

Research Article

A Power Load Forecasting Model Based on FA-CSSA-ELM

Zuoxun Wang , Xinheng Wang, Chunrui Ma, and Zengxu Song

School of Electrical Engineering and Automation, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China

Correspondence should be addressed to Zuoxun Wang; wangzuoxun@126.com

Received 6 March 2021; Revised 30 March 2021; Accepted 10 April 2021; Published 24 April 2021

Academic Editor: Yi Qi

Copyright © 2021 Zuoxun Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate and stable power load forecasting methods are essential for the rational allocation of power resources and grid operation. Due to the nonlinear nature of power loads, it is difficult for a single forecasting method to complete the forecasting task accurately and quickly. In this study, a new combined model for power loads forecasting is proposed. The initial weights and thresholds of the extreme learning machine (ELM) optimized by the chaotic sparrow search algorithm (CSSA) and improved by the firefly algorithm (FA) are used to improve the forecasting performance and achieve accurate forecasting. The early local optimum that exists in the sparrow algorithm is overcome by Tent chaotic mapping. A firefly perturbation strategy is used to improve the global optimization capability of the model. Real values from a power grid in Shandong are used to validate the prediction performance of the proposed FA-CSSA-ELM model. Experiments show that the proposed model produces more accurate forecasting results than other single forecasting models or combined forecasting models.

1. Introduction

Nowadays, power loads have reached almost every corner of human society and have brought great convenience to mankind. And reasonable power loads planning will bring great convenience to human society, and wrong power loads planning will bring great loss to human society. Inaccurate power load forecasting will result in incorrect load planning and layout by the authorities. This will cause huge economic losses and wasted energy. Therefore, accurate forecasting of power loads has been a hot topic in power system planning. Since power loads cannot be stored on a large scale, this leads to low utilization of electric resources. Accurate power load forecasting results can provide correct feedback and decision-making for the power sector. It can also help achieve a reasonable dynamic balance between electricity production and electricity consumption. [1–5].

Initially, a series of traditional methods of forecasting power loads were proposed by many experts and scholars. Traditional methods include trend extrapolation [6, 7], exponential smoothing [8, 9], Kalman filtering [10], and ARIMA [11–13]. These traditional methods of forecasting electrical loads have the advantage of being simple to be

calculated and easy to be implemented. However, traditional methods have the disadvantage of low prediction accuracy, which will cause management departments to be unable to make reasonable and accurate decisions. It also makes it difficult to play a key role in the planning and rational allocation of electrical loads.

With the continuous development of artificial intelligence, many researchers have realized the application potential of intelligent models in dealing with nonlinear and complicated problems. Elman proposed an intelligent model of the ELMAN [14]. Noble proposed an intelligent model of the support vector machine (SVM) [15]. Huang proposed an intelligent model of the extreme learning machine (ELM) [16]. Suykens and his partners proposed an intelligent model of the least squares support vector machine (LSSVM) [17]. Shi and his partners proposed an intelligent model of the recurrent neural networks (RNNs) [18].

Many experts have successfully applied intelligent models to complicated big data and nonlinear problems in power loads forecasting. A method of predicting short-term power loads using SVM had been proposed by Ye and his partners [19]. Using support vector machines can reflect the characteristics of important characteristics of power load to

establish a forecasting model. A new power load forecasting model had been proposed by Wei Li and his partners [20]. The training set was constructed using variational pattern decomposition, and then the decomposed data were fed into the ELM model to construct the prediction model. This forecasting model utilizes an extreme learning machine (ELM) combined with variational mode decomposition (VMD) to forecast power load models. An LSSVM-based model for power loads forecasting had proposed by Xuemei Li and his partners. The model was compared with a back propagation neural network (BPNN) and verified to have better prediction accuracy and generalization ability [21].

Nowadays, the prediction accuracy of a single intelligent forecasting model for power loads is no longer sufficient to meet the normal needs of the power system. So, many scholars have turned their attentions to swarm intelligence optimization algorithms [22]. The research shows that the swarm intelligence optimization algorithm has the characteristics of simple principle, easy realization, strong adaptability, and high efficiency. Therefore, swarm intelligence optimization algorithms are often used to optimize the parameters of a single power load forecasting model by scholars. Swarm intelligence optimization algorithms are mainly derived from the habits of organisms in nature. Although the capacity of a single individual is limited, populations can perform well when they work together. Common swarm intelligence optimization algorithms include the ant colony optimization (ACO) [23], the artificial bee colony algorithm (ABC) [24], the firefly algorithm (FA) [25], the bat algorithm (BA) [26], the cuckoo search (CS) [27], the grey wolf optimization (GWO) [28], the dragonfly algorithm (DA) [29], the whale optimization algorithm (WOA) [30], and the sparrow search algorithm (SSA). And SSA was a new swarm intelligence optimization algorithm proposed by Xue in 2020 [31].

As a result, a series of combinatorial models based on population intelligence optimization algorithms have been proposed by scholars to predict power loads. A method of forecasting short-term electricity loads using WOA optimized long- and short-term memory (LSTM) artificial neural networks was proposed by Haiyan [32]. A chaotic artificial bee colony algorithm to optimize the support vector regression (SVR) short-term power prediction model was proposed by Hong [33]. An improved grey wolf algorithm to optimize support vector machines for short-term power loads forecasting models was proposed by Jiang [34]. It can be found from the above research that the combined forecasting model can well meet the requirements of forecast accuracy and provide correct feedback and information for the power sector.

The SSA algorithm is a new swarm intelligence optimization algorithm, which simulates the foraging and antipredation behavior of sparrows, and is superior to particle swarm optimization (PSO) and GWO algorithms in terms of finding the best performance. The SSA algorithm, like other swarm intelligence algorithms, suffers from poor convergence accuracy and tends to fall into local optima. In

this paper, Tent chaotic mapping is used to initialize the sparrow population. Chaos theory has been applied in many ways, especially to deal with nonlinear problems [35–38]. The initial population can be uniformly distributed in the solution space by using chaotic property. This will help the algorithm converge quickly and jump out of local optimality. And the firefly perturbation strategy is used to update the population position. The global optimization ability and convergence speed of the sparrow search algorithm are improved by using the characteristics of the firefly algorithm.

As a single-layer feedforward neural network (SLFN) [39], ELM has more powerful generalization ability than other traditional neural networks. And ELM is also faster than other neural network models while maintaining learning accuracy. This makes ELM ideal for problems with large amounts of data, such as power load forecasting. Therefore, the FA-CSSA algorithm is used to optimize the initial weights and thresholds of the ELM model. The powerful global search capability of the FA-CSSA algorithm is used to improve the generalization capability of the model and hence the predictive capability of the overall combined power loads forecasting model.

Therefore, this paper addresses the SSA algorithm, the ELM neural network model, and the FA-CSSA algorithm. A new FA-CSSA-ELM electric load forecasting model and the corresponding feedback mechanism for power supply are proposed. And the real load history data of a certain power grid in Shandong is used as the simulation data to verify the prediction performance of the model. In order to better illustrate the excellent performance and accuracy of the FA-CSSA-ELM combined power load forecasting model, in this paper, the prediction results are compared and discussed with those of three single prediction models and two combined prediction models, respectively. The results demonstrate that the FA-CSSA-ELM power load model possesses better prediction accuracy than the other five models.

