processes

MDPI

Article

An Energy System Modeling Approach for Power Transformer Oil Temperature Prediction Based on CEEMD and Robust Deep Ensemble RVFL

Yan Xu [1], Haohao Li [1], Xianyu Meng [2], Jialei Chen [2], Xinyu Zhang [2] and Tian Peng [2,*]

[1] China Yangtze Power Co., Ltd., Wudongde Hydropower Plant, Kunming 651512, China; xu_yan4@ctg.com.cn (Y.X.); li_haohao@ctg.com.cn (H.L.)
[2] Nanjing Nanrui Jibao Engineering Technology Co., Nanjing 211100, China; mengxy@nrec.com (X.M.); chenjialei0408@163.com (J.C.); zxy66181107@163.com (X.Z.)
* Correspondence: husthydropt@126.com; Tel.: +86-150-0523-8786

Abstract

Accurate prediction of transformer oil temperature is crucial for load optimization scheduling and timely early warning of thermal faults in power transformers. This paper proposes a transformer oil temperature prediction method based on Complementary Ensemble Empirical Mode Decomposition (CEEMD), Outlier-Robust Ensemble Deep Random Vector Functional Link Network (ORedRVFL), and error correction. CEEMD is used to decompose the oil temperature data into multiple subsequences, enhancing the regularity and predictability of the data. Regularization and norm improvements are introduced to edRVFL to obtain a more robust ORedRVFL model. The Tent initialization-based Differential Evolution algorithm (TDE) is employed to optimize the model parameters and predict each subsequence. Finally, error correction is applied to the prediction results. Taking the main transformer of a hydropower station in Yunnan, China as an example, the experimental results show that the proposed method improves the prediction accuracy by 5.05% and 4.13% in winter and summer oil temperature predictions, respectively. Moreover, the model's degradation is significantly reduced when random noise is added, which verifies its robustness. This method provides an efficient and accurate solution for transformer oil temperature prediction.

Keywords: transformer oil temperature prediction; CEEMD; robust; error correction; differential evolution

check for updates

Academic Editors: Chang-Hua Lin, Shiue-Der Lu and Hwa-Dong Liu

Received: 15 July 2025
Revised: 2 August 2025
Accepted: 5 August 2025
Published: 6 August 2025

Citation: Xu, Y.; Li, H.; Meng, X.; Chen, J.; Zhang, X.; Peng, T. An Energy System Modeling Approach for Power Transformer Oil Temperature Prediction Based on CEEMD and Robust Deep Ensemble RVFL. *Processes* **2025**, *13*, 2487. https://doi.org/10.3390/pr13082487

Copyright: © 2025 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

Transformers are extremely critical in all aspects of power systems, including generation, transmission, distribution, and utilization of electricity. Their normal operation is the foundation for the efficient functioning and reliable power supply quality of the power system [1]. During operation, transformers generate heat, and high temperatures can accelerate the aging of and even damage to insulation materials, leading to failures. Due to their unique internal structure and environment, the measurement of hot-spot temperatures is subject to random errors, making it difficult to monitor accurately [2]. Accurate oil temperature predictions can assist power system operators in allocating loads more rationally and optimizing the operational efficiency of transformers [3]. Abnormal increases in transformer oil temperature are usually signs of impending failure [4]. By improving the accuracy of predictions, we can detect potential failure risks earlier [5],

thereby taking preventive measures to avoid equipment damage and power outages [6]. Accurate oil temperature predictions can assist power system operators in allocating loads more rationally and optimizing the operational efficiency of transformers [7]. Therefore, proposing an accurate method for accurately forecasting transformer oil temperature holds substantial importance.

During operation, transformers generate heat. Excessively high temperatures can accelerate the aging process of insulating materials, potentially leading to their damage [8]. Due to the unique internal structure and working environment of transformers, random errors may occur in the measurement of hot-spot temperatures inside the transformer, rendering accurate monitoring unfeasible [9]. Therefore, it is of great significance to propose an accurate method for predicting transformer oil temperature [10].

The current approaches to measuring transformer oil temperature primarily consist of direct temperature measurement and thermal circuit modeling, numerical computation, and intelligent model algorithms [11]. Direct temperature measurement involves installing temperature sensors on the transformer to collect temperature data in real time [12]. Although this method provides direct temperature readings, it faces challenges in insulation treatment. Even with the aid of fiber-optic technology, which enables its application, the cost is high and the maintenance is complex [13]. Therefore, despite its ability to provide direct temperature measurements, this method has certain limitations in terms of cost, maintenance, response speed, and accuracy, and is not widely used at present [14].

The thermal equivalent circuit method predicts oil temperature by analyzing the heat generation and dissipation processes within the transformer and constructing a thermal circuit model of the heat transfer process [15,16]. The numerical computation method mainly involves building a physical model of the transformer's interior, employing the finite element method or the finite volume method. These methods discretize continuous equations and then iteratively solve for the temperature field distribution to simulate and calculate the temperature distribution inside the transformer [17]. Both of these methods can accurately predict temperature changes during transformer operation and have certain applications. However, they inevitably face some issues. For example, although the thermal circuit model is physically well-defined and can intuitively reflect the heat transfer and distribution within the transformer, its construction and parameter calculation are relatively complex [18]. Its accuracy depends on appropriate parameters and boundary conditions, requiring a large amount of experimental and field measurement data, which introduces significant uncertainty [19]. The finite element method requires discretization of both time and space domains, while the finite volume method needs continuous iteration and correction during prediction. Although these methods can significantly improve simulation convergence and accuracy, they consume a large amount of computational resources and time and rely heavily on empirical formulas [20].

Based on the aforementioned characteristics, experts have increasingly utilized machine learning models for transformer oil temperature prediction. Zou et al. [21]. verified the effectiveness of Long Short-Term Memory (LSTM) [22] in estimating the upper oil temperature of transformers. Gunda et al. [23] formulated a linear discrete model for the hot-spot temperature of transformer windings based on the Kalman filter algorithm, predicting the hot-spot temperature of oil-immersed transformers under different operating loads and concluding that the winding hot-spot temperature is correlated with season, ambient temperature, and load. Oliveira [24] analyzed the temperature characteristics of oil-immersed transformers and used a BP neural network as the prediction model to accurately predict the hot-spot temperature of a 2000 KVA oil-immersed transformer. Li Ghnatios et al. [25] validated the effectiveness of self-supervised pre-training methods regarding top-oil temperature determination of transformers, proposing a dual-channel

pre-trained temporal attention network model that showed significant results in single time-step prediction of transformer top oil temperature. Juarez-Balderas et al. [26] proposed a transformer hot-spot temperature prediction model based on Artificial Neural Networks (ANNs), which was verified by experimental data through Finite Element Method (FEM) simulations and demonstrated accurate prediction results.

Although the methods proposed by the above scholars have achieved good prediction results, the temperature of transformer oil changes with the variation in the transformer load rate, and thus it has a high degree of uncertainty [27]. Single models are often sensitive to the distribution of historical data and struggle to capture the underlying patterns within this data [28]. Therefore, when the complexity of historical data increases, the prediction accuracy of single models often decreases significantly [29]. As a result, many scholars have combined data preprocessing techniques with prediction models. Feng et al. [30] performed principal component analysis on the main feature quantities to determine the transformer's hot-spot temperature windings, achieving input index reconstruction and obtaining better prediction results. Zhang et al. [31] used Ensemble Empirical Mode Decomposition (EEMD) to reduce data noise and constructed a deep learning model for predicting the hot-spot temperature of transformers using LSTM.

Existing transformer oil temperature forecasting approaches require a significant amount of historical data and are sensitive to data disturbances. However, in practical applications, the amount of data available for modeling and calculation is often limited, and certain training data may be biased due to unknown factors, making it difficult to improve the prediction model through extensive training. To address the above issues, this paper introduces a novel approach for predicting the top oil temperature of transformers, grounded in CEEMD decomposition, ORedRVFL prediction, and error correction:

(1) For the first time, CEEMD is combined with ORedRVFL for oil temperature prediction, addressing the three bottlenecks of traditional methods: noise sensitivity, overfitting to outliers, and blind parameter tuning.
(2) The Huber norm regularization layer is introduced in edRVFL for the first time to suppress the interference of outliers and enhance the model's generalization ability.
(3) Tent chaotic initialization is used instead of random initialization to avoid premature convergence of the DE algorithm and optimize the hyperparameters of the ORedRVFL model, thereby improving the model's prediction accuracy.
(4) A recursive correction mechanism for residual components is established to eliminate cumulative prediction bias, enabling the model to maintain high precision in complex data environments and enhancing its reliability in practical applications.

