

Article

A Combined Model Based on EOBL-CSSA-LSSVM for Power Load Forecasting

Xinheng Wang ¹, Xiaojin Gao ², Zuoxun Wang ^{1,*}, Chunrui Ma ¹ and Zengxu Song ¹

¹ School of Electrical Engineering and Automation, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China; 17862976759@163.com (X.W.); chunruima@126.com (C.M.); songzengxu2021@163.com (Z.S.)

² Science and Technology Service Platform of Shandong Academy of Sciences, Jinan 250000, China; gaoxj@sdas.org

* Correspondence: wangzuoxun@qlu.edu.cn

Abstract: Inaccurate electricity load forecasting can lead to the power sector gaining asymmetric information in the supply and demand relationship. This asymmetric information can lead to incorrect production or generation plans for the power sector. In order to improve the accuracy of load forecasting, a combined power load forecasting model based on machine learning algorithms, swarm intelligence optimization algorithms, and data pre-processing is proposed. Firstly, the original signal is pre-processed by the VMD–singular spectrum analysis data pre-processing method. Secondly, the noise-reduced signals are predicted using the Elman prediction model optimized by the sparrow search algorithm, the ELM prediction model optimized by the chaotic adaptive whale algorithm (CAWOA-ELM), and the LSSVM prediction model optimized by the chaotic sparrow search algorithm based on elite opposition-based learning (EOBL-CSSA-LSSVM) for electricity load data, respectively. Finally, the weighting coefficients of the three prediction models are calculated using the simulated annealing algorithm and weighted to obtain the prediction results. Comparative simulation experiments show that the VMD–singular spectrum analysis method and two improved intelligent optimization algorithms proposed in this paper can effectively improve the prediction accuracy. Additionally, the combined forecasting model proposed in this paper has extremely high forecasting accuracy, which can help the power sector to develop a reasonable production plan and power generation plans.

Keywords: neural network; sparrow search algorithm; whale optimization algorithm; power load forecast; machine learning



Citation: Wang, X.; Gao, X.; Wang, Z.; Ma, C.; Song, Z. A Combined Model Based on EOBL-CSSA-LSSVM for Power Load Forecasting. *Symmetry* **2021**, *13*, 1579. <https://doi.org/10.3390/sym13091579>

Academic Editor: Christos Volos

Received: 6 August 2021

Accepted: 25 August 2021

Published: 27 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

With the improvement in economic and social development, high-quality electric energy supply provides an important guarantee for the efficient and stable development of the whole country. The accuracy of electricity load forecasting is directly related to the economic efficiency and reliability of each energy supply sector. Many important operational decisions such as generation plans, fuel procurement plans, maintenance plans, and energy trading plans are based on electricity load forecasting [1–4].

The intrinsic properties of the electrical load make it fundamentally different from other commodities. This is due to the non-storable nature of the electricity load, which is also influenced by the dynamic balance between supply and demand and the reliability of the intelligent transmission network [5]. Accurate power load forecasting can effectively optimize the allocation of resources in intelligent distribution networks and power systems. For the supply sector of the power system, accurate load forecasting enables rational control of the capacity of generating units and rational dispatch of generating capacity, thus reducing energy wastage and costs. For the management of the power system, mastering the changes in the power load at any given moment allows them to control the power

market information in advance and thus obtain higher economic benefits. In practice, the power sector often does not have access to symmetrical information. This is because inaccurate load forecasts provide the wrong information to the electricity sector, resulting in an asymmetry between the sector and the electricity consumers. The use of asymmetrical load information in production planning not only results in economic losses but also in wasted resources.

In summary, the power sector must ensure a dynamic balance between power demand and supply, while minimizing the waste of resources and economic losses [6]. Therefore, accurate electricity load forecasting is key to achieving this target.

For the above reasons, a novel combinatorial model for electricity load forecasting is proposed by this paper. The model combines a data pre-processing method (VMD–singular spectrum analysis noise reduction method), two novel combinatorial intelligent optimization algorithms (the CAWOA and EOBL-CSSA algorithm), and three independent and efficient forecasting models (ELMAN neural network, LSSVM model, ELM neural network). Firstly, the original signal is pre-processed by the VMD–singular spectrum analysis data pre-processing method. Secondly, the noise-reduced signals are predicted using the Elman prediction model optimized by the SSA algorithm, the ELM prediction model optimized by the CAWOA-ELM algorithm, and the LSSVM prediction model optimized by the EOBL-CSSA-LSSVM algorithm for electricity load data, respectively. Finally, the weighting coefficients of the three prediction models are calculated using the simulated annealing (SA) algorithm and weighted to obtain the prediction results.

2. Literature Review

Due to the rapid development of smart distribution networks, new grid planning and strategy formulations need the support of high-precision power load forecasting. Many researchers have made unremitting efforts to improve the accuracy of power load forecasting in the environment of smart distribution networks. The different forecasting principles can be divided into traditional forecasting methods based on statistics [7,8] and intelligent forecasting methods based on machine learning algorithms [9,10]. Traditional methods of forecasting electrical loads have the advantage of low computational effort and high prediction accuracy for simple linear cases. However, such methods are not sufficient to deal with complex nonlinear load time series and are difficult to meet the needs of modern forecasting. With the development of artificial intelligence (AI) technology [11–13], machine learning is widely used in the field of power load forecasting due to its powerful non-linear processing capability [14].

The powerful adaptive and learning capabilities of machine learning algorithms are well suited for processing non-linear time series. However, different machine learning algorithms have their own limitations, which limit their further development. Currently, swarm intelligence optimization algorithms are widely used in neural networks, machine learning, and other intelligent algorithms for structural optimization [15]. Common swarm intelligence optimization algorithms include: ant colony optimization algorithms (ACO) [16,17] and particle swarm optimization algorithms (PSO) [18,19]. In addition, such novel swarm intelligence optimization algorithms as the whale optimization algorithm (WOA) [15,20] and sparrow search algorithm (SSA) [21] also show amazing potential in processing structure optimization.

Paper [22] has shown that using raw data directly for prediction would significantly confound the results. In order to minimize the forecast error of electrical loads, scholars have carried out many studies on combined forecasting models based on data pre-processing and electrical load forecasting models. Typically, these scholars use wavelet transform (WT) [23], empirical modal decomposition (EMD) [24], variational modal decomposition (VMD) [25], and singular spectrum analysis [26] for noise reduction.

Electricity load is a complex time series with non-linear, highly random fluctuations and time-varying characteristics. The modern requirement for accurate forecasting is difficult to achieve due to the shortcomings of a single forecasting method. Therefore,

the idea of combined forecasting is proposed to compensate for the shortcomings of single forecasting methods to improve the overall forecasting accuracy of the model [27].

Jinliang Zhang and his colleagues [28] proposed a hybrid forecasting model based on improved empirical modal decomposition (IEMD), autoregressive integrated moving average (ARIMA), and wavelet neural network (WNN) based on fruit fly optimization algorithm (FOA) optimization. Jinliang Zhang used the FOA algorithm to optimize the network parameters in the WNN algorithm. This approach can effectively improve the shortcomings of the WNN algorithm in terms of optimization capability. The optimized model is also combined with traditional prediction methods to improve the overall prediction accuracy. Simulation experiments demonstrate the effectiveness of this approach.

Heng Shi and his colleagues [29] proposed a method for household load prediction using deep learning and a new deep recurrent neural network based on clustering. The machine learning algorithm batches the load into the input pool to avoid overfitting. Simulation experiments also demonstrate that the model has a higher prediction accuracy than traditional prediction models.

Kun Xie and his colleagues [30] constructed a hybrid prediction model by optimizing the Elman neural network (ENN) with the PSO algorithm. The PSO algorithm was used to search for the optimal learning rate of the ENN to improve the prediction accuracy of the prediction model. Kun XI also analyzed the effect of neural network parameters on the performance of the network in detail and demonstrated that the optimization of network parameters can effectively improve the performance of the model. Finally, the effectiveness of the method was verified through simulation comparison experiments with the generalized regression neural network (GRNN) and traditional back propagation neural network (BPNN).

Zhihao Shang and his colleagues [31] used the LSSVM, ELM, and GRNN models to construct a combined prediction model. The weight coefficients of the three models were calculated using the WOA algorithm. The model was compared with the traditional BP and ARIMA models by real data. The experimental results verify that the model has excellent prediction accuracy. This is also good evidence that the use of intelligent algorithms to optimize the weighting coefficients of different models can effectively improve the prediction accuracy of the models.

Lizhen Wu and his colleagues [32] proposed a GRU-CNN neural network electricity load forecasting model based on gated recurrent units (GRU) and convolutional neural networks (CNN). Lizhen Wu used the GRU module to specifically process the feature vectors of time series data and the CNN module to process other high-dimensional data feature vectors. The effective combination of the two modules was used to improve the model prediction accuracy. Finally, the prediction performance of the model was verified by comparison experiments with BPNN, GRU, and CNN models.

Zuoxun Wang and his colleagues [33] optimized the ELM neural network with the SSA algorithm improved by firefly perturbation and chaos strategies to construct an electricity load forecasting model. The powerful optimization-seeking capability of the FA-CSSA-ELM algorithm was exploited to optimize two parameters in the ELM algorithm. Finally, a comparison experiment with other competing models was used to verify that the model can improve the prediction accuracy well.

Hairui Zhang and his colleagues [34] proposed a combined prediction model combining VMD, the Jordan neural network, the echo state network and the LSSVM model in order to overcome the shortcomings of the single prediction model. The PSO algorithm was used to optimize the relevant parameters of the neural network. This combined prediction model can make full use of the advantages of different models. Finally, the weight coefficients of the different models were calculated by the SA algorithm to obtain the final prediction results. The simulation experiments of power load forecasting also demonstrated that this model can improve the prediction accuracy of the forecasting model.

In summary, we found that the prediction performance of the combined prediction model is better than that of the single model. Additionally, we believe that there is

still a huge potential for improvement in the selection and improvement of intelligent algorithms and data pre-processing for power load forecasting models, even though several of the above forecasting models have high forecasting accuracy and performance.

3. Materials and Methods

This chapter introduces VMD, singular spectrum analysis, noise reduction methods based on VMD–singular spectrum analysis, Elman neural network model, ELM neural network model, LSSVM model, WOA algorithm, CAWOA algorithm, SSA algorithm, EOBL-CSSA-LSSVM algorithm, and SA algorithm.

3.1. Variational Mode Decomposition

VMD is an adaptive solver and a completely non-recursive method for modal variation and signal processing. VMD can overcome the problems of endpoint effects and modal component confounding that exist in EMD methods. It has been demonstrated in the literature that VMD has a good decomposition effect in dealing with non-linear, non-stationary, and highly complex time series. Additionally, the specific algorithm derivation is shown in the paper [35].

3.2. Singular Spectrum Analysis

Singular spectrum analysis was first applied to oceanographic research in 1978 [36]. Nowadays, singular spectrum analysis is a common method for studying non-linear time series. The basic idea of singular spectrum analysis is to calculate its corresponding Hankel trajectory matrix H from a one-dimensional time series $\{h_i, i = 1, \dots, N\}$, as shown in Equation (1).

$$H = \begin{bmatrix} h_1 & h_2 & \cdots & h_k \\ h_2 & h_3 & \cdots & h_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_L & h_{L+1} & \cdots & h_N \end{bmatrix}, \quad (1)$$

where H is the trajectory matrix; L is the sliding window parameter and $1 < L < N$; K is defined as $N - L + 1$; and the eigenvalues and eigenvectors of their corresponding eigenvectors are combined to reconstruct the new time series.

3.3. VMD–Specular Spectral Analysis Noise Reduction Method

In this paper, a noise reduction approach combining VMD and singular spectrum analysis is proposed. First, the original data are adaptively decomposed and reconstructed by introducing kurtosis calculation to make the variational mode decomposition. This ensures that the reconstructed signal is close to the original signal while removing the high-frequency noise present in it. Then, the reconstructed data are second filtered by an adaptive singular spectrum analysis filter to remove the relatively low-frequency residual noise. Additionally, the kurtosis value $K(x)$ is calculated as shown in Equation (2).

$$K(x) = \frac{E[(x - \eta)^4]}{(E[(x - \eta)^2])^2}, \quad (2)$$

where E is the expectation and η is the mean value of the series x . The calculated kurtosis value $K(x)$ is then compared with the kurtosis threshold set in this paper to select the IMF components that need to be reconstructed. Finally, the signal to be reconstructed is noise-reduced by singular spectrum analysis.