2. Chaotic Sparrow Algorithm Improved by Firefly Algorithm

2.1. Sparrow Search Algorithm. The SSA algorithm is made up of three components: a spotter, a tracker, and a vigilante. Suppose there are N sparrows in a D -dimensional space. Then, the sparrow flock can be expressed as the following equation:

$$X = [x_1, \dots, x_2, \dots, x_N]^T. \quad (1)$$

Then, the position of the i -th sparrow in the D -dimensional search space can be expressed as the following equation:

$$X_i = [x_{i1}, \dots, x_{id}, \dots, x_{iD}], \quad (2)$$

where x_{id} is the position of the i -th sparrow in dimension d . So, the position update formula can be expressed as the following equation:

$$x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{-i}{\partial \cdot T}\right), & R_2 < ST, \\ x_{id}^t + Q \cdot L, & R_2 \geq ST, \end{cases} \quad (3)$$

where t denotes the current number of iterations, T denotes the maximum number of iterations, ∂ is the 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 $[(1/2), 1]$.

It is generally assumed that discoverers make up about 10–20% of the population, with the rest belonging to trackers. The tracker's position update formula can be expressed as the following equation:

$$x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{xw_d^t - x_{id}^t}{i^2}\right), & i > \frac{n}{2}, \\ xb_d^{t+1} + \frac{1}{D} \sum_{d=1}^D (\text{rand}\{-1, 1\} \cdot |x_{id}^t - xb_d^{t+1}|), & i \leq \frac{n}{2}, \end{cases} \quad (4)$$

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 $i > (n/2)$, it means that the population is short of food and needs to go elsewhere to forage. When $i \leq (n/2)$, it means that the tracker is preying near the optimal position xb .

The last guards are used for vigilant reconnaissance of the population and number 10–20% of the total population. Its position update formula can be expressed as the following equation:

$$x_{id}^{t+1} = \begin{cases} xb^t + \beta(x_{id}^t - xb_d^t), & f_i \neq f_g, \\ x_{id}^t + K \left(\frac{x_{id}^t - xw_d^t}{|f_i - f_w| + \mu} \right), & f_i = f_g, \end{cases} \quad (5)$$

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 constant that prevents the denominator from going to zero; here, in this paper, we take $10E - 8$, and f_i is the current fitness, f_g is the best fitness, and f_w is the worst fitness.

The flowchart of sparrow algorithm operation is shown in Figure 1.

2.2. Extreme Learning Machine. The extreme learning machine is an SLFN with faster learning speed and higher generalization capability. Assume that any N different training set (x_j, t_j) , $x_j \in R^d, t_j \in R^m$, the mathematical model of SLFN with n hidden nodes can be defined as

$$\sum_{i=1}^n \beta_i g_i(x_j) = \sum_{i=1}^n \beta_i G_i(a_i, b_i, x_j), \quad j = 1, 2, \dots, N, \quad (6)$$

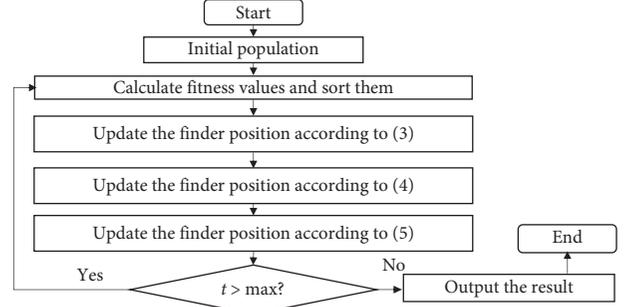


FIGURE 1: Flow chart of SSA algorithm operation.

where a_i is the vector of weights connecting the i -th hidden node to the input node, b_i is the threshold value of the i -th hidden node, β_i is the weight vector connecting the i -th hidden node to the output node, $g_i(x_j) = G_i(a_i, b_i, x_j)$ is the output function of the i -th hidden node, and $g(\bullet)$ is the sigmoid activation function.

Since SLFN can approach these N training samples with zero error, equation (6) can be further defined as the following equation:

$$\sum_{i=1}^n \beta_i G_i(a_i, b_i, x_j) = t_j, \quad j = 1, 2, \dots, N, \quad (7)$$

where t_j is the output function. In addition, equation (7) can compactly express N equations as equation (8) which is given as follows:

$$H\beta = T, \quad (8)$$

$$H = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_n, b_n, x_1) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_n, b_n, x_N) \end{bmatrix},$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}_{n \times m}, \quad (9)$$

$$T = \begin{bmatrix} t_1 \\ \vdots \\ t_N \end{bmatrix}_{n \times N},$$

where H is the hidden layer output matrix of the network. SLFN has been shown to have universal approximation capability, and ELM network training process can be summarized as a nonlinear optimization problem. Its input weight a_i and hidden threshold b_i can be assigned randomly. Training_SSFN is equivalent to finding the least squares solution β for the linear system $H\beta = T$. The mathematical model of the least squares solution can be defined as

$$\hat{\beta} = H^+ T, \quad (10)$$

where H^+ is the Moore–Penrose generalized inverse of the hidden layer output matrix and T is the expected output matrix.

2.3. Other Recommendations. In this paper, the SSA algorithm is optimized by Tent chaotic mapping strategy and firefly perturbation strategy, and an improved firefly chaotic sparrow algorithm is proposed. The FA-CSSA model uses the chaotic properties of the Tent mapping to initialize the population. The chaotic nature of the Tent mapping is used to make the initial population uniformly distributed in the solution space. And the firefly algorithm is used to update the optimal sparrow and sparrow flock position based on the principle that the fireflies with higher brightness in the search space can attract the fireflies with lower brightness to approach. Therefore, the chaotic mapping and firefly disturbance strategy can make up for the shortcomings of the SSA algorithm that it is easy to fall into the local optimum and can enhance the algorithm's global optimization ability and robustness.

2.3.1. Tent Chaos Mapping Strategy. It has been found that the goodness of the initial population profoundly affects the convergence process of the swarm intelligence optimization algorithm [39]. The SSA algorithm is a new swarm intelligence optimization algorithm proposed in 2020. Therefore, the SSA algorithm also suffers from the fact that the initial populations cannot be uniformly distributed in the solution space. This can lead to a lack of population diversity in the processing of the algorithm. So, the SSA algorithm has the disadvantage of low solution efficiency and insufficient global optimization capability when solving complex optimization problems.

As chaos is nonlinear, random, and ergodic [40], it can well allow the initial population to be traversed within the entire space. Therefore, this paper uses the strategy of chaotic mapping to initially optimize the SSA algorithm. In contrast to other types of chaotic mappings, the Tent chaotic mapping has a simple structure and the mapping presents a more uniform density. Tent chaos mapping distribution is shown in Figure 2, and Tent chaos mapping bifurcation diagram is shown in Figure 3. This indicates that Tent chaotic mappings have strong chaotic properties, ergodicity, and iteration speed. Therefore, this paper chooses the Tent chaotic map to avoid the SSA algorithm from falling into the local optimum in the iterative process.

Let the chaotic time series in the space of D dimensions be $x = \{x_n, n = 1, 2, \dots, D\}$, and the Tent chaos mapping can be expressed as the following equation:

$$x_{n+1} = \begin{cases} 2x_n, & 0 \leq x_n < 0.5, \\ 2(1 - x_n), & 0.5 \leq x_n \leq 1. \end{cases} \quad (11)$$

2.3.2. Firefly Disturbance Strategy. In the firefly disturbance strategy [25], the main purpose of the light emitted by fireflies is to act as a light-signal system to attract other individual fireflies. And all fireflies follow the following three points:

- (1) All fireflies are attracted to fireflies that are brighter than them.