2. Methods

2.1. Complementary Ensemble Empirical Mode Decomposition

Complementary Ensemble Empirical Mode Decomposition (CEEMD) [32] reduces reconstruction errors and significantly improves the mode mixing phenomenon found in Empirical Mode Decomposition (EMD) [33]. Ma et al. [34] proposed that, when performing signal decomposition, the noise amplitude is usually set to 0.1 or 0.3 times the signal standard deviation, and an ensemble size is generally chosen between 50 and 200. These parameter settings have achieved good results in prediction tasks. Xiong [35] et al. suggested that, when using CEEMD for signal preprocessing, the choice of noise amplitude and ensemble size has a significant impact on model performance. The noise amplitude is typically between 0.05 and 0.2 times the signal standard deviation, and an ensemble size between 50 and 150 is more appropriate. These parameter ranges have been proven effective in wind power prediction. Therefore, we used these parameters as an initial reference. The illustrative representation of the CEEMD decomposition algorithm is illustrated in Figure 1.

First, the original oil temperature data are corrupted by additive positive and negative Gaussian white noise with a magnitude $o(t)$. The detailed procedure is outlined as follows:

$$\begin{cases} o_i^+(t) = o(t) + \varphi_i^+(t) \\ o_i^-(t) = o(t) + \varphi_i^-(t) \end{cases} \tag{1}$$

where $\varphi_i^+(t)$ and $\varphi_i^-(t)$ represent the bipolar white noise components for the i-th instance, respectively. $o_i^+(t)$ and $o_i^-(t)$ represent the oil temperature data after adding positive and negative white noise, respectively.

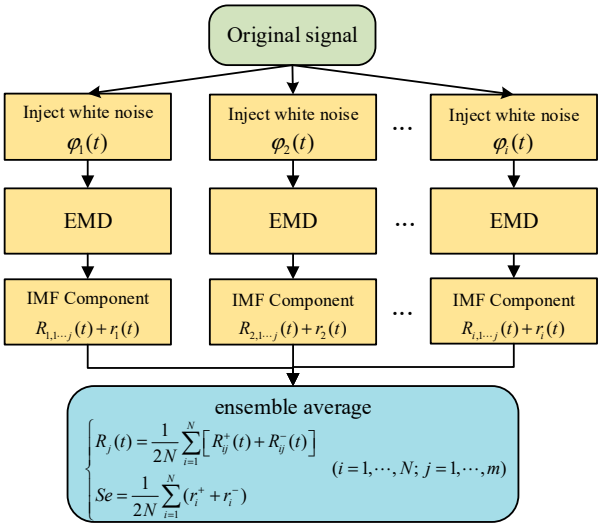


Figure 1. Schematic diagram of CEEMD decomposition algorithm.

Using the EMD algorithm to decompose the signal after adding white noise, the expression formula for the decomposed sequence is as follows:

$$\begin{cases} o_i^+(t) = \sum_{j=1}^{n} R_{ij}^+(t) + r_i^+(t) \\ o_i^-(t) = \sum_{j=1}^{n} R_{ij}^-(t) + r_i^-(t) \end{cases} \tag{2}$$

where $R_{ij}(t)$ represents the j-th Intrinsic Mode Function (IMF) component obtained from the decomposition of the i-th signal, and r_i represents the residue obtained from the dissection of the j-th signal.

Finally, the average value of each IMF component obtained from the decomposition is calculated to obtain the final decomposition result of the CEEMD algorithm. The process unfolds in the following manner:

$$\begin{cases} R_j(t) = \dfrac{1}{2N}\displaystyle\sum_{i=1}^{N}\left[R_{ij}^+(t) + R_{ij}^-(t)\right] \\ Se = \dfrac{1}{2N}\displaystyle\sum_{i=1}^{N}(r_i^+ + r_i^-) \end{cases} \tag{3}$$

where N denotes the number of white noise instances, $R_j(t)$ (for j = 1, . . ., m) represents the j-th component obtained from the final decomposition, and Se is the residue after decomposition.

2.2. Robust Deep Ensemble RVFL

Similar to the ELM model, RVFL is a type of feedforward neural network [35]. The framework structure can be seen in Figure 2. In the RVFL model, the input data from the input layer traverses the nodes in the hidden layer, where it undergoes a nonlinear

mapping, and the concluding result is output through the output function. The RVFL network can be described as in Equation (4):

$$f(x) = \sum_{j=1}^{J} \beta_j g\left(\omega_j^T X + b_j\right) + \sum_{j=J+1}^{J+N} \beta_j x_j \tag{4}$$

where $g(\cdot)$ serves as the activation function, and ω_j represents the weight associated with the j-th hidden node. $X = [x_1, x_2, \ldots, x_N]$ is the input matrix, b_j is the threshold, and β_j is the weight of the output layer.

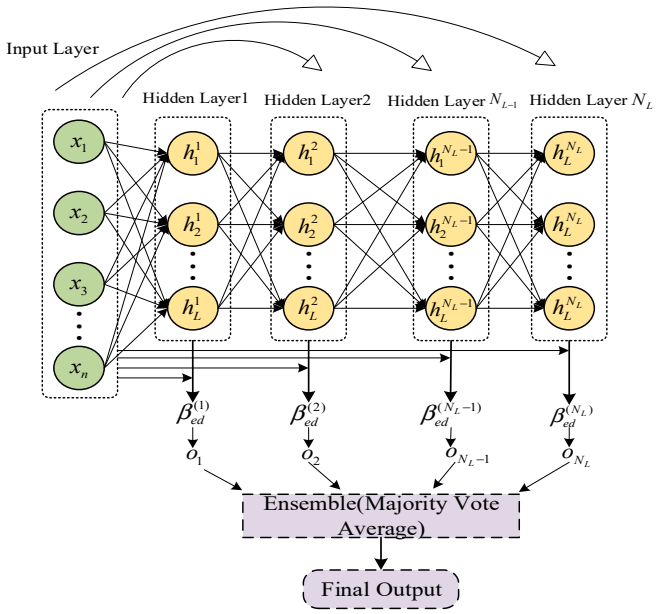


Figure 2. The structure diagram of the RVFL framework.

The edRVFL (Ensemble Deep Random Vector Functional Link) [36] model is an ensemble deep learning framework that combines the advantages of deep learning, ensemble methods, and the IF–THEN properties of fuzzy inference systems (FIS) to generate rich feature representations for training. To tackle the challenge of outliers in the transformer oil temperature data affecting prediction accuracy, this paper proposes a robust edRVFL model. By introducing regularization and balancing the training error and weight relationship through norms, the model reduces the interference of outliers and enhances its robustness. The ORedRVFL network is described as in Equation (5):

$$h_L(x) = \sum_{j=1}^{L} \begin{pmatrix} (H^T H + \dfrac{2}{C\mu I})^{-1} \cdot \\ H^T (y - e_i + \dfrac{\lambda_i}{\mu}) g(\omega_j X + b_j) \end{pmatrix} \tag{5}$$

where L indicates the number of hidden layer nodes, H represents the output matrix of the hidden layer, C is the regularization coefficient, μ is the penalty parameter, e represents the training error, and λ is the extended Lagrange multiplier series.

2.3. Tent Chaotic Differential Evolution Algorithm (TDE)

The Differential Evolution (DE) algorithm optimizes the objective function by simulating the mutation, crossover, and selection operations in biological evolution. The implementation steps of the DE algorithm are as follows:

(1) Initialization

First, initialize the algorithm population. After initialization, each individual is represented as follows:

$$x_{i,G}(i = 1, 2, \ldots, NP) \tag{6}$$

where i denotes the individual index, G represents the generation number, and NP is the population size.

