ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20221007130139
Original scientific paper

Improved Electricity Portfolio Prediction Based on Optimized Ant Colony Algorithm

Li ZEMIN

Abstract: In order to overcome the shortcomings of the traditional single power forecasting method, the article uses LSTM network, GM model and SVR support vector machine regression model to forecast electricity, and also uses ant colony optimization algorithm to build a new combined forecasting model for the three forecasting methods, which takes into account the factors affecting power forecasting more comprehensively and helps to improve the accuracy of power forecasting. The paper also uses the ant colony algorithm to optimize the weights of the single forecasting method, which can effectively avoid the problem of the traditional algorithm falling into the local optimal point, and obtain a more accurate power combination forecasting model. Through application examples, it is verified that the combined forecasting model can effectively improve the accuracy of power forecasting and provide reference for power system planning and operation. The research results show that the combined prediction has a greater improvement in accuracy compared with the single Gray, LSTM network and other predictions.

Keywords: combinatorial prediction; gray prediction; neural networks; support vector machines

1 INTRODUCTION

The forecast of electricity is crucial to power construction planning and macroeconomic development. The analysis and forecast of regional electricity can provide reference for the operation and maintenance personnel of power supply companies to judge whether the electricity sales are abnormal and make corresponding remediation plans. All relevant power supply enterprises can refer to the forecast trend of electricity sales to adjust the power supply plan in time and improve the reliability and efficiency of the power supply structure, which is in line with building a conservation-oriented society and promoting energy conservation. This is in line with the development concept of building a conservation-oriented society and promoting energy conservation and emission reduction. Therefore, the establishment of effective electricity sales forecasting models has always been a hot topic of research in the field of electric power. Therefore, how to improve the forecasting accuracy has always been a concern of researchers. Currently, regression analysis, time series, network models and support vector machines are mainly used for long-term electricity forecasting.

Since the traditional time series is only applicable to data sources with relatively uniform demand changes, linear regression analysis has low prediction accuracy, while nonlinear regression analysis has high computational overhead and complex prediction process. Therefore, a lot of research and improvements have been made at home and abroad for prediction models, such as the gray GM (1, 1) prediction model which has the advantages of requiring less data, not considering the distribution pattern and change trend, and convenient operation, etc. However, the gray prediction model transitions depend on the historical values of electricity, and the prediction accuracy is not high when the data dispersion is relatively large [1]. The support vector machine prediction method can better solve the practical problems of small samples, nonlinearity, and high dimensionality, and has the advantages of simple structure, global optimality, better generalization ability, and higher prediction accuracy [2-3]. Zhang Youquan et al. proposed a power forecasting model based on gray system theory and applied gray forecasting to medium- and long-term models [4]; Liu Qiuhua et al. proposed to combine seasonal index forecasting with gray forecasting to make up for the deficiency that gray forecasting is insensitive to change trends [5]; Lu Hai [6] et al. proposed a load forecasting method based on data-driven concept by extracting different types of load features using least squares. Huang Han [7] et al. used random forest (RF) algorithm to achieve hourly power forecasting by combining historical data of electric load and weather information; RashedulHaq [8] et al. proposed a hybrid short-term electricity forecasting model based on IEMD and T-Copula, combined with deep learning networks to achieve time-specific electricity forecasting. Yang Xiaolei [9] modified random forest for initial electricity and added a gray model for electricity prediction, which effectively reduced the average error. Yongjian Zhang [10] proposed short-term prediction of electricity based on LSTM networks, which improved the accuracy in short-term prediction species.

Although the above forecasts have been actively explored based on the field of electricity forecasting, adapting excellent algorithms such as gray forecasting and support vector machines to the forecasting prediction model, the short- and medium-term electricity is usually influenced by the superposition of various factors such as the users' electricity consumption behaviour, load changes, seasonal changes, holidays, etc., thus causing its time series to show an unstable trend. The commonly used forecasting models such as support vector machine, random forest algorithm and neural network do not realize the refined decomposition of data [11], and the forecasting effect is poor. Meanwhile, the time series of electricity sales contains linear and nonlinear time series components [12]. A single forecasting method cannot analyze and predict each component well, and there are many shortcomings in showing the composite characteristics of the time series.