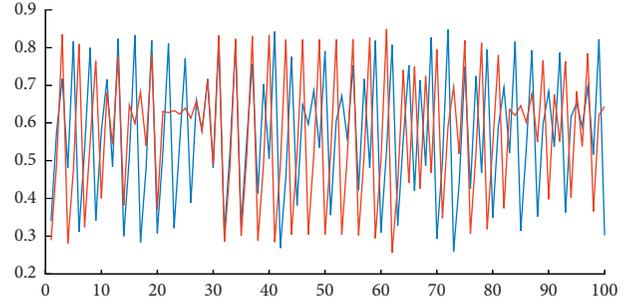


FIGURE 2: Tent chaos mapping distribution.

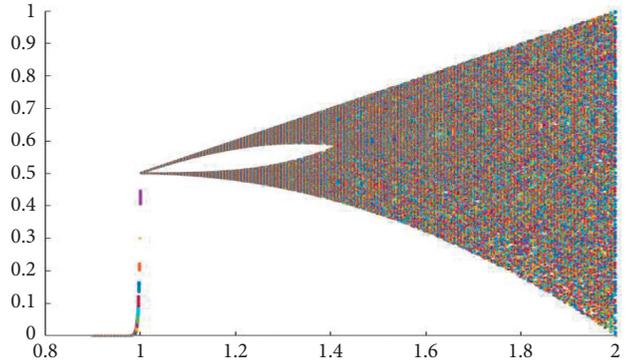


FIGURE 3: Tent chaotic mapping bifurcation diagram.

- (2) The attractiveness of fireflies is directly proportional to their brightness. When a firefly approaches a firefly that is brighter than itself, the firefly's brightness decreases with distance.
- (3) If no brighter firefly is found than the given one, then it will move randomly.

So, the formula for the relative luminosity of fireflies can be expressed as follows:

$$I = I_0 * e^{-\gamma r_{i,j}}. \quad (12)$$

The formula for the attractiveness of fireflies can be expressed as follows:

$$\beta = \beta_0 * e^{-\gamma r_{i,j}^2}. \quad (13)$$

The formula for updating the position of a firefly can be expressed as follows:

$$x_i(t+1) = x_i + \beta \times (x_j - x_i) + \alpha \times (\text{rand} - 0.5), \quad (14)$$

where I_0 is the maximum brightness of the firefly and proportional to the objective function value, γ is the light intensity absorption parameter, $r_{i,j}$ is the distance between fireflies i and j and is the maximum attraction, x_i and x_j are the spatial locations where fireflies i and j are located, respectively, α is a step factor in the range $[0, 1]$, and rand is a random number between $[0, 1]$.

The firefly perturbation strategy is used to update the positions of the optimal sparrows and sparrow flocks to improve the search capability of the algorithm. Finally, the

sparrow positions after the firefly perturbation strategy are compared with the sparrow positions without the firefly perturbation strategy. If the result is better, the sparrow positions are updated.

So, the flow chart of the operation of the FA-CSSA algorithm improved according to the Tent chaos mapping strategy and the firefly perturbation strategy is shown in Figure 4.

3. FA-CSSA-ELM Power Load Forecasting Model and Feedback Mechanism

3.1. The FA-CSSA-ELM Power Load Forecasting Model. The FA-CSSA algorithm is used to optimize the initial weights and thresholds of the ELM model to construct the FA-CSSA-ELM power load prediction model. The FA-CSSA algorithm has strong global search ability, which can improve the generalization ability of the model. And it can further improve the forecasting capability of the FA-CSSA-ELM power load forecasting model.

The specific forecasting steps of the FA-CSSA-ELM power load forecasting model can be expressed as follows:

- (1) Divide the validation data into datasets and test sets.
- (2) Construct the FA-CSSA-ELM prediction model. The SSA algorithm optimized by chaos mapping strategy and firefly disturbance strategy is used to find the optimal initial weight and threshold of the ELM model.
- (3) The real historical data of a certain power grid in Shandong were used to verify the prediction performance of the FA-CSSA-ELM model and other comparison prediction models, and four performance index functions were used as qualitative comparison standards.
- (4) The FA-CSSA-ELM power load forecasting model proposed is applied to the real power load transmission process. The accurate forecasting capability of the FA-CSSA-ELM load forecasting model is used to forecast real power load data. The forecast trends and results are used to provide feedback on the electricity consumption of each region to ensure maximum utilization of the electricity load. This can better achieve the purpose of saving energy and reducing consumption and reducing economic losses.

3.2. The Evaluation Functions. In order to judge the prediction effect of different competitive models more accurately and comprehensively, 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. Moreover, RMSE is highly sensitive to the accuracy of the prediction. MAPE is highly expressive of the prediction. The four evaluation functions are shown in Table 1.

3.3. Power Load Feedback System for Forecasting Models. Typically, the power load transmission process in this paper is shown in Figure 5. Firstly, the power plant transmits the power load through the 220 kV high-voltage transmission line to the first-stage substation for the first power load conversion. Then, the converted power load is transmitted through the 110 kV high-voltage transmission line to the secondary substation for the second power conversion. Finally, the power load of the second conversion will be transmitted to each electricity place. The proposed FA-CSSA-ELM power load prediction model is applied to the power load conversion process of the first-stage substation. Through real-time data update and accumulation in various places, the model can be continuously learned and updated, and the prediction accuracy of the model can be continuously improved, and the dynamic balance of power generation and power supply can be achieved. In this way, relevant departments can accurately predict the changing trend of power load and accurate power load value according to the history of power load. Through accurate prediction, we can not only give reasonable suggestions and guidance to relevant departments but also make the power load distribution more reasonable and maximize the use of power resources.

4. Simulation Experiments

In order to better verify the predictive performance of the FA-CSSA-ELM model proposed in this paper, the combined forecasting model FA-CSSA-ELM is compared with single competing models such as ELMAN, ELM, and SVM. In order to give a more comprehensive picture of the forecasting performance of the proposed FA-CSSA-ELM model, this paper also compares it with the two combined competing models WOA-ELM and PSO-ELMAN. The simulation experimental part consists of two parts: the data description section and the simulation experimental section. The data description section introduces the data used in the simulation experiments, as well as the specific way of dividing the training set and the test set. The experimental part consists of two parts, Experiment I and Experiment II, describing the specific steps of the predicted performance tests and analyzing the results of the simulated experiments.

4.1. Data Description Section

4.1.1. Presentation of Simulation Data. This paper uses real electrical load history data of four weeks from a power grid in Shandong in 2020 as simulation data. In order to predict the electrical load data more accurately, the frequency interval for collection in this paper is 5 minutes. A total of 8064 electrical load history data were measured for 288 electrical load history data per day. The power load time series is shown in Figure 6.

From Figure 6, this paper shows that the power load data are highly nonlinear and regular.

