Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 2, No. 2, 2025

EMD-CPO-GRU-based Transformer Oil Temperature Prediction

Pengbo Han[1, *]

[1]School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China
[*]Corresponding Author: 212309010007@home.hpu.edu.cn

Abstract: To improve the accuracy of transformer oil temperature prediction and ensure the stability and safety of transformers during operation, this paper proposes an innovative oil temperature prediction method—an EMD-CPO-GRU hybrid model based on Empirical Mode Decomposition (EMD), Crested Porcupine Optimization (CPO) algorithm, and Gated Recurrent Unit (GRU). The method first decomposes the oil temperature data using EMD, effectively extracting the nonlinear and non-stationary characteristics of the signal, thereby providing more representative and effective features for subsequent predictions. Next, the CPO algorithm is applied to optimize the key hyperparameters of the GRU model, establishing efficient CPO-GRU sub-models for each modal component to improve the accuracy and robustness of the prediction model. Finally, the prediction results of each sub-model are weighted and integrated to obtain the final oil temperature prediction value. Experimental results show that the EMD-CPO-GRU model outperforms traditional prediction models and other hybrid models in transformer oil temperature prediction tasks. In terms of prediction accuracy, the EMD-CPO-GRU model achieves significant improvement, fully verifying its effectiveness as an efficient and precise transformer oil temperature prediction method. This approach not only provides a reliable basis for real-time monitoring and fault warning of power transformers but also offers new ideas and solutions for similar time-series prediction problems.

Keywords: Transformer oil temperature prediction; Empirical Mode Decomposition; Crested Porcupine Optimization; Gated Recurrent Unit; Hybrid prediction model.

1. Introduction

Power transformers are a crucial component of the electrical transmission system, and their stability plays a vital role in ensuring the reliability of power supply[1]. However, real-time monitoring of abnormal conditions and accurate diagnosis of transformers remain significant challenges[2]. Currently, oil-immersed transformers are the most common type. Transformer oil serves multiple purposes, including insulation, heat dissipation, and arc suppression, while also effectively preventing moisture ingress, thereby protecting the internal components of the transformer. Additionally, transformer oil functions as a coolant, assisting in the removal of electrical arcs caused by high temperatures to ensure safe operation of the equipment[3]. Therefore, changes in transformer oil temperature are an important indicator of the health status of the equipment, directly reflecting the operational state of the transformer. Accurately predicting transformer oil temperature is crucial for ensuring the safe operation of the equipment and optimizing the efficiency of the power system[4].

In the field of transformer oil temperature prediction, traditional statistical forecasting models, such as Autoregressive (AR) models[5], offer fast computational speeds. However, these methods have high data requirements, lack adaptability, and exhibit low robustness[6]. Xi et al.[7]used Support Vector Machine (SVM) with three-phase power load and Particle Swarm Optimization (PSO) for model parameter optimization. However, SVM is sensitive to the choice of kernel function and has high computational complexity when handling large datasets. Xiao et al.[8]proposed a machine learning-based intelligent prediction method, such as a decision forest algorithm learning model with dynamic correlations. However, these methods also have certain limitations. For example, decision forests (such as Random Forests) can handle high-dimensional data and reduce the risk of overfitting, but they suffer from poor model interpretability and difficulty in adjusting a large number of hyperparameters. Rao et al.[9]proposed a transformer oil temperature prediction method based on Bayesian networks and association rules. However, this method also has some drawbacks. For instance, Bayesian networks require high data completeness and accuracy, and missing data or high levels of noise may affect prediction performance. Taheri et al.[10]introduced a model based on basic heat transfer theory, the Electric-Thermal Resistance Model (E-TRM), which constructs a thermal resistance network by assigning thermal resistances to each three-dimensional heat transfer channel to estimate the thermal behavior of the upper oil of internal distribution transformers. However, the thermal resistance network method is typically suited for steady-state heat transfer problems and encounters difficulties in handling transient heat analysis.

To address the aforementioned issues, this paper proposes an innovative transformer oil temperature prediction model, EMD-CPO-GRU. This model combines the Empirical Mode Decomposition (EMD) algorithm with the Crested Porcupine Optimization (CPO) algorithm and the Gated Recurrent Unit (GRU) neural network. Additionally, the EMD-CPO-GRU model was experimentally tested on the ETT-small dataset, with results showing that it outperforms traditional GRU models and other related approaches in terms of prediction accuracy. The study demonstrates that the proposed EMD-CPO-GRU model is an efficient and accurate method for transformer oil temperature prediction, which plays a significant role in ensuring the stable operation and safety of power systems.

