.247%, the experimental RE_{Max} of data set B is 0.439%, and the experimental RE_{Max} of data set is 0.284%.

Some traditional prediction methods are used as contrast experiments. And in TABLE 2, the prediction results are described in detail in the form of data. From it, LSTM and GRU are often used to establish prediction models
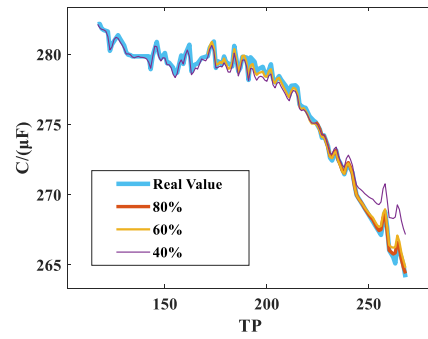
(a)The prediction results of data set A, B and C
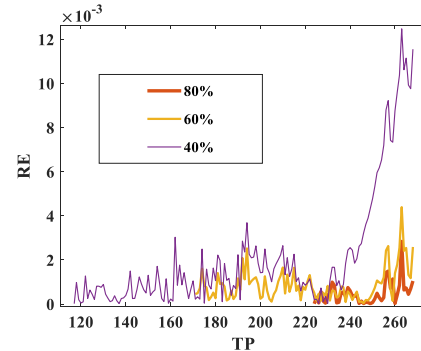
(b) REs of each TP

FIGURE 6. Prediction results of single-step.

TABLE 2. Prediction error of each prediction algorithm in different data sets.

		RMSE	MAE	MAPE
BP	40%	3.0482	2.5534	0.95%
	60%	4.0399	2.8661	1.06%
	80%	5.1207	4.2604	1.55%
LSTM	40%	1.0079	0.72975	0.27%
	60%	6.8808	4.9601	1.84%
	80%	4.6577	3.0743	1.14%
GRU	40%	0.83154	0.66002	0.25%
	60%	4.4113	3.2159	1.19%
	80%	3.159	2.2295	0.82%
GWO-RELM	40%	1.4618	1.1538	0.43%
	60%	2.9993	2.1902	0.81%
	80%	3.1251	2.0573	0.76%
CEEMDAN-GWO-RELM	40%	0.23006	0.1485	0.06%
	60%	0.33486	0.24665	0.09%
	80%	0.9622	0.58812	0.22%

for time series. However, from the prediction results, when the training set is large, its prediction model has a high

TABLE 3. Operation time of fault prediction model.

|		Time (s)
	40%	0.995
BP	60%	1.135
	80%	1.296
	40%	15
GRU	60%	16
	80%	19
	40%	16
LSTM	60%	17
	80%	21
	40%	7.028
CEEMDAN-GWO-RELM	60%	6.573
	80%	6.755

prediction accuracy, and its internal memory unit plays an important role. Though the prediction error is smaller compared with BP and GWO-RELM. When the training set is small, the prediction error of the prediction model in data set B and C is large. The prediction accuracy of the integrated model based on CEEMDAN-GWO-RELM is greater than that of GWO-RELM. It can be seen that the prediction model can obtain greater prediction accuracy after feature decomposition.

The operation time of fault prediction model based on these algorithms is shown in TABLE 3. From it, BP takes least operation time than GRU and LSTM. Because BP is a shallow neural network, its structure is simpler than deep neural networks such as LSTM and GRU, and its operation time is correspondingly shorter. For CEEMDAN-GWO-RELM, it is feature decomposition and parameters optimization that increases the complexity of the model. Though the operation time of CEEMDAN-GWO-RELM is longer than BP, its prediction accuracy is better than BP.

In order to further verify the convergence effect of GWO on the model. The convergence curves of IMF1 and R based on CEEMDAN-GWO-RELM is shown in Fig. 7(a) and 7(b). Each curve is averaged after 20 cycles for each case to avoid accidents. From Fig. 7, with the increase of training set, the fitness value is lower. Moreover, it takes less than 250 iterations in IMF1 and 150 iterations in R to realize the convergence.