In recent years, combined forecasting methods have become an important research direction in the field of forecasting, and combined forecasting models can make greater use of various forecasting sample information, effectively reduce the influence of environmental random factors in individual forecasting models, and improve forecasting accuracy [13]. However, the determination of the weight coefficients of each single prediction model is crucial in the combined prediction model. Considering that the ant colony algorithm has strong global convergence, detailed local search and strong robustness [14], and does not need to resort to the characteristic information of the problem, the ant colony search algorithm is selected in this paper to find the weights of each combined prediction model to improve the prediction accuracy.

2 PROBLEM MODEL

Combination forecasting is a forecasting method that takes the forecasting results obtained from several forecasting methods and weights them appropriately to find the optimal one. Combined forecasting can combine the information provided by each individual forecasting method and integrate the forecasting results from different information sources, which can most effectively improve the forecasting accuracy. In this paper, four forecasting models, namely regression forecasting method, SVR support vector machine forecasting, improved GM (1, 1) and LSTM neural network based forecasting, are selected for medium- and long-term electricity forecasting, and their basic principles are as follows.

Let w_i^1 be the actual electricity of the ith week (i =1, 2, ..., n, n is the number of weeks of predicted electricity), then the time series can be obtained from the n weeks of electricity observations; let f_{ik} be the electricity prediction of the ith week of the kth method (k = 1, 2, ..., K), the prediction error is $e = w_i^1 - f_{ik}$, w_k is the estimated value of the power coefficient of the kth method, and F is the combined forecast value, with:

$$F_i = \sum_{i=1}^{k} w_k f_{ik} \tag{1}$$

The basic principle of the combined prediction model weights is based on the combined prediction model with fixed weight coefficients with minimum sum of squares of errors as follows.

$$\min E = \sum_{i=1}^{n}\left(w_i^1 - \sum_{i=1}^{k} w_k f_{ik}\right)^2 \tag{2}$$

$$\text{s.t.} \sum_{i=1}^{k} w_k = 1 \tag{3}$$

At present, there are more studies on power combination forecasting at home and abroad, mostly using equal-weight average combination forecasting method and variance-covariance preferred combination forecasting method, etc. However, equal-weight average combination forecasting method has equal weights for each model, there is no concept of preference, and the weights derived from variance-covariance preferred combination forecasting method are very unstable. In contrast, the ant colony algorithm has the advantages of fewer parameters, fast computation, and strong global optimization-seeking ability, etc. For this reason, the ant colony algorithm is selected to optimize the weights of the combination prediction model and optimize the combination prediction model [15]. That is, the ant colony algorithm is used to optimize the combination prediction model equation to find out the optimal weight coefficient w_k of the combination prediction model, and then, the optimal combination prediction value with the smallest error sum of squares can be obtained by substituting the weight coefficient w_k into Eq. (1). The specific flow chart is shown in Fig. 1.

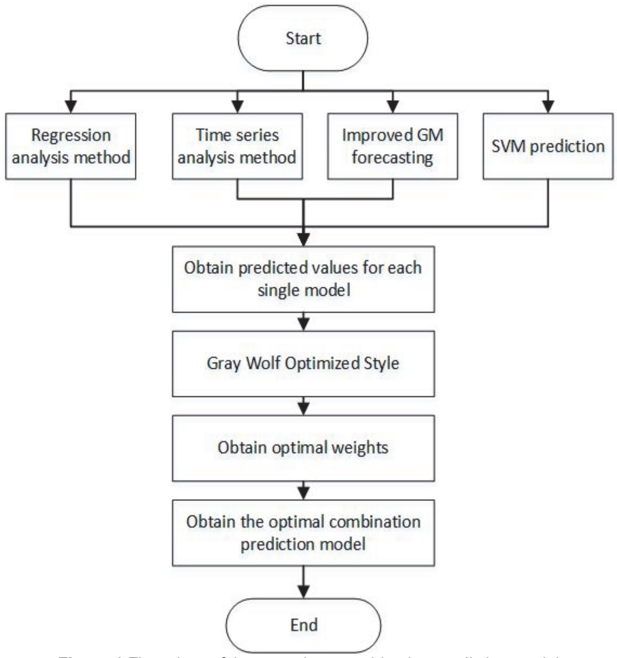

Figure 1 Flow chart of the ant colony combination prediction model

3 ALGORITHM MODEL

3.1 LSTM Neural Network

RNN recurrent neural network can predict time series, which is essentially equivalent to adding a feedback mechanism to a fully connected neural network that can process the previous signal and remember the state after processing, called memory. The structure of RNN recurrent neural network is schematically shown in Fig. 2.