The structure of this paper is as follows: Chapter 1 introduces the importance of transformer oil temperature prediction and the research background, outlining the objectives and significance of this study. Chapter 2 presents the methodology, discussing in detail the basic principles of the Empirical Mode Decomposition (EMD) algorithm, the Crested Porcupine Optimization (CPO) algorithm, and the Gated Recurrent Unit (GRU) neural network. It also focuses on how these three components are effectively integrated to construct the EMD-CPO-GRU-based transformer oil temperature prediction model. Chapter 3 covers the experiments and results analysis, including the design of the experiment, the dataset used, and the prediction results based on the proposed model. A comparison with traditional methods is provided, and the model's performance under different test conditions, along with its advantages, is analyzed. Chapter 4 concludes the paper, summarizing the research findings and reinforcing the effectiveness and accuracy of the proposed EMD-CPO-GRU model in transformer oil temperature prediction. The chapter also suggests potential directions for future research and applications.

2. Methodology

2.1. Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is an innovative signal processing method introduced by Huang et al.[11]. Unlike traditional Fourier and wavelet decompositions, EMD does not require any predefined basis functions. Instead, it performs adaptive decomposition based on the intrinsic time-scale characteristics of the data. This feature distinguishes EMD from decomposition methods that rely on prior harmonic or wavelet basis functions[12]. As a result, EMD is theoretically suitable for decomposing any type of signal, and it exhibits significant advantages, particularly when dealing with non-stationary and nonlinear data[13].

After decomposition, the resulting subsequences consist of a set of intrinsic mode functions (IMFs) and a residual that reflects the overall trend of the signal. The IMFs must satisfy two conditions: the number of local extrema (local maxima and minima) must equal the number of zero-crossings, or their difference must not exceed one; and at any point in time, the mean of the upper envelope, formed by the local maxima, and the lower envelope, formed by the local minima, should be zero[14]. The specific steps of the EMD algorithm are as follows:

1. Calculate the maximum and minimum values of the signal $X(t)$. Use cubic spline interpolation[15] to obtain the upper envelope curve $e_{\max}(t)$ and the lower envelope curve $e_{\min}(t)$, and then compute the mean value of the envelopes, denoted as $m(t)$, i.e., $m(t)=\left(e_{\max}(t)+e_{\min}(t)\right)/2$.

2. Calculate the difference between $h(t)$ and $m(t)$, which gives the residual signal $h(t)=X(t)-m(t)$.

3. Determine whether $h(t)$ satisfies the two conditions for an IMF. If satisfied, record it as the first IMF, denoted as $imf1(t)$.If not, treat $h(t)$ as the original signal and repeat step (1).

4. Extract the first IMF, $imf1(t)$,from the original signal to obtain the residual: $r(t)=X(t)-imf1(t)$.If $r(t)$ is a single oscillation function, continue the process. Otherwise, return to the original signal.

5. Finally, the original signal $X(t)$ can be decomposed into a number of IMFs and a residual:

$$X(t)=\sum_{i=1}^{n} imf_i(t)+r(t) \tag{1}$$

Where n is the number of IMFs.

2.2. Crested Porcupine Optimizer

The Crested Porcupine Optimizer (CPO) is a novel metaheuristic algorithm (intelligent optimization algorithm) proposed by Abdel-Basset et al. in 2024[16]. CPO employs four protection mechanisms[17]: vision, sound, smell, and physical attacks. Among them, the vision and sound strategies correspond to the exploration behavior of CPO, while the smell and physical attack strategies correspond to the exploitation behavior. CPO introduces an innovative strategy—Cyclic Population Reduction (CPR) technique—to maintain population diversity and accelerate convergence. This strategy simulates the phenomenon where only threatened porcupines (CPs) activate their defense mechanisms. During the optimization process, CPO speeds up convergence by removing some CPs from the population and later reintroducing them to enhance diversity, thus avoiding local minima[18], as shown in Equation (2).

$$N=N_{\min}+(N'-N_{\min})*(1-(\frac{t\%\frac{T_{\max}}{T}}{\frac{T_{\max}}{T}})) \tag{2}$$

Where T is the number of iterations, t is the current function value, $T_{\max}$ is the maximum number of iterations, and $N_{\min}$ is the minimum number of individuals in the newly generated population.

1. Exploration Stage: According to the defense behavior of the Crested Porcupine (CP), when the predator is at a greater distance, CP has two defense strategies, namely the vision strategy and the sound strategy. These strategies refer to surveying different areas, focusing on global exploration, as shown in Equation (3).

$$\overrightarrow{x_i^{t+1}}=\overrightarrow{x_i^t}+\tau_1*|2*\tau_2*\overrightarrow{x_{CP}^t}-\overrightarrow{y_i^t}| \tag{3}$$

There, x_{CP} represents the optimal position of the CP, and y_t represents the quantity generated between the current CP and a randomly selected CP, as shown in Equation (4).

$$\overrightarrow{y_i^t}=\frac{\overrightarrow{x_i^t}+\overrightarrow{x_r^t}}{2} \tag{4}$$