In the differential evolution algorithm, let the bounds of the parameter variables be $gx_j^{(L)} < x_j < x_j^U$. The expression for $x_{ji,0}$ is as follows:

$$x_{ji,0} = rand[0, 1] \times (x_j^{(U)} - x_j^{(L)}) + x_j^L \tag{7}$$

(2) Mutation

Following population initialization, the mutation operation is performed. For each candidate solution $x_{i,G}(i = 1, 2, \ldots, NP)$, the mutation vector is generated as follows:

$$v_{i,G+1} = x_{r_1,G} + F \cdot (x_{r_2,G} - x_{r_3,G}) \tag{8}$$

where F is the mutation operator and is a real constant factor between 1 and 2.

(3) Crossover

The crossover process increases the heterogeneity among solution candidates by executing crossover between the mutant vector and the target vector, thereby avoiding randomness and ensuring that the trial vector is altered. The specific process is as follows:

$$u_{i,G+1} = (u_{1_i,G+1}, u_{2_i,G+1}, \ldots, u_{D_i,G+1}) \tag{9}$$

$$u_{ji,G+1} = \begin{cases} v_{ji,G+1} \; if \; \text{rand}(j) <= \text{CR or } j = \text{rnbr}(i) \\ x_{ji,G+1} \; if \; x \neq 0 \text{ and j} \neq \text{rnbr}(i) \end{cases} \tag{10}$$

where $\text{rand}(j)$ denotes the $j - th$ estimate of the stochastic-number generator, $\text{rnbr}(i)$ represents a stochastically drawn sequence, and CR functions as the crossover mechanism.

(4) Selection

The selection operation in the Differential Evolution algorithm is based on the greedy criterion: the trial vector is compared with the target vector $x_{i,G}$. If the trial vector has a higher fitness, it is selected for the next generation; otherwise, the original target vector is retained.

(5) Handling of Boundary Conditions

If, during the mutation process, a solution is generated outside the feasible domain, the mutant vector is redefined within the feasible solution range. The specific process is as follows:

$$u_{ji,G+1} = rand[0, 1] \cdot (x_j^{(U)} - x_j^{(L)}) + x_j^{(L)} \tag{11}$$

where: $i = 1, 2, \ldots, NP; j = 1, 2, \ldots, D$

The random initialization of the original algorithm is prone to population clustering, which affects the convergence speed. In this paper, the Tent map, which has better randomness and uniformity, is used to initialize the individuals in the algorithm. The specific expression is as follows:

$$r_{j+1} = \begin{cases} \psi r_j & r_j < 0.5 \\ \psi(1 - r_j) & r_j \geq 0.5 \end{cases} \tag{12}$$

$$x_{i,G} = lb + (ub - lb) \cdot r_j \tag{13}$$

where r_j denotes a uniformly distributed random value between the specified range [0, 1], ψ is a chaotic parameter in the interval (0, 2], and lb and ub signify the lower and upper extents of the solution range, respectively.

2.4. Error Correction Model

In practical prediction, the inherent errors of models are often overlooked, but the potential information they contain can be mined and utilized [37]. Therefore, this paper proposes to correct the preliminary prediction results to further improve the prediction accuracy. The procedure for error correction is described as follows:

Step 1: Calculate the error sequence. The error sequence is obtained, where o_i represents the $i-th$ true value spanning from the coolant's chill floor to its thermal ceiling sequence, and p_i represents the $i-th$ predicted value from the oil's coolest skin to its searing core sequence. The formulation for the error sequence e_i is as follows:

$$e_i = o_i - p_i \tag{14}$$

Step 2: Train the ELM error model. After reasonably partitioning the error sequence, it is used as the input for training and predicting the error model to obtain the predicted error value for the $i-th$ sample.

Step 3: Add the $i-th$ transformer oil temperature predicted value p_i and the $i-th$ predicted error value y_i to obtain the $i-th$ error-corrected transformer oil temperature predicted value. Finally, the final predicted sequence of transformer oil temperature $\hat{y}_i$ is obtained as shown in Equation (15):

$$\hat{y}_i = p_i + y_i \tag{15}$$

2.5. Construction of the Transformer Top Oil Temperature Prediction Model

The structure of the hybrid transformer top oil temperature forecasting model unveiled in this paper, namely TDE-CEEMD-ORedRVFL-EC, is presented in Figure 3. The implementation process is described as follows:

Step 1: Use the transformer oil temperature data from December 2023 to February 2024 as the winter prediction dataset and the data from June to August 2024 as the summer dataset. Model each of these datasets separately.

Step 2: Decompose the historical transformer oil temperature data using CEEMD to make the sub-sequences more stationary. Then, set the first 70% of each sub-sequence component as the training set and the last 30% as the testing set.

Step 3: Introduce regularization and norms to improve edRVFL to ORedRVFL, enhancing robustness. Input the data from Step 2 into ORedRVFL to generate preliminary predictions.

Step 4: Calculate the prediction error from Step 3, establish a correction model, adjust the initial predictions, and obtain more accurate final results.

Step 5: To rigorously benchmark the model's performance, set ELM, BP, LSTM, edRVFL, ORedRVFL, and CEEMD-ORedRVFL as control group models and conduct comparisons.

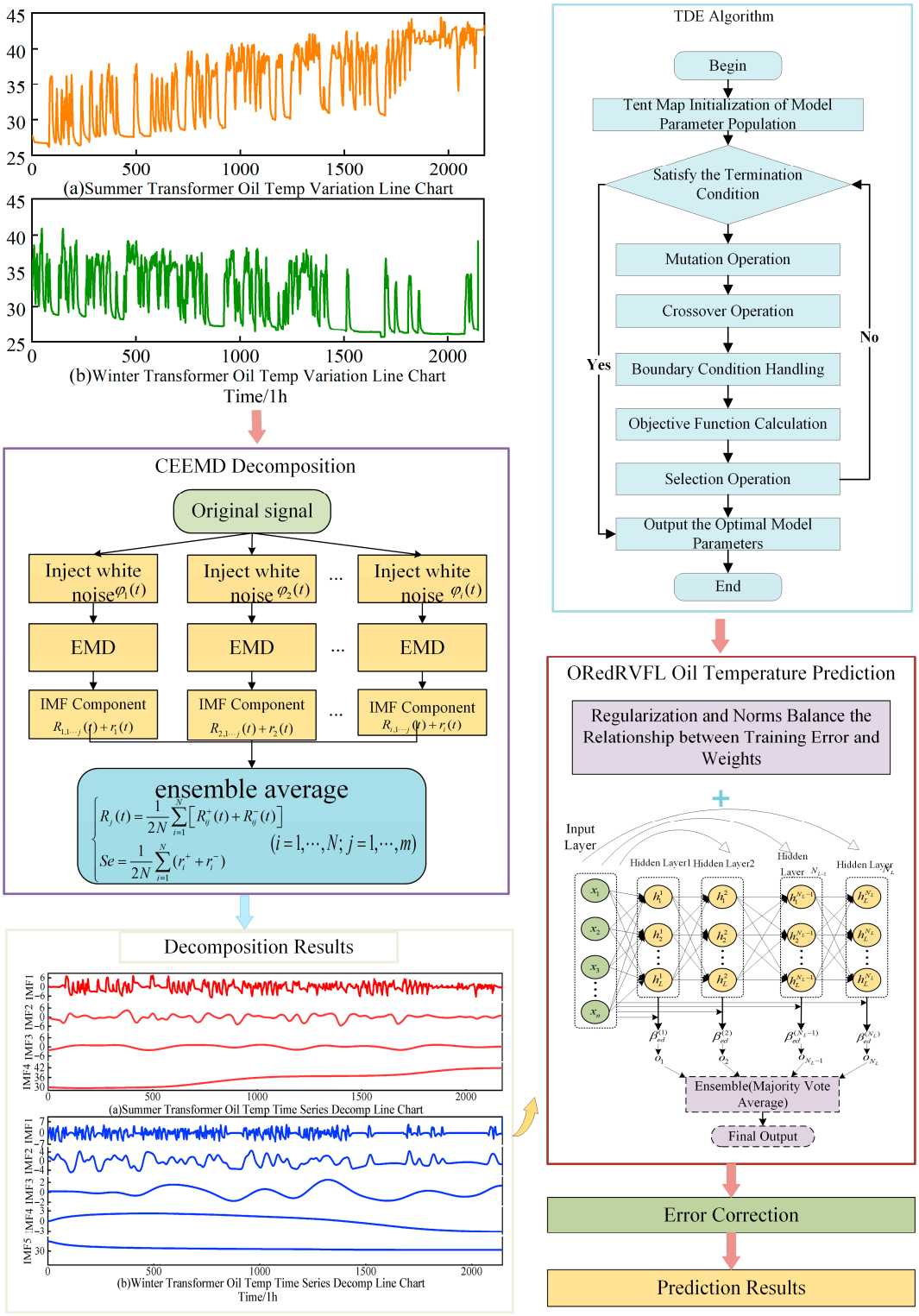


Figure 3. Flowchart of the overall prediction model.