2) RESULTS OF ITERATIVE MULTI-STEPS PREDICTION

On the basis of single-step prediction, the prediction model is used for multi-steps. In practical applications, the problems that may happen in the usage of AEC earlier can be discovered with the implementation of multi-steps. The AEC and its circuit shall be repaired and maintained earlier to ensure the safety and reliability of the system. In this study, the
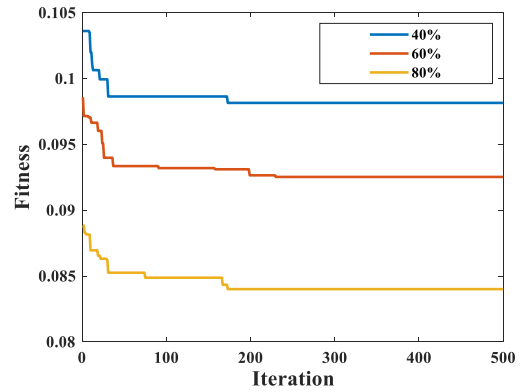
(a)Convergence curve of IMF1
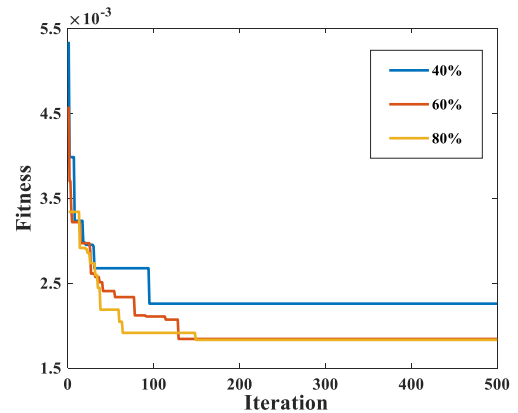
(b)Convergence curve of R

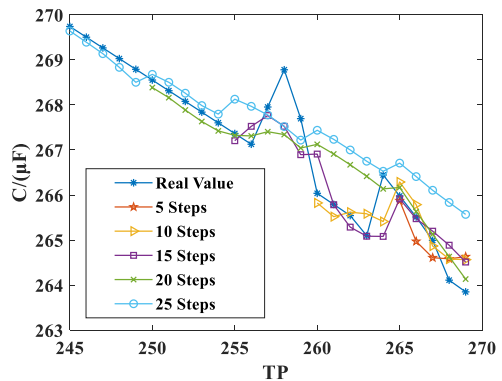**FIGURE 7.** Convergence curves of CEEMDAN-GWO-RELM.

FIGURE 8. Prediction results of different multi-steps.

multi-steps prediction is set ahead of 5, 10, 15, 20 and 25, respectively.

In Fig. 8, the prediction results are shown as lines, where the gray represents the real value and other prediction results are colored differently. short-term prediction of AEC time series can be effectively and accurately conducted based on CEEMDAN-GWO-RELM.
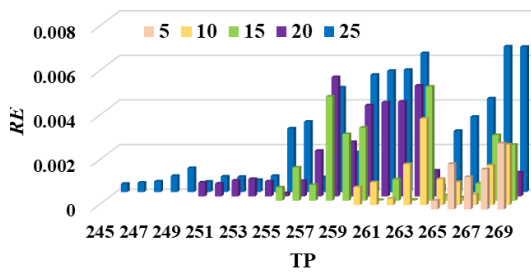
FIGURE 9. Prediction REs of different multi-steps.

TABLE 4. Prediction errors of different multi-steps.

Steps	RMSE	MAE	MAPE
5	0.50742	0.45733	0.17%
10	0.48438	0.39713	0.15%
15	0.6433	0.47042	0.18%
20	0.66901	0.49854	0.19%
25	0.91304	0.69557	0.26%

*RE*s at each TP is shown in Fig. 9. It can be seen from the Fig. 9 that when the prediction is made 25 steps ahead that is 150 hours ahead of schedule, the RE_{max} at each TP is not exceed 0.7%. And the $RMSE_{max}$ is 0.91304, MAE_{max} is 0.69557, and $MAPE_{max}$ is 0.26%, which the results are as shown in TABLE 4. This model not only has high prediction accuracy in single step prediction. Moreover, when the advance steps are large, the *RE* and prediction error will also increase. This model not only in single step prediction, it also has high prediction accuracy in the multi-steps prediction. In addition, short-term prediction through recursive multi-steps prediction can reduce the number of recursion, reduce the cumulative error, and further improve the prediction accuracy of short-term prediction.