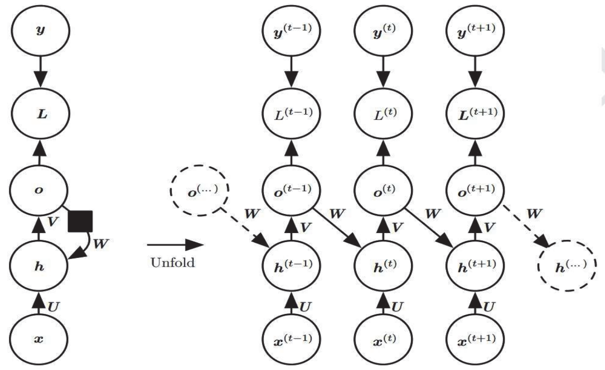

Figure 2 Schematic diagram of the basic structure of RNN structure NN neural network

The LSTM neural network is an improvement on the RNN recurrent neural network, which was proposed in 1997 [16-18]. The LSTM neural network is composed of multiple recurrent units recursively connected to each other, and the structure is schematically shown in Fig. 3. The recurrent unit of the LSTM long and short-term memory neural network includes three gate structures inside, which are the forgetting gate, the input gate and the output gate.

The function of the forgetting gate is to select the part of the previous memory state that should be retained, which is usually implemented by the tanh function; the input gate can decide which information to update the long-term memory state with, and this valve filters out the useful information to feed the network, which is usually implemented by the Sigmoid function; the last one is the output valve, which is placed after the output of the network and can automatically extract the important part of the output information, which is also implemented by the Sigmoid function. The inventors of long and short-term memory neural networks have experimentally verified the superiority of LSTM over RNN in their paper, mainly in the following two points: 1. the three gating structures introduced by LSTM can learn long-term memory information and solve the long-time dependence problem; 2. the activation function in LSTM is a combination of sigmoid function and tanh function, which keeps the gradient almost constant when back-propagating. The combination of the sigmoid function and the tanh function in the LSTM makes the gradient remain almost constant when deriving, which avoids gradient disappearance or explosion and greatly accelerates model convergence [18-20].

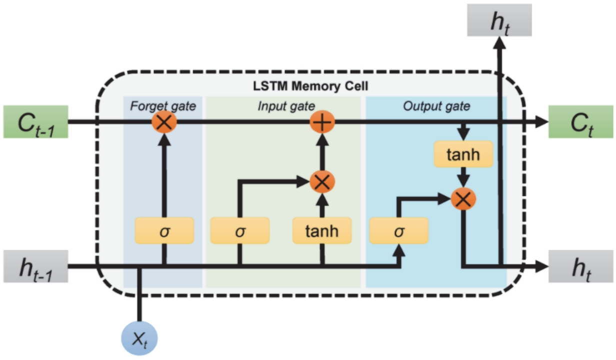

Figure 3 Schematic diagram of the basic structure of RNN structure NN neural network

In this study, Python 3.9.12 and Keras are used to implement the LSTM neural network, and extensive experiments are done to tune the parameters of the LSTM model on a dataset of high-pressure users across the industry. The main parameters to be adjusted are the prediction time window, i.e., sequence length, the number of neurons in the hidden layer, and the parameters α and β for data smoothing.

3.2 Gray Prediction Method

The GM (1, 1) model is one of the most commonly used gray models and consists of a first-order differential equation containing only a single variable. The gray model is generated by accumulating the original data, so that the generated data columns have an exponential growth pattern. After the generation of the series, the differential equation model is established, the time response function of the differential equation is obtained, and the gray prediction model of the original series is obtained by cumulative reduction. The gray prediction model of the original series is:

$$x_1^{(0)}(k+1)=x_1^{(1)}(k+1)-x_1^{(1)}(k)=\left(1-e^{a}\right)\left(x^{(0)}(1)-\frac{u}{a}\right)e^{-ak},\ k=0,1,2 \tag{4}$$

$x_1^{(1)}$ is the predicted value of the cumulative generated series; $x_1^{(0)}$ is the predicted value of the original series; a and u are the parameters of the first order differential equation.