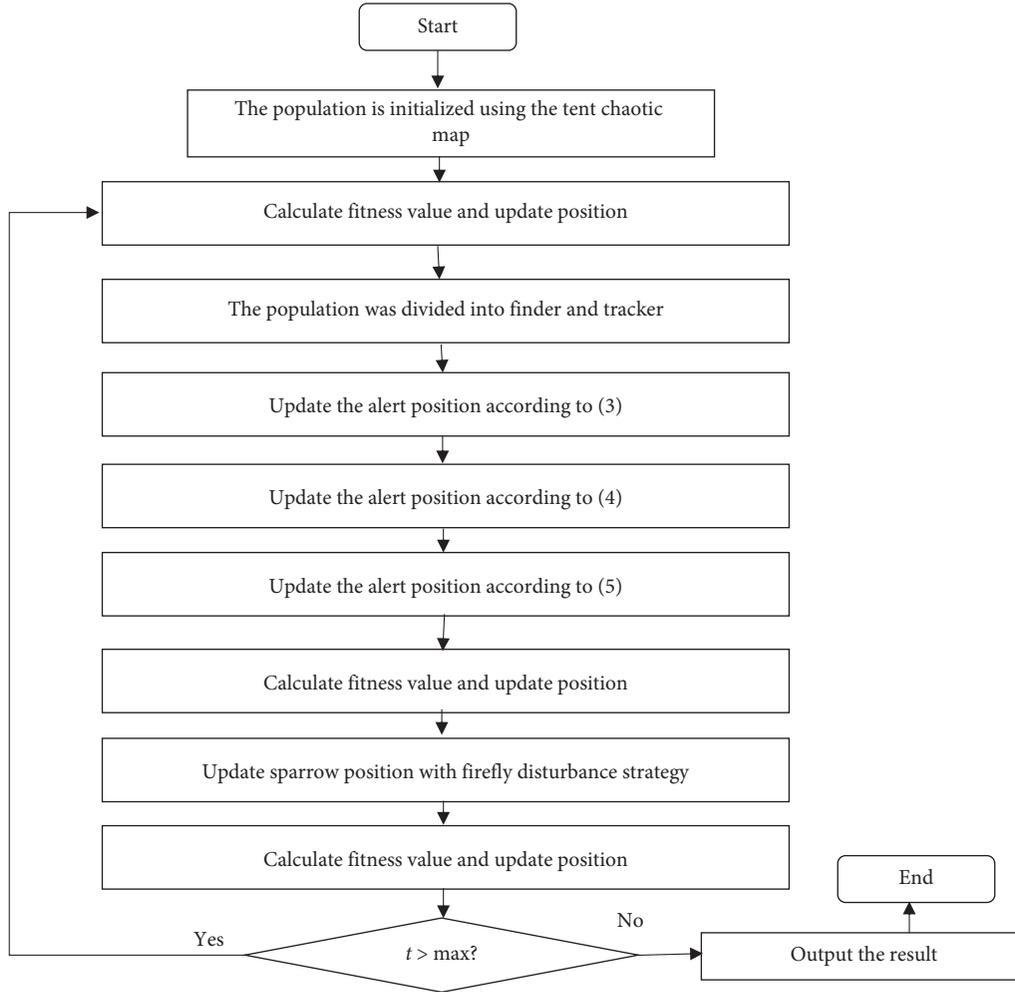


FIGURE 4: Flow chart of the FA-CSSA algorithm operation.

TABLE 1: Four types of evaluation functions.

Metrics	Definition	Equation
RMSE	The square root of average of the error squares	$RMSE = \sqrt{(1/N) \sum_{i=1}^N (t_i - p_i)^2}$
MAPE	The average of absolute percentage error	$MAPE = (100\%/N) \sum_{i=1}^N (t_i - p_i/t_i) $
MSE	The square root of the mean of the sum of squares of the errors	$MSE = (1/N) \sum_{i=1}^N (t_i - p_i)^2$
MAE	The average value of the absolute error between the observed value and the true value	$MAE = (1/N) \sum_{i=1}^N t_i - p_i $

t_i : i -th sample of expected output; p_i : i -th sample of predicted output; N : sample size.

4.1.2. Division of the Dataset. The dataset is divided into two sections: the training set and the test set. The training set is used to learn and train the model, and the test set is used to verify the training effect of the model. In order to make the distribution of power loads more rational and the forecasts more accurate, in this paper, the measured 8064 real power load history data of a power grid in Shandong Province were divided into 7 time series of data subsets. The 7 data subsets are created in the chronological order from Monday to Sunday. Each time series was recorded every 5 minutes for a total of 4 days. Each day has 288 data, and each set has 1152 data. By dividing the data in this way, the prediction units in this paper have been refined from months or weeks to a specific day. This not only improves the accuracy and

relevance of the model predictions but also provides more reasonable suggestions for the allocation of power loads.

This paper divides the 8064 historical power load data into 7 subsets from Monday to Sunday. So, each subset has 1152 power load history data. In this paper, the data from the first three weeks are used as the test set data and the data from the last week are used as the validator data. For example, the test set for the first subset is the data for each Monday of the first three weeks, and the validator data are the data for Monday of the fourth week. The test set for the second subset is the data from Tuesday of the previous three weeks, and the validator data are the data from Tuesday of the fourth week. The remaining subsets of test and validation sets are divided according to this pattern.

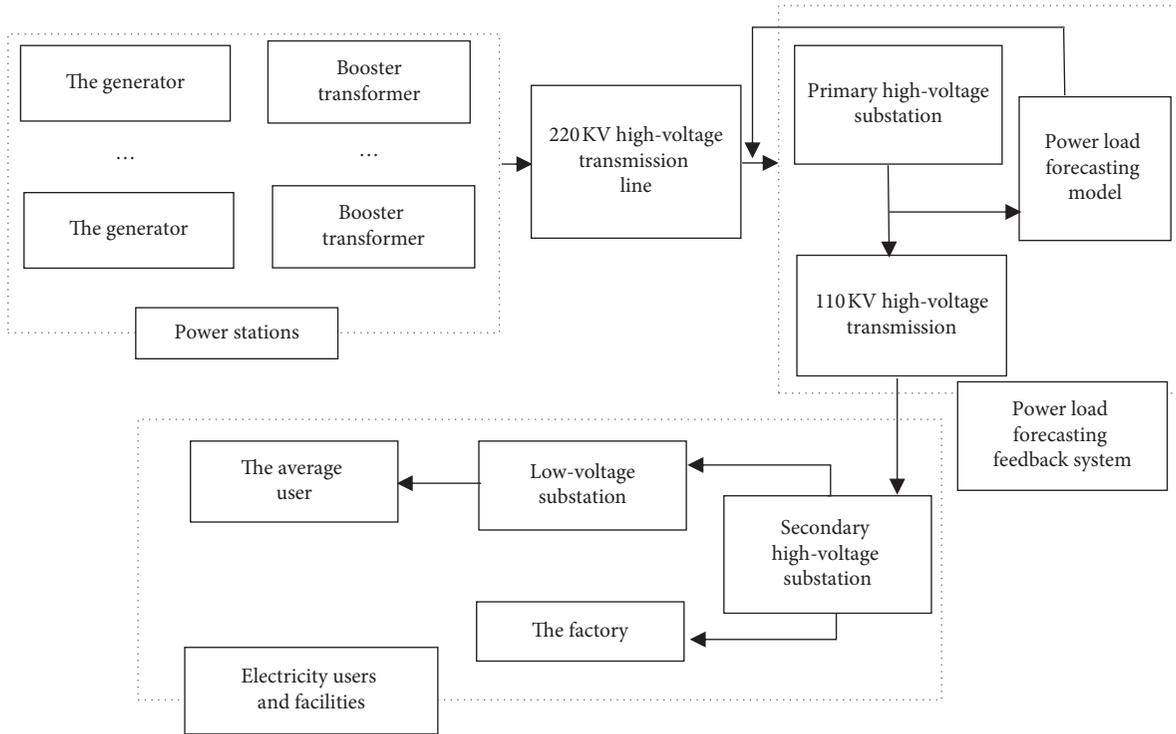


FIGURE 5: Electric load transfer process.

