electronics

MDPI

Article

An Analytical Approach for IGBT Life Prediction Using Successive Variational Mode Decomposition and Bidirectional Long Short-Term Memory Networks

Kaitian Deng [1], Xianglian Xu [1,*], Fang Yuan [2,*], Tianyu Zhang [1], Yuli Xu [1], Tunzhen Xie [1], Yuanqing Song [1] and Ruiqing Zhao [1]

[1] School of Automation, Wuhan University of Technology, Wuhan 430070, China; 348798@whut.edu.cn (K.D.); 348850@whut.edu.cn (T.Z.); 348839@whut.edu.cn (Y.X.); tunzhen_xie@whut.edu.cn (T.X.); songyuan_q@whut.edu.cn (Y.S.); 334989@whut.edu.cn (R.Z.)

[2] School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China

* Correspondence: xuxianglian@whut.edu.cn (X.X.); coderyuan@whu.edu.cn (F.Y.)

Abstract: The precise estimation of the operational lifespan of insulated gate bipolar transistors (IGBT) holds paramount significance for ensuring the efficient and uncompromised safety of industrial equipment. However, numerous methodologies and models currently employed for this purpose often fall short of delivering highly accurate predictions. The analytical approach that combines the Pattern Optimization Algorithm (POA) with Successive Variational Mode Decomposition (SVMD) and Bidirectional Long Short-term Memory (BiLSTM) network is introduced. Firstly, SVMD is employed as an unsupervised feature learning method to partition the data into intrinsic modal functions (IMFs), which are used to eliminate noise and preserve the essential signal. Secondly, the BiLSTM network is integrated for supervised learning purposes, enabling the prediction of the decomposed sequence. Additionally, the hyperparameters of BiLSTM and the penalty coefficients of SVMD are optimized utilizing the POA technique. Subsequently, the various modal functions are predicted utilizing the trained prediction model, and the individual mode predictions are subsequently aggregated to yield the model's definitive final life prediction. Through case studies involving IGBT aging datasets, the optimal prediction model was formulated and its lifespan prediction capability was validated. The superiority of the proposed method is demonstrated by comparing it with benchmark models and other state-of-the-art methods.

Keywords: IGBT life prediction; successive variational mode decomposition; pelican optimization algorithm; bidirectional long short-term memory

check for updates

Citation: Deng, K.; Xu, X.; Yuan, F.; Zhang, T.; Xu, Y.; Xie, T.; Song, Y.; Zhao, R. An Analytical Approach for IGBT Life Prediction Using Successive Variational Mode Decomposition and Bidirectional Long Short-Term Memory Networks. *Electronics* **2024**, *13*, 4002. https://doi.org/10.3390/electronics13204002

Received: 10 September 2024
Revised: 3 October 2024
Accepted: 8 October 2024
Published: 11 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The IGBT, as the pivotal device for power conversion, boasts swift switching speeds and minimal driving power requirements [1]. Currently, it is widely applied in high-reliability scenarios such as aerospace, traction systems, industrial control, and energy storage [2]. Given the intricate operational environments in these domains and the accumulation of fatigue stemming from the IGBT's repetitive functioning, damage, and failure are increasingly likely to occur in these devices, ultimately rendering the entire system dysfunctional and leading to substantial economic losses. Consequently, the investigation of IGBT reliability is of paramount importance, and predicting their remaining operational lifespan is imperative [3].

According to the current research landscape, the primary approaches for predicting the remaining useful life (RUL) of IGBT can be broadly grouped into three categories: data-driven RUL prediction methods, RUL prediction methods grounded in failure physics models, and RUL prediction methods utilizing analytical models [4]. For RUL prediction methods based on failure physics models, it is imperative that the failure mechanisms of

the component, system geometry, material properties, external loads, and other pertinent physical characteristics be comprehended in order to establish an accurate mathematical model [5]. Alternatively, RUL prediction methods rooted in analytical models involve fitting aging data to establish the mathematical relationship between IGBT switching cycles and electrical parameters, subsequently constructing an aging prediction model. However, this approach necessitates high data accuracy and poses greater challenges in establishing the mathematical relationships [6]. The data-driven prediction approach entails the collection of a vast quantity of operational data and characteristic parameters pertaining to the RUL of IGBT, which are subsequently integrated with methodologies such as machine learning and deep learning for the purpose of analyzing extensive datasets on the aging process of IGBT [7]. The input parameters, including cycle counts, currents, and voltages, are utilized to feed the signals into the algorithm. Subsequently, models like neural networks are employed to discern patterns and rules within the IGBT aging data, thereby facilitating scientific predictions to be made [8].

Currently, the remaining life of IGBT is predominantly forecasted by researchers using data-driven prediction methodologies [9], which encompass, but are not limited to, time series analysis, neural networks, support vector machines, particle swarm optimization, and their respective combinations [10–12]. Notably, within the realm of deep learning models, LSTM networks are highlighted in the literature [13] for their distinctive architecture and exceptional prowess in tackling time series prediction challenges. Y. Wang et al. [14] introduced a whale optimization algorithm (WOA) to refine the LSTM-based lifetime prediction model, mitigating the subjectivity of manual parameter tuning. To delve into the original data, Y. Qin et al. [15] extracted and fused principal components of degraded features, integrating LSTM with macro- and microscopic attention for enhanced lifetime prediction. Owing to its superior performance, the LSTM network approach has garnered widespread adoption in fault diagnosis [16] and lifetime prediction [17], resolving the gradient vanishing issue of recurrent neural networks (RNNs) and optimizing temporal long-term dependency handling [18]. However, LSTM's forward propagation confines prediction solely to the last time step's learned features, neglecting information from other sequence steps. Conversely, the BiLSTM network leverages overlapping forward and backward propagation, adeptly capturing bidirectional temporal characteristics of IGBT aging data, given its intricate and lengthy degradation process [19]. J. Cui et al. [20] refined BiLSTM's structure with Squeeze and Excitation (SE) to enhance prediction accuracy. To mitigate noise interference, L. Wan et al. [21] employed complete ensemble empirical mode decomposition, reconstructing components through fine-to-coarse and phase space reconstruction, which were then fed into a CNN-BiLSTM fusion model for prediction. J. Yang combines Multiple Self Attention (MHSA) Mechanisms with Bidirectional Long Short-term Memory (BiLSTM) Networks to enhance the weight of key information and reduce unnecessary parts in order to increase prediction accuracy [22]. H. Wu utilizes a Bidirectional Long Short-term Memory Network that is optimized through the improved Newton–Raphson genetic algorithm (NRGA), and the innovative periodic attention mechanism (PAM) is added to focus on the periodic characteristics of the data [23]. H. Li presented the BiGRU model, which integrates bidirectional feature learning capabilities by combining the gated recurrent unit (GRU) structure with a bidirectional transmission mechanism, thereby enhancing its ability to capture time series features in the data [24]. Addressing IGBT degradation's nonlinear, non-smooth, and complex time series nature, numerous studies utilized empirical mode decomposition (EMD) and complete ensemble empirical mode decomposition (CEEMD) for denoising. Z. Huang employs Variational Mode Decomposition (VMD) to decompose performance data into intrinsic modes, elucidating critical multiscale dynamics vital for the complex degradation processes in fuel cells [25]. W. Li extracts features representing motor bearing faults by an improved complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and Variational Mode Decomposition (VMD)-based signal reconstruction method for fault diagnosis [26]. Yet these methods are often grappled with by issues of mode mixing and incomplete decom-

position [27]. In contrast, Successive Variational Modal Decomposition (SVMD) boasts robustness to initial mode center frequency values, adaptively decomposing data into multiple modes, strengthening data characteristics, inhibiting mode mixing, and revealing intrinsic connections, thereby enhancing prediction model accuracy in capturing complex sequence fluctuations [28]. With a solid theoretical foundation, SVMD has found applications in wind speed prediction, fault diagnosis, and related fields [29], proving adept at handling IGBT degradation's intricate raw data. Notably, the BiLSTM network and SVMD performance are significantly influenced by parameters, and the rationality of these parameters greatly impacts prediction outcomes. Manually selected initial parameters often yield unstable and inaccurate results, necessitating the optimization of BiLSTM and SVMD hyperparameters through group intelligence algorithms for more convincing outcomes.

To address the aforementioned limitations, a fusion lifetime prediction analytical approach, rooted in the POA and combined with SVMD and BiLSTM is advocated. Its prediction performance is substantiated using openly accessible datasets. Initially, data from the aging dataset are extracted and undergo preprocessing, during which the penalty factor of SVMD is optimized by POA. The refined feature data are then dissected into multiple modes by SVMD, a process that efficiently filters noise while preserving crucial signals. In parallel, the parameters of the BiLSTM model, inclusive of learning rate and regularization, are also optimized via POA. Subsequently, the decomposed multimodal data serves as input to the optimized BiLSTM network, which predicts the lifespan of each modality individually, exploring the temporal correlations inherent in degrading features. Finally, the predicted values of each sub-modal sequence are aggregated to yield the ultimate prediction. Comprehensive experiments, encompassing multiple hybrid models for both horizontal and longitudinal comparisons, are undertaken to affirm the feasibility and superiority of the approach introduced.

The following is a summary of the contributions made by this paper:

(1) Aging data are analyzed using Successive Variational Mode Decomposition (SVMD) to extract multimodal features, enhancing the prediction of lifespan by capturing subtle aging changes more effectively than traditional methods.

(2) This study employs the Pelican Optimization Algorithm (POA) combined with a Bidirectional Long Short-term Memory (BiLSTM) Network to optimize network parameters, while simultaneously optimizing the penalty factor in Successive Variational Mode Decomposition (SVMD), thereby enhancing the effectiveness and accuracy of data decomposition, improving learning efficiency, and fitting complex nonlinear relationships.