V. CONCLUSION

The establishment of fault prediction model for AEC is conducive to maintaining the normal operation of AEC and the circuit where it set, resulting in improving the reliability of system operation. In this paper, an integrated model for fault prediction of AEC degradation data from accelerated aging tests under thermal stress based on CEEMDAN-GWO-RELM is designed. The main conclusions from the experiments are: (1) Compared with EEMD and VMD, CEEMDAN applied to AEC sequence feature decomposition has better decomposition effect and smaller reconstruction error after decomposition. It effectively overcomes the nonlinear problem of time series caused by local regeneration phenomenon during AEC degradation, decreasing the difficulty of model prediction and improving the prediction accuracy of GWO-RELM. (2) After the comparison and analysis with other neural networks commonly used, the proposed integrated prediction model has higher prediction accuracy, and the RE_{max} in the single-step prediction of each TP with 40%

data set as the training set is not higher than 1.6%. In the short-term prediction of 25-steps ahead, its *RMSE* is 0.91304 and the *RE*s of each TP point is less than 0.653%, which has a high prediction accuracy.

However, in the operating circuit, AECs are also affected by electrical stress. Moreover, the electrical stress has a more serious influence on the regeneration of AEC. This further increases the difficulty of fault prediction for AECs and the circuit. Therefore, thermal and electrical stress aging experiments will be further conducted on AECs and circuit to study the failure mechanism and fault prediction methods in practical engineering applications.

REFERENCES

[1] F. H. Gandoman, A. Ahmadi, P. V. D. Bossche, J. Van Mierlo, N. Omar, A. E. Nezhad, H. Mavalizadeh, and C. Mayet, "Status and future perspectives of reliability assessment for electric vehicles," *Rel. Eng. Syst. Saf.*, vol. 183, pp. 1–16, Mar. 2019.

[2] Y. Li, A. Zhang, J. Huang, and Z. Xu, "An approach based on transfer learning to lifetime degradation rate prediction of the dry-type transformer," *IEEE Trans. Ind. Electron.*, vol. 70, no. 2, pp. 1811–1819, Feb. 2023.

[3] C. Bhargava, P. K. Sharma, M. Senthilkumar, S. Padmanaban, V. K. Ramachandaramurthy, Z. Leonowicz, F. Blaabjerg, and M. Mitolo, "Review of health prognostics and condition monitoring of electronic components," *IEEE Access*, vol. 8, pp. 75163–75183, 2020.

[4] H. Wang and F. Blaabjerg, "Power electronics reliability: State of the art and outlook," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 6, pp. 6476–6493, Dec. 2021.

[5] A. Moradzadeh, B. Mohammadi-Ivatloo, K. Pourhossein, and A. Anvari-Moghaddam, "Data mining applications to fault diagnosis in power electronic systems: A systematic review," *IEEE Trans. Power Electron.*, vol. 37, no. 5, pp. 6026–6050, May 2022.

[6] J. Fan and W. Yuan, "Review of parametric fault prediction methods for power electronic circuits," *Eng. Res. Exp.*, vol. 3, no. 4, Dec. 2021, Art. no. 042002.

[7] M. Wang, W. Niu, and Y. Zhao, "Remaining useful life prediction of capacitor based on genetic algorithm and particle filter," in *Proc. IEEE/ACIS 20th Int. Fall Conf. Comput. Inf. Sci. (ICIS Fall)*, Xian, China, Oct. 2021, pp. 91–95.

[8] S. Rathnapriya and V. Manikandan, "Remaining useful life prediction of analog circuit using improved unscented particle filter," *J. Electron. Test.*, vol. 36, no. 2, pp. 169–181, Apr. 2020.