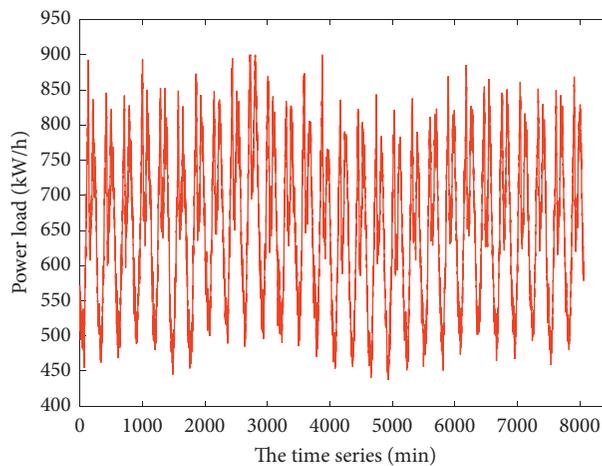


FIGURE 6: Real power load data for a grid in Shandong.

4.2. Simulation Experiments Section. The experimental simulation part introduces the prediction effect comparison between the FA-CSSA-ELM prediction model proposed in this paper and other competitive models.

4.2.1. Experiment I. The purpose of Experiment I is to compare the performance of the FA-CSS-ELM model with that of the single prediction model. And the single prediction models include ELMAN, ELM, and SVM. In order to make the data more accurate and representative, the data in the tables of this paper are calculated by averaging 20 operations. The metrics of the four evaluation functions compared with

the single competing model are shown in Table 2 (the best data are highlighted in this paper).

A comparison of the data for the four indicators from the FA-CSSA-ELM model proposed in this paper with the single-competition model is shown in Table 2. The FA-CSSA-ELM model is the most effective, followed by the ELM single-competition model. And the SVM single prediction model is the least effective. Through data comparison, we can find that the FA-CSA-ELM model is superior to the single prediction model in all indicators. For a more visual observation, a histogram of the mean of the evaluation functions for these seven datasets is also plotted in this paper to represent it. And the histogram is shown in Figure 7.

TABLE 2: Experiment I: comparison of data with the evaluation function of a single competitive model.

Data	MSE	MAPE	RMSE	MAE
Monday				
ELMAN	1.5202	3.1632	3.1632	20.5353
ELM	0.9305	1.9381	1.9381	12.6597
SVM	5.7933	13.0061	13.0061	80.9685
FA-CSSA-ELM	0.77477	1.6099	1.6099	9.8266
Tuesday				
ELMAN	1.4905	3.1364	3.1364	20.2276
ELM	6.307	2.2570	2.2570	13.8804
SVM	5.7933	13.9705	13.9705	88.3321
FA-CSSA-ELM	0.8733	1.8208	1.8208	12.3912
Wednesday				
ELMAN	2.4602	5.1642	5.1642	33.698
ELM	1.3248	2.7226	2.7226	17.6342
SVM	5.307	10.9705	10.9705	67.3321
FA-CSSA-ELM	1.061	2.2299	2.2299	14.6999
Thursday				
ELMAN	1.8634	3.977	3.977	25.2004
ELM	1.3149	2.8083	2.8083	18.0642
SVM	6.3128	14.2424	14.2424	89.8527
FA-CSSA-ELM	1.118	2.3729	2.3729	15.6980
Friday				
ELMAN	1.3959	3.0618	3.0618	19.5877
ELM	1.0001	2.201	2.201	13.9655
SVM	5.7571	12.6769	12.6769	80.5392
FA-CSSA-ELM	0.7942	1.8056	1.8056	11.7099
Saturday				
ELMAN	1.674	3.646	3.646	23.0146
ELM	1.1409	2.4056	2.4056	15.4727
SVM	5.7579	13.1629	13.1629	83.5783
FA-CSSA-ELM	1.0377	2.2249	2.2249	14.5872
Sunday				
ELMAN	1.6556	3.536	3.536	22.5124
ELM	1.9696	2.4904	2.4904	15.9062
SVM	5.7579	13.1629	13.1629	83.5783
FA-CSSA-ELM	1.0326	2.1783	2.1783	14.0861

A comparison of the different competition model performance metric functions in Figure 7 shows that the FA-CSSA-ELM improved 72.29%, 99.8%, and 478.2% in MSE metrics compared with the other three single prediction models ELMAN, ELM, and SVM, respectively. The FA-CSSA-ELM improved by 80.3%, 18%, and 540% in the MAPE metric compared with the remaining three single prediction models ELMAN, ELM, and SVM, respectively. In the RMSE metric, the improvement is 77.8%, 14%, and 488% compared with the other three single forecasting models ELMAN, ELM, and SVM, respectively. In terms of MAE metric, the improvement is 77.2%, 15.6%, and 517% for ELMAN, ELM, and SVM, respectively. From the comparison data, the FA-CSSA-ELM model proposed in this paper is much more effective than the three representative single forecasting models compared.

4.2.2. Experiment II. The purpose of Experiment II is to compare the FA-CSSA-ELM model with other representative combinatorial competition models. The combined

prediction models include WOA-ELM and PSO-ELMAN. In this paper, four performance indicators are used to verify the superiority of the model. And the evaluation function pairs of the three competitive models are shown in Table 3 (the best data are marked in bold in this paper). For a more intuitive view, a histogram of the mean values of the four evaluation functions for these seven datasets is also plotted in this paper. The histogram is shown in Figure 8.

It is shown in Figure 8 and Table 3 that the FA-CSSA-ELM prediction model proposed has superiority in all evaluation indicators, and it is more stable and has accurate prediction results in this paper. The PSO-ELMAN combined model, on the other hand, has the least satisfactory evaluation indexes and the lowest prediction accuracy. Although the WOA-ELM competition model also has excellent prediction results, it still does not surpass the FA-CSSA-ELM prediction model in the comparison of various evaluation indicators. Compared with the WOA-ELM model and PSO-ELMAN model, the MSE index of the FA-CSSA-ELM model increased by