(3) A comprehensive POA-VMD-POA-BiLSTM prediction approach has been developed, integrating multimodal data decomposition, parameter optimization, and neural network prediction.

(4) The core principles and technological framework of the proposed IGBT life prediction methodology can also be applied in a multitude of other fields, including manufacturing and industrial equipment, the energy sector, medical and healthcare equipment, as well as automotive manufacturing, among others.

2. Basic Theory

2.1. Bidirectional Long Short-Term Memory Model

LSTM, as an optimization of the recurrent neural network, inherits the memorability function of RNN and introduces several crucial gates: the forgetting gate, the input gate, and the output gate [30]. These gates are implemented to refine the transmission path of the signal sequence, thereby effectively mitigating the long-term dependency issue encountered in standard RNNs and averting gradient vanishing and gradient explosion. Specifically, the input gate is operated upon the input feature parameters, while the forgetting gate determines whether the data within the candidate memory cells should be retained or discarded. The output gate, on the other hand, dictates whether the feature parameters are to be outputted. The memory cells themselves function as intermediate variables, storing crucial information throughout the process [31].

In Figure 1, G= $[G_1, G_2, G_3, \ldots, G_k]$ is the given time series, and G_t is the input of the current moment. The LSTM model uses the output of the previous moment h_{t-1} and the input of the current moment G_t to compute the current moment forgetting gate f_t, the input gate i_t, and the output gate o_t, and then f_t and i_t act together on the previous moment memory unit C_{t-1}, which updates the current moment memory state C_t according to the computational formula, and, finally, the output gate o_t passes the current moment memory state to the current moment hidden state h_t [32].

$$i_t = \sigma(w_{xi}G_t + w_{hi}h_{t-1} + b_i) \tag{1}$$

$$f_t = \sigma(w_{xf}G_t + w_{hf}h_{t-1} + b_f) \tag{2}$$

$$o_t = \sigma(w_{xo}G_t + w_{ho}h_{t-1} + b_o) \tag{3}$$

$$\widetilde{c} = \tanh(w_{xc}G_t + w_{hc}h_{t-1} + b_c) \tag{4}$$

$$C_t = f_t \odot C_{t-1} + \widetilde{c} \odot i_t \tag{5}$$

$$h_t = o_t \odot \tanh(C_t) \tag{6}$$

where i_t is the input gate at time step t; f_t is the forget gate at time step t; o_t is the output gate at time step t; h_t is the hidden state at time step t; σ is the Sigmoid function; w_{xi}, w_{xf}, and w_{xo} are the weight parameters of the input gate, the forget gate, and the output gate at time step t; w_{hi}, w_{hf}, and w_{ho} are the weight parameters of the input gate, the forget gate, and the output gate at time step $t-1$, respectively; b_i, b_f, and b_o are the bias parameters of the time step t; and C_t is the memory element of time step t moment's memory element.

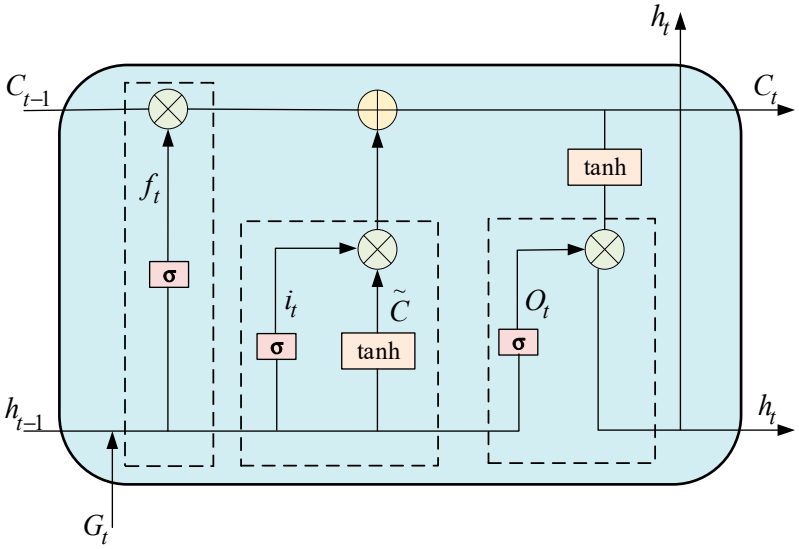


Figure 1. Structure diagram of LSTM.

BiLSTM consists of two layers of LSTM with the same number of neurons; the forward layer and the backward layer are used to compute the sequential and reverse order information of the time series [33], respectively, and, finally, the output of the forward layer $\overrightarrow{h}_t$ and the output of the backward layer $\overleftarrow{h}_t$ are linearly superimposed in accordance with a certain weighting $\overrightarrow{w}$ and $\overleftarrow{w}$ to obtain the final output y_t, and the mathematical expression is shown in the following equation:

$$\overrightarrow{h}_t = LSTM(G_t, \overrightarrow{h}_{t-1}) \tag{7}$$

$$\overleftarrow{h}_t = LSTM(G_t, \overleftarrow{h}_{t-1}) \tag{8}$$

$$y_t = \overrightarrow{w}\overrightarrow{h}_t + \overleftarrow{w}\overleftarrow{h}_t + b_y \tag{9}$$

where $LSTM$ is the computation process of traditional LSTM; G_t is the load feature information extracted by CNN; $\overrightarrow{h}_t$ is the hidden state vector of forward LSTM, which consists of neuron parameters output by the network after deep learning; $\overleftarrow{h}_t$ is the hidden state vector of reverse LSTM; y_t is the hidden state vector of BiLSTM, which consists of $\overrightarrow{h}_t$ and $\overleftarrow{h}_t$, and contains bidirectional temporal information; $\overrightarrow{w}$ and $\overleftarrow{w}$ are forward output weights and backward output weights, respectively, of BiLSTM, which is the bias optimization parameter b_y and are the forward and backward output weights of BiLSTM, respectively, which is the bias optimization parameter. Figure 2 Shows the BiLSTM network structure diagram.

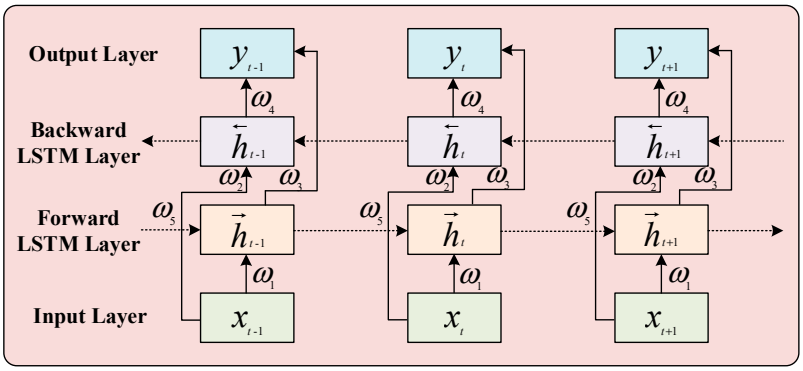


Figure 2. BiLSTM network structure diagram.

2.2. Successive Variational Mode Decomposition Method

Successive Variational Modal Decomposition (SVMD), an efficient approach for signal processing and data analysis, is represented as an optimized extension of the Variational Mode Decomposition (VMD) method [34]. The fundamental principle of Successive Variational Modal Decomposition (SVMD) involves a complex initial signal being decomposed into modal functions using Variational Modal Decomposition. In each iterative step, the modal functions are updated by minimizing the discrepancy between the signal and these functions, with the process reiterated until convergence Is achieved. The resultant modal functions, once converged, can be utilized to reconstruct the original signal with enhanced features, facilitating further analysis and processing as required.

In the experiment, SVMD decomposes the IGBT transient peak voltage degradation data $f(t)$ into two parts, i.e., the first L transient peak voltage component $u_L(t)$ and the residual voltage $f_r(t)$, i.e.,

$$f(t) = u_L(t) + f_r(t) \tag{10}$$

where $f_r(t)$ is the voltage component other than $u_L(t)$, which consists of two parts: the sum of the previously extracted IGBT transient peak voltage component and the unprocessed voltage component, as follows

$$f_r(t) = \sum_{i=1}^{L-1} u_i(t) + f_u(t) \tag{11}$$

To obtain the first voltage component in $f(t)$, the iterative expression for SVMD is as follows:

$$\hat{u}_L^{n+1}(\omega) = \frac{\hat{f}(\omega) + \alpha^2(\omega - \omega_L^n)^4 \hat{u}_L^n(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{[1 + \alpha^2(\omega - \omega_L^n)^4]\left[1 + 2\alpha(\omega - \omega_L^n)^2 + \sum_{i=1}^{L-1} \frac{1}{\alpha^2(\omega - \omega_i)^4}\right]} \tag{12}$$

where α is the balancing parameter for the data fidelity constraint, its value is usually very large, so the equation for updating ω_L will be approximated as follows:

$$\omega_L^{n+1} = \frac{\int_0^\infty \omega \left|\hat{u}_L^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_L^{n+1}(\omega)\right|^2 d\omega} \tag{13}$$

The Lagrange multipliers λ are denoted as follows:

$$\hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau\left[\hat{f}(\omega) - \left(\hat{u}_L^{n+1}(\omega) + f_u^{n+1}(t) + \sum_{i=1}^{L-1} u_i^{n+1}(\omega)\right)\right] \tag{14}$$

Finally, the modified expression is obtained using the double ascent method as follows:

$$\begin{cases} y(\omega) = \dfrac{\alpha^2(\omega - \omega_L^n)^4\left[\hat{f}(\omega) - \hat{u}_L^{n+1}(\omega) - \sum\limits_{i=1}^{L-1}\hat{u}_i(\omega) + \dfrac{\hat{\lambda}(\omega)}{2}\right] - \sum\limits_{i=1}^{L-1}\hat{u}_i(\omega)}{1 + \alpha^2(\omega - \omega_L^{n+1})^4} \\ \hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau\left[\hat{f}(\omega) - \left(\hat{u}_L^{n+1}(\omega) + y(\omega) + \sum\limits_{i=1}^{L-1} u_i^{n+1}(\omega)\right)\right] \end{cases} \tag{15}$$

where τ is the update parameter and ω_L is the center frequency [34].