The CP uses the sound method to create noise and threaten the predator. When the predator approaches, the sound of the CP increases, as shown in Equation (5).

$$\overrightarrow{x_i^{t+1}} = (1-\overrightarrow{U_1}) * \overrightarrow{x_i^t} + \overrightarrow{U_1} * (\overrightarrow{y} + \tau_3 * (\overrightarrow{x_{r1}^t} - \overrightarrow{x_{r2}^t})) \quad (5)$$

2. Exploitation Stage: According to the defense behavior of the Crested Porcupine (CP), when the predator approaches, CP has two defense strategies, namely the smell attack strategy and the physical attack strategy. These strategies focus on the development of promising areas, dedicated to local exploitation searches. During the smell attack phase, the CP secretes a foul odor that spreads throughout the surrounding area to prevent the predator from approaching, as shown in Equation (6).

$$\overrightarrow{x_i^{t+1}} = (1-\overrightarrow{U_1}) * \overrightarrow{x_i^t} + \overrightarrow{U_1} * (\overrightarrow{x_{r1}}' + S_i' * (\overrightarrow{x_{r2}}' - \overrightarrow{x_{r3}}') - \tau_3 * \vec{\delta} * \gamma_t * S_i') \quad (6)$$

δ is the parameter controlling the search direction, γ_t is the defense factor, and S_t is the smell diffusion factor, as shown in Equation (7).

$$\vec{\delta} = \begin{cases} +1, i\overrightarrow{frand} \leq 0.5 \\ -1, Else \end{cases}$$

$$\gamma_t = 2 * rand * (1 - \frac{t}{t_{\max}})^{\frac{t}{t_{\max}}} \quad (7)$$

$$S_t^i = \exp(\frac{f(x_t^i)}{\sum_{k=1}^{N} f(x_t^k) + \varepsilon})$$

3. Physical Attack Stage: When the predator is very close, the CP attacks it with short and thick quills, as shown in Equation (8).

$$\overrightarrow{x_i^{t+1}} = \overrightarrow{x_{CP}^t} + (\alpha(1-\tau_4)+\tau_4) * (\delta * \overrightarrow{x_{CP}^t} - \overrightarrow{x_i^t}) - \tau_5 * \delta * \gamma_t * \overrightarrow{F_i^t} \quad (8)$$

Where α is the sensitivity coefficient[19].

2.3. Gated Recurrent Unit

The Gated Recurrent Unit (GRU) network[20] is an improved and simplified version of the Long Short-Term Memory (LSTM) network. In the GRU network, the forget gate and input gate of the LSTM are redesigned and integrated into a single update gate. In this paper, the role of the update gate is to determine the extent to which previous state information is retained. The larger the value of the update gate, the more information is retained. Compared to the LSTM network, the GRU network has the following advantages: fewer training parameters, faster learning speed, and the ability to achieve higher prediction performance in many application scenarios[21]. The structural unit of the GRU network is shown in Figure 1, and its computational formula can be expressed as:

$$r_t = \sigma\left(x_t W_{xr} + H_{t-1} W_{hr} + b_r\right) \quad (9)$$

$$z_t = \sigma\left(x_t W_{xz} + H_{t-1} W_{hz} + b_z\right) \quad (10)$$

$$\widetilde{H_t} = \tanh\left(x_t W_{hx} + R_t \odot H_{t-1} W_{hh} + b_h\right) \quad (11)$$

$$H_t = \left(1 - Z_t\right) \odot H_{t-1} + Z_t \odot \widetilde{H_t} \quad (12)$$

Where σ is an activation function, with the output range between 0 and 1; H_{t-1} contains the past information; R_t is the reset gate; ⊙ is the element-wise product; $\widetilde{H_t}$ is the candidate hidden state; and Z_t is the update gate[22].