When the gray prediction model is less accurate when the data dispersion is large, for this reason, it is necessary to improve the gray prediction model. The ways to improve the gray prediction are generally transforming the original series, selecting the initial values, improving the gray model, and improving the technical methods. In order to reduce the unsmoothness of the data and strengthen the approximate trend of the original series, the sliding average method is selected to improve the gray prediction model by smoothing the data to improve the fit of the predicted data. Let the original series be $\{x^0(t)\}$, t =1, 2, …, n, then the sliding average calculation formula is:

$$x_1'^{(0)}(t)=\frac{x^0(t-1)+2x^0(t)+x^0(t+1)}{4} \tag{5}$$

For the 2 end points, the following equation should be used.

$$x'^{(0)}(t)=\frac{3x^0(1)+x^0(2)}{4} \tag{6}$$

$$x'^{(0)}(n)=\frac{x^0(n-1)+3x^0(n)}{4} \tag{7}$$

In the above equation: x^0 is the original series value; x'^0 is the value after sliding average of the original series. The sliding average of the original data increases the weight of the current data and avoids excessive fluctuation of the values, which makes the predicted data fit better.

3.3 Support Vector Machine Prediction Model

Support vector machine method is a relatively new intelligent algorithm recently proposed, which is a small-sample learning method based on statistical learning theory, i.e., VC dimensional theory, and structural risk minimization principle, and has the characteristics of requiring fewer parameters to be determined and having a globally optimal unique solution in theory. In addition, the support vector machine is more suitable for medium- and long-term power forecasting because of its good generalization ability under small samples. The support vector machine regression function is:

$$f\left(x,\beta,\beta^*\right)=\sum_{i=1}^{N}\left(\beta_i-\beta_i^*\right)K\left(x_i,x_j\right)+b \tag{8}$$

Where: x is the training sample; β_i and β_i^* are Lagrange multipliers; b is the bias constant; $K\left(x_i,x_j\right)$ is the kernel function, which needs to satisfy the Mercer condition, and the most commonly used Gaussian kernel function $K\left(x_i,x_j\right)=exp\left(-\left\|x_i-x_j\right\|^2/\left(2\sigma^2\right)\right)$ is generally selected, thus: The larger ε is, the smaller the number of support vectors in the trained model, the lower the accuracy of regression prediction; the smaller ε is, the higher the accuracy of regression prediction, but the support vector will increase and the generalization ability may be weaker. The penalty factor C is used to control the trade off between model complexity and approximation error. The larger the value, the better the fit to the data, but if C is too large, it will easily cause overlearning on the training data and lead to lower generalization ability. The kernel width σ is related to the range of the learning sample input space; the larger the sample, the larger its value; the smaller the sample space, the smaller its value. Therefore, when constructing the support vector machine model, the parameters of the support vector machine need to be selected reasonably and appropriately to improve the learning and generalization ability of the model.

3.4 Ant Colony Optimization Algorithm Solving

Ant colony algorithm is a bionic optimal swarm intelligence algorithm that simulates the foraging principle of ants. At the beginning, it was applied to solve the Traveling Traders Problem (TSP), where an ant needs to visit all n cities and each city only once, and finally return to the starting point with the task of finding the shortest path.

(1) Colony initialization. Set the colony size (number of ants), the pheromone importance factor, the heuristic function importance factor, the pheromone volatility factor, the total number of pheromone releases, and the maximum number of iterations.