3. Data Preprocessing

3.1. Dataset Introduction

The hourly-level transformer oil temperature data from December 2023 to February 2024 (winter) and June to August 2024 (summer) of the main transformer at a hydropower station in Yunnan, China, were selected to verify the model. The sampling interval was 1 h. There were 2145 data points for the winter dataset and 2173 for the summer dataset.

Figure 4 illustrates the variation in the oil temperature data sequences for the two seasons, and Table 1 provides their maximum, minimum, and average values.

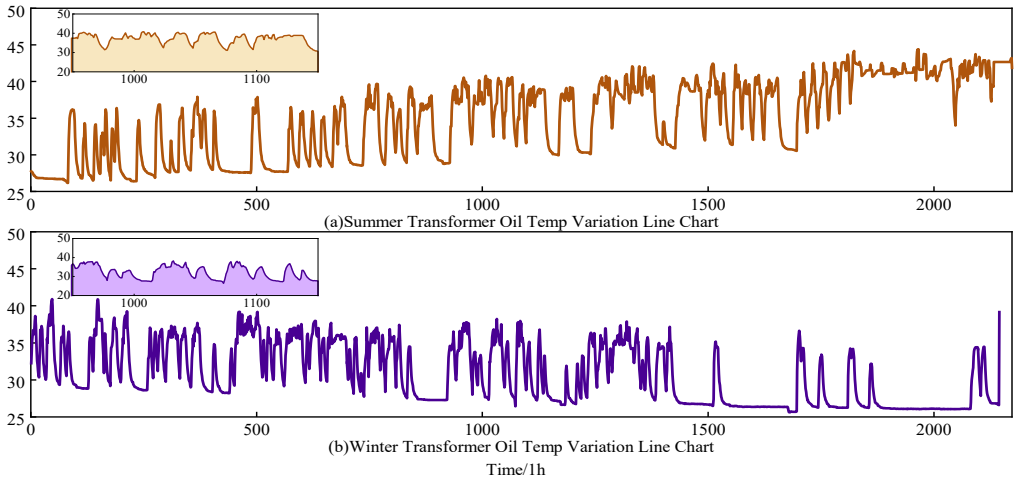


Figure 4. Historical transformer oil temperature data.

Table 1. Transformer oil temperature history data information.

Data Name	Summer	Winter
Maximum Value	44.381	40.922
Minimum Value	26.166	25.631
Average Value	34.9936	30.8591

3.2. Data Decomposition and Partitioning

The decomposition results of the oil temperature data by the CEEMD algorithm are presented in Figure 5, where Figure 5a shows the winter decomposition result and Figure 5b shows the summer decomposition result. Under the decomposition of the CEEMD algorithm, the winter oil temperature data are segmented into four IMF components, and the summer oil temperature data are broken down into five IMF elements. The frequency of the transformer oil temperature data components decomposed by the CEEMD algorithm gradually decreases, the regularity is enhanced, and the trend becomes more stable.

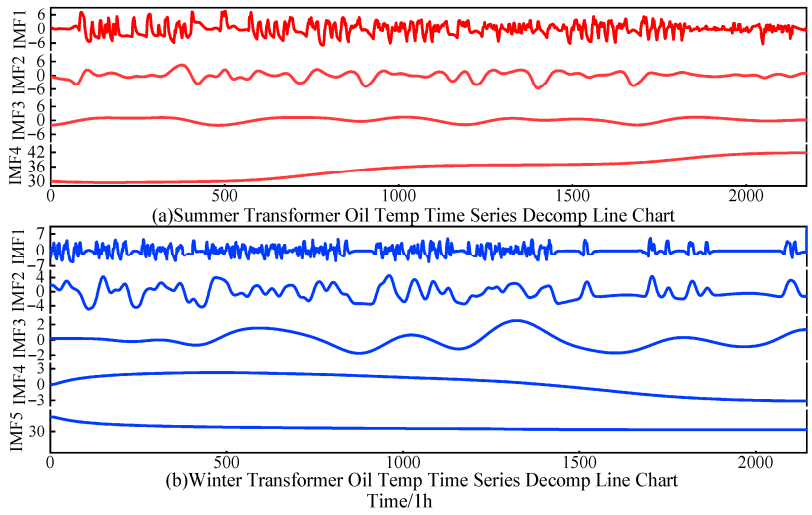


Figure 5. Line graph of transformer oil temperature data components.

3.3. Random Noise Data

To verify the prediction performance and robustness of the ORedRVFL model under random noise and complex data conditions, this paper randomly generated 150 and 152 random noises (with an amplitude of ±5, accounting for 10% of the length of the training dataset) in the winter and summer datasets, respectively, to simulate the interference of transformer oil temperature sensors leading to inaccurate measurements. Figure 6 shows the line charts for the two sets of noise.

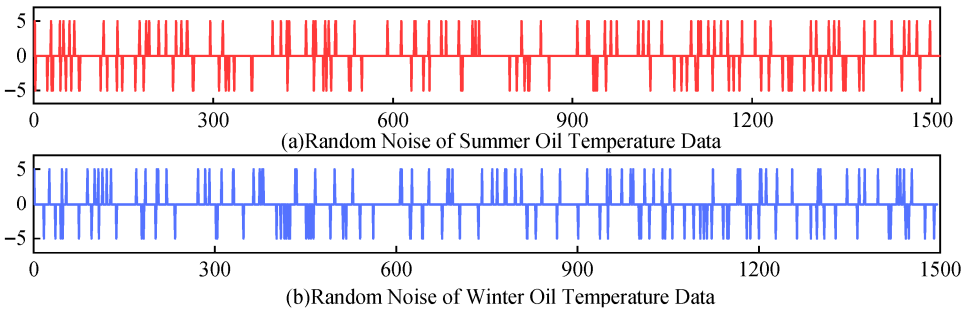


Figure 6. Plot of random noise.

4. Experimental Results and Analysis

4.1. The Performance Comparison of Algorithms

The Tent mapping initialization method is based on the good randomness and uniformity characteristics of the Tent mapping, which can generate a more uniformly distributed initial population, thereby significantly enhancing the algorithm's global search ability and performance. The TDE algorithm has been compared with four other algorithms, namely, Artificial Bee Colony Optimization (ASO), Particle Swarm Optimization (PSO), Harris Hawks Optimization (HHO), and Differential Evolution (DE) on three benchmark functions (F1, F9, and F14). Performance comparison results of the algorithms are given in Table 2, with the best performance results and the corresponding algorithm are indicated in bold. The convergence curves of the five algorithms are given in Figure 7. Experimental results also show that the TDE algorithm with Tent mapping initialization performs excellently on multiple test functions, especially in complex multimodal and high-dimensional optimization problems, where it demonstrates faster convergence speed and higher solution quality.

Table 2. Performance comparison chart for the algorithms.