The detailed steps of the VMD–singular spectrum analysis of noise reduction method can be summarized as follows:

1. Define the relevant parameters in VMD: the number of modes K and the penalty factor ∂ ;

2. Input time series $f(t)$. The time series $f(t)$ is decomposed into K components of IMF by using VMD, and the decomposed components are denoted as $F = \{u_1, \dots, u_k, \dots, u_K\}$;
3. The kurtosis values of the K components of IMF obtained from the VMD decomposition are calculated according to Equation (2) and noted as $Ku = \{Ku_1, \dots, Ku_k, \dots, Ku_K\}$;
4. Calculate the kurtosis value for each IMF component and search for IMF components greater than the kurtosis threshold;
5. The IMF components obtained in step four are linearly reconstructed to obtain the VMD noise-reduced time series $f(t)'$;
6. For the VMD processed signal $f(t)'$, choose a suitable window length parameter L to lag-arrange the original time series to construct the Hankel matrix $H_{L \times K}$;
7. Singular spectrum analysis method is used to denoise Hankel matrix $H_{L \times K}$, and $H = U \Sigma V^T$ is obtained. U and V represent the associated left and right singular spectral vectors, respectively. Firstly, the covariance matrix $S = HH^T$ of the Hankel matrix H is calculated, and then the covariance matrix S is decomposed into eigenvalues to obtain $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ and the corresponding eigenvector U_1, U_2, \dots, U_L . At this point, $U = [U_1, U_2, \dots, U_L]$ and $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_L} \geq 0$ are singular spectra of the original sequence, which can be expressed as $H = \sum_{m=1}^L \sqrt{\lambda_m} U_m V_m^T$, $V_m = \frac{H^T U_m}{\sqrt{\lambda_m}}, m = 1, \dots, L$, where λ_i is the corresponding eigenvalue and U_i is a time-empirical orthogonal function that can reflect the evolutionary pattern of the time series;
8. Divide the L components of the trajectory matrix H into C disjoint subsets representing the different trend components;
9. Reconstruct the time series. Calculate the projection a_i^m of the hysteresis matrix H_i onto U_m , $a_i^m = H_i U_m = \sum_{j=1}^L H_{i+j} U_m, 0 \leq i \leq N - L$, where a_i^m reflects the weight of H_i at the time of the original series h_{i+1}, \dots, h_{i+L} , and N is the length of the time series. Finally, the time series is reconstructed by means of the time-empirical orthogonal function U_i and the weights a_i^m . The specific reconstruction process can be defined as Equation (3):

$$h_i^k = \begin{cases} \frac{1}{i} \sum_{j=1}^i a_{i-j}^k U_{k,j}, & 1 \leq i \leq L - 1 \\ \frac{1}{L} \sum_{j=1}^L a_{i-j}^k U_{k,j}, & L \leq i \leq N - L + 1 \\ \frac{1}{N-i+1} \sum_{j=i-N+L}^L a_{i-j}^k U_{k,j}, & N - L + 2 \leq i \leq N \end{cases} \quad (3)$$

3.4. Elman Neural Network Model

The Elman neural network with great computational power. Each layer of the Elman neural network is independent of each other, which is why it is widely used in the field of power load forecasting [37].

Additionally, the non-linear state space expression of the Elman network can be defined as:

$$y(k) = g(\omega^3 x(k)), \quad (4)$$

$$x(k) = f(\omega^1 x_c(k) + \omega^2 (u(k-1))), \quad (5)$$

$$x_c(k) = x(k-1), \quad (6)$$

where y is the output vector; x is the unit vector; u is the input vector; x_c is the state vector; ω^3 is the weight of the output layer and the intermediate layer; ω^2 is the weight of the input and intermediate layers; ω^1 is the weight of the take-up layer and the intermediate layer; g and f is the transfer function.

3.5. Extreme Learning Machine Neural Network

Elman is a neural network with powerful generalization capabilities. However, the initial weights and thresholds of ELM are randomly assigned [38,39].

Suppose there are Q learning samples, $\{(x_l, y_l)\}_{Q_l=1}$ and $x_l \in R^\pi, y_l \in R^\psi$. Assume that the number of hidden layer neurons is M . The standard form is shown in Equation (7):

$$\sum_{j=1}^M a_j(\omega_j x_l + b_j) = f_M(x) \quad (7)$$

According to the zero-error approximation principle, the existence of b_j, a_j, ω_j reduces the normalized form to Equation (8).

$$Ha = Y, \quad (8)$$

where H is the output matrix, Y is the desired output matrix, and a is the output weight matrix found using the solved least squares method. Finally, the solution is continued using $a' = H^+Y$. The specific mathematical model is shown in paper [38].

3.6. Least Square Support Vector Machine

LSSVM can convert complex quadratic programming problems into linear systems of equations problems and improve the speed of problem solving by replacing inequality constraints with equation constraints in SVM optimization problems [40,41].

Let the input and output of the n training samples be $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), i = 1, 2, \dots, n$, where x_i is the input vector and y_i is the output vector. The optimal decision function is constructed in the high-dimensional feature space using a non-linear mapping function, as shown in Equation (9):

$$f(x) = \omega^T \varphi(x) + b, \quad (9)$$

where ω^T is the weight vector and b is a constant. The optimization objective can be expressed as Equation (10):

$$\min \frac{1}{2} \|\omega\| + \frac{1}{2} \lambda \sum_{i=1}^n \xi_i^2 \quad (10)$$

The constraint can be expressed as Equation (11):

$$y_i = \omega^T \varphi(x_i) + b + \xi, i = 1, 2, \dots, n \quad (11)$$

where λ is the penalty factor and ξ is the relaxation factor. Because the LSSVM can transform a quadratic programming problem into a problem of solving the above system of linear equations, the prediction model can be summarized as Equation (12).

$$y_i = b + \sum_{i=1}^n a_i k(x, x_i) \quad (12)$$

The specific mathematical derivation is shown in paper [40].

3.7. Sparrow Search Algorithm

The sparrow search algorithm is a swarm intelligence optimization algorithm proposed by Xue in 2020 [21]. The algorithm is mainly based on the foraging and anti-predation behavior of sparrows. The SSA has high search accuracy, fast convergence, high stability, and robustness compared to other population intelligence optimization algorithms. Additionally, the SSA has been successfully applied in the field of path planning [42] and structural optimization of micro-grids [43].

Sparrows in the SSA algorithm are classified as discoverers, joiners, and vigilantes. The discoverer is responsible for finding food for the entire population and providing

foraging directions for the joiners, so the discoverer's foraging search is larger than that of the joiners. During each iteration, the discoverer iterates according to Equation (13):

$$x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{-i}{\vartheta \cdot T}\right), R_2 < ST \\ x_{id}^t + Q \cdot L, R_2 \geq ST \end{cases}, \quad (13)$$

where t denotes the current number of iterations, T denotes the maximum number of iterations, ϑ is a random number between $[0, 1]$, Q is a random number subject to a normal distribution, L is a matrix of $1 \times d$ whose elements are all 1, R_2 denotes a guard value, ranging from $[0, 1]$, and ST is a safe value, ranging from $[\frac{1}{2}, 1]$.

During the foraging process, some joiners will always monitor the finder, and when the finder finds better food, the joiner will compete with it. If successful, it will immediately obtain that finder's food, otherwise the joiner will update its position according to Equation (14).

$$x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{xw_d^t - x_{id}^t}{i^2}\right), i > \frac{n}{2} \\ xb_d^{t+1} + \frac{1}{D} \sum_{d=1}^D \left(\text{rand}\{-1, 1\} \cdot |x_{id}^t - xb_d^{t+1}| \right), i \leq \frac{n}{2} \end{cases}, \quad (14)$$

where xw_d^t denotes the worst position in dimension d of the t -th iteration and xb_d^{t+1} denotes the best position. When $i > \frac{n}{2}$, it means that the population is short of food and needs to go elsewhere to forage. When $i \leq \frac{n}{2}$, it means that the tracker is preying near the optimal position xb .

The number of sparrows aware of danger in a sparrow population is 10–20% of the total, and the location of these sparrows is randomly generated and continuously updated according to Equation (15) for the location of the vigilantes.

$$x_{id}^{t+1} = \begin{cases} xb^t + \beta(x_{id}^t - xb_d^t), f_i \neq f_g \\ x_{id}^t + K \left(\frac{x_{id}^t - xw_d^t}{|f_i - f_w| + \mu} \right), f_i = f_g \end{cases}, \quad (15)$$

where β is the step control parameter, a normally distributed random number with a mean of 0 and a variance of 1; K is a random number between $[-1, 1]$; μ is a very small number, just in case the denominator is 0; f_i is the current fitness; f_g is the best fitness; and f_w is the worst fitness.

3.8. The EOBL–CSSA Algorithm

In this paper, the chaotic mapping strategy and the elite opposition-based learning strategy are used to optimize the SSA algorithm. Firstly, the initial population is initialized by Tent chaotic mapping to improve the quality of the initial solution and enhance the global search capability of the algorithm. An elite back-learning mechanism is then introduced on top of the chaotic sparrow algorithm to extend the global search capability of the algorithm. Before extending the neighborhood of the current best individual, reverse learning is performed on it to generate a reverse search population within its search boundary, guiding the algorithm to approach the solution space containing the global optimum, thus improving the algorithm's balancing and exploration capabilities as well as its ability to jump out of local extremes.

3.8.1. Tent Chaotic Mapping Strategy

The random and ergodic nature of chaos can effectively maintain the diversity of the population, helping the algorithm to jump out of the local optimum and improving the global search capability of the algorithm. It has been documented that the algorithm's ability to find an optimum is influenced by the ergodicity of the chaotic mapping [44]. The more uniform the chaotic mapping, the faster the convergence of the algorithm. As

shown in Figure 1, we can observe that the Tent chaotic mapping is well distributed and has even better traversal performance.

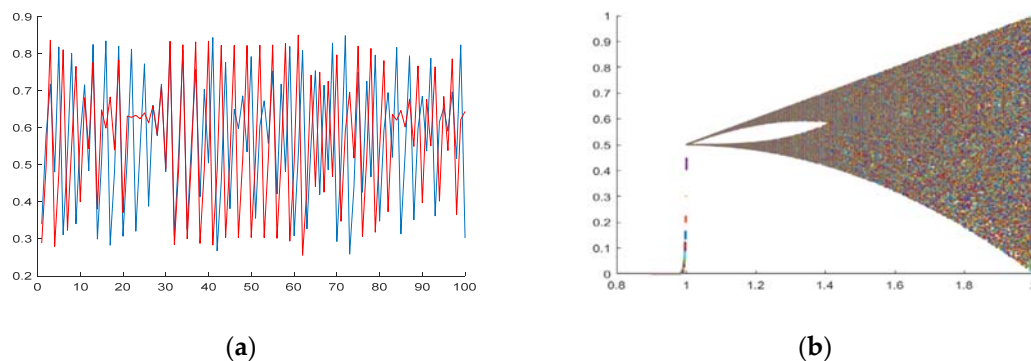


Figure 1. Tent chaos mapping. (a) Tent chaos mapping distribution; (b) Tent chaotic mapping bifurcation diagram.

Suppose that in a space of dimension D , where $x = \{x_n, n = 1, 2, \dots, D\}$. Additionally, the Tent chaos mapping can be expressed as Equation (16):

$$x_{n+1} = \begin{cases} 2x_n, & 0 \leq x_n < 0.5 \\ 2(1 - x_n), & 0.5 \leq x_n \leq 1 \end{cases} \quad (16)$$

3.8.2. Elite Opposition-Based Learning Strategies

Paper [45] introduces the concept of backward learning for the first time. This method generates a reverse solution individual of the current individual in the fetching region and will select the better of the two into the next generation. It is further shown that the reverse solution has a probability of being closer to the global optimum than the current population by about 50%. The elite inverse strategy has been used in group intelligence optimization algorithms such as PSO [46], Harris Hawk Optimization (HHO) [47] and the Dragonfly Algorithm (DA) [48].

In this paper, the global search capability of the algorithm is extended by introducing an elite opposition-based learning mechanism. The strategy takes the top 10% of individuals in terms of sparrow fitness as the elite solution, while obtaining the dynamic boundaries of the elite sparrows. Before neighborhood expansion is performed on the current best sparrow individual, backward learning is performed on it to generate a reverse search population within its search boundary, guiding the algorithm to approach the solution space containing the global optimum. The sparrow position is updated by comparing the sparrow adaptation values before and after the update. If the comparison result is better, the previous sparrow is replaced. In summary, the elite opposition-based learning strategy can be a good way to improve the algorithm's balance and exploration capabilities, thus helping the algorithm to jump out of local extremes. The elite opposition-based learning strategy can be illustrated by the following three definitions:

Theorem 1. Reverse solution [49]: Let there exist a real number x on the interval $[a, b]$. Then, the reverse solution of x is defined as $x' = a + b - x$. Suppose there exists a certain N -dimensional point $p = (x_1, \dots, x_i, \dots, x_N)$ on an R -domain and $x_i \in [a_i, b_i]$. Then, define $p' = (x'_1, \dots, x'_i, \dots, x'_N)$ to be the inverse of p , where $x'_i = k(a_i + b_i) - x_i$; k is a generalized coefficient and is a uniform random number belonging within $[0, 1]$.