(2) Construct the solution space. The individual ants are randomly placed at different departure points, and the next city to be visited is determined for each ant according to the transfer probability formula, and a path is available after visiting all cities.

$$P_{ij}^{k} = \frac{\left(\tau_{ij}^{\alpha}\right)\left(\eta_{ij}^{\beta}\right)}{\sum_{Z\in akkowedx}\left(\tau_{ij}^{\alpha}\right)\left(\eta_{ij}^{\beta}\right)} \tag{9}$$

(3) Update the pheromone. The computational path length is updated, and the pheromone concentration is updated according to the pheromone iteration formula, while the optimal solution is recorded.

$$\tau_{ij}\left(t+1\right) = \left(1-\rho\right)\cdot\tau_{ij}\left(t\right) + \Delta\tau_{ij} \tag{10}$$

(4) Termination condition judgment. If the termination condition is satisfied (the maximum number of iterations is reached or the minimum is satisfied), the individual with the maximum fitness obtained in the evolutionary process is output as the optimal solution and the calculation is terminated; otherwise, it goes to step 1 and continues to iterate.

3.5 Power Prediction Combination Model Solving Process

According to the main operation process of ant colony algorithm, the process of solving the power combination prediction model based on ant colony algorithm is mainly divided into the following four steps, and the solution process is shown in Fig. 4.

(1) Initialize the population, set the number of populations, learning factors and other parameters, set the minimum value $F_{\min}$ and the maximum number of iterations of the algorithm $dc_{\max}$.

(2) Performing individual fitness evaluation to provide a basis for future ant colony direction operation.

(3) updating the weight value corresponding to the single prediction method, substituting the power value predicted by the historical time single prediction method into the combined prediction model, obtaining the predicted value of the combined prediction model for each historical time power under the weight, calculating the error between the predicted value and the actual value for each historical time power, recorded as F_{dc} and comparing it with the minimal value $F_{\min}$ to determine whether to record the weight value of the single prediction method.

(4) To carry out the program termination judgment, if dc is greater than the maximum number of iterations $dc_{\max}$, the solution of the ant colony algorithm is terminated, and the recorded weight values are brought into Eq. (1) to obtain the combined power prediction model.

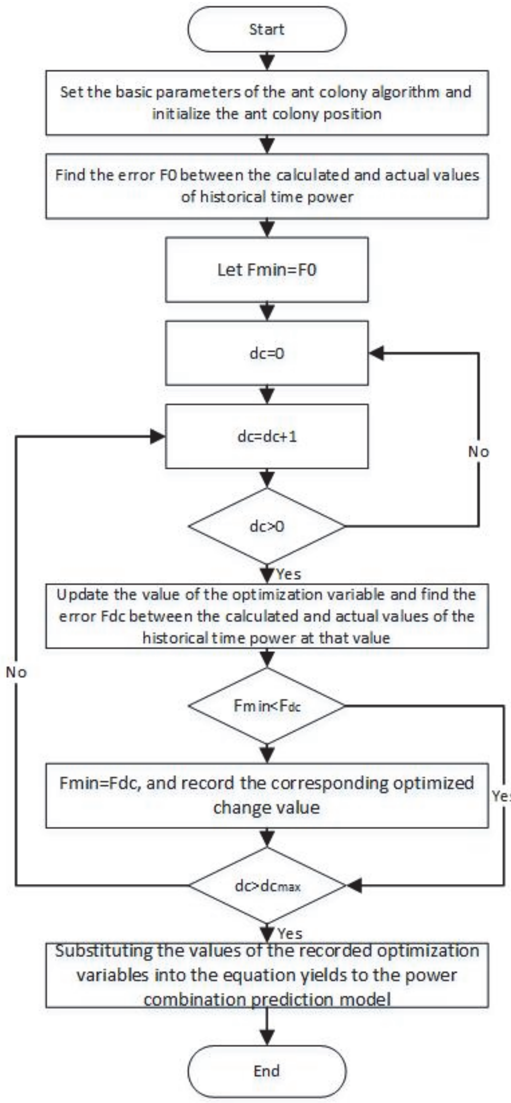

Figure 4 Solving process of power combination prediction method based on ant colony algorithm

4 EXPERIMENTAL VERIFICATION
4.1 Lab Environment

The experimental environment of this paper is 64-bit operating system, x64-based processor i7-10875H, system version is Windows 10, main frequency is 2.30 GHz, running memory is 16.00 G DDR4-2400, programming language is python 3.9.12, storage is 1T solid state drive, development platform is PyCharm2021, ant colony. The algorithm part size = 100, and the maximum number of iterations is set to 40.