2.3. Pelican Optimization Algorithm

POA is the Pelican Optimization Algorithm, which achieves global exploration and local exploitation of the optimal solution by simulating the pelican flock's approaching prey behavior and surface flight behavior, respectively, and achieves a proportional balance between exploration and exploitation to obtain the optimal solution. Figure 3 shows the optimization process of POA [35].

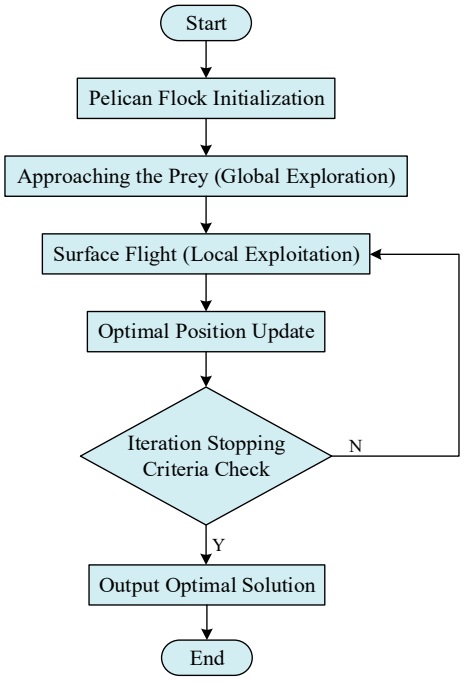


Figure 3. Optimization process of POA.

The optimization search steps of POA are as follows.

Step 1: Pelican flock initialization. Each pelican in the pelican flock represents a candidate solution, and the mathematical expression for the initialization of the pelican flock is as follows:

$$\begin{gathered} x_{i,j} = l_j + rand(u_j - l_j) \\ i = 1, 2, \ldots, N; j = 1, 2, \ldots, m; \end{gathered} \tag{16}$$

where $x_{i,j}$ is the value of the j variable of the i candidate solution; $rand$ is the random number in [0, 1]; u_j and l_j are the upper and lower limits of the j variable; N is the number of pelican groups, i.e., the number of candidate solutions; and m is the number of variables of the problem to be solved.

The first i candidate solution X_i can be expressed as follows:

$$X_i = [x_{i1}, x_{i2}, \cdots, x_{im}] \tag{17}$$

Step 2: Approaching the prey to achieve global exploration. The POA randomly generates the prey location in the solution space, and the pelican flock moves toward the prey location with the following mathematical expression:

$$x_{i,j}^{(1)} = \begin{cases} x_{i,j} + rand(p_j - Ix_{i,j}) & F_p < F_i \\ x_{i,j} + rand(x_{i,j} - p_j) & F_p \geq F_i \end{cases} \tag{18}$$

where $x_{i,j}^{(1)}$ is the updated value of the j variable of the i candidate solution; p_j is the value of the j variable of the prey; I is a random number equal to 1 or 2; F_p and F_i are the values of the fitness function of the prey and the i pelican, respectively.

If the fitness function value of the first i candidate solution is improved after updating its position, replace $x_{i,j}$ with $x_{i,j}^{(1)}$; otherwise, $x_{i,j}$ is unchanged.

Step 3: Surface flight to achieve local exploitation. POA to further approximate the optimal solution, the pelican position is updated using the following equation:

$$x_{i,j}^{(2)} = x_{i,j} + R(1 - t/T)(2rand - 1)x_{i,j} \tag{19}$$

where R is a constant, usually 0.2; t and T are the current iteration number and the maximum iteration number, respectively. If the fitness function value is improved after updating the position, then replace $x_{i,j}$ with $x_{i,j}^{(2)}$; otherwise, $x_{i,j}$ is unchanged.

Step 4: Optimal position update. After the positions of all candidate solutions are updated to the optimal solution for the current iteration number, their fitness function values are calculated. Following this, the algorithm advances to the subsequent iteration, with this sequence continuing until the iteration stop condition is fulfilled. The optimal solution attained during the iterative process of the algorithm is then designated as the optimal solution to the given problem.

2.4. An Analytical Pipeline for IGBT Life Prediction Based on POA-SVMD-POA-BiLSTM

In the field of time series data analysis, new analytical methods and technologies are continually explored by researchers. Furthermore, the study of time series data is no longer limited to a single domain but demonstrates a trend toward cross-domain integration. Typically, the five steps involved in its analysis and utilization are as follows:

Step 1: Data are collected from reliable sources, often containing parameters that vary over time. To prepare these data for modeling, it is proportionally divided into training, validation, and testing sets. Subsequently, the training and validation sets undergo a cleaning process where outliers are removed, missing values are handled, and necessary standardization is applied. Finally, the cleaned data are transformed into time series vectors to facilitate subsequent model processing.

Step 2: Appropriate feature parameters are selected based on the actual situation, extracted from the dataset, and then decomposed and reconstructed using suitable methods in order to have the noise removed and the signal enhanced.

Step 3: An appropriate time series model is selected based on the characteristics of the data and the purpose of analysis, and then the processed data are input into it for training.

Step 4: The prediction results are analyzed using error metrics including Root Mean Square Error (RMSE), Mean Relative Error (MRE), and Mean Absolute Percentage Error (MAPE) to assess the predictive performance of the model, and compare it with other models and methods to verify its superiority.

Step 5: The trained model is applied to practical situations, and continuous tuning and optimization are performed on the model based on the feedback and newly acquired data, with the goal of improving its prediction accuracy and generalization capability.

Figure 4 shows the process of processing the time series data.

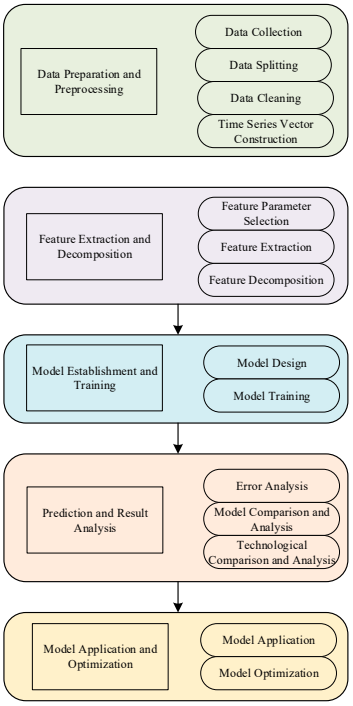


Figure 4. Diagram of the process of processing time series data.

In this study, a relevant prediction model based on the BiLSTM network was developed to tackle the challenges posed by IGBT aging time series data. First, the IGBT aging time series is decomposed into various modes using SVMD, whereupon a BiLSTM model, utilizing POA optimization, is constructed for each of the decomposed modes. Subsequently, the prediction outcomes corresponding to each mode are integrated to attain the final lifetime prediction. Figure 5 presents the flowchart of the POA-SVMD-POA-BiLSTM IGBT lifetime prediction.

The prediction steps of the approach developed are specified below:

(1) The full life cycle data of IGBT aging is proportionally segmented into training, validation, and test sets. These sets are specifically employed for the purposes of training the IGBT life prediction approach, validating its performance, and predicting the future life of the IGBT, respectively. Additionally, both missing value imputation and outlier detection are applied to ensure data quality. Simultaneously, the individual sequence components within the training and validation sets undergo standardization to normalize their scales, and from these standardized components, time series vectors are constructed by employing a sliding time window method to facilitate the modeling process. And 70% of the data in the dataset are used for training, while 30% of the data are employed as the validation of the prediction. The data normalization calculation formula is as follows:

$$x_{norm,t} = \frac{x_{init,t} - x_{\min,t}}{x_{\max,t} - x_{\min,t}} \tag{20}$$

where $x_{norm,t}$ is the normalized data; $x_{\min,t}$ is the minimum value in the original sequence; $x_{\max,t}$ is the maximum value in the original sequence.

(2) G. Sonnenfeld et al. found that the peak collector–emitter turn-off voltage exhibits a notable decreasing trend as the aging time increases, and this trend can be utilized to predict the RUL of the IGBT [36]. Therefore, it is chosen as the characteristic parameter of the prediction approach. In the feature decomposition section, the time series data are decomposed utilizing Sequential Variational Mode Decomposition (SVMD), resulting in its breakdown into multiple Intrinsic Mode Functions (IMFs), from which effective modes are subsequently selected. The penalty coefficient α serves as the pivotal setting parameter for SVMD, regulating the bandwidth. By selecting a suitable value for this coefficient, one can significantly enhance the accuracy of model predictions. This parameter is optimized in the article using the POA, thus eliminating the subjective effect of manually setting the parameter. Additionally, for the selection of the SVMD fitness function, since the envelope entropy reflects the sparse characteristics of the signal, a smaller value of the envelope entropy aids in identifying optimal parameter settings when the signal exhibits less noise and more feature information. Consequently, the envelope entropy is chosen as the fitness function for SVMD, and minimizing the envelope entropy becomes the objective of the search for the optimal.