Theorem 2. Optimization based on reverse solutions [49]: Let the problem to be optimized be the minimum problem and the fitness function be set to f . If there exists some feasible solution x , then its reverse feasible solution is x' . If $f(x') < f(x)$ holds, then replace x with x' .

Theorem 3. *Elite Reverse Solution [49]: Let $x_{best}^t = (x_1^t, \dots, x_i^t, \dots, x_N^t)$ be the reverse solution of the current group of elite individuals $x_{best} = (x_1, \dots, x_i, \dots, x_N)$ in some N-dimensional space, where $x_i^t = k(a_i + b_i) - x_i$; $x_i \in [a_i, b_i]$; k is a generalized coefficient and is a uniform random number belonging to the interior of $[0, 1]$.*

The flow of the chaotic sparrow search algorithm based on the elite opposition-based learning strategy can be summarized as follows:

1. Initialize the population and the number of iterations using the Tent chaos mapping formula $x_{n+1} = \begin{cases} 2x_n, 0 \leq x_n < 0.5 \\ 2(1 - x_n), 0.5 \leq x_n \leq 1 \end{cases}$. Initialize the initial ratio of predators and joiners;
2. Calculate fitness values and ranking based on the results;
3. Update predator location based on $x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{-i}{\partial \cdot T}\right), R_2 < ST \\ x_{id}^t + Q \cdot L, R_2 \geq ST \end{cases}$;
4. Update joiner location based on $x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{xw_d^t - x_{id}^t}{\rho}\right), i > \frac{n}{2} \\ xb_d^{t+1} + \frac{1}{D} \sum_{d=1}^D (rand\{-1, 1\} \cdot |x_{id}^t - xb_d^{t+1}|), i \leq \frac{n}{2} \end{cases}$;
5. Update guard positions according to $x_{id}^{t+1} = \begin{cases} xb^t + \beta(x_{id}^t - xb_d^t), f_i \neq f_g \\ x_{id}^t + K \left(\frac{x_{id}^t - xw_d^t}{|f_i - f_w| + \mu} \right), f_i = f_g \end{cases}$;
6. Calculate fitness values and update sparrow positions;
7. Find the reverse solution of all current solutions according to the defining formula $x' = a + b - x$ for the reverse solution;
8. Those individuals whose fitness value of the original solution is greater than the fitness value of the reverse solution are selected according to the elite selection formula $f(x') < f(x)$ and form an elite group;
9. On the new search space constructed by the elite population, the reverse solution is then found for individuals whose original solution fitness value is less than the reverse solution fitness value according to equation $x_i^t = k(a_i + b_i) - x_i$. If the algorithm converges to the global optimal solution, the search interval formed by the elite population must converge to the region where the optimal solution is located. This makes full use of the effective information of the elite population to generate the inverse solution on the dynamically defined interval formed by the elite population, guiding the search closer to the optimal solution;
10. Calculate the fitness value and update the sparrow individuals and locations. Compare the sparrow individuals and locations before and after the update and compare the results with each other to see if the results are better. If the result is better, replace the previous sparrow;
11. Determine whether the stop condition is met. If the condition is met, exit and output the result. Otherwise, repeat steps 2–10.

3.9. Whale Optimization Algorithm

The whale optimization algorithm is a new nature-inspired optimization algorithm proposed by Mirjalili in 2016 [20]. The WOA simulates the social behavior of humpback whales, using random or optimal search agents to model the special hunting behavior of humpback whales, and introduces a bubble attack strategy based on this. The WOA converges quickly around the optimal value and has good global optimization capability. Paper [20] systematically illustrates the mathematical model of the whale optimization algorithm.

However, the WOA has the disadvantages of uneven initial population distribution, low convergence accuracy, and insufficient global optimization capability when solving complex problems. Therefore, this paper proposes a chaotic adaptive whale algorithm on this basis.

3.10. Chaotic Adaptive Whale Optimization Algorithm

The population is initialized by the chaotic properties of the Sine chaos mapping to ensure a uniform distribution of this whale in the solution space. In addition, studies have shown that larger thresholds facilitate global exploration and smaller thresholds enhance local exploitation. Therefore, this paper introduces adaptive inertia weights to improve the convergence accuracy and global optimization capability of the algorithm. The algorithm is made to have larger inertia weights in the early iterations and smaller inertia weights in the later iterations.

3.10.1. Sine Chaotic Mapping Strategy

The mathematical model of Sine chaotic self-mapping is defined as Equation (17):

$$X_{n+1} = \frac{a}{4} \sin(\pi X_n) \quad (17)$$

Additionally, when $a \in [0, 4]$, the algorithm is in a chaotic state. This ensures that the whales are evenly distributed in the solution space after a certain number of iterations. The Sine chaos mapping distribution is shown in Figure 2.

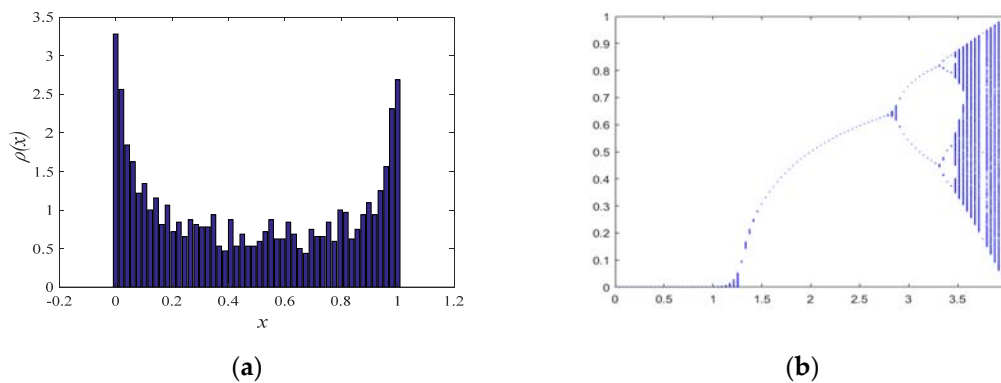


Figure 2. Sine chaos mapping. (a) Probability density diagram of Sine chaotic mapping; (b) Bifurcation diagram of Sine chaotic mapping.

3.10.2. Adaptive Inertia Weights

The adaptive inertia weight is introduced to balance the global exploration capability and local exploitation capability of the algorithm. This can be a good way to improve the optimization capability of the algorithm. The mathematical model of adaptive inertia weight ω is defined as Equation (18):

$$\omega = \frac{1}{5} + \frac{1}{0.4 - \exp\left(\frac{f_{fit}(x)}{s}\right)^r} \quad (18)$$

where $f_{fit}(x)$ is the fitness value of whale x , r is the current number of iterations, and s is the best adaptation value for the whale population in the first iteration of the calculation. Using the dynamic nonlinearities of ω to obtain the new whale position, the optimized formula can be expressed as Equations (19) and (22):

$$X(t+1) = \omega \cdot X_d(t) - A \cdot |CX_d(t) - X(t)| \quad (19)$$

$$X(t+1) = \omega \cdot X_d(t) + (|X_d(t) - X(t)|) \cdot e^{bh} \cdot \cos(2\pi h) \quad (20)$$

$$X(t+1) = \begin{cases} \omega \cdot X_d(t) - A \cdot |CX_d(t) - X(t)| & p \leq P_o \\ \omega \cdot X(t) + (|X_d(t) - X(t)|) \cdot e^{bh} \cdot \cos(2\pi h) & p > P_o \end{cases} \quad (21)$$

$$X(t+1) = \omega \cdot X_R(t) - A \cdot |C \cdot X_R(t) - X(t)| \quad (22)$$

where $|X_d(t) - X(t)|$ is the distance between the optimal solution and the current position, b is the constant of the shape of the logarithmic spiral, h is the random number between $[-1, 1]$, $X(t)$ is the current position of the whale, $X_d(t)$ is the prey position, which is the optimal solution, and t is the number of iterations. A and C are coefficient vectors, and $P_o = \frac{1}{2}$ is selection probability.

3.11. Simulated Annealing

The SA algorithm can be divided into two parts: the *Metropolis* criterion and the annealing process. The specific mathematical model is shown in paper [50]. In this paper, the SA algorithm is used to determine the proportion of weights for the prediction results of the three models, and the initial weights of all three models are set to $1/3$. Finally, the prediction results of each prediction model are used to determine the final prediction results using Equation (23).

$$Y = \omega_1 Y_1 + \omega_2 Y_2 + \omega_3 Y_3 \quad (23)$$

where Y is the final prediction result; Y_1 , Y_2 , and Y_3 are the Elman prediction model, CAWOA-ELM prediction model, and EOBL-CSSA-LSSVM prediction model, respectively; and ω_1 , ω_2 , and ω_3 are the best weighting coefficients for the above three prediction models, respectively. The size of the weights represents the degree of influence the model has on the overall portfolio forecasting model. The larger the weight, the greater the contribution of the model to the portfolio model. The smaller the weight, the smaller the model's contribution to the portfolio model. The specific operational flow of the SA algorithm to determine the optimal weighting coefficients for the three prediction models is shown in Figure 3.

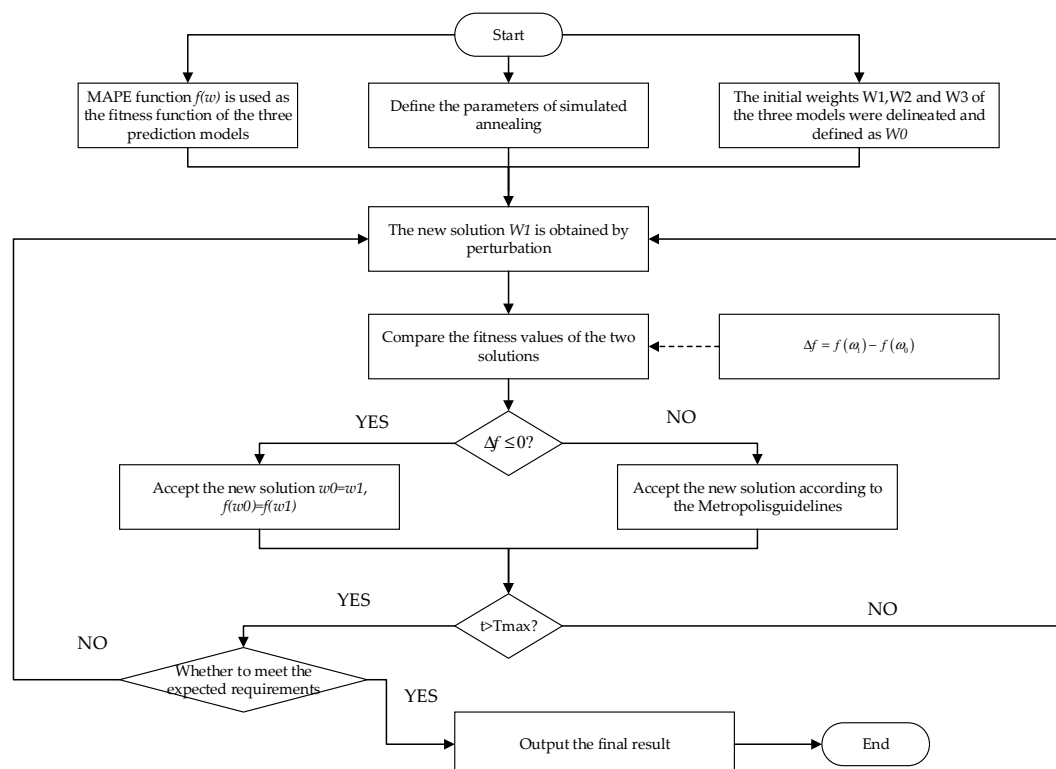


Figure 3. Flow chart of SA algorithm to determine the optimal weight coefficients of three prediction models.

3.12. The Combined Forecasting Models

This paper introduces a novel combined prediction model based on machine learning algorithms, swarm intelligence optimization algorithms, and data pre-processing. As

shown in Figure 4, the prediction model can be simplified into three modules: Module A, Module B, and Module C.