Benchmark Function	Algorithm	Average Value	Standard Deviation	Optimal Solution
f_1	ASO	7.89×10^{-5}	0.000249401	2.94×10^{-15}
	PSO	0.85220533	0.330388366	0.362945255
	HHO	11566.65089	3885.164694	5835.221608
	DE	3.89×10^{-15}	3.36×10^{-15}	5.90×10^{-16}
	TDE	$\mathbf{6.16} \times 10^{-62}$	$\mathbf{1.53} \times 10^{-61}$	$\mathbf{2.51} \times 10^{-75}$
f_9	ASO	42.21042681	6.548604671	35.82220426
	PSO	56.32590351	14.01800727	41.41953517
	HHO	297.6034322	19.8717627	268.7495013
	DE	5.357551215	5.168711573	1.59×10^{-12}
	TDE	**0**	**0**	**0**
f_{14}	ASO	1.794378737	1.023738232	0.998003838
	PSO	1.494226808	0.843210368	0.998003838
	HHO	1.295816676	0.669811166	0.998003838
	DE	5.598097203	4.530763325	0.998003839
	TDE	**1.097406545**	**0.314338957**	**0.998003838**

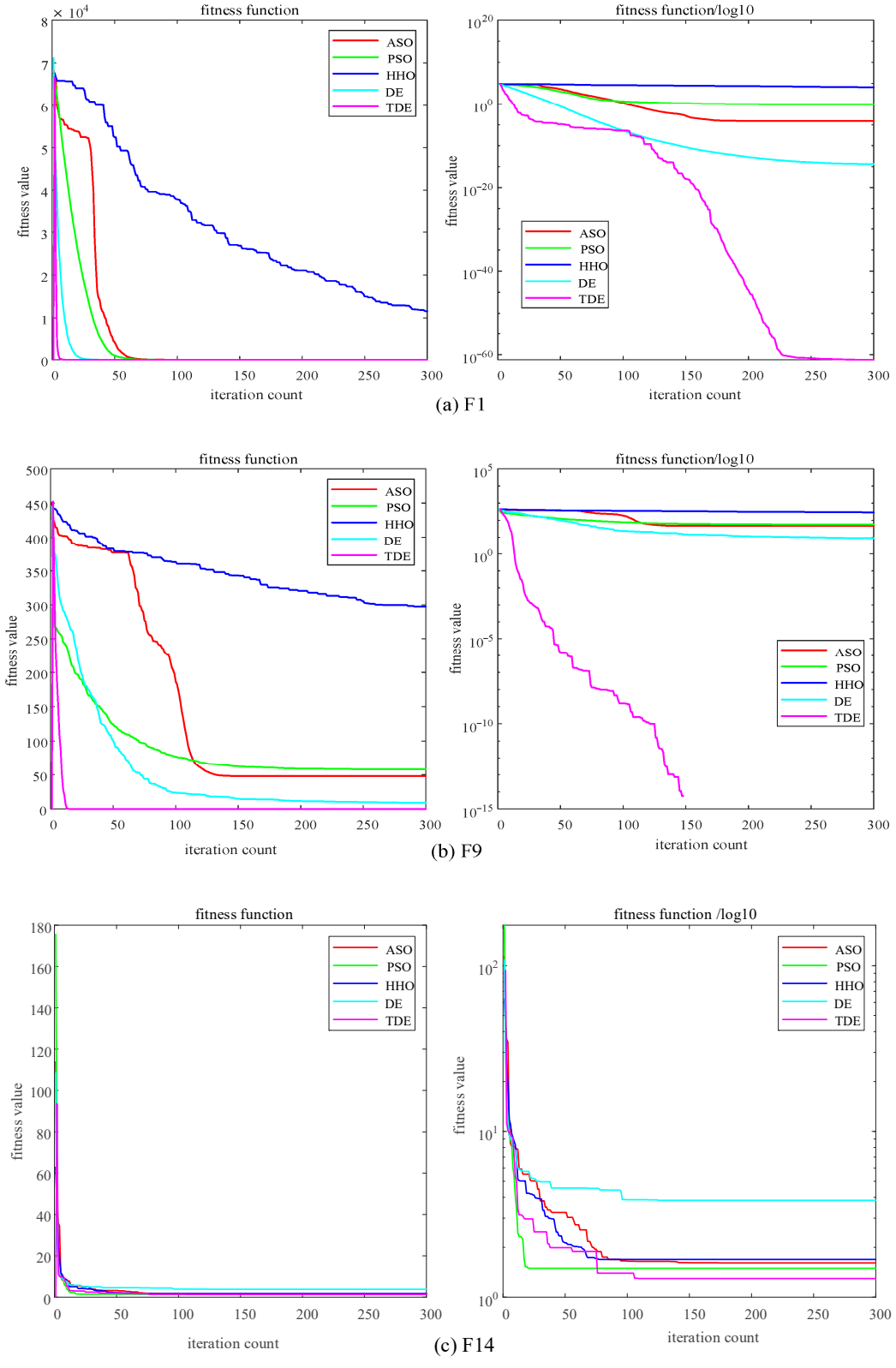


Figure 7. Comparison of Algorithm Convergence Speed.

From Figure 7, it can be seen that the TDE (purple curve) demonstrates the best convergence performance across all test functions. At the end of the iterations, the fitness value of TDE is the lowest, indicating that it has found the optimal solution or a solution close to the optimal one. This suggests that TDE has superior global search capabilities compared to other algorithms. Additionally, the curve of TDE shows good stability during

the iteration process, with no significant fluctuations, indicating that the algorithm is relatively stable during the search process and less likely to get stuck in local optima. The performance comparison chart of the algorithms is shown in Table 2.

4.2. Evaluation Metrics

In the field of data prediction, evaluation metrics serve as the fundamental yardsticks for gauging both model performance and the fidelity of forecasted outcomes. The evaluation metrics used in this paper and their calculation formulas are detailed in Table 3.

Table 3. Evaluation indicators and their formulas.

Assessment Indicators	Formula
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}$
Mean Absolute Error (MAE)	$\frac{1}{n}\sum_{i=1}^{n}\lvert x_i - y_i\rvert$
Mean Absolute Percentage Error (MAPE)	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{x_i - y_i}{x_i}\right\rvert \times 100\%$
Correlation Coefficient (R)	$\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2\sum_{i=1}^{n}(y_i-\overline{y})^2}}$

where x_i captures the actual observed value, $\overline{x}$ represents the average of the ground-truth values, y_i captures the model's forecasts value, $\overline{y}$ represents the average of the predicted values, and n denotes the sample count.

4.3. Experimental Design and Result Analysis

To validate the forecasting capability of the CEEMD-ORedRVFL-EC model, this paper conducted predictions of the real hourly transformer top oil temperature for both summer and winter using MATLAB 2023a, supplied by MathWorks, Inc., with its headquarters located in Natick, MA, USA. and compared the results with those from LSTM, BP, ELM, edRVFL, ORedRVFL, and CEEMD-ORedRVFL as control groups. Table 4 presents each model's individual prediction errors, and Figure 8 shows the fitting line charts of some models.

Table 4. Evaluation indexes of transformer oil temperature prediction results.

Season	Model	RMSE	MAE	R	MAPE
Summer	ELM	0.98499	0.78857	0.96842	0.01959
	BP	0.90876	0.70660	0.97154	0.01764
	LSTM	0.80409	0.58828	0.97539	0.01472
	edRVFL	0.76446	0.46345	0.97497	0.01180
	ORedRVFL	0.75688	0.47819	0.97529	0.01216
	CEEMD-ORedRVFL	0.71868	0.45684	0.97730	0.01167
	CEEMD-ORedRVFL-EC	0.71304	0.45326	0.97762	0.01159
Winter	ELM	0.87662	0.43601	0.93748	0.01489
	BP	0.76791	0.47134	0.95918	0.01688
	LSTM	0.74059	0.30142	0.95684	0.01028
	edRVFL	0.70246	0.19226	0.95971	0.00625
	ORedRVFL	0.69605	0.21843	0.96084	0.00724
	CEEMD-ORedRVFL	0.66732	0.23765	0.96607	0.00796
	CEEMD-ORedRVFL-EC	0.65980	0.21371	0.96619	0.00708