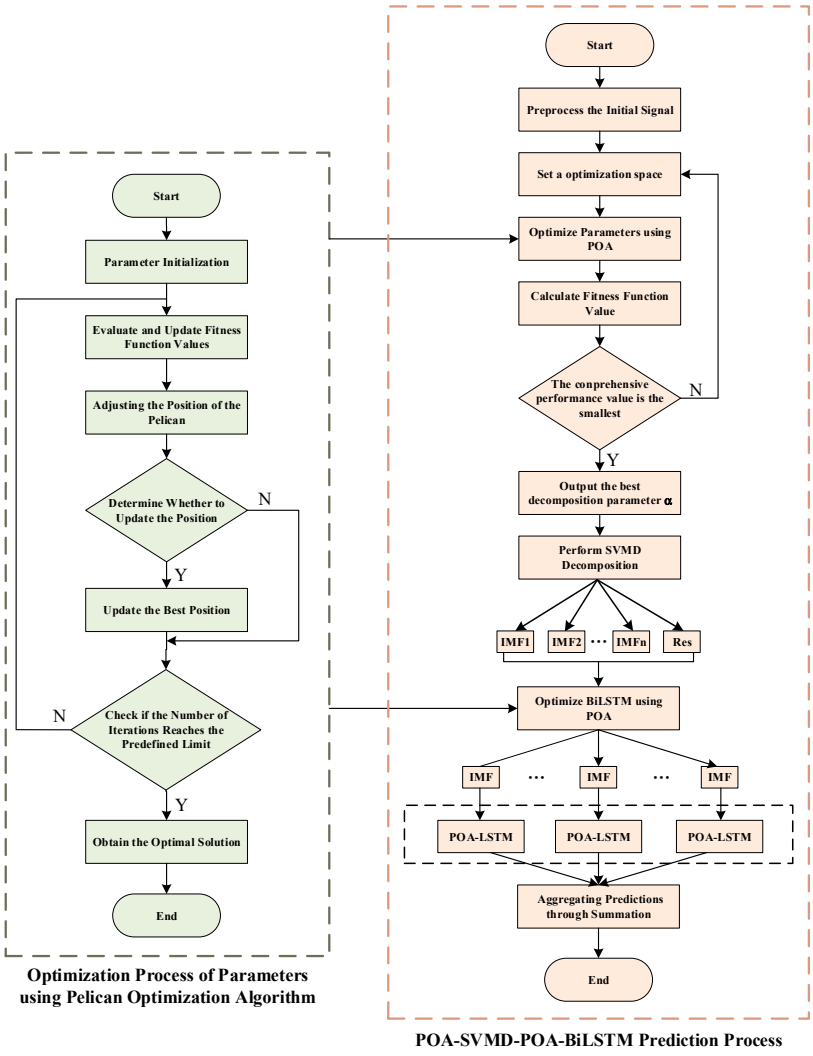


Figure 5. The flowchart of POA-SVMD-POA-BiLSTM IGBT life prediction.

The envelope entropy E_p of the $x(i)$ (i = 1, 2, . . ., N) signal with time sampling length N can be expressed as follows:

$$p_i = \frac{a(i)}{\sum_{i=1}^{N} a(i)} \tag{21}$$

$$E_p = -\sum_{j=1}^{N} p_i \lg p_i \tag{22}$$

where $a(i)$ is the envelope of the signal $x(i)$ after demodulation by the Hebert transform, and p_i is the result of the normalized probability distribution of $a(i)$.

(3) In the modeling part, the BiLSTM network is first initialized by the POA, with the optimization objectives being the number of neurons, the learning rate of the optimizer, and the width of the hidden layer. In this case, the distribution of the contemporary function is computed at random, and the model adapts the distribution of the contemporary function following the method employed for function selection, thereby assessing whether the termination criteria are fulfilled. Upon satisfaction, the highest-quality hyperparameter value is returned, and this principle dictates the construction of a BiLSTM network model equipped with the optimal hyperparameters.

(4) Multiple IMF components after POA-SVMD are input into the POA-BiLSTM network for temporal feature mining of the time series of the data, respectively, and, finally, the prediction results of each subsequence are inverted normalized and then superimposed and summed to obtain the final lifetime prediction results. The superiority of this model is highlighted through the analysis of prediction errors using error indicators like RMSE, MRE, and MAPE, as well as comparisons with various other models.

3. Experimental Process

3.1. Data Source

The foundational data for validating the aging prediction approach are provided by the accelerated aging test dataset for IGBT, published by the National Aeronautics and Space Administration (NASA) Ames Research Center [36]. The accelerated aging test dataset, specifically for the IRG4BC30KD IGBT model featuring a rated voltage of 600 V and a rated current of 15 A in a TO220 package, was collected with a 0.2 Ω resistor connected in series with the IGBT collector as a load, powered by a 4 V regulated power supply. In this experiment, a drive signal with a voltage of 10 V, a duty cycle of 40%, and a switching frequency of 10 kHz was applied to the gate of the IGBT, and the parameters of gate-emitter voltage, collector–emitter voltage, and collector–emitter current were recorded. After 172 min of testing and 418 turn-on and turn-off processes, the huge cyclic hermos-mechanical stresses within the module are generated by the junction temperature fluctuation of the IGBT, arising from the incompatibility of thermal expansion coefficients between various packaging materials. Subsequently, these stresses cause the degradation of the bond wires, solder layers, and other packaging materials, alongside the internal thermal impedance. Ultimately, this degradation results in a further escalation of temperature, with a latch-up effect being triggered, ultimately leading to the failure of the IGBT [36].

During each turn-on and turn-off process, the changes in collector voltage, heat sink temperature, and gate current were comprehensively recorded by the test instrument. Table 1 presents the key parameters of accelerated aging test.

Table 1. Key parameters of accelerated aging test.

Experimental Condition	Parameterization
Switching Frequency	10 kHZ
Gate Voltage	10 V
Duty Cycle	40%

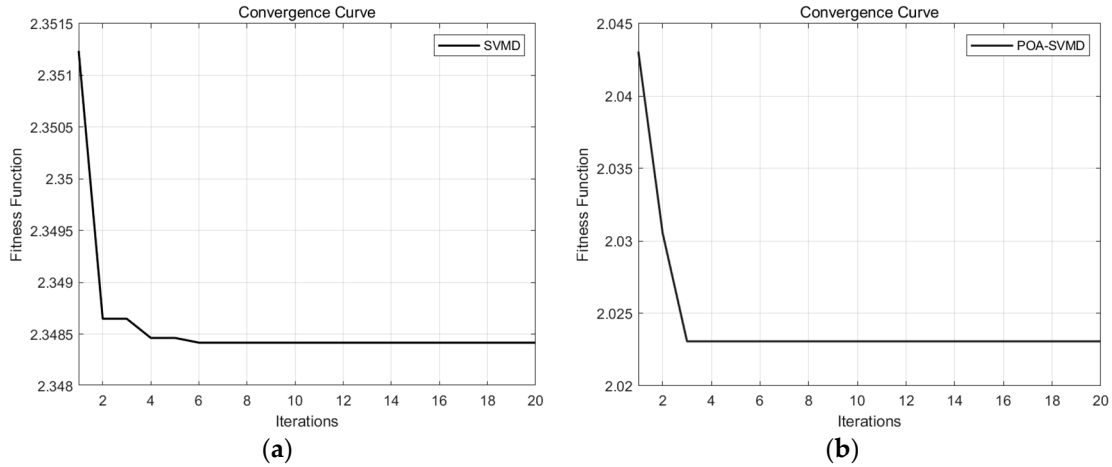
G. Sonnenfeld et al. found that the peak collector–emitter turn-off voltage exhibits a notable decreasing trend as the aging time increases, and this trend can be utilized to predict the RUL of the IGBT [36]. Therefore, the 418 transients $V_{ce-off-peak}$ of IGBT in the 418 switching cycles of the aforementioned dataset are selected as the model samples, and the data are standardized, with the time series vector constructed by employing a sliding time window method of 6. The sliding time window method with a window length of 6 is used to construct the time series vector of $[x_{norm,1}, x_{norm,2}, x_{norm,3}, x_{norm,4}, x_{norm,5}, x_{norm,6}]$, and then predict $x_{norm,7}$, so as to obtain 412 data samples. Subsequently, 70% of the data in the dataset are used for training, while 30% of the data are employed as the validation of the prediction.

3.2. Parameter Optimization Experiments and Settings

3.2.1. Decomposition Experiment of POA-SVMD

The α value for SVMD of the aging dataset is optimized by the POA, with 20 experimental tests conducted, 20 populations set, and 20 iterations performed. The range for α is specified as [20, 5000], and the fitness function selected aims to minimize the envelope entropy, ensuring a more coherent and passive formulation.

Figure 6 shows the fitness curves of SVMD and POA-SVMD, respectively. From Figure 6a, the optimal solution was found by SVMD after 6 iterations, with an envelope entropy value of 2.3484, and, consistently across 20 experiments, the number of iterations required remained within the range of 5 to 10 times. It is observable that, once SVMD initially identifies the interval of the solution set, the subsequent iteration is passively geared towards refining the search for the optimum within this interval, thereby facilitating the likelihood of encountering local optimal solutions. From Figure 6b, it can be observed that the envelope entropy value of 2.023 is converged by the optimal solution of POA-SVMD after merely 3 iterations, and, in 20 experiments, the number of iterations required remains within 5 times. This underscores that POA-SVMD exhibits superior stability, convergence accuracy, and speed compared to SVMD, with the iteration process effectively avoiding the emergence of local optimal solutions.



(**a**) (**b**)

Figure 6. Fitness curve during decomposition: (**a**) fitness curve during optimization with SVMD; (**b**) fitness curve during optimization with POA-SVMD.

The decomposition results of the voltage degradation data by POA-SVMD are shown in Figure 7. The IGBT transient peak voltage degradation data can be observed to have been decomposed into four voltage components containing different features, among which component 1, component 3, component 2, and component 4 are extracted in order of frequency from low to high. After the above experimental optimization and validation of the results, it is finally concluded that α = 1963 is the optimal solution.