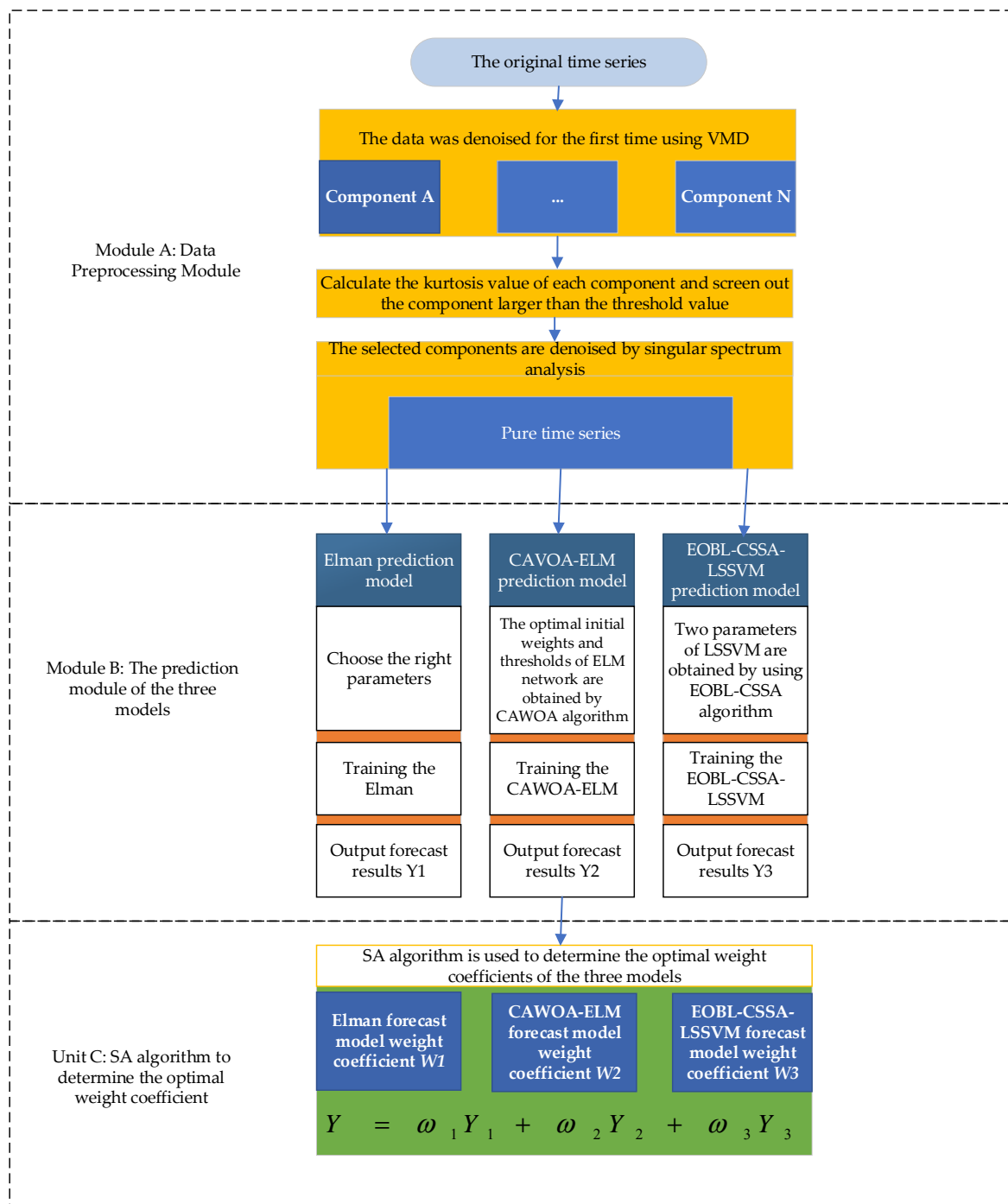


Figure 4. The specific operation process of the new combination prediction model proposed in this paper.

Module A represents the data processing module. In real life, raw electrical load data inevitably contain data noise. This data noise can significantly interfere with the learning ability of the model. Therefore, we use VMD–spectral analysis to obtain pure signals.

Module B represents the process of obtaining forecasts using three independent forecasting models. Divide the pure data obtained through module A into the input set and the output set. Firstly, the ELMAN prediction model is trained with the dataset and

the optimal number of hidden layers is determined by simulation. The prediction results are then recorded as Y_1 . Secondly, the CAWOA-ELM model is trained with the dataset to obtain the prediction result Y_2 . The appropriate initial weights and thresholds for the ELM model are selected by the CAWOA algorithm. Thirdly, the EOBL-CSSA-LSSVM model is trained using the dataset to obtain the prediction result Y_3 . Then, the EOBL-CSSA algorithm is used to optimize the two parameters in the LSSVM: the penalty factor gam and RBF kernel parameter sig .

Module C represents the process of determining the best weighting coefficients for each set of prediction models by using the SA algorithm. Module C represents the weighting calculation process. The SA algorithm is used to determine the best weighting coefficients for each set of predictions obtained from module B predictions. Additionally, the weight coefficients obtained for each set are multiplied with the prediction results of each of the three prediction models. Finally, the final prediction results are obtained by summing.

3.13. Forecast Feedback System for Electricity Load Forecasting Models

In real life, the way in which the electricity load forecasting model is applied is shown in Figure 5. The generator converts the voltage to 220 KV through the booster transformer. The electrical load is then transmitted to the primary high-voltage substation via the 220 KV high-voltage transmission line. The voltage is converted from 220 KV to 110 KV at the primary high-voltage substation and then transmitted to the secondary high-voltage substation via the 110 KV high-voltage transmission line. Finally, the secondary high-voltage substation converts the electrical load to the voltage required by the factory or the general user for the daily supply of electricity. Due to the nature of the electrical load, it cannot be stored on a large scale. If too much of the electrical load is transmitted, it can lead to a waste of resources. If too little electrical load is transmitted, it can lead to inadequate power supply and inconvenience to the population.

Based on the above problems, this paper applies the power load forecasting model to the primary high-voltage substation phase. Firstly, historical data of the power load in the area are collected through the relevant power department. The historical data are applied to the combined forecasting model proposed in this paper. The predictive model learns continuously to accurately predict the values and trends of the power loads on the 110 KV high-voltage transmission lines in the coming days. Secondly, the forecast results are fed back to the relevant authorities to provide accurate and reasonable feedback to the power sector. Finally, the power sector obtains symmetric information through the high accuracy of the combined forecasting model proposed in this paper. This symmetric information will not only help the power sector to maintain a dynamic balance between power supply and consumption, but also to reduce the waste of resources.

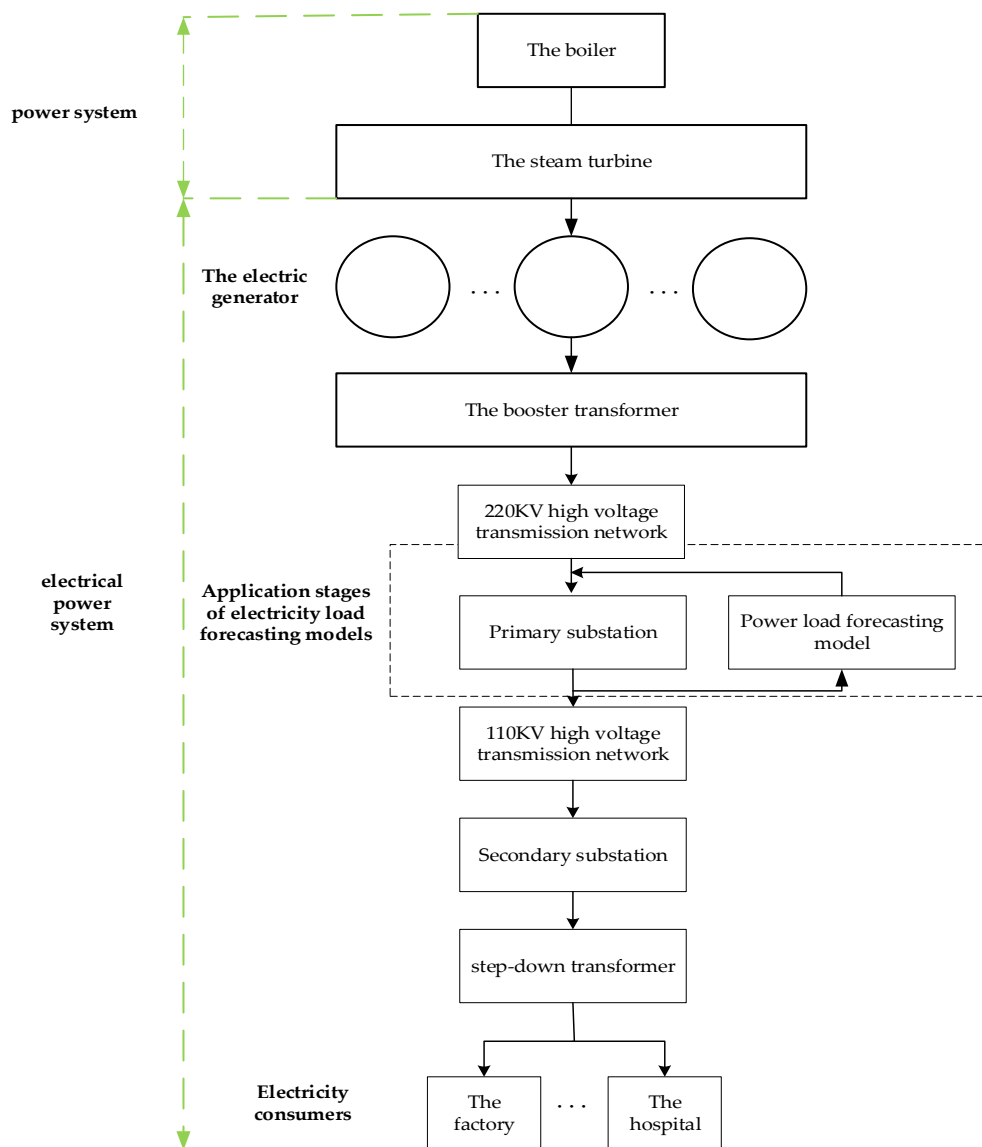


Figure 5. Feedback flow chart for the application of electricity load forecasting models.

4. The Simulation Experiments

4.1. The Datasets

To verify the forecasting performance and applicability of the combined model proposed in this paper, real electricity load data for 20 weeks from 00:00 on 30 April 2007 to 24:00 on 12 September 2007 for the five administrative states of the south-eastern Australian grid are used as the experimental data for the simulation. The experimental data are collected using a 30-min measurement interval. In other words, a total of 48 sets of experimental data are collected each day, making a total of 6720 sets of experimental data. In this paper, electricity load data are divided into seven data subsets in the order of Monday to Sunday. That is, the electricity load data for each Monday of the 20 weeks are stored in the Monday data subset, and so on for the rest, creating a total of seven data subsets from Monday to Sunday. Then, a combined model is built for each data subset for prediction. In this paper, 912 data points from the first 19 weeks of the 20 weeks of data from each data subset are used as the training set and 48 data points from the last week as the test set to verify the prediction performance of the proposed combination model. In this way, each data subset is used to predict the data of the next week using the data of the previous week.

4.2. Error Evaluation Indicators

Because of the uncertainty and randomness of electrical loads, errors are inevitable in any forecasting method. In this paper, the Root Mean Square Error (*RMSE*), Mean Absolute Percentage Error (*MAPE*), Mean Square Error (*MSE*), and Mean Absolute Error (*MAE*) are used to verify the results. The four evaluation functions are shown in Table 1. In addition, the statistics of the four error evaluation metrics are presented in numerical form to indicate the prediction results of the forecasting model. The lower the value of the statistic, the better the prediction performance.

Table 1. The four error evaluation indicators.

Metrics	The Formula
<i>RMSE</i>	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (t_i - p_i)^2}$
<i>MAPE</i>	$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left \frac{t_i - p_i}{t_i} \right $
<i>MSE</i>	$MSE = \frac{1}{N} \sum_{i=1}^N (t_i - p_i)^2$
<i>MAE</i>	$MAE = \frac{1}{N} \sum_{i=1}^N t_i - p_i $

where t_i is the i -th sample of expected output; p_i is the i -th sample of predicted output; and N is the sample size.

4.3. Electricity Load Simulation Experiments

4.3.1. Electricity Load Forecasting Data Pre-Processing Experiments

Any noise removal method can only reduce the noise present in the signal as much as possible but cannot remove it completely. In order to verify the noise reduction performance of the VMD–singular spectrum analysis method, in this paper, the data pre-processing simulation comparison experiments are divided into two groups, namely, the VMD–singular spectrum analysis and the singular spectrum analysis method comparison experiments, and the VMD–singular spectrum analysis and the original un-noised method comparison experiments.

In order to better compare the de-noising performance of the VMD–singular spectrum analysis method proposed, the paper is compared with the singular spectrum analysis method used in paper [34] in the comparative test of the de-noising effect.