4.2 Data Processing and Pre-Processing

With the construction of the electric power industry, the growth rate of business data in the electric power industry is increasing and various data have been accumulated for many years. From the daily electricity consumption data of high-voltage users in public stations, the algorithm is applied to make a whole month forecast of electricity consumption for many months, and then compared with the known data to judge the accuracy of the forecast. When forecasting, the whole area is divided into several regions, and then the electricity consumption of each region is forecasted separately, and finally aggregated into the total electricity consumption of the whole region. When calculating the error, we calculate the error of the aggregated total electricity forecast for the whole area.

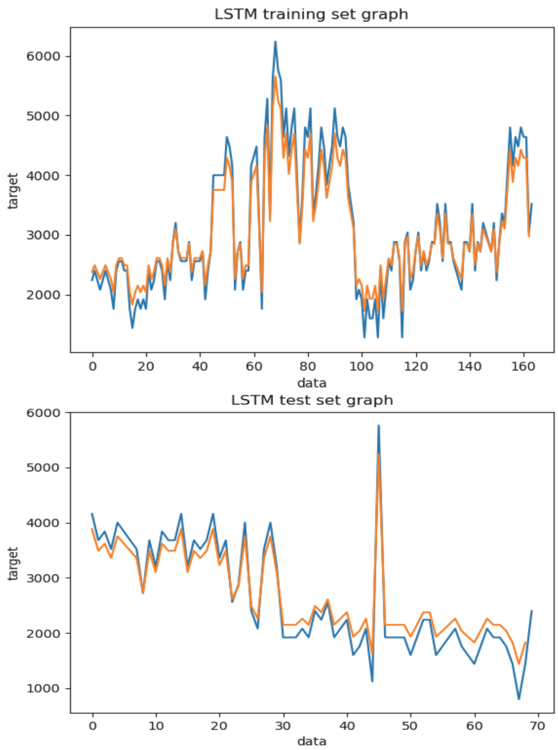

Figure 5 Solving process of power combination prediction method based on ant colony algorithm

Before forecasting the time series composed of daily electricity, the data are first smoothed. The purpose of smoothing is to replace some extremely large or small or abnormal daily electricity values with reasonable electricity values. Anomalous values are defined as power values that are less than α ($0 < \alpha < 1$) times the daily average power or greater than β ($\beta > 1$) times the daily average power. Since daily power can be considered as a 7 day period, anomalous power values can be replaced with power around 7 or 14 days.

4.3 Analysis of Experimental Results

In this paper, the gray prediction algorithm, SVR support vector machine optimization algorithm, LSTM neural network algorithm and regression prediction method were run, in which the first three were taken as the first 70% of the training set and the last 30% as the test set, with the number of iterations being 100 and the data being 240 days of electricity data from industrial parks.