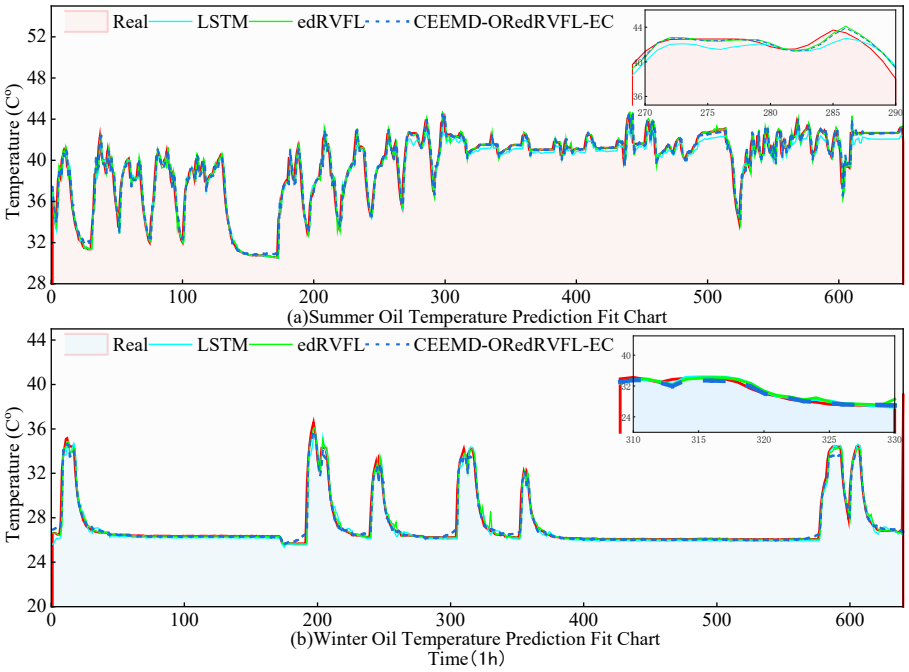


Figure 8. Fitted plot of transformer oil temperature prediction results.

From the metrics in Table 4, the results reveal that the forecast performance of the proposed model is better in winter than in summer. This is due to the higher volatility of electricity usage in summer, which leads to more complex and variable transformer oil temperature data, thereby validating the authenticity of this data. The RMSE results show that the edRVFL model outperforms all single-model control groups in terms of prediction accuracy. The introduction of robust methods in the ORedRVFL model leads to better prediction accuracy than that of the edRVFL model. In both summer and winter predictions, the RMSE of the ORedRVFL model's performance improves by 0.992% and 0.913%, respectively, relative to the edRVFL model, proving that the ORedRVFL model demonstrates an advantage in dealing with complex data. Compared with the ORedRVFL model, the CEEMD-ORedRVFL model exhibits marked improvements in prediction accuracy, with RMSE reductions of 5.047% and 4.128% in winter and summer, respectively. This demonstrates that the decomposition algorithm markedly enhances the prediction performance of the ORedRVFL model. The CEEMD-ORedRVFL-EC model, which incorporates an error-correction model grounded in the CEEMD-ORedRVFL model, further improves the prediction accuracy. Its prediction metrics are improved by 0.785% and 0.501% in summer and winter oil temperature predictions, respectively, compared with the CEEMD-ORedRVFL model. This proves that the application of the error-correction approach can further boost the predictive reliability of the CEEMD-ORedRVFL thermal profile of the transformer oil prediction model. As can be seen from Figure 8, the predictive outcomes of the CEEMD-ORedRVFL-EC model have the highest degree of overlap with the true data curve.

To verify the robustness of the ORedRVFL model, this paper conducted simulation experiments on the summer and winter transformer oil temperature data with added random noise (the random noise information is described in Section 3.3). The experimental process was the same as that without added noise, and the result metrics and fitting results are presented in Table 5 and Figure 9, respectively.

Table 5. Evaluation metrics for oil temperature prediction results with the addition of random noise.

Season	Model	RMSE	MAE	R	MAPE
Summer	ELM	1.23618	0.97025	0.95088	0.02391
	BP	1.19375	0.98887	0.96861	0.02436
	LSTM	0.94845	0.75448	0.96998	0.01883
	edRVFL	0.87346	0.56784	0.96918	0.01434
	ORedRVFL	0.77231	0.50869	0.97470	0.01289
	CEEMD-ORedRVFL	0.77126	0.50902	0.97489	0.01289
	CEEMD-ORedRVFL-EC	0.77079	0.49497	0.97467	0.01255
Winter	ELM	1.10196	0.66029	0.90071	0.02298
	BP	0.93037	0.66356	0.94384	0.02381
	LSTM	0.92479	0.54066	0.93744	0.01898
	edRVFL	0.85924	0.47096	0.94070	0.01629
	ORedRVFL	0.71107	0.24781	0.95968	0.00832
	CEEMD-ORedRVFL	0.70846	0.26364	0.95973	0.00894
	CEEMD-ORedRVFL-EC	0.70462	0.23483	0.96130	0.00784

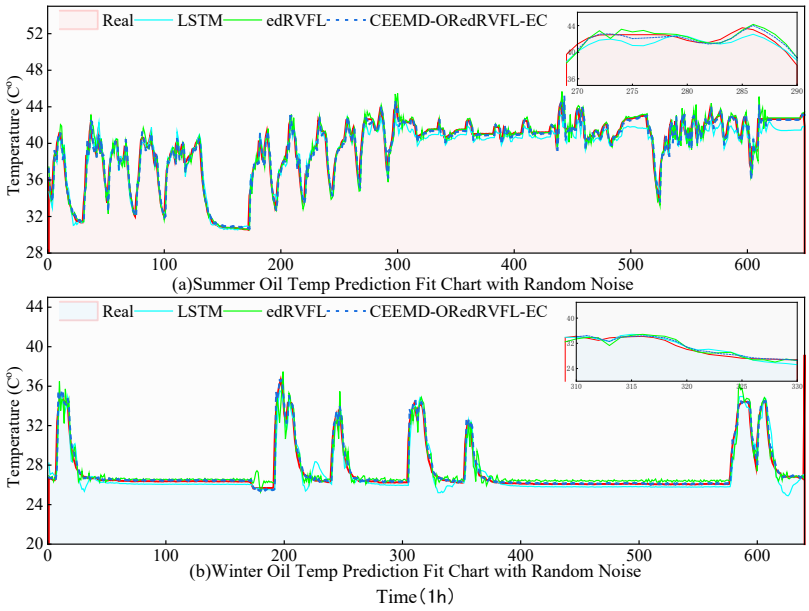


Figure 9. Fitted plot of oil temperature prediction results with the addition of random noise.

Table 4 presents the prediction error results for the summer and winter transformer oil temperature predictions after adding random noise, and Figure 8 shows the fitting line charts of the main comparison models for the summer and winter transformer oil temperature predictions after adding random noise. The metrics in Table 4 indicate that the prediction results of all models deteriorated to varying degrees after the addition of random noise. Specifically, the BP model deteriorated by 31.36% and 21.16% in summer and winter oil temperature predictions, respectively; LSTM deteriorated by 17.95% and 24.87%; edRVFL deteriorated by 14.26% and 22.32%; while ORedRVFL deteriorated by only 2.04% and 2.16%. This demonstrates that the degradation of the ORedRVFL model is significantly lower compared to other models. Thanks to the incorporation of decomposition and error correction, the CEEMD-ORedRVFL-EC model shows improved accuracy over the ORedRVFL model. As can be clearly seen from Figure 8, the prediction result curve of the CEEMD-ORedRVFL-EC model almost overlaps with the true values, while the single-model control groups all exhibit significant deviations. This demonstrates that the

robustness enhancement of the proposed CEEMD-ORedRVFL-EC model in this paper is highly effective.

Figure 10 shows the bar charts of the RMSE metrics for each model in both seasons with and without added noise, where (a) and (b) are for summer, and (c) and (d) are for winter. Figure 10 suggests that the developed model yields the lowest RMSE, with prediction results closer to the actual values. Comparing the figures with and without noise, it is observable that the RMSE of the control group models significantly increases after adding noise, while the proposed model only manifests a marginal increase. This demonstrates that the proposed model is capable of consistently delivering accurate predictions even in the presence of random noise, showing strong robustness and the ability to handle complex oil temperature data.