Model A is the combined forecasting model proposed in paper [34] combining the Jordan, ESN, and LSSVM models. Model B is the novel combined forecasting model proposed in this paper combining the Elman, CAWOA-ELM, and EOBL-CSSA-LSSVM models. In addition, the two combined models are broken down into six single forecasting models for observation and the *MAPE* values are used as quantitative data. The comparison of noise reduction effect data is shown in Tables 2–4.

Table 2. VMD–singular spectrum analysis is used in power load denoising reduction experiments (*MAPE*).

Time	Jordan	PSO-ESN	SA-LSSVM	Combination Model A	Elman	CAWOA-ELM	EOBL-CSSA-LSSVM	Combination Model B
Monday	2.88	1.36	1.02	1.01	1.27	0.97	1.01	0.61
Tuesday	3.26	1.85	1.11	1.03	3.28	1.82	1.09	1.00
Wednesday	2.05	1.13	1.07	0.74	1.42	1.12	1.12	0.81
Thursday	1.79	1.65	1.65	1.41	1.70	1.31	1.00	0.87
Friday	1.21	1.27	0.90	0.62	1.33	0.79	0.64	0.58
Saturday	1.93	1.93	1.06	0.89	1.51	1.03	0.72	0.57
Sunday	2.13	2.39	1.68	1.25	1.85	1.42	0.95	0.79
Average	2.18	1.53	1.21	0.98	1.76	1.20	0.79	0.74

Table 3. Singular spectrum analysis is used in power load noise reduction experiments (*MAPE*).

Time	Jordan	PSO-ESN	SA-LSSVM	Combination Model A	Elman	CAWOA-ELM	EOBL-CSSA-LSSVM	Combination Model B
Monday	2.70	2.04	1.21	1.14	2.01	1.12	1.09	0.82
Tuesday	1.44	2.79	1.39	1.21	2.79	2.47	1.32	1.23
Wednesday	2.03	1.13	2.42	1.02	1.65	1.22	0.97	0.89
Thursday	1.83	1.72	1.26	1.19	2.42	1.37	1.07	0.91
Friday	1.44	1.30	1.03	0.98	1.27	1.26	1.71	1.20
Saturday	2.41	2.48	1.55	1.32	1.86	1.68	1.11	1.08
Sunday	2.26	2.17	1.46	1.29	1.88	1.15	1.23	1.07
Average	2.02	1.95	1.47	1.16	1.98	1.47	1.21	1.02

Table 4. Raw data are used for power load noise reduction experiments (*MAPE*).

Time	Jordan	PSO-ESN	SA-LSSVM	Combination Model A	Elman	CAWOA-ELM	EOBL-CSSA-LSSVM	Combination Model B
Monday	5.02	2.45	1.09	1.32	4.14	1.35	1.02	0.94
Tuesday	6.88	4.07	2.44	2.37	5.02	3.01	1.67	1.53
Wednesday	4.94	1.51	2.80	1.66	3.71	1.74	2.35	1.67
Thursday	3.12	2.95	1.51	1.78	4.45	2.27	1.52	1.48
Friday	5.10	1.66	2.36	1.89	4.88	1.64	2.80	1.57
Saturday	5.47	2.37	4.72	2.35	6.02	2.28	4.42	2.15
Sunday	4.71	1.48	3.16	1.56	6.78	2.79	3.33	2.44
Average	5.03	2.36	2.58	1.85	5.01	2.15	2.44	1.68

Through the comparative experimental analysis of the denoising effect of the VMD–singular spectrum analysis and the singular spectrum analysis in Tables 2 and 3, we can clearly observe that the combined model A has a different degree of reduction in all subsets except for the Thursday subset, where the *MAPE* value is slightly increased. Among them, the highest reduction in *MAPE* value is 0.43%, the lowest is 0.13%, and the average prediction error is reduced by 0.18%. For the combined model B, the noise reduction effect of the VMD–spectral analysis was reduced to varying degrees compared to the noise removal effect of the singular spectrum analysis. The highest reduction in *MAPE* value was 0.62%, the lowest was 0.04%, and the overall average reduction was 0.28%.

In addition, we observed that for the six single combination models, *MAPE* values decreased to varying degrees for most of the subsets, except for some single models that showed an increase in *MAPE* values for individual time subsets. Especially for the SA-LSSVM model and the EOBL-CSSA-LSSVM model, the prediction errors are reduced by 0.26% and 0.42%, respectively. Therefore, we can obtain the following conclusion: from the perspective of the overall noise reduction effect, the noise reduction effect of VMD–singular spectrum analysis is better than that of the singular spectrum analysis noise reduction method.

Through the VMD–singular spectrum analysis in Tables 2 and 4 and the comparative experimental analysis of the original data without denoising, we can observe that, regardless of the combination model A or the combination model B, the effect of noise reduction processing is more obvious, especially for the Elman prediction model and the Jordan prediction model. At the same time, we also observed that almost all the data pre-processed by VMD–singular spectrum analysis in the combined model B have different degrees of reduction in the *MAPE* value of the unprocessed original data. This also shows that the VMD–singular spectrum noise reduction method does have a positive effect on improving the prediction performance of the combined model. In addition, the highest value of the noise reduction effect is on Sunday, and its *MAPE* value is reduced by 1.65%. The lowest value of noise reduction effect is on Monday, and its *MAPE* value is reduced by 0.33%. This may be because the data model of Sunday is more complicated, so the denoising effect is obvious, while the data model of Monday is relatively stable and the noise

reduction effect is relatively gentle. In summary, we can draw the conclusion that for the new combined prediction model proposed in this article, the VMD–singular spectrum analysis denoising method can indeed effectively improve the prediction performance of the combined prediction model.

4.3.2. Performance Analysis of Simulation Experiment of Elman Electric Load Forecasting Model

Secondly, Elman forecasting models are built for each of the seven subsets of temporal data divided as described above. In addition, in order to obtain the best predictive performance for the whole combination model, it is necessary to ensure that each sub-model of the combination model achieves the best predictive performance. Since the prediction performance of the Elman network is affected by the number of hidden layers, we ensure that the Elman model achieves the best performance by selecting the appropriate number of hidden layers and choosing a range between 15 and 50. The optimal number of hidden layers for the Elman model is shown in Table 5.

Table 5. Optimal number of hidden layers in Elman model.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Hidden layer	25	27	25	23	26	32	29
MAPE (%)	1.27	3.28	1.42	1.70	1.33	1.51	1.85

As shown in Table 5, the optimal number of hidden layers in the Elman model from the Monday subset to the Friday subset is very similar, and the value fluctuates around 25. Additionally, Monday and Wednesday have the same optimal number of hidden layers. Interestingly, we also found that *MAPE* values are very similar in all the models except Tuesday. Although the optimal number of hidden layers for Saturday and Sunday is 32 and 37, respectively, the *MAPE* values for these two subsets are also relatively similar. This may be due to the similarity of the data patterns from Monday to Friday, even though the magnitude of the change in *MAPE* values is more pronounced on Thursday. We also found that the *MAPE* values of the ELMAN model fluctuate considerably, and the maximum value is 2.01% of the minimum value.

4.3.3. Performance Analysis of CAWOA-ELM Power Load Forecasting Model

Thirdly, this paper builds seven CAWOA-ELM models for each of the seven data subsets divided above. The performance of the ELM model depends on the initial weights and thresholds. Therefore, we use the CAWOA algorithm to determine the ELM model. In addition, this paper also improves the prediction accuracy of the model by selecting the best hidden layer neurons.

Table 6 shows the optimal number of hidden layers and *MAPE* values for different datasets. As can be seen from Table 6, the prediction performance of the CAWOA-ELM model is excellent, and the fluctuation of *MAPE* values is relatively small. The difference between the maximum and minimum values of the *MAPE* is 1.03%. The difference between the maximum and minimum *MAPE* values is 1.03%, and the *MAPE* value for Friday is 0.78%.

Table 6. The optimal number of hidden layer neurons in CAWOA-ELM model.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Hidden layer	25	24	25	29	42	38	32
MAPE (%)	0.97	1.81	1.12	1.31	0.78	1.03	1.42

4.3.4. Performance Analysis of the EOBL-CSSA-LSSVM Electricity Load Forecasting Model

Finally, seven EOBL-CSSA-LSSVM prediction models are built for each of the seven temporal data subsets divided above. Additionally, the novel EOBL-CSSA algorithm is used to optimize the penalty factor gam and the RBF kernel parameter sig in LSSVM, and the results of the optimization search are shown in Table 7. Then, the two parameters obtained by the optimization of the EOBL-CSSA algorithm are assigned to the LSSVM prediction model.

Table 7. The optimization results of EOBL-CSSA algorithm for LSSVM model.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
gam	7896.90	10,000.00	9881.20	10,000.00	5712.00	7822.40	6127.30
Sig	153.97	1260.20	31.67	41.10	150.35	30.77	14.44
MAPE (%)	1.01	1.09	1.12	1.00	0.64	0.72	0.95

4.3.5. SA Algorithm Optimizes the Weighting Factors of Three Electricity Load Forecasting Models

Finally, the SA algorithm is used to solve for the optimal weights for each predictive model in the linear combination formulation. The parameters of the SA algorithm were set as follows: the Markov chain length is 50, the decay parameter is 0.998, and the step factor is 0.01. Moreover, the initial value of the combined weights for all three models was set to 0.33. The optimal weight proportions for each model are shown in Table 8. From Table 8, we find some interesting phenomena: firstly, the average weight of the three models shows that the EOBL-CSSA-LSSVM model has the highest weight and contribution, accounting for approximately 60% of the weight. The Elman model has the lowest weight and contribution, about 10%, while the CAWOA-ELM model also plays a key role, about 30%. Additionally, as shown in the Monday subset, $W2$ has a weight of 0.5714, while $W3$ has a weight of 0.5714. This is a good indication that the prediction accuracy of the combined model is influenced in large part by the CAWOA-ELM model. In the Wednesday subset, the weights of the three models are very close to each other, which is a good indication that the three models make similar contributions to the combined model.

Table 8. The optimized combination weights by the SA algorithm for load forecasting.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Average
$W1$	0.1175	0.0402	0.2774	0.1761	0.0815	0.0281	0.0386	0.1085
$W2$	0.5714	0.0961	0.3613	0.3125	0.4473	0.1068	0.0971	0.2846
$W3$	0.3111	0.8637	0.3613	0.5114	0.4712	0.8651	0.8643	0.6069
MAPE	0.61	1.00	0.81	0.87	0.58	0.57	0.79	0.74

Among them, $W1$ represents the weight of the Elman model, $W2$ represents the weight of the CAWOA-ELM model, and $W3$ represents the weight of the EOBL-CSSA-LSSVM model.

5. Analysis of Experimental Results of Power Load Simulation

In order to illustrate more effectively the points of the combined prediction model proposed in this paper, we divided the simulation experimental results into several groups of experiments for analysis. Figure 6 shows the prediction results and prediction trends for the three combined prediction models and the six individual prediction models. Figure A1 in Appendix A shows an enlarged view of the prediction results of the competitive model. The three combined forecasting models include the combined forecasting model proposed in this paper, the combined forecasting model proposed in paper [34], and the FA-CSSA-ELM forecasting model proposed in paper [33]. The six individual forecasting models include the Jordan model, the Elman model, the PSO-ESN, the SA-LSSVM, the CAWOA-

ELM, and the EOBL-CSSA-LSSVM forecasting model. Additionally, the different color curves represent the prediction results of different prediction models.