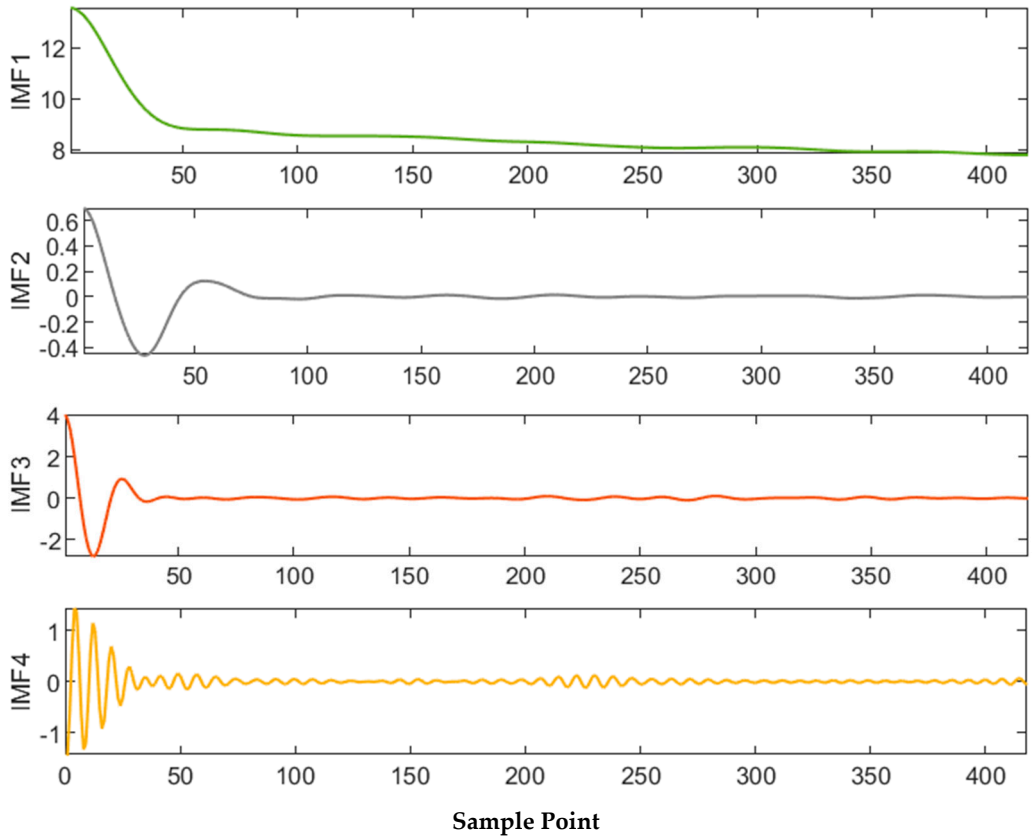


Figure 7. The modal component obtained by decomposing the detected signal.

In the modeling prediction comparison between POA-SVMD-POA-BiLSTM and SVMD-POA-BiLSTM, Table 2 reveals that the error associated with the POA-SVMD-POA-BiLSTM model, in which SVMD is optimized using POA, is notably smaller than that of the SVMD-POA-BiLSTM model, which employs a manually set penalty factor. This underscores the effectiveness of optimizing the penalty factor of SVMD through POA in enhancing the accuracy of lifetime prediction.

Table 2. Evaluation metrics for prediction error for both models.

Predictive Model	Prediction Error		
	RMSE/V	MAE/V	MAPE/%
POA-SVMD-POA-BiLSTM	0.0266	0.0131	0.0081
SVMD-POA-BiLSTM	0.0337	0.0177	0.0101

3.2.2. Optimization Process of POA-BiLSTM

The model was trained, validated, and tested on a Matlab®R2019a with Intel Core i5-10200H CPU, 8 GB RAM, and Microsoft Windows 10 64-bit operating system. In the training of the bidirectional long and short-term memory network, the Mean Square Error loss (MSE) is chosen to be optimized as the loss function, and the model parameters are dynamically adjusted and updated with the assistance of the Adam optimizer, aimed at minimizing the value of this loss function. The iterations are then repeated until the change in the loss value between consecutive iterations becomes sufficiently small. The main parameters of the IGBT life prediction model are shown in Table 3.

Table 3. Parameters of the BiLSTM.

Parameters	Value
Training Set	70%CM
Validation Set	30%CM
Maximum Optimization Time	3600 s
Maximum Number of Iterations	500
Abandonment Rate	0.3

To form the optimal combination of hyperparameters corresponding to the input data, the Pelican algorithm is utilized to optimize the model's hyperparameters, including the number of hidden layers, the number of neurons in the hidden layer, the learning rate, and the regularization parameter. The optimization search intervals for these relevant hyperparameters are presented in Table 4.

Table 4. Optimization-seeking intervals for hyperparameters.

Hyperparameterization	Optimization Interval
Number of Implicit Layers	[1, 5]
Number of Neurons in Hidden Layer	[50, 300]
Learning Rate	[0.005, 0.1]
Regularization Parameter	$[1 \times 10^{-5}, 1 \times 10^{-1}]$

Furthermore, the model parameters of BiLSTM are optimized by employing the POA optimization, which is then combined with the Adam optimizer for model training. This approach effectively enhances the performance and generalization ability of the deep learning model. The optimal hyperparameter combinations that have been optimized using the POA are shown in Table 5.

Table 5. Optimal solutions for hyperparameters.

Hyperparameterization	Value
Number of Implicit Layers	3
Number of Neurons in Hidden Layer	156
Learning Rate	0.019
Regularization Parameter	0.0812

In order to ensure the rigor of the results, the parameter settings, excluding the hyperparameters, were maintained consistently across models for subsequent comparative evaluations.

3.3. Model Evaluation Indicators

To evaluate the prediction performance of the POA-SVMD-POA-BiLSTM network model, the Root Mean Square Error (RMSE), the mean absolute error (MAE), and the Mean Absolute Percentage Error (MAPE) are employed as metrics. The calculation formulas for these metrics are outlined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{23}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{24}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{25}$$

where y_i and $\hat{y}_i$ denote the actual value of the ith test and the algorithmic model output, respectively, n is the total number of tests, and $\overline{y}_i$ denotes the average value of the algorithmic model output. And, where the smaller the value of each error evaluation index, the better the model prediction effect.

4. Experiment Process and Results

To verify the feasibility of the proposed POA-SVMD-POA-LSTM aging prediction model and to mitigate the constraints of a single prediction model, multiple models are adopted for horizontal and longitudinal comparisons, thereby enabling the evaluation and analysis of the proposed model.

4.1. Horizontal Comparison of Models

To validate the effectiveness of the SVMD sequence decomposition algorithm within the proposed POA-SVMD-POA-BiLSTM aging prediction model, two alternative algorithms, LMD (Local Mean Decomposition) and EEMD (Ensemble Empirical Modal Decomposition), are utilized to decompose the same initial data. Consequently, the POA-LMD-POA-BiLSTM model, the POA-BiLSTM model (assuming direct application without decomposition), and the POA-EEMD-POA-BiLSTM model are established, and a series of prediction experiments are conducted. Subsequently, the prediction error results are presented as follows:

As demonstrated in Figure 8 and Table 6, notable differences are exhibited in the prediction accuracy of BiLSTM models constructed utilizing varying sequence decomposition methods, and significantly distinct prediction effects are yielded by models that incorporate sequence decomposition techniques and those that do not. Taking SVMD as an illustrative example, the original sequence is divided by this method into trend and fluctuation sequences, with predictions for each being modeled individually before their summation. The model that incorporates SVMD has the trend sequence's characteristics learned from it, enabling the capture of the overall trend at the subsequent time point within the time series, and, subsequently, the fluctuation sequence's characteristics are learned, thereby enhancing the granularity of the predicted values.

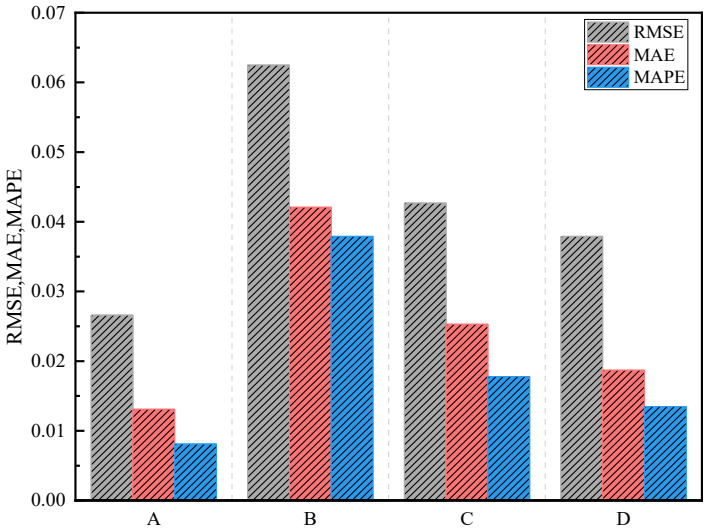


Figure 8. RMSE, MAE, and MAPE for different models. (A: POA-SVMD-POA-BiLSTM; B: POA-BiLSTM; C: POA-LMD-POA-BiLSTM; D: POA-EEMD-POA-BiLSTM).

A comparative analysis of the graphs has revealed that the RMSE values for the POA-SVMD-POA-BiLSTM, POA-LMD-POA-BiLSTM, and POA-EEMD-POA-BiLSTM models are 0.0266, 0.0427, and 0.0379, respectively, suggesting that the hybrid BiLSTM model augmented with SVMD is outperformed by none other than those utilizing LMD and EEMD.