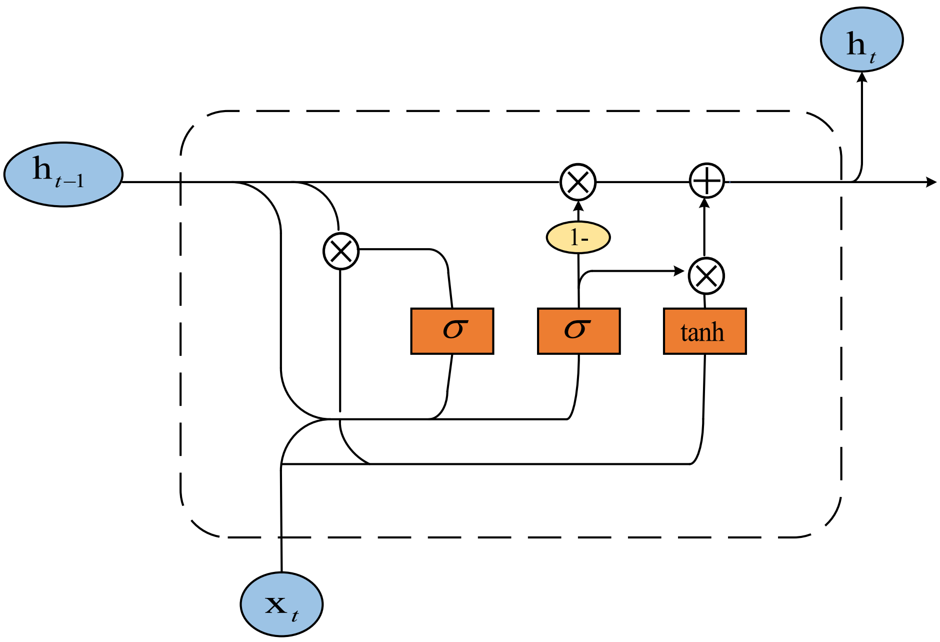

Figure 1. Structure Diagram of the Gated Recurrent Unit Network.

2.4. EMD-CPO-GRU-Based Oil Temperature Prediction Model

The oil temperature sequence of power transformers exhibits complex and non-stationary characteristics. To improve prediction accuracy, this paper proposes a hybrid prediction model architecture, EMD-CPO-GRU (as shown in Figure 2). The specific implementation steps are as follows:

1. Signal Decomposition: Use Empirical Mode Decomposition (EMD) to decompose the original load time series into multiple subsequences, extracting the nonlinear features within[23].

2. Data Preprocessing: Divide each subsequence into a training set and a validation set, and normalize the data. Since the GRU model is sensitive to the data scale and the original data has a large magnitude difference, normalizing the data to the range [0, 1] effectively avoids large fluctuations from masking small variations, thereby improving the accuracy of load prediction. This also helps accelerate the convergence speed of the GRU model.

3. Model Optimization and Training: Construct the GRU network and use the Crested Porcupine Optimization (CPO) algorithm to determine the optimal hyperparameter combination to learn the complex mapping relationship between input and output variables for each subsequence.

4. Denormalization and Prediction: Use the trained GRU network to predict each subsequence, and then perform denormalization to restore the results to their real values.

5. Result Summary: Summarize the prediction results of all subsequences to obtain the final load prediction result.

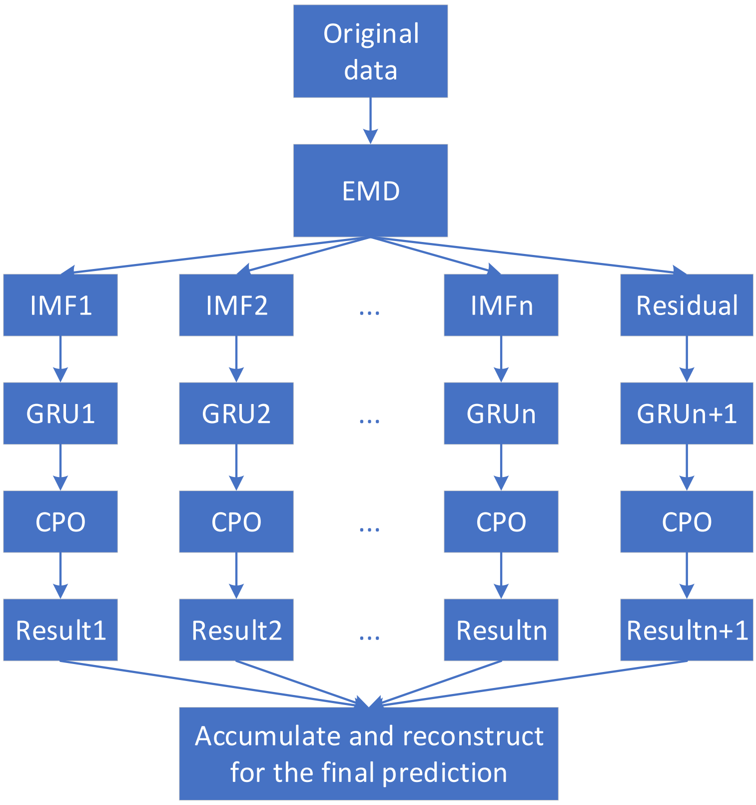

Figure 2. EMD-CPO-GRU Hybrid Prediction Model Architecture

3. Results and Discussion

3.1. Experimental Dataset

This study utilizes the publicly available power transformer dataset, specifically the ETT-small dataset[24]. The dataset contains operational data from two power transformers in a region of China over a two-year period. Data points recorded at a 15-minute frequency are labeled as ETT-m1 and ETT-m2, while variants with an hourly time granularity are named ETT-h1 and ETT-h2. Each data point includes eight features, including the date, the target variable (oil temperature), and six load values (HUFL, HULL, MUFL, MULL, LUFL, and LULL). Missing values in the dataset were cleaned and filled using data preprocessing methods. This dataset is of great value to time series research due to its authenticity and the richness of the data. In this experiment, ETT-h1 dataset was used for analysis[25]. A portion of the ETT-h1 dataset is shown in Table 1:

Table 1. Partial Data of ETT-h1 Dataset

date	HUFL	HULL	MUFL	MULL	LUFL	LULL	OT
2016/7/1 0:00:00	5.827000141143799	2.009000062942505	1.5989999771118164	0.4620000123977661	4.203000068664552	1.3400000333786009	30.5310001373291
2016/7/1 1:00:00	5.692999839782715	2.075999975204468	1.4919999837875366	0.4259999990463257	4.142000198364259	1.371000051498413	27.78700065612793
2016/7/1 2:00:00	5.1570000648498535	1.741000056266785	1.2790000438690186	0.3549999892711639	3.776999950408936	1.218000054359436	25.04400062561035
…	…	…	…	…	…	…	…
2018/6/26 19:00:00	10.11400032043457	3.549999952316284	6.183000087738037	1.5640000104904177	3.716000080108643	1.462000012397766	9.56700038909912

3.2. Data Decomposition Using EMD

Using the EMD method, the original oil temperature data from the ETT-h1 dataset were decomposed into multiple IMF components and a residual component[26]. Each IMF component represents different fluctuation characteristics and changing trends within the oil temperature sequence. Figure 3 shows the decomposition results for the four datasets, with these components displayed in order from high frequency to low frequency. By analyzing these charts, it can be observed that the first few IMF components exhibit relatively random fluctuations with no obvious regularity, indicating that oil temperature changes are significantly influenced by sudden events or other factors. The middle IMF components show relatively regular fluctuation patterns, which are similar to the daily fluctuations of the original oil temperature sequence. These components reflect stable changes related to daily operational cycles and are less affected by factors such as weather. The last few IMF components are low-frequency components with longer periods and smoother fluctuations, reflecting the slow changes in oil temperature over time, potentially related to long-term weather variations and other factors. The final residual component represents the overall trend of the oil temperature data, capturing long-term variation characteristics.

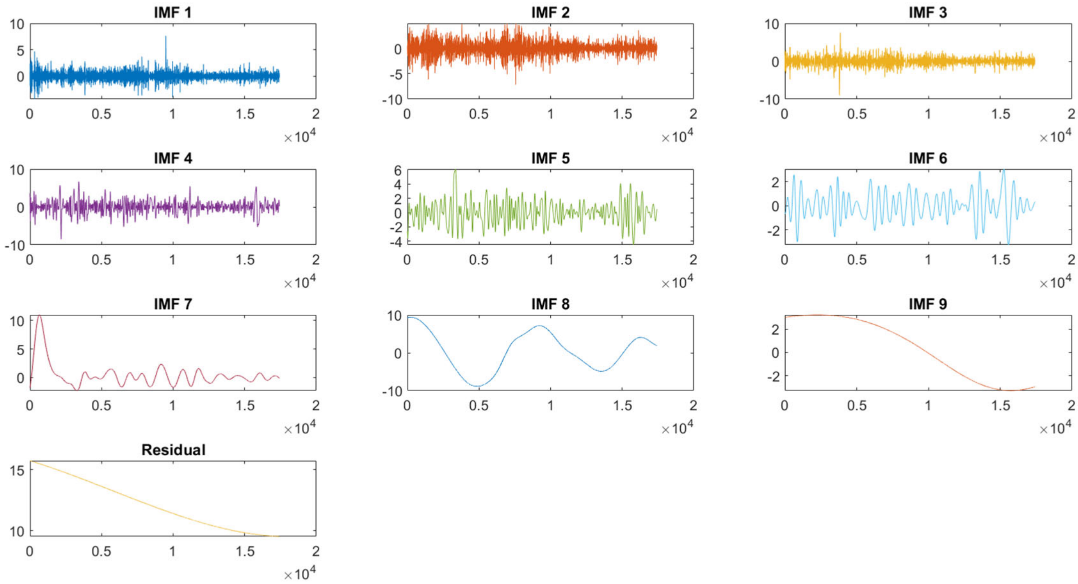

Figure 3. Decomposition Results of the ETT-h1 Dataset

3.3. Optimizing GRU Hyperparameters Using CPO

The Crested Porcupine Optimizer (CPO) is an efficient population-based optimization algorithm known for its simplicity, robustness, and fast convergence. It can effectively avoid local optima and efficiently find the global optimum solution[27]. In this study, the CPO algorithm is applied to optimize the hyperparameters of the GRU model to enhance its performance in load prediction tasks.