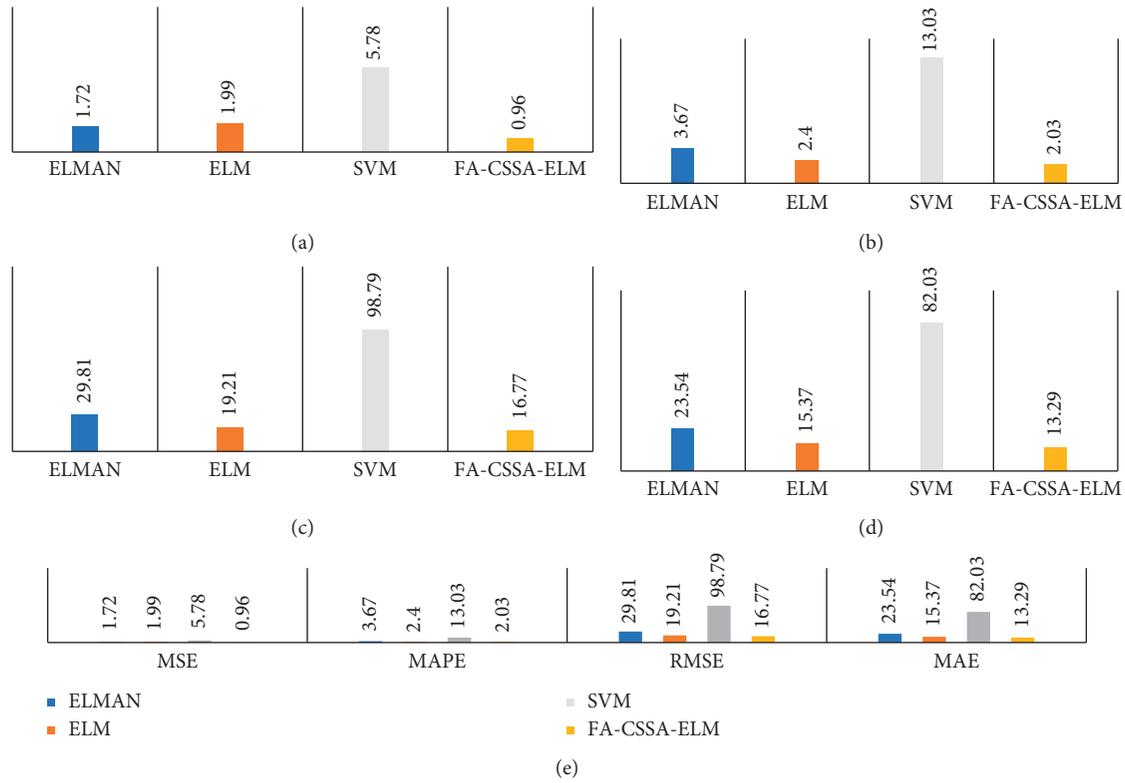


FIGURE 7: (a) Comparison chart of MSE data for four competitive models. (b) Comparison chart of MAPE data for four competitive models. (c) Comparison chart of RMSE data for four competitive models. (d) Comparison chart of MAE data for four competitive models. (e) Comparison diagram of MSE, MAPE, RMSE, and MAE of ELMAN, ELM, SVM, and FA-CSSA-ELM models.

TABLE 3: Experiment II: compared with the evaluation function of the combinatorial competition model.

Data	MSE	MAPE	RMSE	MAE
Monday				
WOA-ELM	0.82163	1.6683	13.9436	10.8239
PSO-ELMAN	1.5443	3.2796	26.2078	21.1148
FA-CSSA-ELM	0.77477	1.6099	13.8068	9.8266
Tuesday				
WOA-ELM	0.9408	1.9509	15.9538	12.9269
PSO-ELMAN	1.7102	3.6173	29.0235	22.9117
FA-CSSA-ELM	0.8733	1.8208	15.4092	12.3912
Wednesday				
WOA-ELM	1.1136	2.3014	18.9082	14.8435
PSO-ELMAN	1.7956	3.8768	30.4726	24.5816
FA-CSSA-ELM	1.061	2.2299	18.8049	14.6999
Thursday				
WOA-ELM	1.1359	2.3608	19.2711	15.3207
PSO-ELMAN	1.8188	3.8585	30.8664	24.3114
FA-CSSA-ELM	1.118	2.3559	18.7718	15.6980
Friday				
WOA-ELM	0.85777	1.8753	14.5568	11.8553
PSO-ELMAN	1.7447	3.9528	30.1052	24.7635
FA-CSSA-ELM	0.7942	1.8056	13.8266	11.7099
Saturday				
WOA-ELM	1.43121	2.3145	19.5556	15.0731
PSO-ELMAN	1.732	3.619	39.3924	22.9106
FA-CSSA-ELM	1.0377	2.2249	18.4252	14.5872
Sunday				
WOA-ELM	1.0807	2.2217	18.3403	14.09331
PSO-ELMAN	1.9889	4.3066	33.7536	27.6812
FA-CSSA-ELM	1.0326	2.1783	18.3226	14.0861

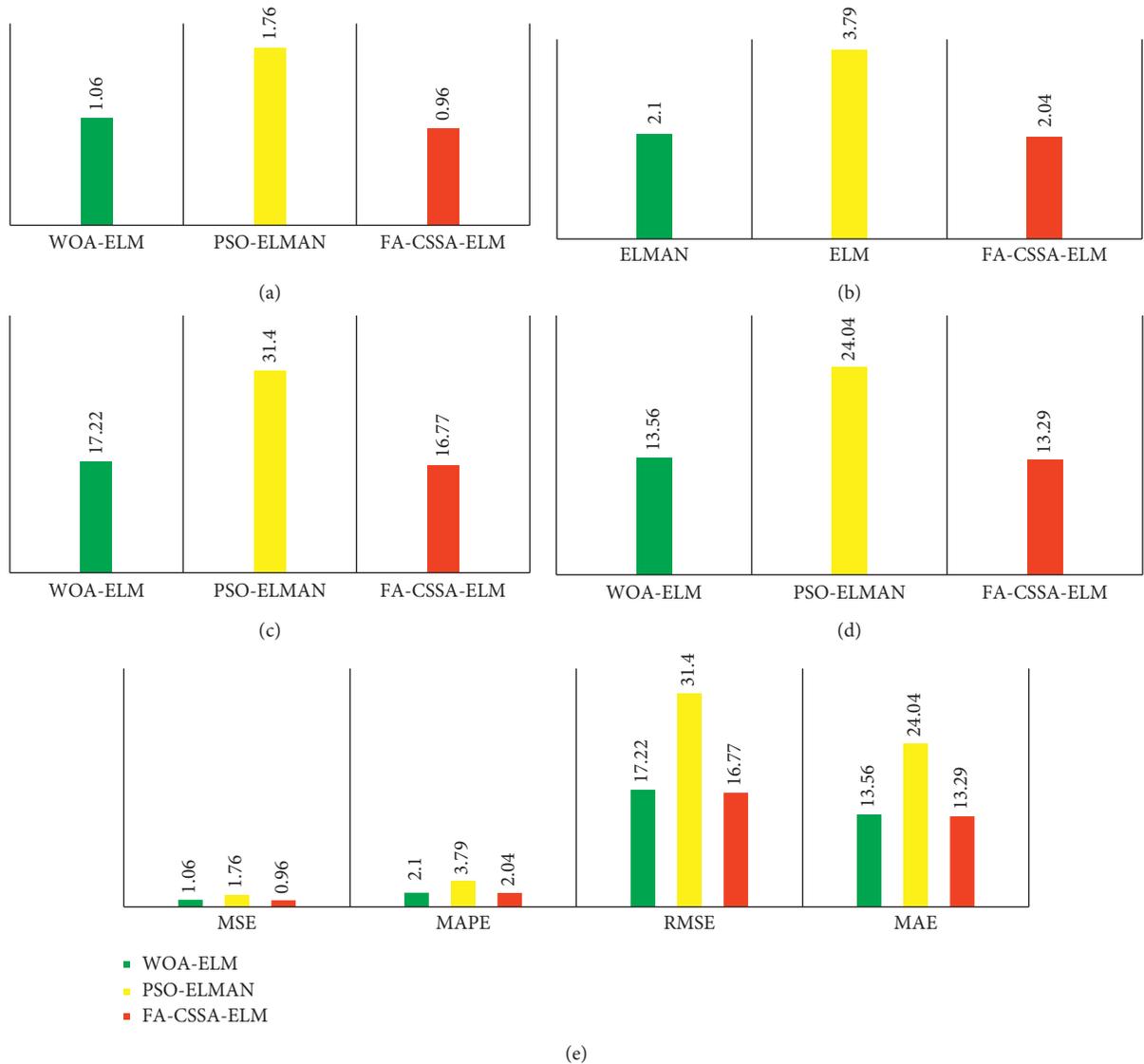


FIGURE 8: (a) Comparison chart of MSE data for three competitive models. (b) Comparison chart of MAPE data for three competitive models. (c) Comparison chart of RMSE data for three competitive models. (d) Comparison chart of MAE data for three competitive models. (e) Comparison diagram of MSE, MAPE, RMSE, and MAE of WOA-ELM, PSO-ELMAN, and FA-CSSA-ELM models.