Notably, the integration of SVMD is found to enhance the performance of the proposed hybrid model to a greater extent than that of LMD and EEMD, thereby emphasizing the necessity and significance of employing data decomposition techniques. Furthermore, when compared to the POA-EEMD-POA-BiLSTM model, which demonstrates relatively superior prediction outcomes, the presented prediction model is seen to have undergone reductions in RMSE, MAE, and MAPE by 0.0113, 0.0056, and 0.0054, respectively. This observation underscores that lower values of these metrics correspond to increased prediction accuracy, thereby validating the superior predictive capability that has been imparted to the model developed in this research.

Table 6. Evaluation index based on POA-BiLSTM combined with different decomposition methods.

Predictive Model	Prediction Error		
	RMSE/V	MAE/V	MAPE/%
POA-SVMD-POA-BiLSTM	0.0266	0.0131	0.0081
POA-BiLSTM	0.0625	0.0421	0.0379
POA-LMD-POA-BiLSTM	0.0427	0.0253	0.0178
POA-EEMD-POA-BiLSTM	0.0379	0.0187	0.0135

To provide a more comprehensive view of the prediction error distribution, a box plot of residuals is constructed and presented in Figure 9. This box plot is presented as a depiction of the predicted residual values for the four distinct models. Notably, when juxtaposed with the other models, the IGBT aging prediction model, grounded in POA-SVMD-POA-BiLSTM, is found to exhibit the narrowest range of residuals, featuring residuals that are more centrally concentrated and an overall residual mean that is closer to 0. In summary, upon being subjected to decomposition utilizing SVMD, the original time series is endowed with heightened robustness compared to alternative methods, ultimately contributing to a more coherent and substantial improvement in the overall prediction accuracy of the model.

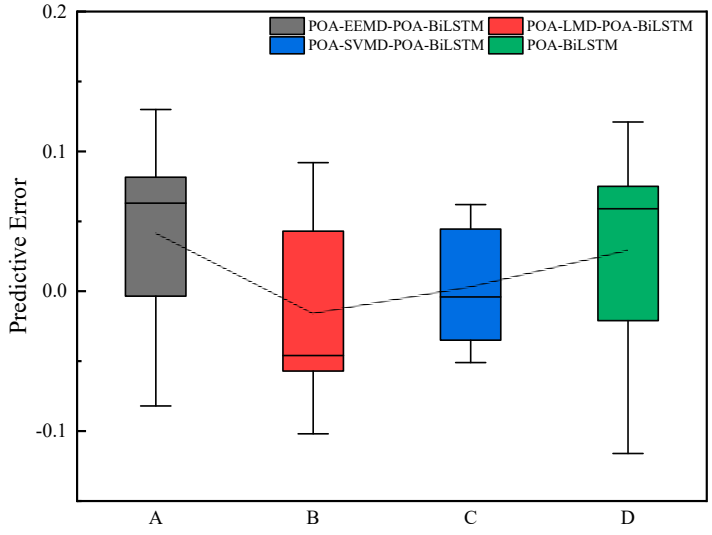


Figure 9. Box diagram of forecasting error absolute values (A: POA-EEMD-POA-BiLSTM; B: POA-LMD-POA-BiLSTM; C: POA-SVMD-POA-BiLSTM; D: POA-BiLSTM).

4.2. Longitudinal Comparison of Models

To underscore the superiority of the POA in optimizing the BiLSTM network within the proposed POA-SVMD-POA-BiLSTM aging prediction model, the GWO (Gray Wolf Optimization) Algorithm and SOA (Seagull Optimization Algorithm) are incorporated for the formulation of the POA-SVMD-GWO-BiLSTM, POA-SVMD-SOA-BiLSTM, and POA-SVMD-BiLSTM aging prediction models. These models are then subjected to a longitudinal

comparison with the introduced models, ensuring a more coherent and comprehensive analysis of their performance. The aim of this comparison is to explore the impact that the employment of diverse algorithms, for optimizing or non-optimizing neural networks, has on experimental prediction outcomes, given consistent input data.

To uphold the rigor of experimental findings, the four modal components derived from SVMD subsequent to POA optimization are uniformly selected as neural network inputs for prediction. First, both the POA-SVMD-BiLSTM and POA-SVMD-POA-BiLSTM models are individually fed with each modal component, and the subsequent prediction accuracy is compared. This approach emphasizes the necessity of POA optimization for BiLSTM networks. Furthermore, a comparative analysis is conducted among the aforementioned four models, with a focus on their overall prediction performance, thereby highlighting the precision achieved through the optimization facilitated by the POA. Shown below are the diverse component prediction plots, resulting from the input of each modal component into both the POA-SVMD-BiLSTM and the POA-SVMD-POA-BiLSTM models.

From Figure 10, the position by the SVMD optimized with POA displays a reduced deviation from the true values in comparison to the original SVMD. Furthermore, the prediction accuracy for the true values is heightened, leading to a more accurate prediction effect. Consequently, the utilization of POA to optimize the penalty factor of SVMD can be deemed necessary, and the incorporation of this optimization significantly enhances the prediction effect of the entire model.

In addition, the final overall prediction results of the POA-SVMD-POA-BiLSTM model, POA-SVMD-BiLSTM model, POA-SVMD-GWO-BiLSTM model, POA-SVMD-SOA-BiLSTM model, and POA-SVMD-BiLSTM model are shown in Figure 11.

From Figure 11, it can be observed that the overall trend in the predicted curve of the modal input network model, after being decomposed by SVMD, is the same as that of the measured value curve. Notably, the prediction curve of POA-SVMD BiLSTM exhibits the most pronounced fluctuations, deviating significantly from the true value curve and resulting in the lowest prediction accuracy. The remaining three curves reveal that the optimization of BiLSTM hyperparameters through algorithmic means effectively enhances the model's prediction accuracy. Furthermore, the utilization of the POA for hyperparameter optimization results in notably higher prediction accuracy compared to the other two algorithms, thereby establishing the POA-SVMD-POA-BiLSTM model as the one with the optimal prediction performance. The RMSE, MAE, and MAPE for the four prediction models are presented in Table 7.

As observed from Table 7, among the four distinct models, the three models that incorporate optimization of the BiLSTM network display smaller error evaluation indices than the POA-SVMD-BiLSTM model devoid of such optimization, thereby underscoring the necessity of hyperparameter optimization. Among the models employing three varying algorithms for hyperparameter optimization, the proposed method is found to exhibit the smallest error and the optimal prediction effect.

The overall error evaluation metrics, when calculated for these four models, reveal that the POA-SVMD-POA-BiLSTM-based IGBT aging prediction model undergoes a notable reduction of 25.69% in RMSE, 36.41% in MAE, and 44.52% in MAPE when juxtaposed with the POA-SVMD-BiLSTM model, thus signifying the enhancement in generalization capability attained by the proposed model. Furthermore, the POA-SVMD-POA-BiLSTM model surpasses both the POA-SVMD-GWO-BiLSTM and POA-SVMD-SOA-BiLSTM models, exhibiting reductions of 14.74% and 11.65% in RMSE, 19.63% and 10.88% in MAE, and 33.61% in MAPE, respectively, emphasizing its closer alignment between predicted and measured values. These comparative outcomes underscore that POA surpasses the other three algorithms in prediction accuracy, as evidenced by its lower loss values and improved stability. Consequently, the proposed model stands out as exhibiting superior prediction accuracy compared to other optimized prediction models, offering a significant reference point and applicability in predicting the remaining life of IGBT.