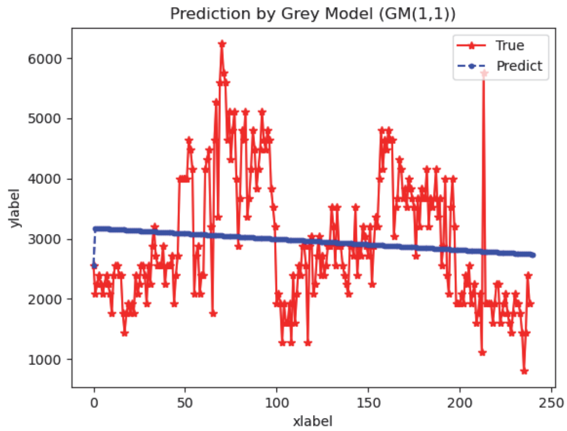

Figure 6 Graph of grey prediction results

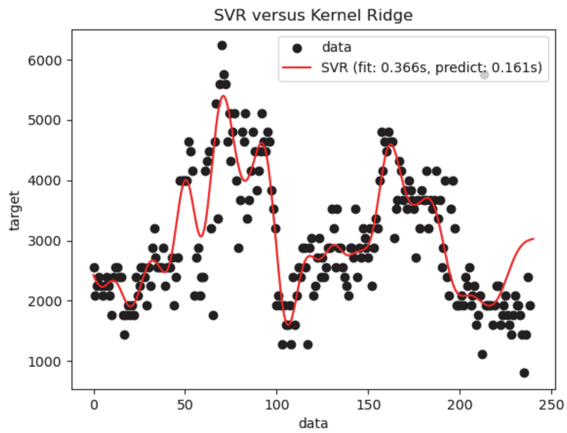

Figure 7 SVR support vector machine prediction graph

From the results of the three high-precision algorithms, the gray prediction model has lower prediction accuracy due to the linear condition constraint in the longer time; the SVR has fair performance error after doing discrete optimization but cannot take into account the discrete points with great deviation, while the LSTM makes accurate prediction in the overall trend and effectively predicts the trend of power change in days, but is vulnerable to the influence of discrete points so that the prediction value cannot control the error. However, it is susceptible to the influence of discrete points and thus the prediction value cannot control the error. We choose the last 72 days as the degree of electricity error calculation, and bring the optimized value of the last 72 days into the ant colony optimization equation, and the combined prediction model can be optimized as:

$$F = 0.140 f_1 + 0.613 f_2 + 0.247 f_3 \tag{11}$$

where f_1 is the gray forecast value, f_2 is the LSTM forecast value, f_3 is the SVR optimization vector machine forecast value. After making the combined forecast processing, we will have the new 60 days of electricity for the data set brought into the combined forecast model. Forecast can be obtained.

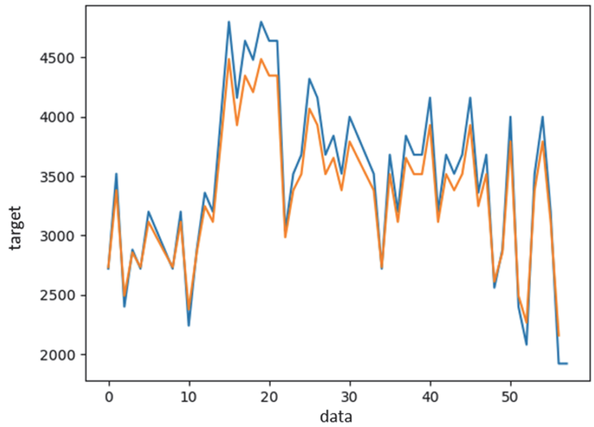

Figure 8 Predicted Electricity Diagram of Combined Model

It can be seen that the prediction error is reduced to 12.67014%, compared with 53% for gray prediction, 17% for LSTM, and 34% for SVR, which reduces the larger error and greatly improves the accuracy of power prediction.

5 CONCLUSION

The improved GM(1,1), LSTM neural network model and SVR short-term power optimization combination forecasting model proposed in this paper combine the advantages of the improved GM(1, 1) model, SVR model and LSTM neural network model, which can better deal with the effects of periodic and stochastic factors in the short and medium-term power forecasting, Improved forecast accuracy and more accurate forecast trends compared to actual electricity compared to single forecast methods, and is more accurate and suitable for short-term forecasting compared with a single forecasting method. The ant colony search algorithm has the characteristics of strong global search capability and detailed local search. The proposed Ant colony optimization algorithm determines the weights of each model in the combined forecasting model, which overcomes the limitations of weight solving in the combined forecasting model and can better exploit the improved effect of combined forecasting. However, the combined forecasting model is also affected by the discrete and excessive accuracy of the single method in the short-term power model, and the forecasts in some nodes are not as good as expected, due to the difference in accuracy between the method forecasts, resulting in a greater impact by the higher accuracy methods, and also limiting the methods with lower errors, which should continue to strengthen the combined search for excellence while seeking more excellent single methods.

Acknowledgments

This work was supported by Inner Mongolia Power Co., Ltd.

6 REFERENCES

[1] Chen, J., Yang, Z., Zhou, Z., Jiang, W., & Lin, X. (2020). Research on the application scenarios of smart energy meters based on the new generation of smart energy measurement system. *Guangxi Electric Power*, *43*(03), 16-21. https://doi.org/10.16427/j.cnki.issn1671-8380.2020.03.003

[2] Li, Y., Lu, J., & Jia, Z. (2020). Exploration of computer operating system scheduling methods. *Decision Exploration*, *2020*(02), 80.