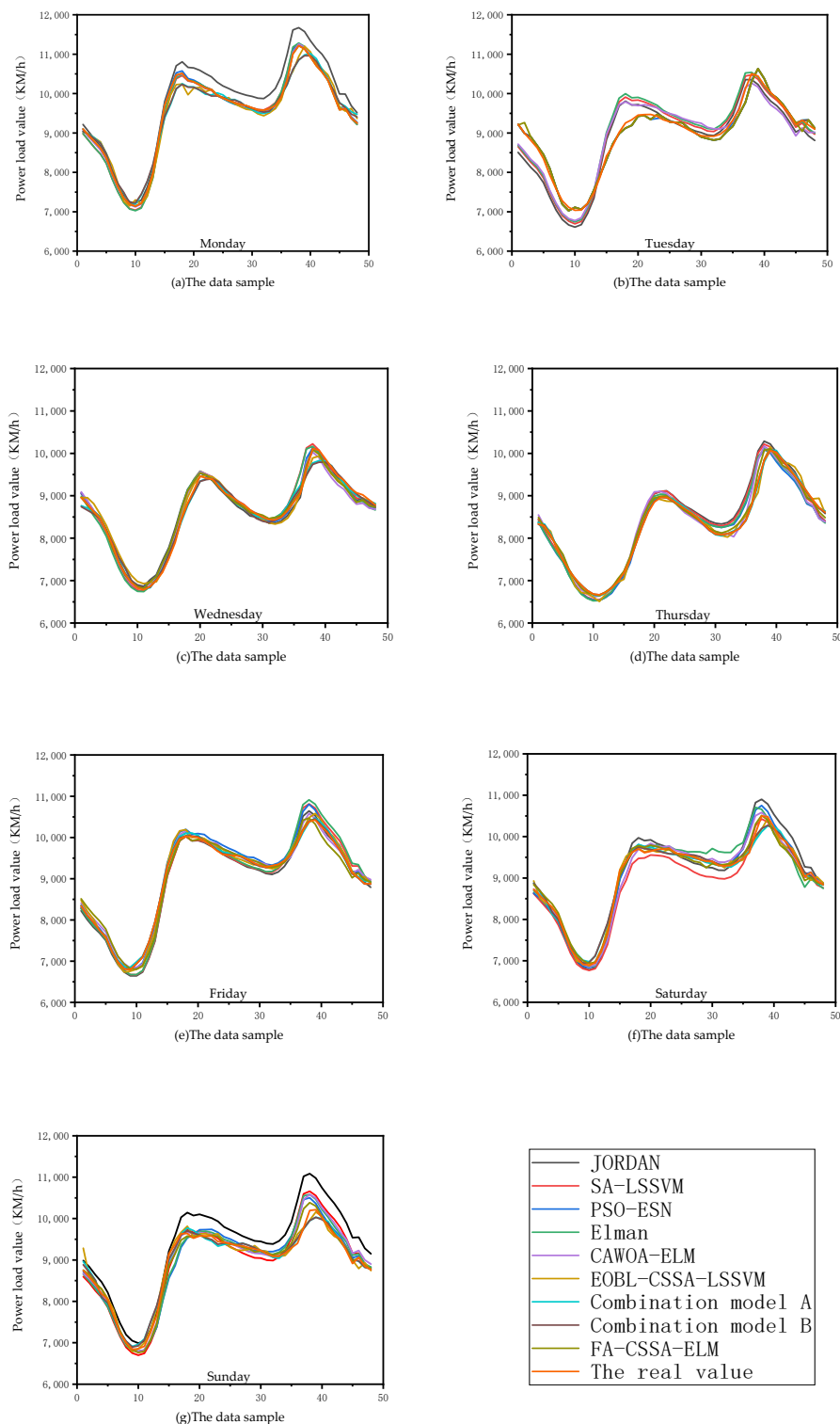


Figure 6. The final daily power load forecast value of different forecast models. (a) Describes the prediction results of different prediction models in the Monday subset; (b) Describes the prediction results of different prediction models in the Tuesday subset; (c) Describes the prediction results of different prediction models in the Wednesday subset; (d) Describes the prediction results of different prediction models in the Thursday subset; (e) Describes the prediction results of different prediction models in the Friday subset; (f) Describes the prediction results of different prediction models in the Saturday subset; (g) Describes the prediction results of different prediction models in the Sunday subset.

We conducted a comparative analysis between the combined prediction model proposed in this paper and the six individual prediction models. As shown in Figure 6, we found that the prediction results of the combined prediction model proposed in this paper are closer to the real historical data. This is followed by the prediction results of the EOBL-CSSA-LSSVM model and the PSO-ESN prediction model. Although the predictive performance of both the EOBL-CSSA-LSSVM model and the PSO-ESN model is excellent, the overall volatility of the PSO-ESN model is higher than that of the EOBL-CSSA-LSSVM model. However, the overall volatility of the PSO-ESN model is higher than that of the EOBL-CSSA-LSSVM model, especially on weekdays. Therefore, we believe that the EOBL-CSSA-LSSVM model has a better predictive performance. In addition, the prediction curves of Jordan's model and Elman's model are the furthest away from the real data and have the largest variation. Therefore, we conclude that the prediction performance of the Jordan and Elman models is relatively poor. In summary, the combined prediction model proposed in this paper showed the best prediction accuracy and prediction performance compared to the six individual prediction models.

We also conducted a comparative analysis between the combined prediction model proposed in this paper and the three combined prediction models. As shown in Figure 6, among the three combined prediction models, the FA-CSSA-ELM prediction model has the lowest prediction accuracy and the prediction curve is the farthest away from the real data. In addition, the prediction curves of both the combined prediction model proposed in paper [34] and the prediction model proposed in this paper are very close to the distance of the real data. However, on Thursday, the prediction curves of the model proposed in this paper are closer to the real data. Therefore, from an overall perspective, the prediction model proposed in this paper has the best prediction accuracy.

In order to more intuitively verify the predictive performance of the combined models proposed in this paper, we gave the evaluation values of the different prediction models in numerical form in conjunction with the four error evaluation metrics shown in Table 1, as shown in Table 9. In particular, the *RMSE*, *MSE*, and *MAE* can reflect the prediction accuracy, while the *MAPE* shows a high prediction expressiveness. In addition, we also compared the prediction performance of different prediction models more visually in the form of bar charts, as shown in Figure 7. From the values of the indicators presented in Table 9 and Figure 7, we found that the *RMSE*, *MAE*, *MAPE*, and *MSE* values of the combined forecasting model proposed in this paper have the best expressions compared to other competing models. We also found a satisfactory prediction result: the *MAPE* values for all data subsets were less than or equal to 1%, even with a minimum value of 0.57%. Although the mean *MAPE* value of the prediction model proposed in paper [34] is very close to 1%, its maximum and minimum *MAPE* values are 1.52% and 0.62%, respectively. This also indicates that our proposed new combined forecasting model has the best forecasting accuracy and forecasting performance.

Table 9. Evaluation index statistics of different power load competition models.

	Time	Jordan	Elman	PSO-ESN	SA-LSSVM	CAWOA-ELM	EOBL-CSSA-LSSVM	Combination Model A	Combination Model B	FA-CSSA-ELM
<i>MSE</i>	Monday	49.31	21.95	19.38	22.50	16.83	15.62	15.62	10.13	17.09
	Tuesday	50.82	31.45	18.89	51.50	30.13	18.88	17.88	15.22	24.87
	Wednesday	25.51	20.90	17.27	23.45	18.24	17.24	12.10	14.30	16.63
	Thursday	31.57	28.97	27.35	27.65	22.65	17.95	26.24	12.75	13.35
	Friday	18.34	21.68	15.25	23.84	12.30	11.13	10.91	8.72	22.03
	Saturday	33.65	38.12	21.12	30.00	19.96	13.00	14.91	7.95	20.85
	Sunday	35.47	26.43	31.64	33.06	26.86	19.04	21.12	14.28	17.83
<i>MAPE</i>	Monday	2.89	1.36	1.03	1.27	0.97	1.01	1.01	0.61	1.00
	Tuesday	3.27	1.85	1.09	3.28	1.82	1.09	1.03	1.00	1.49
	Wednesday	2.05	1.13	1.07	1.42	1.12	1.12	0.74	0.81	1.02
	Thursday	1.79	1.65	1.65	1.71	1.31	1.01	1.52	0.87	0.94
	Friday	1.21	1.28	0.91	1.33	0.78	0.64	0.62	0.58	1.30
	Saturday	1.93	1.93	1.06	1.51	1.03	0.72	0.90	0.58	1.03
	Sunday	2.13	1.57	1.69	1.85	1.43	0.95	1.26	0.79	0.90

Table 9. Cont.

	Time	Jordan	Elman	PSO-ESN	SA-LSSVM	CAWOA-ELM	EOBL-CSSA-LSSVM	Combination Model A	Combination Model B	FA-CSSA-ELM
RMSE	Monday	328.56	152.10	134.26	155.88	116.57	112.84	112.84	0.61	118.41
	Tuesday	352.07	235.70	130.85	356.81	230.71	130.82	120.82	74.36	150.76
	Wednesday	128.23	142.80	119.71	162.48	126.37	119.43	80.12	110.42	115.23
	Thursday	218.69	200.71	189.48	191.58	156.93	124.33	197.43	89.76	92.49
	Friday	127.07	150.20	105.69	165.16	85.22	77.12	66.16	91.33	152.65
	Saturday	233.16	264.07	146.34	207.83	138.25	90.09	91.26	57.08	144.48
	Sunday	245.72	183.08	219.18	229.02	186.09	131.90	146.34	54.06	123.56
MAE	Monday	266.73	125.41	98.21	118.17	91.83	87.25	87.25	61.54	96.07
	Tuesday	281.34	158.95	95.95	286.55	155.51	95.95	95.95	96.79	109.62
	Wednesday	99.76	97.21	91.59	122.54	98.53	97.31	66.95	91.55	87.40
	Thursday	150.87	140.69	141.31	145.57	114.11	82.80	144.30	76.96	79.00
	Friday	106.02	115.95	84.13	121.27	68.89	57.95	71.79	67.95	117.4595
	Saturday	179.19	183.82	96.56	141.89	92.43	64.50	79.84	62.89	94.38
	Sunday	191.45	139.69	153.12	167.76	130.51	85.42	96.56	79.10	79.42

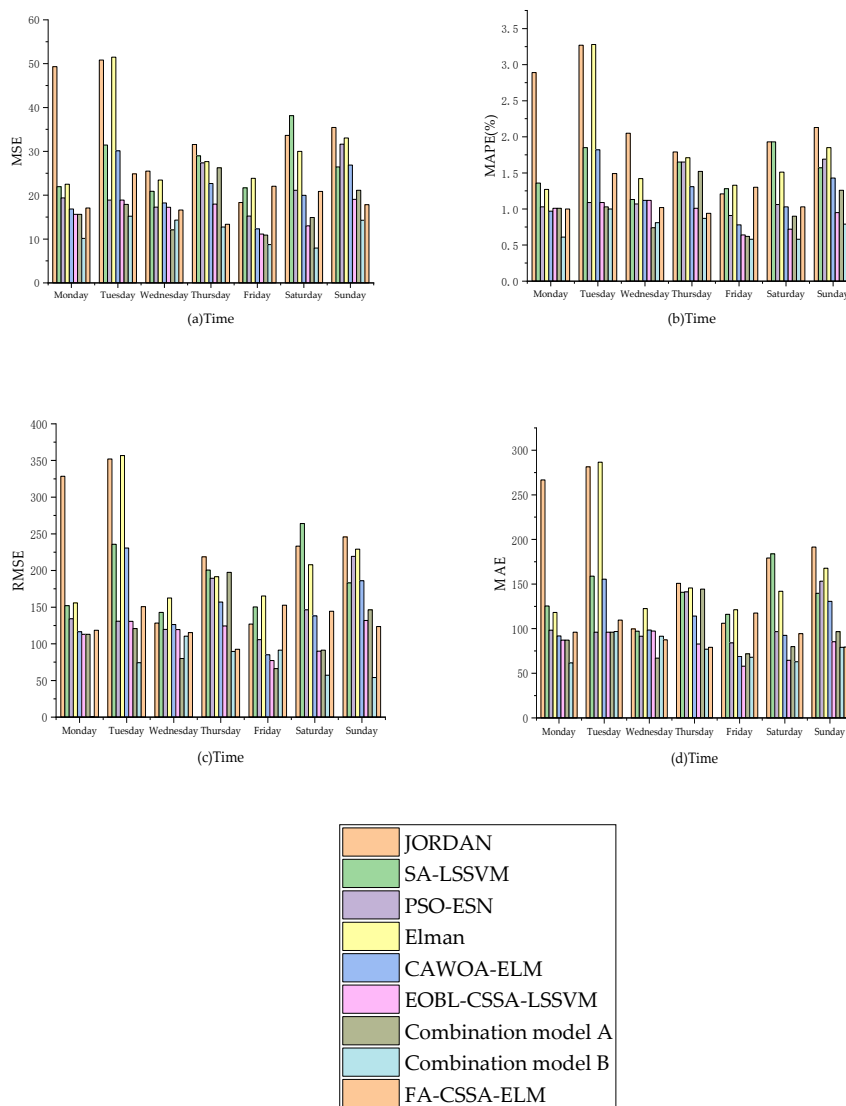


Figure 7. The statistics of evaluation indexes of different power load competition models. (a) Comparison of *MSE* values for different competition forecasting models; (b) Comparison of *MAPE* values for different competition forecasting models; (c) Comparison of *RMSE* values for different competition forecasting models; (d) Comparison of *MAE* values for different competition forecasting models.