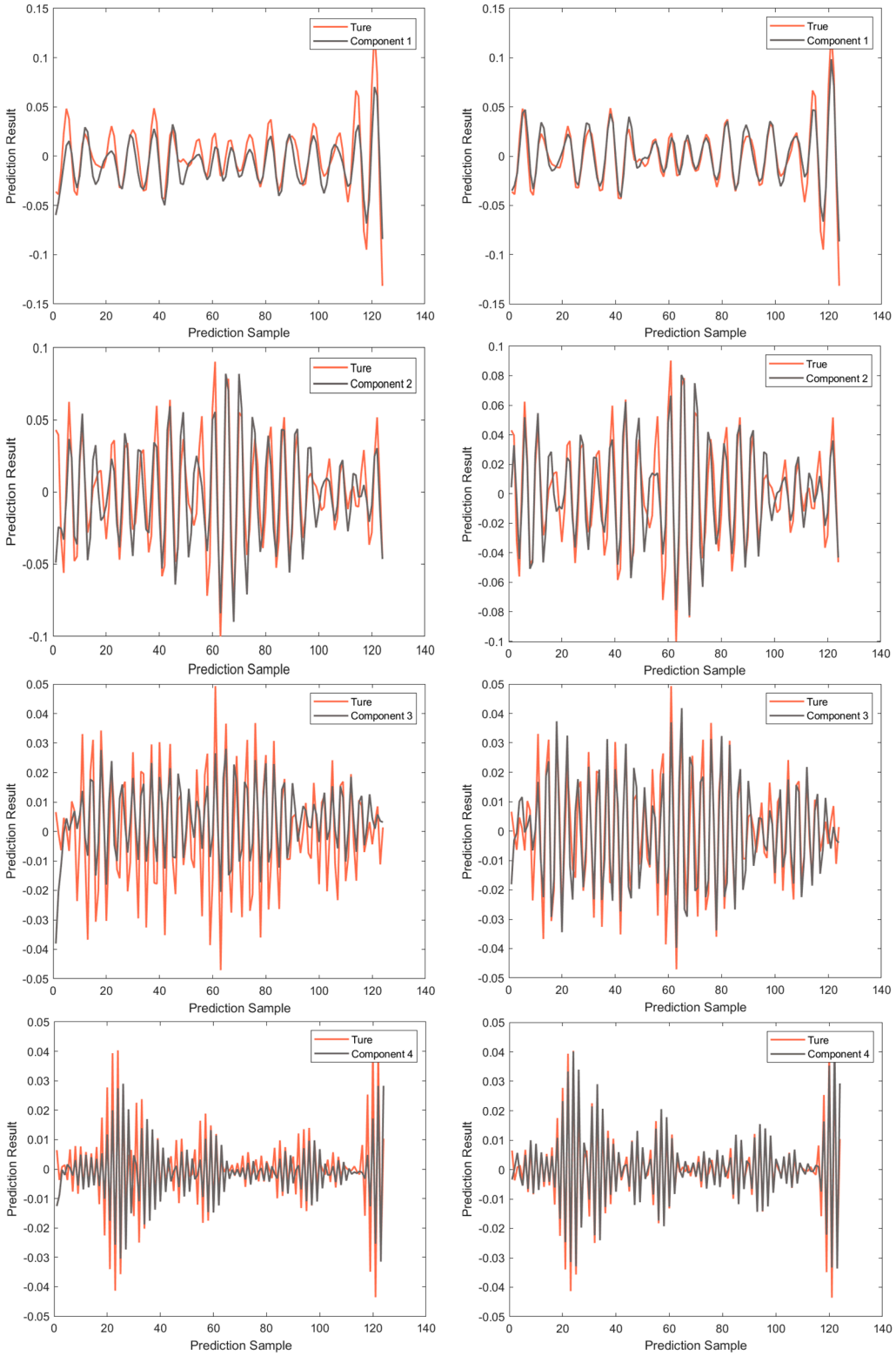


Figure 10. The life prediction results of each modal component: (**a**) predicted values of the components of the POA-SVMD-BiLSTM model; (**b**) predicted values of the components of the POA-SVMD-POA-BiLSTM model.

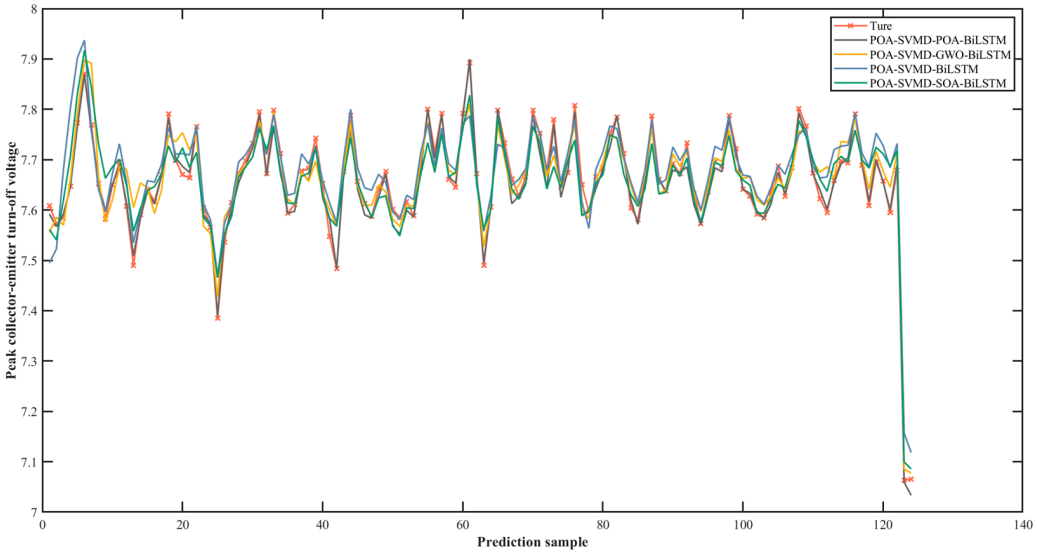


Figure 11. Predictions compared to different models.

Table 7. Evaluation index based on POA-SVMD combined with different prediction methods.

Predictive Model	Prediction Error		
	RMSE/V	MAE/V	MAPE/%
POA-SVMD-POA-BiLSTM	0.0266	0.0131	0.0081
POA-SVMD-BiLSTM	0.0358	0.0206	0.0146
POA-SVMD-GWO-BiLSTM	0.0312	0.0163	0.0122
POA-SVMD-SOA-BiLSTM	0.0297	0.0147	0.0109

5. Conclusions

As industrialization progresses, the considerable escalation in demand for power electronic devices, which encompass inverters and frequency converters, has prompted growing concern about the reliability and safety of their operation. Being the pivotal constituent of power electronic devices, the aging and failure of IGBT are the primary causes of equipment malfunctions, seriously affecting the operational safety of industrial devices under complex working conditions. Consequently, the development of an IGBT lifetime prediction approach holds immense significance, for it not only furnishes guidance for operational maintenance but also facilitates the mitigation of unplanned downtime, markedly bolstering the system's sustained operational capability and reliability. Furthermore, it is imperative to emphasize that the significance of this IGBT lifetime prediction approach transcends the realm of IGBT alone. Rather, it serves as a paradigm that ought to be extended and adapted to various other domains. By harnessing the principles and methodologies underpinning this approach, we can develop predictive maintenance strategies for a broad spectrum of electronic components and systems, thereby enhancing their operational reliability and safety across industries.

To tackle these challenges, an approach for predicting the residual life of IGBT utilizing POA-optimized SVMD-BiLSTM parameters was proposed. The precision in forecasting the RUL of IGBT was improved by merging SVMD data decomposition capabilities with BiLSTM's proficiency in handling time series data. In addition, the aging prediction model, being generic in nature, can be adapted to other areas where the health and longevity of a component, system, or structure are critical for ensuring performance, safety, and cost-effectiveness, rendering the proposed analysis approach of great significance. The main contributions are summarized as follows:

(1) The IGBT aging time series data are processed as signal data using SVMD, and the raw data are decomposed into various frequency-based patterns from low to high frequencies, thereby extracting the frequency characteristics through unsupervised feature learning. The decomposed modalities are input into the BiLSTM network for supervised learning, further enhancing its forecasting performance.

(2) The optimization of the hyperparameters of the Bidirectional Long Short-term Memory Network and the penalty factors of Successive Variational Modal Decomposition, facilitated by the Particle Optimization Algorithm, contributes significantly to the overall effectiveness of our proposed approach. This optimization process ensures that the model is adeptly adapted to the specific characteristics of the input data, thereby enhancing its generalization capability and robustness.

(3) The model, proposed for predicting the remaining life of IGBT utilizing POA and SVMD-BiLSTM, was compared horizontally and longitudinally with numerous other models, showcasing high prediction accuracy and stability. Notably, the efficient and adaptive realization of the POA-SVMD data preprocessing method is achieved, effectively segregating the detected signal components. When compared to alternative time series decomposition methods, the prediction accuracy of the hybrid model utilizing this method is observed to be superior, with a significant reduction in prediction errors. Furthermore, upon the optimization of the BiLSTM network's hyperparameters through the utilization of POA, the RMSE of the fusion model undergoes a 25.69% decrease, the MAE diminishes by 36.41%, and the MAPE declines by 44.52%. These enhancements are experimentally corroborated, emphasizing the improved predictive performance and outstanding generalization capability of the present approach when compared to others.

Inevitably, this paper remains subject to certain shortcomings that can be addressed for improvement. Firstly, the model's training data are confined to specific types of IGBT, potentially limiting its effectiveness in accurately predicting the lifespan of disparate IGBT types. Secondly, the establishment of the lifespan prediction model is based solely on a single feature quantity, overlooking other influential factors that may significantly affect the residual life. Therefore, for subsequent research endeavors, it is envisioned that transfer learning and related techniques will be utilized to enhance the model's universality and general applicability, while the correlation coefficient analysis method will be adopted to perform optimal parameter selection among multiple feature parameters. With these strategies being adopted, it is intended that meaningful contributions be made to the field by this research, extending its impact beyond academic discourse to foster practical progress and innovation, and enabling its application across diverse domains.

Supplementary Materials: The following supporting information can be downloaded at: https://data.nasa.gov/download/7wwx-fk77/application/zip. The research data in the Supplementary Materials are public data provided by the research institute of NASA.

Author Contributions: Conceptualization, K.D.; Data Curation, K.D.; Formal Analysis, R.Z.; Investigation, Y.S.; Methodology, X.X.; Project Administration, F.Y.; Software, T.Z.; Supervision, X.X.; Validation, K.D., Y.X. and T.X.; Visualization, K.D.; Writing—Original Draft, K.D.; Writing—Review and Editing, X.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The research data can be found in the Supplementary Materials.

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

References

1. Zhan, C.; Zhan, C.; Zhu, L.; Zhu, L.; Wang, W.; Wang, W.; Zhang, Y.; Zhang, Y.; Ji, S.; Ji, S.; et al. Multidimensional Mission-Profile-Based Lifetime Estimation Approach for IGBT Modules in MMC–HVdc Application Considering Bidirectional Power Transfer. *IEEE Trans. Ind. Electron.* **2023**, *70*, 7290–7300. [CrossRef]
2. Luo, W.; Zhang, D.; Xue, Y.; Yang, Z. Data-driven Prognostic for IGBT Failure Precursor Parameter Based on Informer. In Proceedings of the 2022 5th International Conference on Computing and Big Data (ICCBD), Shanghai, China, 16–18 December 2022; pp. 62–66. [CrossRef]