To achieve this, the CPO algorithm optimizes three key hyperparameters: learning rate, the number of hidden layer nodes, and the regularization coefficient[28]. During the optimization process, the value ranges for these parameters are clearly defined: the learning rate ranges from [0.0001, 0.01], the number of hidden layer nodes ranges from [8, 128], and the regularization coefficient ranges from [0.00001, 0.01]. By searching for the optimal values within these ranges, the CPO algorithm can dynamically adjust the hyperparameters, significantly improving the prediction accuracy of the GRU model.

3.4. Model Evaluation Metrics

To verify the accuracy of the prediction results, the Mean Squared Error (MSE), Mean Absolute Error (MAE), and R^2 are used for evaluation. Their calculation methods are as follows[29]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_{pre,i} - y_{act,i}\right)^2 \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left| y_{pre,i} - y_{act,i} \right| \tag{14}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{pre,i} - y_{act,i}\right)^2}{\sum_{i=1}^{n}\left(y_{act,i} - \overline{y}_{act}\right)^2} \tag{15}$$

Where n is the number of data points; $y_{act,i}$ is the actual value at time step of i; $y_{pre,i}$ is the predicted value at time step of i; and $\overline{y}_{act}$ is the average of the actual values.

MSE and MAE are better when smaller, indicating that the model performs well and the difference between predicted values and actual values is minimal. Ideally, they should approach 0, reflecting high prediction accuracy. An R^2 value of 1 indicates perfect fitting, meaning the model can explain all the variance in the data. R^2 should ideally approach 1, indicating a higher degree of fit between the model and the data.

3.5. Experimental Results

To validate the prediction performance advantages of the proposed model on the same dataset, multiple comparison experiments were conducted[30]. First, the performance of the single LSTM and GRU networks was compared to highlight the unique advantages of the proposed model. Next, a comparison was made with the VMD-GRU model. Third, the performance of different optimization algorithms combined with GRU was evaluated, including GWO-GRU and CPO-GRU, to determine the best optimization strategy. Finally, the combined CNN-GRU model was compared. All models were implemented under the same conditions, and Table 2 summarizes the results of the three evaluations metrics for all models.

Table 2. Evaluation Metric Results for Different Models

Model	MAE	MSE	R^2
LSTM	0.547	0.588	0.956
GRU	0.478	0.447	0.967
VMD-GRU	0.222	0.498	0.963
GWO-GRU	0.469	0.442	0.961
CPO-GRU	0.489	0.478	0.964
LSTM-GRU	0.592	0.678	0.949
EMD-CPO-GRU	0.367	0.221	0.984

Based on the results in the table, the EMD-CPO-GRU model demonstrates significant advantages in all three evaluations metrics: MAE (Mean Absolute Error), MSE (Mean Squared Error), and R^2. Its MAE value is 0.367, which is much lower than that of other models, indicating that the difference between the predicted and actual values is minimal, and the prediction accuracy is high. At the same time, the EMD-CPO-GRU model's MSE value is 0.221, the lowest among all models, showing that the model effectively suppresses large errors, particularly excelling in handling high-error situations. Furthermore, the model's R^2 value reaches 0.984, significantly higher than other models (for example, LSTM at 0.956, GRU at 0.967). The R^2 value close to 1 indicates a high degree of fit to the data, allowing the model to accurately capture the temperature variation trends of the transformer oil, thus significantly improving prediction accuracy. The prediction results of different models are shown in Figure 4.