6. Conclusions

Nowadays, accurate electrical loads not only help the power sector to make rational work plans and production decisions, but also to reduce the waste of resources and economic losses. Based on the above problems, this paper proposes a combined power load forecasting model based on machine learning, swarm intelligence optimization algorithms, and data pre-processing. The combined model is based on the Elman model, the ELM model, and the LSSVM model. Additionally, two improved swarm intelligence optimization algorithms (CAWOA-CSSA and EOBL-CSSA algorithms) proposed in this paper are used to optimize the parameters of the ELM and LSSVM models, respectively. Then, the SA algorithm is used to calculate and assign the weighting coefficients of the three models. Finally, the final prediction results are obtained by weighted summation. By combining models, the advantages of machine learning algorithms, swarm intelligence optimization algorithms, and data pre-processing can be combined to reduce the shortcomings of a single model. In addition, through several sets of simulations and analysis of the experimental results, we obtained the following conclusions.

1. We found that the noise reduction effect of the VMD–singular spectrum analysis method proposed in this paper is obvious and can effectively improve the prediction accuracy of the prediction model. The average *MAPE* value of the combined prediction model is reduced by 0.28% and the prediction accuracy is improved by 27.4% when compared to the singular spectrum analysis method. The average *MAPE* value of the combined prediction model is reduced by 0.94% and the prediction accuracy is improved by 55% when compared to the non-denoised method;
2. The two improved swarm intelligence optimization algorithms proposed in this paper can improve the prediction accuracy of the combined prediction model. Through Table 8, we can find that the weight proportion of the EOBL-CSSA-LSSVM model is 60% and the weight proportion of the CAWOA-ELM model is 30%, which is good proof that our proposed EOBL-CSSA algorithm and CAWOA-CSSA algorithm play an important role in the combined prediction model;
3. The simulation experiments and the analysis of the experimental results show that the combined forecasting model proposed in this paper has the best performance in terms of forecasting performance and forecasting accuracy. The combined model has the best performance in terms of *MSE*, *MAPE*, *RMSE*, and *MAE*. The prediction model proposed in paper [34] and the FA-CSSA-ELM model proposed in paper [33] are the next best, while the Jordan model has the worst prediction performance. Compared with the Jordan model, the *MSE* value of the combined model is reduced by 67%, the *MAPE* value is reduced by 66%, the *RMSE* value is reduced by 69%, and the *MAE* value is reduced by 57%. Compared with the FA-CSSA-ELM model, the combined model's *MSE* value is reduced by 38%, the *MAPE* value is reduced by 32%, the *RMSE* value is reduced by 45%, and the *MAE* value is reduced by 18%. Compared with paper [34], the *MSE* value of the combined model is reduced by 31%, the *MAPE* value is reduced by 32%, the *RMSE* value is reduced by 33%, and the *MAE* value is reduced by 15%.

In short, the combined forecasting model proposed in this paper has strong forecasting performance and forecasting accuracy. Additionally, the VMD–singular spectrum analysis method proposed in this paper has an obvious denoising effect. In addition, the EOBL-CSSA algorithm and CAWOA algorithm proposed in this paper can also effectively improve the deficiencies of the machine learning model. Although the predictive performance of our proposed predictive model is excellent, it can provide effective feedback and information for the power sector. However, we have not considered too many practical factors such as weather and holidays. Therefore, we will consider the impact of other factors on power load forecasting in future research.

Author Contributions: Conceptualization, X.W., X.G. and Z.W.; methodology, X.W., X.G. and Z.W.; software, X.W., X.G. and Z.W.; validation, X.W., C.M. and Z.S.; formal analysis, X.G., C.M. and Z.S.; investigation, X.W. and X.G.; resources, Z.W., X.G. and Z.S.; writing—original draft preparation, X.W. and Z.S.; writing—review and editing, X.W., X.G. and Z.W.; visualization, X.W., C.M. and Z.S.; supervision, X.W. and Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science, Education, and Industry Integration Innovation Pilot Project of Qilu University of Technology (Shandong Academy of Sciences), grant number 2020KJC-ZD04; the Postgraduate Tutors' Guidance Ability Improvement Project of Shandong Province, grant number SDYY17076; and the Empirical Research on Innovation of Cultivation Model of Control Graduate Students Based on System Synergy Theory, grant number SDYY18151.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: The study did not involve humans.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Appendix A

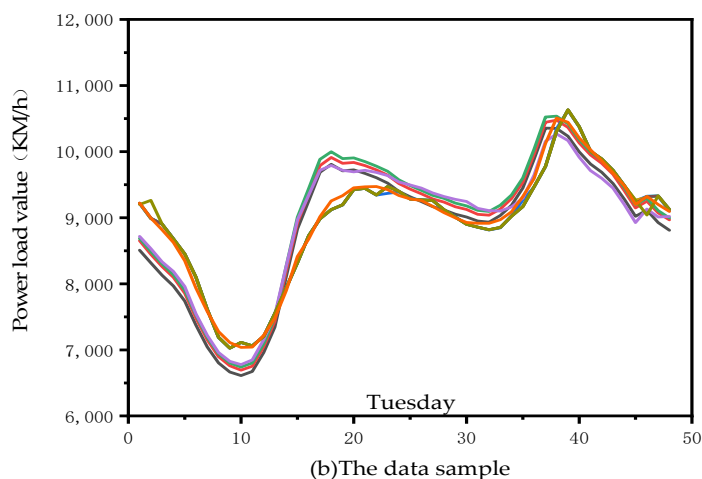
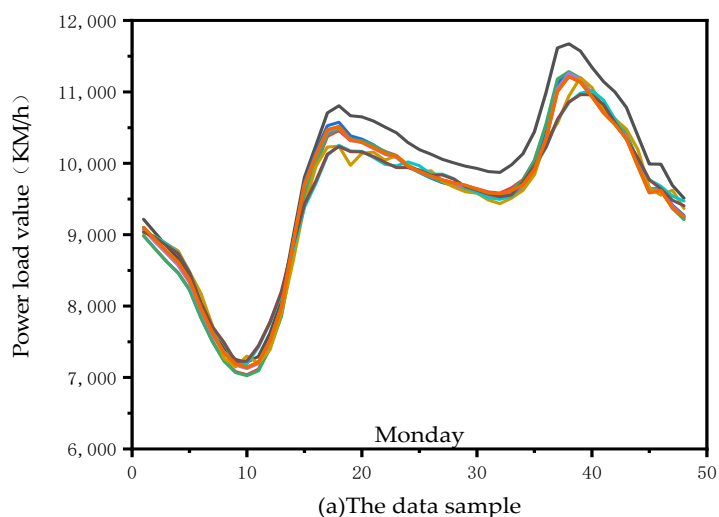


Figure A1. Cont.

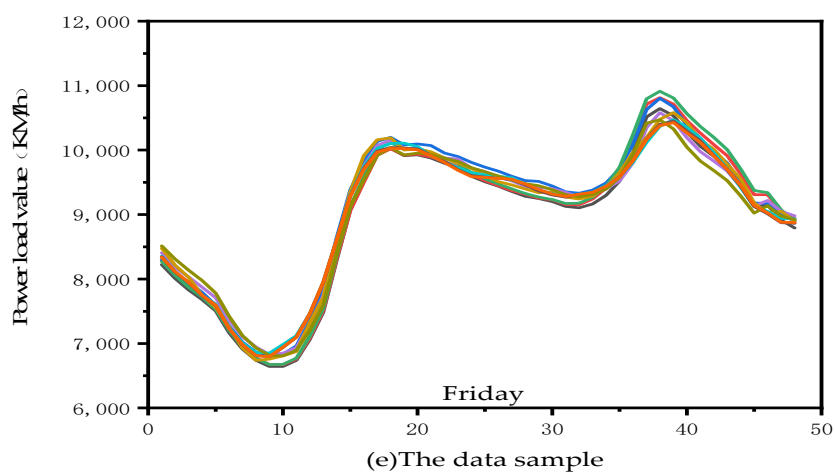
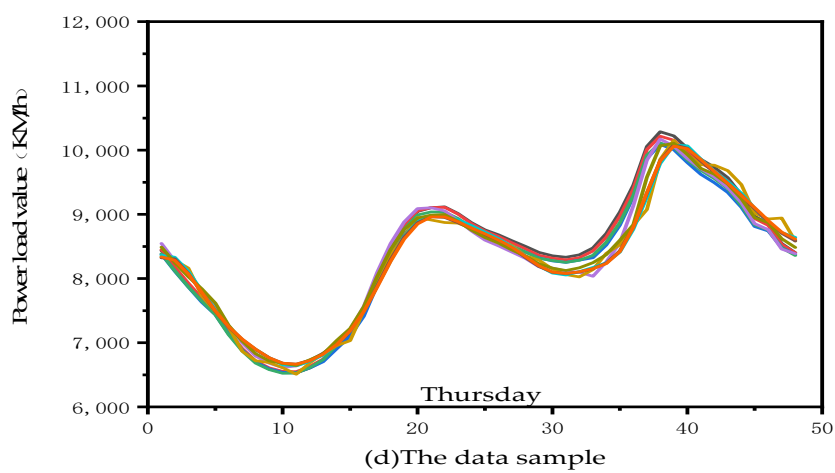
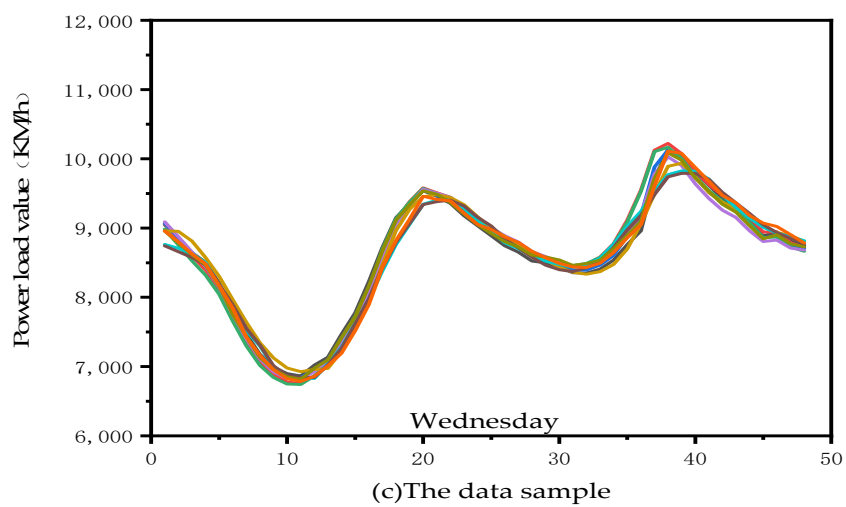


Figure A1. Cont.

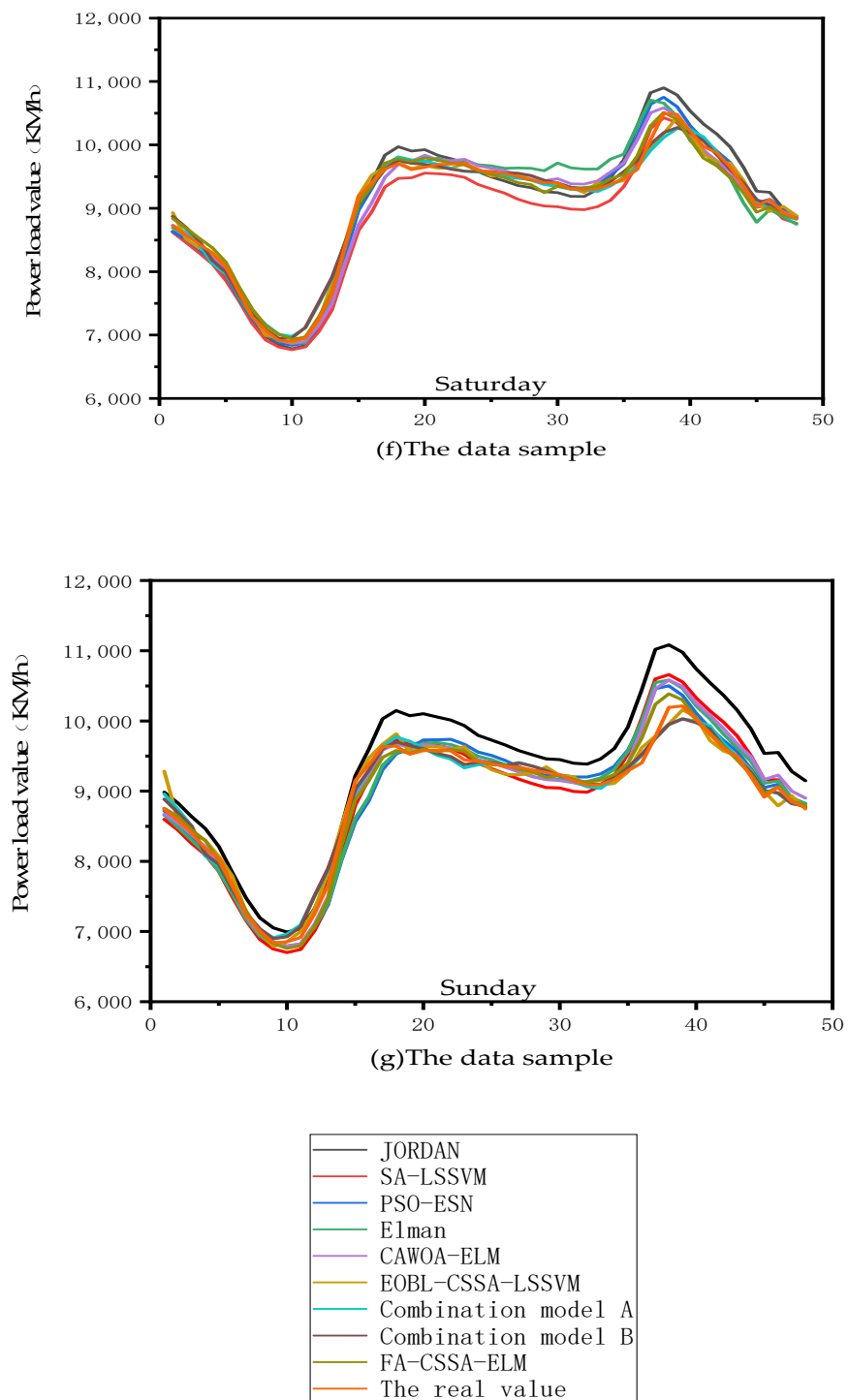


Figure A1. The final daily power load forecast value of different forecast models. (a) Describes the prediction results of different prediction models in the Monday subset; (b) Describes the prediction results of different prediction models in the Tuesday subset; (c) Describes the prediction results of different prediction models in the Wednesday subset; (d) Describes the prediction results of different prediction models in the Thursday subset; (e) Describes the prediction results of different prediction models in the Friday subset; (f) Describes the prediction results of different prediction models in the Saturday subset; (g) Describes the prediction results of different prediction models in the Sunday subset.