3. Musallam, M.; Yin, C.; Bailey, C.; Johnson, M. Mission Profile-Based Reliability Design and Real-Time Life Consumption Estimation in Power Electronics. *IEEE Trans. Power Electron.* **2015**, *30*, 2601–2613. [CrossRef]
4. Xu, K.; Xie, S.; Yan, Y.; Zhang, Z.; Zhang, B.; Qian, Q. Fault detection method for IGBT open-circuit faults in the modular multilevel converter based on predictive model. In Proceedings of the 2017 IEEE Energy Conversion Congress and Exposition (ECCE), Cincinnati, OH, USA, 1–5 October 2017; pp. 4190–4195. [CrossRef]
5. Abuelnaga, A.; Narimani, M.; Bahman, A.S. A Review on IGBT Module Failure Modes and Lifetime Testing. *IEEE Access* **2021**, *9*, 9643–9663. [CrossRef]
6. Wu, Q.; Chen, Y.; Luo, H.; Zhang, J.; Li, W.; He, X.; Fujishima, N.; Nishio, H.; Sumida, H. Identification Method for Various Failure Modes with Shared Kelvin and Power Wires Configuration in IGBT Power Modules. In Proceedings of the 2022 International Power Electronics Conference (IPEC-Himeji 2022-ECCE Asia), Himeji, Japan, 15–19 May 2022; pp. 133–138. [CrossRef]
7. Yang, X.; Zhang, Y.; Wu, X.; Liu, G. Failure Mode Classification of IGBT Modules Under Power Cycling Tests Based on Data-Driven Machine Learning Framework. *IEEE Trans. Power Electron.* **2023**, *38*, 16130–16141. [CrossRef]
8. A review of data-driven prognostic for IGBT remaining useful life. *Chin. J. Electr. Eng.* **2018**, *4*, 73–79. [CrossRef]
9. He, C.; Yu, W.; Zheng, Y.; Gong, W. Machine Learning Based Prognostics for Predicting Remaining Useful Life of IGBT—NASA IGBT Accelerated Ageing Case Study. In Proceedings of the 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Xi'an, China, 15–17 October 2021; pp. 1357–1361. [CrossRef]
10. Rakitskiy, A. Efficient Algorithms for Time Series Prediction Method. In Proceedings of the 2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), Yekaterinburg, Russia, 11–13 November 2022; pp. 120–123. [CrossRef]
11. Wang, X.; Zhang, Y. Multi-Step-Ahead Time Series Prediction Method with Stacking LSTM Neural Network. In Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 28–31 May 2020; pp. 51–55. [CrossRef]
12. Wang, M.; Lei, Y.; Liang, Y.; Lv, X.; Mo, W. A Support Vector Machine Model based on Mixed Time Series Prediction. In Proceedings of the 2021 IEEE International Conference on Data Science and Computer Application (ICDSCA), Dalian, China, 29–31 October 2021; pp. 47–51. [CrossRef]
13. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [CrossRef] [PubMed]
14. Wang, Y.; Li, N.; Zhao, W.; Guo, S.; Shen, M.; Li, X. Research on Life Prediction of Inverter IGBT Based on WOA Optimized LSTM Model. In Proceedings of the 2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Dali, China, 11–13 October 2023; pp. 1475–1479. [CrossRef]
15. Qin, Y.; Xiang, S.; Chai, Y.; Chen, H. Macroscopic–Microscopic Attention in LSTM Networks Based on Fusion Features for Gear Remaining Life Prediction. *IEEE Trans. Ind. Electron.* **2020**, *67*, 10865–10875. [CrossRef]
16. Zhu, D.; Song, X.; Yang, J.; Cong, Y.; Wang, L. A Bearing Fault Diagnosis Method Based on L1 Regularization Transfer Learning and LSTM Deep Learning. In Proceedings of the 2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE), Chengdu, China, 19–21 March 2021; pp. 308–312. [CrossRef]
17. Zhongyi, C.; Wenbo, H.; Rui, G.; Kai, Z.; Fuzhi, Z.; Zezhou, W. Remaining useful lifetime prediction based on the fusion of Attention-LSTM and Wiener process. In Proceedings of the 2023 Global Reliability and Prognostics and Health Management Conference (PHM-Hangzhou), Hangzhou, China, 12–15 October 2023; pp. 1–4. [CrossRef]
18. Ding, Y.; Jia, M.; Miao, Q.; Cao, Y. A novel time–frequency Transformer based on self–attention mechanism and its application in fault diagnosis of rolling bearings. *Mech. Syst. Signal Process.* **2022**, *168*, 108616. [CrossRef]
19. Liu, Y.-J.; Lee, Y.-D.; Lee, C.-Y.; Cheng, C.-C.; Hou, P.-Y.; Chen, Y.-F. A Comparative Analysis of LSTM and BILSTM Network-Based Methods in PV Power Prediction. In Proceedings of the 2022 IET International Conference on Engineering Technologies and Applications (IET-ICETA), Changhua, Taiwan, 14–16 October 2022; pp. 1–2. [CrossRef]
20. Cui, J.; Wang, Y.; Cui, X.; Jiang, L.; Liu, D.; Du, W.; Tang, X.; Wang, J. Prediction of Aeroengine Remaining Useful Life Based on SE-BiLSTM. In Proceedings of the 2022 34th Chinese Control and Decision Conference (CCDC), Hefei, China, 15–17 August 2022; pp. 1781–1786. [CrossRef]
21. Wan, L.; Chen, K.; Li, Y.; Wu, Y.; Wang, Z.; Li, C. A Novel Remaining Useful Life Prediction Method Based on CEEMDAN-IFTC-PSR and Ensemble CNN/BiLSTM Model for Cutting Tool. *IEEE Access* **2022**, *10*, 2182–2195. [CrossRef]
22. Yang, J.; Zhang, X.; Liu, S.; Yang, X.; Li, S. Rolling Bearing Residual Useful Life Prediction Model Based on the Particle Swarm Optimization-Optimized Fusion of Convolutional Neural Network and Bidirectional Long–Short-Term Memory–Multihead Self-Attention. *Electronics* **2024**, *13*, 2120. [CrossRef]
23. Wu, H.; Liu, H.; Jin, H.; He, Y. Ultra-Short-Term Photovoltaic Power Prediction by NRGA-BiLSTM Considering Seasonality and Periodicity of Data. *Energies* **2024**, *17*, 4739. [CrossRef]
24. Li, H.; Chen, C.; Wei, J.; Chen, Z.; Lei, G.; Wu, L. State of Health (SOH) Estimation of Lithium-Ion Batteries Based on ABC-BiGRU. *Electronics* **2024**, *13*, 1675. [CrossRef]
25. Huang, Z.; Zhang, D.; Wang, X.; Huang, X.; Wang, C.; Liao, L.; Dong, Y.; Hou, X.; Cao, Y.; Zhou, X. Machine Learning Prediction of Fuel Cell Remaining Life Enhanced by Variational Mode Decomposition and Improved Whale Optimization Algorithm. *Mathematics* **2024**, *12*, 2959. [CrossRef]

26. Li, W.; Fan, N.; Peng, X.; Zhang, C.; Li, M.; Yang, X.; Ma, L. Fault Diagnosis for Motor Bearings via an Intelligent Strategy Combined with Signal Reconstruction and Deep Learning. *Energies* **2024**, *17*, 4773. [CrossRef]
27. Ho, R.; Hung, K. A Comparative Investigation of Mode Mixing in EEG Decomposition Using EMD, EEMD and M-EMD. In Proceedings of the 2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Penang, Malaysia, 18–19 April 2020; pp. 203–210. [CrossRef]
28. Wang, C.; Jia, P.; Zhou, F.; Wang, R.; Tian, X. Planetary Gearbox Fault Diagnosis Based On a Multi-Convolutional Neural Network with SVMD and Feature Fusion Under Variable Speed Conditions. In Proceedings of the 2023 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), Xi'an, China, 2–4 November 2023; pp. 1–6. [CrossRef]
29. Sheng, A.; Xie, L.; Zhou, Y.; Wang, Z.; Liu, Y. A Hybrid Model Based on Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, GRU Network and Whale Optimization Algorithm for Wind Power Prediction. *IEEE Access* **2023**, *11*, 62840–62854. [CrossRef]
30. Dzulhijjah, D.A.; Kusrini, K.; Yuana, K.A. Comparative Analysis of RNN, LSTM, Bi-LSTM Performance for Location and Time Entity Recognition in Forest Fire Texts. In Proceedings of the 2024 2nd International Conference on Software Engineering and Information Technology (ICoSEIT), Bandung, Indonesia, 28–29 February 2024; pp. 181–186. [CrossRef]
31. Bao, R.; Zhou, Y.; Jiang, W. FL-CNN-LSTM: Indoor Air Quality Prediction Using Fuzzy Logic and CNN-LSTM Model. In Proceedings of the 2022 2nd International Conference on Electrical Engineering and Control Science (IC2ECS), Nanjing, China, 16–18 December 2022; pp. 986–989. [CrossRef]
32. Zheng, X.; Li, X. Wind Electricity Power Prediction Based on CNN—LSTM Network Model. In Proceedings of the 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE), Jinzhou, China, 18–20 August 2023; pp. 231–236. [CrossRef]
33. Fadili, Y.; El Yamani, Y.; Kilani, J.; El Kamoun, N.; Baddi, Y.; Bensalah, F. An Enhancing Timeseries Anomaly Detection Using LSTM and Bi-LSTM Architectures. In Proceedings of the 2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM), Leeds, UK, 23–25 July 2024; pp. 1–6. [CrossRef]
34. Huang, Z.; Li, L.; Ding, G. A Daily Air Pollutant Concentration Prediction Framework Combining Successive Variational Mode Decomposition and Bidirectional Long Short-Term Memory Network. *Sustainability* **2023**, *15*, 10660. [CrossRef]
35. Zhang, J.; Min, B.-W.; Gu, H.; Wu, G.; Wu, W. Rheological Behavior of SiO_2 Ceramic Slurry in Stereolithography and Its Prediction Model Based on POA-DELM. *Materials* **2024**, *17*, 4270. [CrossRef] [PubMed]
36. Sonnenfeld, G.; Goebel, K.; Celaya, J.R. An agile accelerated aging, characterization and scenario simulation system for gate controlled power transistors. In Proceedings of the 2008 IEEE AUTOTESTCON, Salt Lake City, UT, USA, 8–11 September 2008; pp. 208–215. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.