10.2% and 87.3%, respectively. The MAPE index of the FA-CSSA-ELM model is increased by 3% and 86% compared with the WOA-ELM model and PSO-ELMAN model, respectively. Compared with WOA-ELM and PSO-ELMAN, the RMSE index of the FA-CSSA-ELM model increased by 2.6% and 89%, respectively. The MAE index of the FA-CSSA-ELM model compared with the WOA-ELM model and PSO-ELMAN model increased by 2.1% and 80.9%, respectively.

And this paper can also be more intuitively analyzed from the comparison graphs of the prediction effects of the

six different competing models shown in Figure 9. The combined power load forecasting model is better than the single power load forecasting model in both accuracy and stability. Among the combined models, the FA-CSSA-ELM model proposed in this paper is the most superior. The FA-CSSA-ELM forecasting model outperforms the rest of the competing models in all evaluation metrics and has strong forecasting accuracy and stability. Therefore, the FA-CSSA-ELM power load forecasting model proposed in this paper can give accurate power forecasts and correct feedback to the authorities concerned.

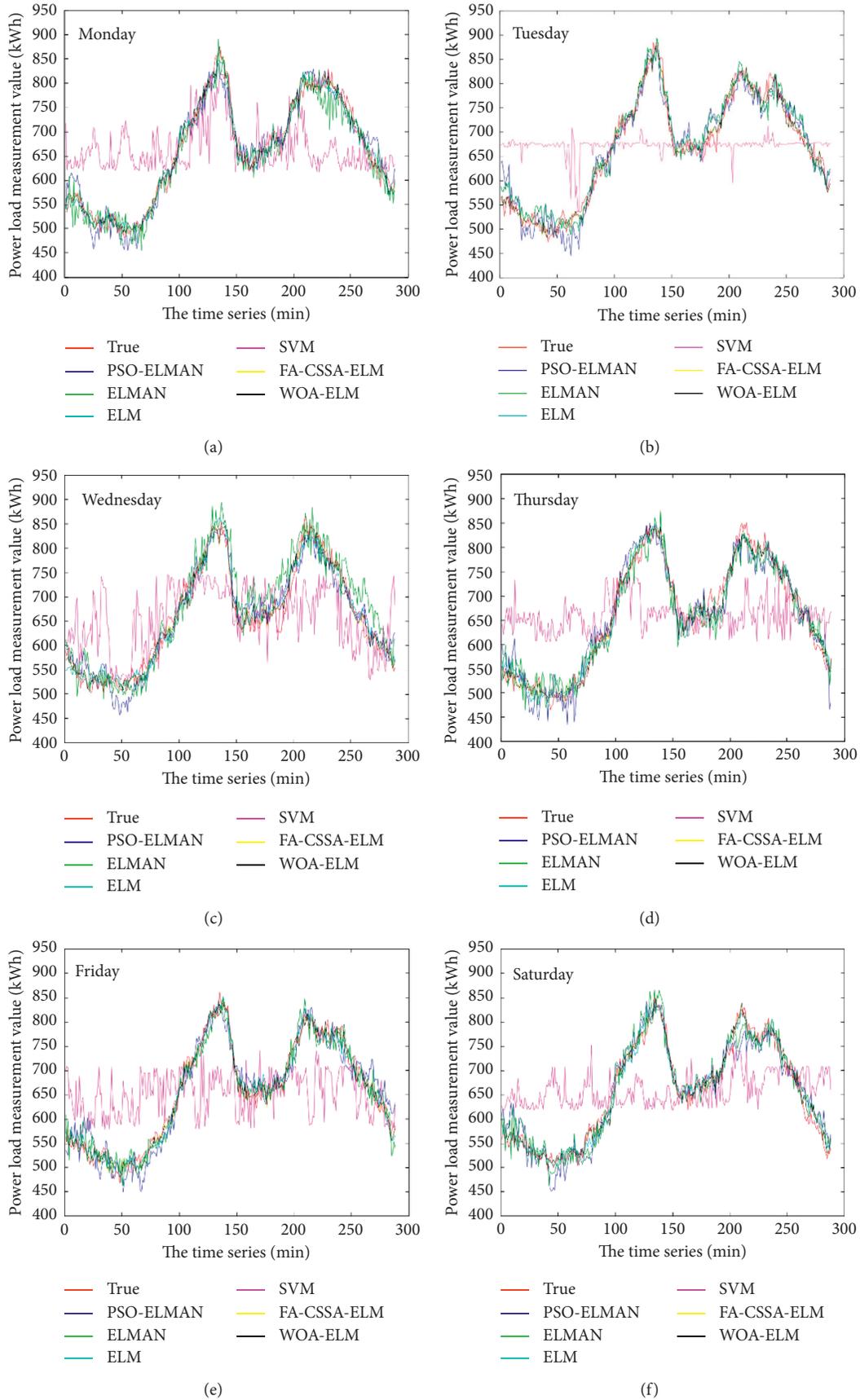


FIGURE 9: Continued.

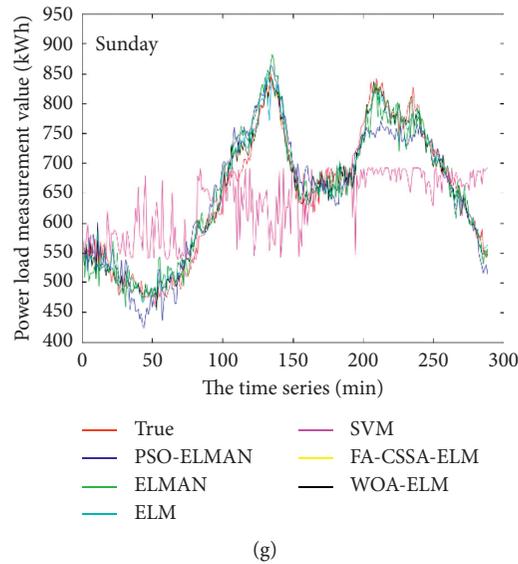


FIGURE 9: Forecast line chart of six competing models from Monday to Sunday.

5. Conclusion

In this study, we propose an FA-CSSA-ELM power load forecasting model to predict power loads accurately. First, this paper optimizes the SSA algorithm using the Tent chaos mapping strategy and the firefly perturbation strategy. Then, the constructed FA-CSSA algorithm is used to optimize the initial thresholds and weights of the ELM model. Finally, the FA-CSSA-ELM power load forecasting model is compared with three single forecasting models and two combined forecasting models. The forecasting effect is visualized through the 4 evaluation functions and the corresponding histograms. The following conclusions can be obtained through the simulation experimental validation of real power load data from a power grid in Shandong.