References

1. Khan, A.N.; Iqbal, N.; Ahmad, R.; Kim, D.-H. Ensemble Prediction Approach Based on Learning to Statistical Model for Efficient Building Energy Consumption Management. *Symmetry* **2021**, *13*, 405. [\[CrossRef\]](#)
2. Montoya, O.D.; Grisales-Noreña, L.F.; Gil-González, W.; Alcalá, G.; Hernandez-Escobedo, Q. Optimal Location and Sizing of PV Sources in DC Networks for Minimizing Greenhouse Emissions in Diesel Generators. *Symmetry* **2020**, *12*, 322. [\[CrossRef\]](#)
3. Zhao, L.; Zhou, X. Forecasting Electricity Demand Using a New Grey Prediction Model with Smoothness Operator. *Symmetry* **2018**, *10*, 693. [\[CrossRef\]](#)
4. Jin, X.-B.; Zheng, W.-Z.; Kong, J.-L.; Wang, X.-Y.; Bai, Y.-T.; Su, T.-L.; Lin, S. Deep-Learning Forecasting Method for Electric Power Load via Attention-Based Encoder-Decoder with Bayesian Optimization. *Energies* **2021**, *14*, 1596. [\[CrossRef\]](#)
5. Shieh, H.-L.; Chen, F.-H. Forecasting for Ultra-Short-Term Electric Power Load Based on Integrated Artificial Neural Networks. *Symmetry* **2019**, *11*, 1063. [\[CrossRef\]](#)
6. Tan, M.; Yuan, S.; Li, S.; Su, Y.; Li, H.; He, F. Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning. *IEEE Trans. Power Appar. Syst.* **2019**, *35*, 2937–2948. [\[CrossRef\]](#)
7. Aprillia, H.; Yang, H.-T.; Huang, C.-M. Statistical load forecasting using optimal quantile regression random forest and risk assessment index. *IEEE Trans. Smart Grid.* **2020**, *12*, 1467–1480. [\[CrossRef\]](#)
8. Aboubacar, A.; El Machkouri, M. Recursive kernel density estimation for time series. *IEEE Trans. Inf. Theory* **2020**, *66*, 6378–6388. [\[CrossRef\]](#)
9. Jordan, M.I.; Mitchell, T.M.J.S. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [\[CrossRef\]](#)
10. Borges, F.; Pinto, A.; Ribeiro, D.; Barbosa, T.; Pereira, D.; Magalhães, R.; Barbosa, B.; Ferreira, D. An unsupervised method based on support vector machines and higher-order statistics for mechanical faults detection. *IEEE Lat. Am. Trans.* **2020**, *18*, 1093–1101. [\[CrossRef\]](#)
11. Diyan, M.; Khan, M.; Nathali Silva, B.; Han, K. Scheduling Sensor Duty Cycling Based on Event Detection Using Bi-Directional Long Short-Term Memory and Reinforcement Learning. *Sensors* **2020**, *20*, 5498. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Diyan, M.; Silva, B.N.; Han, K. A multi-objective approach for optimal energy management in smart home using the reinforcement learning. *Sensors* **2020**, *20*, 3450. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Deng, Y.; Lu, H.; Zhou, W. Security event-triggered control for markovian jump neural networks against actuator saturation and hybrid cyber attacks. *J. Franklin Inst.* **2021**, in press. [\[CrossRef\]](#)
14. Liu, J.; Liu, X.; Le, B.T. Rolling force prediction of hot rolling based on GA-MELM. *Complexity* **2019**, *2019*, 11. [\[CrossRef\]](#)
15. Jin, Q.; Xu, Z.; Cai, W. An Improved Whale Optimization Algorithm with Random Evolution and Special Reinforcement Dual-Operation Strategy Collaboration. *Symmetry* **2021**, *13*, 238. [\[CrossRef\]](#)
16. Dorigo, M.; Birattari, M.; Stutzle, T. Ant colony optimization. *IEEE Comput. Intell. Mag.* **2006**, *1*, 28–39. [\[CrossRef\]](#)
17. Sangeetha, V.; Krishankumar, R.; Ravichandran, K.S.; Cavallaro, F.; Kar, S.; Pamucar, D.; Mardani, A. A Fuzzy Gain-Based Dynamic Ant Colony Optimization for Path Planning in Dynamic Environments. *Symmetry* **2021**, *13*, 280. [\[CrossRef\]](#)
18. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks(ICNN), Perth, Australia, 27 November–1 December 1995; pp. 1942–1948.
19. Zhao, Q.; Li, C. Two-stage multi-swarm particle swarm optimizer for unconstrained and constrained global optimization. *IEEE Access* **2020**, *8*, 124905–124927. [\[CrossRef\]](#)
20. Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [\[CrossRef\]](#)
21. Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [\[CrossRef\]](#)
22. Zhang, W.; Qu, Z.; Zhang, K.; Mao, W.; Ma, Y.; Fan, X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. *Energy Convers. Manag.* **2017**, *136*, 439–451. [\[CrossRef\]](#)
23. Wang, Z.; Xu, L. Fault Detection of the Power System Based on the Chaotic Neural Network and Wavelet Transform. *Complexity* **2020**, *2020*, 15.
24. Malik, H.; Alotaibi, M.A.; Almutairi, A. A new hybrid model combining EMD and neural network for multi-step ahead load forecasting. *J. Intell. Fuzzy Syst.* **2021**, 1–16. [\[CrossRef\]](#)
25. Jin, Y.; Guo, H.; Wang, J.; Song, A. A hybrid system based on LSTM for short-term power load forecasting. *Energies* **2020**, *13*, 6241. [\[CrossRef\]](#)
26. Jain, S.; Panda, R.; Tripathy, R.K. Multivariate sliding-mode singular spectrum analysis for the decomposition of multisensor time series. *IEEE Sens. Lett.* **2020**, *4*, 1–4.
27. Moghram, I.; Rahaman, S. Analysis and evaluation of five short-term load forecasting techniques. *IEEE Trans. Power Syst.* **1989**, *4*, 1484–1491. [\[CrossRef\]](#)
28. Zhang, J.; Wei, Y.-M.; Li, D.; Tan, Z.; Zhou, J. Short term electricity load forecasting using a hybrid model. *Energy* **2018**, *158*, 774–781. [\[CrossRef\]](#)
29. Shi, H.; Xu, M.; Li, R. Deep learning for household load forecasting—A novel pooling deep RNN. *IEEE Trans. Smart Grid* **2017**, *9*, 5271–5280. [\[CrossRef\]](#)
30. Xie, K.; Yi, H.; Hu, G.; Li, L.; Fan, Z. Short-term power load forecasting based on Elman neural network with particle swarm optimization. *Neurocomputing* **2020**, *416*, 136–142. [\[CrossRef\]](#)

31. Shang, Z.; He, Z.; Song, Y.; Yang, Y.; Li, L.; Chen, Y. A Novel Combined Model for Short-Term Electric Load Forecasting Based on Whale Optimization Algorithm. *Neural Process. Lett.* **2020**, *52*, 1207–1232. [[CrossRef](#)]
32. Wu, L.; Kong, C.; Hao, X.; Chen, W. A short-term load forecasting method based on GRU-CNN hybrid neural network model. *Math. Probl. Eng.* **2020**, *2020*, 10. [[CrossRef](#)]
33. Wang, Z.; Wang, X.; Ma, C.; Song, Z. A Power Load Forecasting Model Based on FA-CSSA-ELM. *Math. Probl. Eng.* **2021**, *2021*, 14.
34. Zhang, H.; Yang, Y.; Zhang, Y.; He, Z.; Yuan, W.; Yang, Y.; Qiu, W.; Li, L. A combined model based on SSA, neural networks, and LSSVM for short-term electric load and price forecasting. *Neural. Comput. Appl.* **2021**, *33*, 773–788. [[CrossRef](#)]
35. Han, Y.; Jing, Y.; Li, K.; Dimirovski, G.M. Network traffic prediction using variational mode decomposition and multi-reservoirs echo state network. *IEEE Access* **2019**, *7*, 138364–138377. [[CrossRef](#)]
36. Vautard, R.; Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D* **1989**, *35*, 395–424. [[CrossRef](#)]
37. Xu, L.; Yu, X.; Gulliver, T.A. Intelligent outage probability prediction for mobile IoT networks based on an IGWO-elman neural network. *IEEE Trans. Veh. Technol.* **2021**, *70*, 1365–1375. [[CrossRef](#)]
38. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
39. Shen, Q.; Ban, X.; Guo, C. Urban Traffic Congestion Evaluation Based on Kernel the Semi-Supervised Extreme Learning Machine. *Symmetry* **2017**, *9*, 70. [[CrossRef](#)]
40. Zhou, S. Sparse LSSVM in primal using Cholesky factorization for large-scale problems. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 783–795. [[CrossRef](#)]
41. Tian, Z. Short-term wind speed prediction based on LMD and improved FA optimized combined kernel function LSSVM. *Eng. Appl. Artif. Intell.* **2020**, *91*, 103573. [[CrossRef](#)]
42. Liu, G.; Shu, C.; Liang, Z.; Peng, B.; Cheng, L. A modified sparrow search algorithm with application in 3d route planning for UAV. *Sensors* **2021**, *21*, 1224. [[CrossRef](#)]
43. Wang, P.; Zhang, Y.; Yang, H. Research on Economic Optimization of Microgrid Cluster Based on Chaos Sparrow Search Algorithm. *Comput. Intell. Neurosci.* **2021**, *2021*, 18.
44. Zhang, C.; Ding, S. A stochastic configuration network based on chaotic sparrow search algorithm. *Knowl.-Based Syst.* **2021**, *220*, 106924. [[CrossRef](#)]
45. Tizhoosh, H.R. Opposition-based learning: A new scheme for machine intelligence. In Proceedings of International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents (CIMCA-IAWTIC'06), Vienna, Austria, 28–30 November 2005; pp. 695–701.
46. Kang, Q.; Xiong, C.; Zhou, M.; Meng, L. Opposition-based hybrid strategy for particle swarm optimization in noisy environments. *IEEE Access* **2018**, *6*, 21888–21900. [[CrossRef](#)]
47. Sihwail, R.; Omar, K.; Ariffin, K.A.Z.; Tubishat, M. Improved harris hawks optimization using elite opposition-based learning and novel search mechanism for feature selection. *IEEE Access* **2020**, *8*, 121127–121145. [[CrossRef](#)]
48. Bao, X.; Jia, H.; Lang, C. Dragonfly algorithm with opposition-based learning for multilevel thresholding Color Image Segmentation. *Symmetry* **2019**, *11*, 716. [[CrossRef](#)]
49. Seif, Z.; Ahmadi, M.B. An opposition-based algorithm for function optimization. *Eng. Appl. Artif. Intell.* **2015**, *37*, 293–306. [[CrossRef](#)]
50. Bertsimas, D.; Tsitsiklis, J. Simulated annealing. *Stat. Sci.* **1993**, *8*, 10–15. [[CrossRef](#)]