[9] H. Liu, T. Claeys, D. Pissoort, and G. A. E. Vandenbosch, "Prediction of capacitor's accelerated aging based on advanced measurements and deep neural network techniques," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 9019–9027, Nov. 2020.

[10] J.-H. Park, J.-H. Seok, K.-M. Cheon, and J.-W. Hur, "Machine learning based failure prognostics of aluminum electrolytic capacitors," *Korean Soc. Manuf. Process Eng.*, vol. 19, no. 11, pp. 94–101, Nov. 2020.

[11] A. L. Mesquita, V. P. Pinto, and L. R. Rodrigues, "Detection and fault prediction in electrolytic capacitors using artificial neural networks," in *Proc. Int. Conf. Appl. Technol.* Cham, Switzerland: Springer, Apr. 2020, pp. 287–298.

[12] D. K. B. Kulevome, H. Wang, and X. Wang, "A bidirectional LSTM-based prognostication of electrolytic capacitor," *Prog. Electromagn. Res. C*, vol. 109, pp. 139–152, 2021.

[13] S. Zhao, C. Zhang, and Y. Wang, "Lithium-ion battery capacity and remaining useful life prediction using board learning system and long short-term memory neural network," *J. Energy Storage*, vol. 52, Aug. 2022, Art. no. 104901.

[14] J. Ma, D. Xia, H. Guo, Y. Wang, X. Niu, Z. Liu, and S. Jiang, "Metaheuristic-based support vector regression for landslide displacement prediction: A comparative study," *Landslides*, vol. 19, no. 10, pp. 2489–2511, Jun. 2022.

[15] J. Ma, D. Xia, Y. Wang, X. Niu, S. Jiang, Z. Liu, and H. Guo, "A comprehensive comparison among metaheuristics (MHs) for geohazard modeling using machine learning: Insights from a case study of landslide displacement prediction," *Eng. Appl. Artif. Intell.*, vol. 114, Sep. 2022, Art. no. 105150.

[16] K. Feng, J. C. Ji, Y. Zhang, Q. Ni, Z. Liu, and M. Beer, "Digital twin-driven intelligent assessment of gear surface degradation," *Mech. Syst. Signal Process.*, vol. 186, Mar. 2023, Art. no. 109896.

[17] Z. Wang, J. Qu, X. Fang, H. Li, T. Zhong, and H. Ren, "Prediction of early stabilization time of electrolytic capacitor based on ARIMA-Bi_LSTM hybrid model," *Neurocomputing*, vol. 403, pp. 63–79, Aug. 2020.

[18] Z. Yang, Y. Wang, and C. Kong, "Remaining useful life prediction of lithium-ion batteries based on a mixture of ensemble empirical mode decomposition and GWO-SVR model," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.

[19] S. H. I. Yong-Sheng, S. H. I. Meng-zhuo, D. I. N. G. En-Song, H. O. N. G. Yuan-Tao, and O. U. Yang, "Combined prediction method of lithium-ion battery life based on CEEMDAN-LSTM," *Chin. J. Eng.*, vol. 43, no. 7, pp. 985–994, 2021.

[20] Q. Ni, J. C. Ji, K. Feng, and B. Halkon, "A fault information-guided variational mode decomposition (FIVMD) method for rolling element bearings diagnosis," *Mech. Syst. Signal Process.*, vol. 164, Feb. 2022, Art. no. 108216.

[21] C. Zhang, S. Zhao, and Y. He, "An integrated method of the future capacity and RUL prediction for lithium-ion battery pack," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2601–2613, Mar. 2022.

[22] C. Wu, H. Sun, and Z. Zhang, "Stages prediction of the remaining useful life of rolling bearing based on regularized extreme learning machine," *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, vol. 235, no. 22, pp. 6599–6610, Nov. 2021.

[23] H. Soliman, H. Wang, and F. Blaabjerg, "A review of the condition monitoring of capacitors in power electronic converters," *IEEE Trans. Ind. Appl.*, vol. 52, no. 6, pp. 4976–4989, Nov. 2016.