[3] Niu, D., Ying, S., Zhao, L., et al. (2002). Power load forecasting technology and its applications. *Beijing: China Electric Power Press*, 15-29.

[4] Jin, Z., Chen, K., Yan, W., et al. (2008). Research and application of medium- and long-term power load forecasting based on support vector machines. *Journal of Shanghai University of Technology*, *2008*(2), 130-131.

[5] Teruaki, S., Takahiro, T., Bruce, B., & Hideaki, K. (2006). Black hole production in tachyonic preheating. *Journal of Cosmology and Astroparticle Physics*, *2006*. https://doi.org/10.1088/1475-7516/2006/04/001

[6] Zhang, Y. (1999). A medium- and long-term power demand forecasting model based on gray system theory. *Power Grid Technology*, *23*(8), 4.

[7] Liu, Q. (2006). A monthly electricity forecasting model based on seasonal index and gray forecasting. *Journal of Nanjing Engineering College: Natural Science Edition*, *4*(1), 6.

[8] Lu, H., Luo, F. Z., Yang, X., et al. (2020). A data-driven approach to daily load curve prediction for electricity. *Journal of Power Systems and Automation*, *32*(1), 42-49.

[9] Huang, H., Sun, K., & Liu, D. (2018). Research on hourly load forecasting of power system based on random forest. *Smart Power*, *46*(5), 8-14.

[10] Haq, M. R. & Ni, Z.. (2019). A new hybrid model for short-term electricity load forecasting. *IEEE Access*, *2019*, 125413-125423. https://doi.org/10.1109/ACCESS.2019.2937222

[11] Yang, X. L., Yao, J. F., Xu, J., & Xu, C. (2020). Electricity forecasting based on grey prediction and random forest combination algorithm. *Rural Electrification*, *2020*(07), 13-17.

[12] Zhang, Y. J., Zhou, S. H., Shen, C. H. H., & Zhou, C. H. S. (2021). Short-term power forecasting based on LSTM neural network. *Power Big Data*, *24*(08), 9-18.

[13] Zou, P., Yao, J., Kong, W., et al. (2017). Medium- and long-term power load forecasting based on multivariate time series inversion self-memory model. *Journal of Power Systems and Automation*, *29*(10), 98-105.

[14] Huang, Y. (2018). *Short-term load forecasting and nonlinear economic dispatch for hybrid water-fire systems*. Wuhan, Huazhong University of Science and Technology.

[15] Li, J., Zhao, F., & Liu, F. (2011). Combinatorial forecasting of medium- and long-term power loads. *Journal of power systems and their automation*, *23*(4).

[16] Xiong, W. P. & Zeng, B. Q. (2010). A comparative study of several bionic optimization algorithms. *Computer Technology and Development*, *10*(3), 9-12.

[17] Zhao, P. J. & Liu, S. (2009). A hybrid frog-jumping algorithm for solving complex function optimization problems. *Computer Application Research*, *26*(7), 2435-2437.

[18] Zhou, G., Luo, P., & He, Q. (2018). Predicting compositional time series viaautoregressive Dirichlet estimation. *Science China(information sciences)*, *61*(09), 300-305. https://doi.org/10.1007/s11432-017-9335-5

[19] Ma, H., Leng, S., & Chen, L. (2018). Data-based prediction and causality inference of nonlinear dynamics. *Science China(mathematics)*, *61*(3), 403 -420. https://doi.org/10.1007/s11425-017-9177-0

[20] Wanq, Q., Huang, Y., Jia, W., et al. (2020). FACLSTM: ConvLSTM with focused attention for scene text

recognition. *Sciece China (information sciences)*, *63* (02), 35-48. https://doi.org/10.1007/s11432-019-2713-1
[21] Xu, J., Xu, W. H., Zeng, X., et al. (2018). Innovation and application of power forecasting methods based on big data mining. *Power Big Data*, *21*(10), 57-61.
[22] Li, Z. & Wu, T. (2019). An efficient network traffic prediction method based on PF-LSTM network. *Computer Application Research*, *36*(12), 3833-3836.

Contact information:

Li ZEMIN, PhD
(Corresponding author)
Inner Mongolia Power (Group) Company Limited Electricity Marketing Service and Operation Management Branch
E-mail: lizemin_nmdl@sina.com