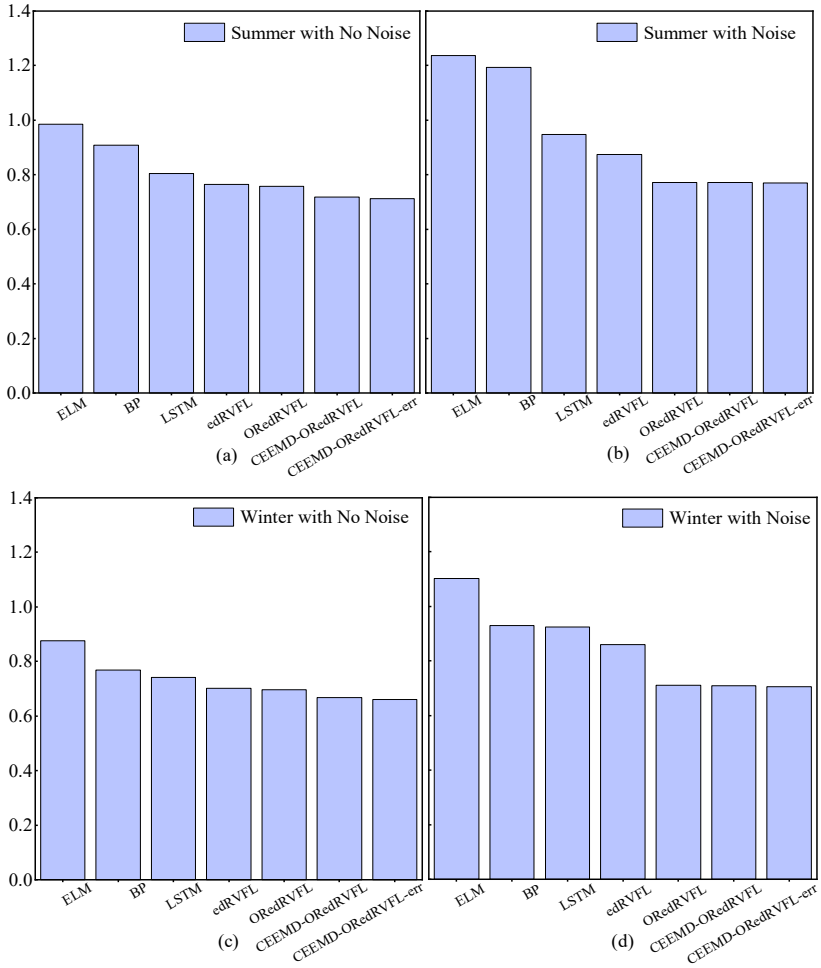


Figure 10. Histogram of RMSE error. (**a**) RMSE for Summer without Noise (**b**) RMSE for Summer with Noise (**c**) RMSE for Winter without Noise (**d**) RMSE for Winter with Noise.

5. Diebold-Mariano Test

The Diebold-Mariano test (DM test) is a non-parametric method for comparing the predictive accuracy of two forecasting models. It assesses whether there is a significant difference in predictive performance by comparing the forecast errors of the two models, with the null hypothesis being that the predictive accuracy of the two models is the same [38]. This test is applicable to a variety of loss functions and in this study, we primarily used the Mean Squared Error (MSE).

After conducting the DM test on the oil temperature dataset with added random noise, comparing our proposed model (CEEMD-ORedRVFL-EC) with several other models, the results are shown in Table 6.

Table 6. Significance Analysis of the Proposed Model CEEMD-ORedRVFL-EC Compared with Other Benchmark Models.

Season	Model	DM
Summer	CEEMD-ORedRVFL	0.5599
	ORedRVFL	2.9853
	edRVFL	4.8496
	LSTM	7.6922
	BP	13.736
	ELM	11.8441
Winter	CEEMD-ORedRVFL	1.3716
	ORedRVFL	2.2922
	edRVFL	2.5265
	LSTM	2.7642
	BP	6.254
	ELM	6.2713

In both summer and winter data, the DM values of the proposed model compared with each benchmark model are greater than 0, intuitively demonstrating its superiority in predictive performance. In the summer data, the DM values of the proposed model compared with ORedRVFL, edRVFL, LSTM, BP, ELM, and other models are all greater than 2.576, proving that its prediction accuracy is significantly higher than these models at the 5% significance level. In the winter data, the DM values of the proposed model compared with most benchmark models are greater than 1.96, indicating that its prediction accuracy is significantly better than these models at the 1% significance level. Overall, the proposed model has higher prediction accuracy than other models, proving that it has significant and outstanding performance advantages in the oil temperature prediction task and is the best-performing model currently.

Overall, the proposed model has significantly higher prediction accuracy than other models in the vast majority of cases, indisputably proving that the proposed model has significant and outstanding performance advantages in the oil temperature prediction task and is the best-performing model currently.

6. Conclusions

Aiming at the issues of low real-time performance and weak anti-interference capability in transformer oil temperature measurement and prediction, this paper presents a transformer-oil temperature forecasting approach rooted in the CEEMD-ORedRVFL-EC model. Using the main transformer's historical oil-temperature data at a hydropower station in Yunnan, China, through experimental research conducted as a case study, the following insights are derived:

(1) The CEEMD algorithm decomposes the oil temperature sequence into multiple subsequences with different frequencies, significantly enhancing the regularity and predictability of the data. Experiments show that, after the introduction of the decomposition algorithm, the model's prediction accuracy for winter and summer oil temperatures increased by 5.05% and 4.13%, respectively.

(2) The introduction of regularization and norm improvements to edRVFL resulted in the ORedRVFL model, which exhibited significantly reduced degradation when subjected to random noise. This validates its robustness and anti-interference capability.

(3) The error correction mechanism further improved prediction accuracy, enabling the model to more accurately reflect the actual changes in transformer oil temperature.
(4) Based on the experimental results, the proposed model's predictive accuracy surpasses that of other control group models, achieving more accurate transformer oil temperature prediction.

Author Contributions: Y.X.: software, visualization, writing—original draft; H.L.: data curation, methodology, software; X.M.: supervision, writing—review and editing; J.C.: visualization, writing—review and editing; X.Z.: visualization, writing—review and editing; T.P.: writing—review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by China Yangtze Power Co., Ltd. (CYPC) Sponsored Project (Z522302017).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Yan Xu and Haohao Li were employed by China Yangtze Power Co., Ltd. Authors Xianyu Meng, Jialei Chen, Xinyu Zhang and Tian Peng were employed by Nanjing Nanrui Jibao Engineering Technology Co. The China Yangtze Power Co. and Nanrui Jibao Engineering Technology Co. had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Xie, S.; Dai, D.; Dong, W.; He, Y. Comprehensive Condition Evaluation of Distribution Transformer Considering Internal Operation, External Environment, and Load Operation for Business Expansion. *Energies* **2025**, *18*, 2456. [CrossRef]
2. Chen, J.; Tao, Z.; Gao, D.; Zhang, Y. A transformer state evaluation method based on multi-source information fusion. In Proceedings of the Journal of Physics: Conference Series, Shenyang, China, 28–30 March 2025; Volume 3043, p. 012142.
3. Yang, M.; Han, C.; Zhang, W.; Fang, G.; Jia, Y. A short-term power prediction method based on numerical weather prediction correction and the fusion of adaptive spatiotemporal graph feature information for wind farm cluster. *Expert Syst. Appl.* **2025**, *274*, 126979. [CrossRef]
4. Yang, M.; Xu, C.; Bai, Y.; Ma, M.; Su, X. Investigating black-box model for wind power forecasting using local interpretable model-agnostic explanations algorithm. *CSEE J. Power Energy Syst.* **2023**, *11*, 227–242.
5. Chen, S.; Wang, Y.; Yang, M.; Xiao, X.; Gomis-Bellmunt, O. Forced oscillation in hybrid system with grid-following and grid-forming inverters. *Energy* **2025**, *319*, 134915. [CrossRef]
6. Wang, Y.; Chen, S.; Yang, M.; Liao, P.; Xiao, X.; Xie, X.; Li, Y. Low-frequency oscillation in power grids with virtual synchronous generators: A comprehensive review. *Renew. Sustain. Energy Rev.* **2025**, *207*, 114921. [CrossRef]
7. Li, N.; Zhang, C.; Liu, Y.; Zhuo, C.; Liu, M.; Yang, J.; Zhang, Y. Single-degree-of-freedom hybrid modulation strategy and light-load efficiency optimization for dual-active-bridge converter. *IEEE J. Emerg. Sel. Top. Power Electron.* **2024**, *12*, 3936–3947. [CrossRef]
8. Grabko, V.; Tkachenko, S.; Palaniuk, O. Determination of temperature distribution on windings of oil transformer based on the laws of heat transfer. *ScienceRise* **2021**, *5*, 3–13. [CrossRef]
9. Zhang, P.; Hu, K.; Yang, Y.; Yi, G.; Zhang, X.; Peng, R.; Liu, J. Research on Prediction of Dissolved Gas Concentration in a Transformer Based on Dempster–Shafer Evidence Theory-Optimized Ensemble Learning. *Electronics* **2025**, *14*, 1266. [CrossRef]
10. Torres-Bermeo, P.; López-Eugenio, K.; Del-Valle-Soto, C.; Palacios-Navarro, G.; Varela-Aldás, J. Sizing and Characterization of Load Curves of Distribution Transformers Using Clustering and Predictive Machine Learning Models. *Energies* **2025**, *18*, 1832. [CrossRef]
11. Yang, L.; Chen, L.; Zhang, F.; Ma, S.; Zhang, Y.; Yang, S. A Transformer Oil Temperature Prediction Method Based on Data-Driven and Multi-Model Fusion. *Processes* **2025**, *13*, 302. [CrossRef]
12. Monteiro, C.S.; Rodrigues, A.V.; Viveiros, D.; Linhares, C.; Mendes, H.; Silva, S.O.; Marques, P.V.; Tavares, S.M.; Frazão, O. Optical fiber sensors for structural monitoring in power transformers. *Sensors* **2021**, *21*, 6127. [CrossRef]
13. Liu, J.; Hou, Z.; Liu, B.; Zhou, X. Mathematical and Machine Learning Innovations for Power Systems: Predicting Transformer Oil Temperature with Beluga Whale Optimization-Based Hybrid Neural Networks. *Mathematics* **2025**, *13*, 1785. [CrossRef]
14. Lin, L.; Qiang, C.; Zhang, H.; Chen, Q.; An, Z.; Xu, W. Review of Studies on the Hot Spot Temperature of Oil-Immersed Transformers. *Energies* **2024**, *18*, 74. [CrossRef]
15. Zhang, N.; Zhao, G.; Zou, L.; Wang, S.; Ning, S. Monitoring of Transformer Hotspot Temperature Using Support Vector Regression Combined with Wireless Mesh Networks. *Energies* **2024**, *17*, 6266. [CrossRef]