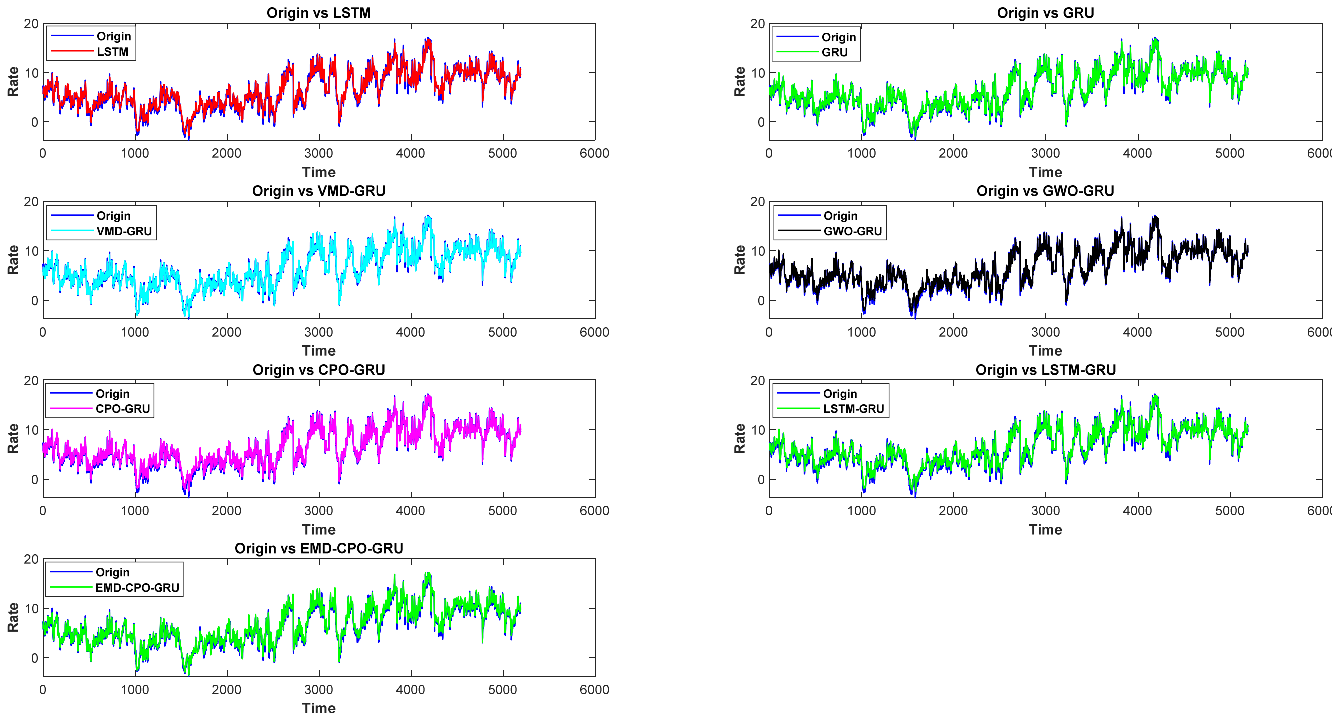

Figure 4. Fitting Plot of Prediction Results for Different Models

From the figure, it can be observed that the prediction results of the EMD-CPO-GRU model fit the original data exceptionally well. The green curve closely follows the blue curve, indicating that the model can accurately capture the fluctuation trends in oil temperature changes with minimal error. In contrast, traditional models like LSTM and GRU show significant deviations in their predictions, especially in areas with sharp data fluctuations, where they fail to accurately track the changes in the original data. Models like VMD-GRU, GWO-GRU, CPO-GRU, and LSTM-GRU, while performing well in certain time periods, still do not match the overall fitting performance of the EMD-CPO-GRU model, exhibiting some degree of error.

4. Conclusion

As a crucial power transmission and transformation equipment in the power system, power transformers generate a wealth of oil temperature time series data during operation[31]. Analyzing and mining these data can effectively monitor the operational status of transformers, thereby ensuring the stability and reliability of the power supply. To this end, this paper proposes an oil temperature prediction method based on Empirical Mode Decomposition (EMD), Crested Porcupine Optimization (CPO) algorithm, and Gated Recurrent Unit (GRU). The method takes transformer oil temperature time series data as the research object, treating it as a target for signal processing. By decomposing the raw data using EMD, it extracts frequency and amplitude features, accurately capturing the nonlinear characteristics and complex patterns in oil temperature changes, thereby significantly improving the model's accuracy in predicting short-term fluctuations (especially sudden oil temperature changes). During the prediction process, the CPO optimization algorithm is used to optimize the key parameters of the GRU model, further enhancing its ability to predict future oil temperature trends. This method provides a new technological pathway for monitoring the operational status of power transformers and holds significant practical application value and research significance.

References

[1] Jiménez-Navarro M J, Martínez-Ballesteros M, Martínez-Álvarez F, et al. A new deep learning architecture with inductive bias balance for transformer oil temperature forecasting[J]. Journal of Big Data, 2023, 10(1): 80.

[2] He Q, Si J, Tylavsky D J. Prediction of top-oil temperature for transformers using neural networks[J]. IEEE Transactions on Power Delivery, 2000, 15(4): 1205-1211.

[3] Chen H, Meng L, Xi Y, et al. Gru based time series forecast of oil temperature in power transformer[J]. Distributed Generation & Alternative Energy Journal, 2023: 393-412.

[4] Yang L, Chen L, Zhang F, et al. A Transformer Oil Temperature Prediction Method Based on Data-Driven and Multi-Model Fusion[J]. Processes, 2025, 13(2): 302.

[5] Wong C S, Li W K. On a mixture autoregressive model[J]. Journal of the Royal Statistical Society Series B: Statistical Methodology, 2000, 62(1): 95-115.

[6] Wang Z, Zhao B, Ji W, et al. Short-term load forecasting method based on GRU-NN model[J]. Automation of Electric Power Systems, 2019, 43(5): 53-58.

[7] Xi Y, Lin D, Yu L, et al. Oil temperature prediction of power transformers based on modified support vector regression machine[J]. International journal of emerging electric power systems, 2023, 24(3): 367-375.