[24] Y.-R. Chen, H. Liao, and J.-F. Chen, "Life prediction of aluminum electrolytic capacitors under normal operation," in *Proc. IEEE 4th Int. Future Energy Electron. Conf. (IFEEC)*, Nov. 2019, pp. 1–8.

[25] K. Laadjal, M. Sahraoui, A. J. M. Cardoso, and A. M. R. Amaral, "Online estimation of aluminum electrolytic-capacitor parameters using a modified Prony's method," *IEEE Trans. Ind. Appl.*, vol. 54, no. 5, pp. 4764–4774, Sep. 2018.

[26] F. Zhou, Z. Huang, and C. Zhang, "Carbon price forecasting based on CEEMDAN and LSTM," *Appl. Energy*, vol. 311, Apr. 2022, Art. no. 118601.

[27] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.

[28] Z. Zhang, L. L. Cao, W. H. Lin, J. K. Sun, X. M. Feng, and Q. Liu, "Improved prediction model for BOF end-point manganese content based on IPSO-RELM method," *Chin. J. Eng.*, vol. 41, no. 8, pp. 1052–1060, 2019.

[29] Y. Zhang, W. Zhang, Z. Guo, and S. Zhang, "An effective wind speed prediction model combining secondary decomposition and regularised extreme learning machine optimized by cuckoo search algorithm," *Wind Energy*, vol. 25, no. 8, pp. 1406–1433, Aug. 2022.

[30] Y. J. Zhang, K. Cheng-Long, L. Ya-Qian, F. Xing-Hu, Z. Jin-Xiao, W. Ming-Xue, and Y. Liu-Zhen, "Rapidly detection of total nitrogen and phosphorus content in water by surface enhanced Raman spectroscopy and GWO-SVR algorithm," *Spectrosc. Spectral Anal.*, vol. 41, no. 10, pp. 3147–3152, Oct. 2021.

[31] J. Yu, "State of health prediction of lithium-ion batteries: Multiscale logic regression and Gaussian process regression ensemble," *Rel. Eng. Syst. Saf.*, vol. 174, pp. 82–95, Jun. 2018.

LINGHUI MENG received the B.E. degree in information and computing science and the Ph.D. degree in electrical engineering from Beijing Jiao-tong University, Beijing, China, in 2011 and 2017, respectively.

Previously, he was a Visiting Research Scholar with the Center for Advanced Life Cycle Engineering, University of Maryland, USA, from 2015 to 2016. He has been working with the China Electronic Product Reliability and Environmental Testing Research Institute, Guangzhou, Guangdong, China. He is also with the Science and Technology on Reliability Physics and Application of Electronic Component Laboratory, Guangzhou. His research interests include fault diagnosis, prognostics, and health management of power electronics systems and devices.

QUAN SUN (Member, IEEE) received the Ph.D. degree in measuring and testing technologies and instruments from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2018.

He is currently working as an Associate Professor with the School of Automation, Nanjing Institute of Technology, Nanjing. His current research interests include condition monitoring, the fault diagnosis and prognostics of power electronic systems, and deep learning for intelligent health management.
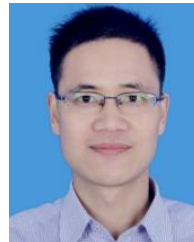
ZHENWEI ZHOU received the Ph.D. degree in operational research and cybernetics from the University of Chinese Academy of Sciences, in 2012. He is currently a Senior Engineer with the China Electronic Product Reliability and Environmental Testing Research Institute, Guangzhou, Guangdong, China. His research interests include fault diagnosis, prognostics, health management, and statistical reliability of electronics systems and devices.

LICHEN YANG was born in Jiangyin, Jiangsu, China, in 1997. He received the B.S. degree from the Department of Automation, Nanjing Institute of Technology, Nanjing, China, in 2019. He is currently pursuing the M.S. degree in mechanical engineering with the Nanjing Institute of Technology.

His current research interests include the fault diagnosis and deep learning of power electronics.

• • •