16. Zhang, S.; Zhou, H. Transformer Fault Diagnosis Based on Multi-Strategy Enhanced Dung Beetle Algorithm and Optimized SVM. *Energies* **2024**, *17*, 6296. [CrossRef]
17. Ma, C.; Zhang, C.; Yao, J.; Zhang, X.; Nazir, M.S.; Peng, T. Enhancement of wind speed forecasting using optimized decomposition technique, entropy-based reconstruction, and evolutionary PatchTST. *Energy Convers. Manag.* **2025**, *333*, 119819. [CrossRef]
18. Sun, H.; Zhai, W.; Wang, Y.; Yin, L.; Zhou, F. Privileged information-driven random network based non-iterative integration model for building energy consumption prediction. *Appl. Soft Comput.* **2021**, *108*, 107438. [CrossRef]
19. Han, M.; Fan, M.; Zhao, X.; Ye, L. Knowledge-based hyper-parameter adaptation of multi-stage differential evolution by deep reinforcement learning. *Neurocomputing* **2025**, *648*, 130633. [CrossRef]
20. Hu, M.; Chion, J.H.; Suganthan, P.N.; Katuwal, R.K. Ensemble deep random vector functional link neural network for regression. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *53*, 2604–2615. [CrossRef]
21. Zou, D.; Xu, H.; Quan, H.; Yin, J.; Peng, Q.; Wang, S.; Dai, W.; Hong, Z. Top-Oil Temperature Prediction of Power Transformer Based on Long Short-Term Memory Neural Network with Self-Attention Mechanism Optimized by Improved Whale Optimization Algorithm. *Symmetry* **2024**, *16*, 1382. [CrossRef]
22. Qi, J.; Pang, Z.; Liu, Z.; Du, Y. Method of transformer top oil temperature forecasting based on grey-autoregressive differential moving average model. In Proceedings of the International Conference on Computer Network Security and Software Engineering (CNSSE 2023), Sanya, China, 10–12 February 2023; Volume 12714, pp. 268–273.
23. Gunda, S.K.; Dhanikonda, V.S.S.S.S. Discrimination of transformer inrush currents and internal fault currents using extended kalman filter algorithm (Ekf). *Energies* **2021**, *14*, 6020. [CrossRef]
24. Oliveira, M.; Medeiros, L.; Kaminski, A., Jr.; Falcão, C.; Beltrame, R.; Bender, V.; Marchesan, T.; Marin, M. Thermal-hydraulic model for temperature prediction on oil-directed power transformers. *Int. J. Electr. Power Energy Syst.* **2023**, *151*, 109133. [CrossRef]
25. Ghnatios, C.; Kestelyn, X.; Denis, G.; Champaney, V.; Chinesta, F. Learning data-driven stable corrections of dynamical systems—Application to the simulation of the top-oil temperature evolution of a power transformer. *Energies* **2023**, *16*, 5790. [CrossRef]
26. Juarez-Balderas, E.A.; Medina-Marin, J.; Olivares-Galvan, J.C.; Hernandez-Romero, N.; Seck-Tuoh-Mora, J.C.; Rodriguez-Aguilar, A. Hot-spot temperature forecasting of the instrument transformer using an artificial neural network. *IEEE Access* **2020**, *8*, 164392–164406. [CrossRef]
27. Jiang, H.; Hu, W.; Xiao, L.; Dong, Y. A decomposition ensemble based deep learning approach for crude oil price forecasting. *Resour. Policy* **2022**, *78*, 102855. [CrossRef]
28. Chen, H.; Huang, H.; Zheng, Y.; Yang, B. A load forecasting approach for integrated energy systems based on aggregation hybrid modal decomposition and combined model. *Appl. Energy* **2024**, *375*, 124166. [CrossRef]
29. Lyu, Z.; Wan, Z.; Bian, Z.; Liu, Y.; Zhao, W. Integrated Digital Twins System for Oil Temperature Prediction of Power Transformer Based On Internet of Things. *IEEE Internet Things J.* **2025**, *12*, 13746–13756. [CrossRef]
30. Feng, D.; Wang, Z.; Jarman, P. Evaluation of power transformers' effective hot-spot factors by thermal modeling of scrapped units. *IEEE Trans. Power Deliv.* **2014**, *29*, 2077–2085. [CrossRef]
31. Zhang, X.; Zhang, C.; He, R.; Ma, C.; Yao, J.; Nazir, M.S.; Peng, T. A Pyramidal Attention-based Transformer model Based on Improved Differential Innovation Search Algorithm and Feature Extraction for Solar Radiation Prediction Considering Relevant Factors. *Renew. Energy* **2025**, *253*, 123666. [CrossRef]
32. Yeh, J.-R.; Shieh, J.-S.; Huang, N.E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Adv. Adapt. Data Anal.* **2010**, *2*, 135–156. [CrossRef]
33. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [CrossRef]
34. Ma, H.; Zhang, C.; Peng, T.; Nazir, M.S.; Li, Y. An integrated framework of gated recurrent unit based on improved sine cosine algorithm for photovoltaic power forecasting. *Energy* **2022**, *256*, 124650. [CrossRef]
35. Xiong, J.; Peng, T.; Tao, Z.; Zhang, C.; Song, S.; Nazir, M.S. A dual-scale deep learning model based on ELM-BiLSTM and improved reptile search algorithm for wind power prediction. *Energy* **2023**, *266*, 126419. [CrossRef]
36. Huang, C. International joint conference on neural networks. In Proceedings of the IJCNN 2010, Barcelona, Spain, 18–23 July 2010.
37. Malik, A.K.; Gao, R.; Ganaie, M.; Tanveer, M.; Suganthan, P.N. Random vector functional link network: Recent developments, applications, and future directions. *Appl. Soft Comput.* **2023**, *143*, 110377. [CrossRef]
38. Hua, L.; Zhang, C.; Peng, T.; Ji, C.; Nazir, M.S. Integrated framework of extreme learning machine (ELM) based on improved atom search optimization for short-term wind speed prediction. *Energy Convers. Manag.* **2022**, *252*, 115102. [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.