[8] Xiao F, Yang G, Hu W. Research on intelligent diagnosis method of oil temperature defect in distribution transformer based on machine learning[J]. 2019.

[9] Rao W, Zhu L, Pan S, et al. Bayesian network and association rules-based transformer oil temperature prediction[C]//Journal of Physics: Conference Series. IOP Publishing, 2019, 1314(1): 012066.

[10] Taheri A A, Abdali A, Rabiee A. Indoor distribution transformers oil temperature prediction using new electro-thermal resistance model and normal cyclic overloading strategy: an experimental case study[J]. IET Generation, Transmission & Distribution, 2020, 14(24): 5792-5803.

[11] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 1998, 454(1971): 903-995.

[12] Boudraa A O, Cexus J C. EMD-based signal filtering[J]. IEEE transactions on instrumentation and measurement, 2007, 56(6): 2196-2202.

[13] Lahmiri S. Comparing variational and empirical mode decomposition in forecasting day-ahead energy prices[J]. IEEE Systems Journal, 2015, 11(3): 1907-1910.

[14] Da Silva I D, Moura M C, Lins I D, et al. Non-stationary demand forecasting based on empirical mode decomposition and support vector machines[J]. IEEE Latin America Transactions, 2017, 15(9): 1785-1792.

[15] McKinley S, Levine M. Cubic spline interpolation[J]. College of the Redwoods, 1998, 45(1): 1049-1060.

[16] Abdel-Basset M, Mohamed R, Abouhawwash M. Crested Porcupine Optimizer: A new nature-inspired metaheuristic[J]. Knowledge-Based Systems, 2024, 284: 111257.

[17] Zang J, Cao B, Hong Y. Research on the Fiber-to-the-Room Network Traffic Prediction Method Based on Crested Porcupine Optimizer Optimization[J]. Applied Sciences, 2024, 14(11): 4840.

[18] Fan Y, Ma Z, Tang W, et al. Using Crested Porcupine Optimizer Algorithm and CNN-LSTM-Attention Model Combined with Deep Learning Methods to Enhance Short-Term Power Forecasting in PV Generation[J]. Energies, 2024, 17(14): 3435.

[19] He Q, Li S. Prediction of Subgrade Settlement of High-speed Railway Based on BP Neural Network Optimized by Crested Porcupine[C]//2024 5th International Conference on Information Science, Parallel and Distributed Systems (ISPDS). IEEE, 2024: 356-359.

[20] Dey, R.; Salem, F.M. Gate-variants of gated recurrent unit (GRU) neural networks. In Proceedings of the 2017 IEEE 60th International Mid-West Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6–9 August 2017; IEEE: Piscatway, NJ, USA, 2017;

[21] Fantini D G, Silva R N, Siqueira M B B, et al. Wind speed short-term prediction using recurrent neural network GRU model and stationary wavelet transform GRU hybrid model[J]. Energy Conversion and Management, 2024, 308: 118333.

[22] Tian Q, Luo W, Guo L. Water quality prediction in the Yellow River source area based on the DeepTCN-GRU model[J]. Journal of Water Process Engineering, 2024, 59: 105052.

[23] Qin Q, Wang Z, Li B, et al. The HWAM-EMD-GRU Forecasting Model for Short-Term Passenger Flow in an Airport Light Rail Transit Line[J]. Urban Rail Transit, 2024, 10(2): 178-187.

[24] Zhou H, Zhang S, Peng J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. 2021, 35(12): 11106-11115.

[25] Zhang Y, Lu Y, Zhang W, et al. Big data analysis method based on time series temperature field data of substation transformer[C]//Journal of Physics: Conference Series. IOP Publishing, 2024, 2703(1): 012064.

[26] Fang T, Zheng C, Wang D. Forecasting the crude oil prices with an EMD-ISBM-FNN model[J]. Energy, 2023, 263: 125407.

[27] Liu H, Zhou R, Zhong X, et al. Multi-strategy enhanced crested porcupine optimizer: CAPCPO[J]. Mathematics, 2024, 12(19): 3080.

[28] Liu S, Jin Z, Lin H, et al. An improve crested porcupine algorithm for UAV delivery path planning in challenging environments[J]. Scientific Reports, 2024, 14(1): 20445.

[29] Bandi A, Adapa P V S R, Kuchi Y E V P K. The power of generative ai: A review of requirements, models, input–output

formats, evaluation metrics, and challenges[J]. Future Internet, 2023, 15(8): 260.

[30] Lara-Benítez P, Carranza-García M, Riquelme J C. An experimental review on deep learning architectures for time series forecasting[J]. International journal of neural systems, 2021, 31(03): 2130001.

[31] Chen H, Meng L, Xi Y, et al. Power transformer oil temperature prediction based on empirical mode decomposition-bidirectional long short-term memory[J]. Journal of algorithms & computational technology, 2023, 17: 17483026231176196.