By comparing the FA-CSSA-ELM power load forecasting model with three typical single power load forecasting models, this paper finds that the FA-CSSA-ELM improved 72.29%, 99.8%, and 478.2% in MSE metrics compared with the other three single prediction models ELMAN, ELM, and SVM, respectively. The FA-CSSA-ELM improved by 80.3%, 18%, and 540% in the MAPE metric compared with the remaining three single prediction models ELMAN, ELM, and SVM, respectively. In the RMSE metric, the improvement is 77.8%, 14%, and 488% compared with the other three single forecasting models ELMAN, ELM, and SVM, respectively. In terms of MAE metric, the improvement is 77.2%, 15.6%, and 517% for ELMAN, ELM, and SVM, respectively. From the comparison data, the FA-CSSA-ELM model proposed in this paper has a much better prediction effect than the three representative single prediction models.

It can be found from Experiment I that the performance of the FA-CSSA-ELM prediction model is far superior to that of the three single prediction models. Then, WOA-ELM and PSO-ELMAN combination models are compared through Experiment II. Compared with the WOA-ELM

model and PSO-ELMAN model, the MSE index of the FA-CSSA-ELM model increased by 10.2% and 87.3%, respectively. The MAPE index of the FA-CSSA-ELM model is increased by 3% and 86% compared with the WOA-ELM model and PSO-ELMAN model, respectively. Compared with WOA-ELM and PSO-ELMAN, the RMSE index of the FA-CSSA-ELM model increased by 2.6% and 89%, respectively. The MAE index of the FA-CSSA-ELM model compared with that of the WOA-ELM model and PSO-ELMAN model increased by 2.1% and 80.9%, respectively. This further indicates that the FA-CSSA-ELM power load forecasting model has higher forecasting accuracy and performance. The FA-CSSA-ELM prediction model can provide more accurate feedback to relevant departments. Relevant departments can guide the rational layout of the power system through accurate feedback so as to reduce the waste of power load and the economic loss of the industry.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

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

References

- [1] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, 1987.
- [2] A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1535–1547, 1990.
- [3] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using support vector Machines: a study on EUNITE competition 2001," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1821–1830, 2004.
- [4] J.-F. Chen, W.-M. Wang, and C.-M. Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting," *Electric Power Systems Research*, vol. 34, no. 3, pp. 187–196, 1995.
- [5] K. Lang, M. Zhang, Y. Yuan, and X. Yue, "Short-term load forecasting based on multivariate time series prediction and weighted neural network with random weights and kernels," *Cluster Computing*, vol. 22, no. S5, pp. 12589–12597, 2019.
- [6] J. J. Xia, H. Qi, and Z. Q. Wang, "Combination forecasting of power load based on polynomial trend extrapolation and ARIMA model," *Advanced Materials Research*, vol. 546–547, pp. 357–362, 2012.
- [7] Y. Wu, Z. Pan, X. Luo, J. Gao, and Y. Zhang, "A hybrid forecasting method of electricity consumption based on trend extrapolation theory and LSSVM," in *Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 2333–2337, Xi'an, China, October 2016.
- [8] W. Christiaanse, "Short-term load forecasting using general exponential smoothing," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 2, pp. 900–911, 1971.
- [9] J. W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *Journal of the Operational Research Society*, vol. 54, no. 8, pp. 799–805, 2003.
- [10] H. Akçay and T. Filik, "Short-term wind speed forecasting by spectral analysis from long-term observations with missing values," *Applied Energy*, vol. 191, pp. 653–662, 2017.
- [11] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1035–1042, 2005.
- [12] T. Jakaša, I. Androćec, and P. Sprčić, "Electricity price forecasting — ARIMA model approach," in *Proceedings of the 2011 8th International Conference on the European Energy Market (EEM)*, pp. 222–225, Zagreb, Croatia, May 2011.
- [13] Y. Wang, J. Wang, G. Zhao, and Y. Dong, "Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: a case study of China," *Energy Policy*, vol. 48, pp. 284–294, 2012.
- [14] J. L. Elman, "Finding structure in time," *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.
- [15] W. S. Noble, "What is a support vector machine?" *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.
- [16] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 2–501, 2006.
- [17] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [18] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting-A novel pooling deep RNN," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5271–5280, 2018.
- [19] N. Ye, Y. Liu, and Y. Wang, "Short-term power load forecasting based on SVM," *World Automation Congress*, vol. 2012, pp. 47–51, Puerto Vallarta, Mexico, 2012.
- [20] W. Li, C. Quan, X. Wang, and S. Zhang, "Short-term power load forecasting based on a combination of VMD and ELM," *Polish Journal of Environmental Studies*, vol. 27, no. 5, pp. 2143–2154, 2018.
- [21] X. Li, Lv Jin-hu, L. Ding, G. Xu, and J. Li, "Building cooling load forecasting model based on LS-SVM," in *Proceedings of the 2009 Asia-Pacific Conference on Information Processing*, pp. 55–58, Shenzhen, China, July 2009.
- [22] F. Yang, P. Wang, Y. Zhang, L. Zheng, and J. Lu, "Survey of swarm intelligence optimization algorithms," in *Proceedings of the 2017 IEEE International Conference on Unmanned Systems (ICUS)*, pp. 544–549, Beijing, China, October 2017.
- [23] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28–39, 2006.
- [24] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," *Applied Mathematics and Computation*, vol. 214, no. 1, pp. 108–132, 2009.
- [25] X. S. Yang and X. He, "Firefly algorithm: recent advances and applications," *International Journal of Swarm Intelligence*, vol. 1, no. 1, pp. 36–50, 2013.
- [26] X. S. Yang and A. Hossein Gandomi, "Bat algorithm: a novel approach for global engineering optimization," *Engineering Computations*, vol. 29, no. 5, pp. 464–483, 2012.
- [27] R. Rajabioun, "Cuckoo optimization algorithm," *Applied Soft Computing*, vol. 11, no. 8, pp. 5508–5518, 2011.
- [28] M. Seyedali, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
- [29] M. Seyedali, "Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Computing and Applications*, vol. 27, no. 4, pp. 1053–1073, 2016.
- [30] M. Seyedali and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [31] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 22–34, 2020.
- [32] H. Wang, X. Lv, and X. Luo, "Short-term load forecasting of power grid based on improved WOA optimized LSTM," in *Proceedings of the 2020 5th International Conference on Power and Renewable Energy (ICPRE)*, pp. 54–60, Shanghai, China, September 2020.
- [33] W.-C. Hong, "Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm," *Energy*, vol. 36, no. 91, pp. 5568–5578, 2011.
- [34] F. Jiang, Z. Peng, and J. He, "Short-term load forecasting based on support vector regression with improved grey wolf optimizer," in *Proceedings of the 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)*, pp. 807–812, Xiamen, China, March 2018.
- [35] X. Yi, R. Guo, and Y. Qi, "Stabilization of chaotic systems with both uncertainty and disturbance by the UDE-based control method," *IEEE Access*, vol. 8, pp. 62471–62477, 2020.
- [36] L. Liu, B. Li, and R. Guo, "Consensus control for networked manipulators with switched parameters and topologies," *IEEE Access*, vol. 9, pp. 9209–9217, 2021.

- [37] T. Hou, Y. Liu, and F. Deng, "Finite horizon H_2/H_∞ control for SDEs with infinite Markovian jumps," *Nonlinear Analysis: Hybrid Systems*, vol. 34, pp. 108–120, 2019.
- [38] R. Xu and F. Zhang, " ϵ -Nash mean-field games for general linear-quadratic systems with applications," *Automatica*, vol. 114, Article ID 108835, 2020.
- [39] R. L. Haupt and S. E. Haupt, *Practical genetic algorithms*, Wiley, Hoboken, NJ, USA, 2004.
- [40] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.