Short-term power load forecasting based on I-GWO-KELM algorithm

Xiaoyu Chen[1,*], *Xiangli* Dong[1], and *Li* Shi[2]

[1]Shandong Xiehe University, Computer College, Jinan, China
[2]Jinan Bodor CNC Machine Co.,Ltd. Jinan, China

Abstract. In this paper, I-GWO-KELM algorithm is used for short-term power load forecasting. Normalize the power data and meteorological data of the short-term power load, and use GWO to optimize the regularization coefficient of KELM and the RBF kernel parameters. To apply the model to short-term power load forecasting to obtain simulations for the next 24 hours and 168 hours curve. Experiments show that the improved model I3-GWO-KELM proposed in this paper has the best effect. The improvement of GWO in this paper is effective and feasible. In the application of short-term power load forecasting, the IGWO-KELM model is more accurate than the ELM and KELM models.

1 Introduction

Power load forecasting ([1]) is to determine the power load data at a specific time in the future based on many factors, such as the operating characteristics of the system, capacity expansion decisions, natural conditions and social influences. Short-term power load forecasting is one of the important tasks of the power sector ([2]). According to the past and present of the power load, it is estimated that its future value. Accurate load forecasting can economically and reasonably arrange the start and stop of power generation units within the power grid, and maintain the safety and stability of power grid operation. To ensure the normal production and life of the society, effectively reduce the cost of power generation, and improve economic and social benefits.

In the short-term load forecasting, it is necessary to fully study the load variation law of power grid and analyse the relevant factors of load variation ([3]), especially the relationship between weather factors, day type and short-term load variation. Neural network and swarm intelligence optimization algorithm have good prediction effect ([4] - [5]). Grey wolf optimization method based on kernel extreme learning machine is adopted in this paper. The second part and the third part will introduce in detail.

2 Basic theory

This part mainly introduces the extreme learning machine and the improved grey wolf optimization algorithm. For detailed formulas, please refer to the literature ([6] - [7]).

* Corresponding author: buct_cxy@163.com

2.1 Extreme learning machine

The ELM algorithm needs to predetermine the number of neurons in the hidden layer, and randomly assign input weights and deviations, to generate the connection weights of the input layer and the hidden layer and the threshold of the hidden layer neurons. In the ELM training process, only the number of hidden layer neurons needs to be set to obtain the unique optimal solution.

Compared with traditional feedforward neural networks, ELM has the advantages of fast learning speed and good generalization performance. When the output matrix of the hidden layer in the hidden layer is unknown, under Mercer condition, the activation function adopts radial basis function, which is called KELM ([8]).

2.2 Grey wolf optimizer

GWO is a swarm intelligence optimization method that imitates the habits of grey wolves in nature. Grey wolf has two kinds of social behaviors ([9]): one is a strict hierarchy, and the other is a group hunting. Mapped to the algorithm, the grey wolf represents the dynamic process of finding the best target, and the prey is the target that needs to be optimized.

From the literature, we can see that there are three kinds of GWO defects: poor population diversity, slow convergence speed in the later stage, and easy to fall into local optimum. This paper aims at the two shortcomings of slow convergence speed and easy to fall into local optimum in the later stage of GWO, and proposes a convergence factor based on adjustment. The following is the convergence method of $\boldsymbol{a}$ and the method based on the improved search mechanism.

2.2.1 Change the convergence of parameter $\boldsymbol{a}$

The GWO convergence process is not completely linear convergence, while the parameter $\boldsymbol{a}$ in the original GWO is linear convergence. The parameter $\boldsymbol{a}$ can be changed from linear convergence to nonlinear convergence by improving the parameters. The formula is follows:

$$\boldsymbol{a}=\ln\left(e^2-\frac{e^2-1}{t_{max}}\cdot t\right) \tag{1}$$

where e is the base, and t is the current iteration numbers. the formula is updated with each update of the grey wolf, and is the maximum iteration number. The formula has achieved control of global search performance and local search performance.

2.2.2 Improved search mechanism

By improving the search mechanism, GWO is given a weighted distance, and the weighted sum of the best position is used to update the position. GWO can converge faster in the early stage, thus affecting the later convergence speed.

In the original grey wolf optimization algorithm, α , β and δ individuals are selected through fitness sorting, these guide the position of ω hunting. In the process of searching for prey, grey wolves are interested in directions $\vec{A}$ and obstacles approaching prey $\vec{C}$, which will affect the position vector $\vec{X}$ of each wolf. So $\vec{A}$ and $\vec{C}$ are key to determining the direction of the grey wolf. Using the activity trend of grey wolves, weighted and position updating were used to make the grey wolves converge faster in the early stage, and accelerate the convergence speed in the late stage. The formula is as follows:

$$\begin{cases} \vec{A}_1 = 2\vec{a}\cdot\vec{r}_1 - \vec{a} \\ \vec{A}_2 = 2\vec{a}\cdot\vec{r}_1 - \vec{a} \\ \vec{A}_3 = 2\vec{a}\cdot\vec{r}_1 - \vec{a} \end{cases}, \quad \begin{cases} \vec{C}_1 = 2\vec{r}_2 \\ \vec{C}_2 = 2\vec{r}_2 \\ \vec{C}_3 = 2\vec{r}_2 \end{cases} \tag{2}$$

$$w_1 = A_1 * C_1,\ \ w_2 = A_2 * C_2,\ \ w_3 = A_3 * C_3 \tag{3}$$

$$\vec{X}(t+1) = \frac{w_1 * \vec{X}_1 + w_2 * \vec{X}_2 + w_3 * \vec{X}_3}{(w_1 + w_2 + w_3)} \tag{4}$$

where parameter $\boldsymbol{a}$ is the improved nonlinear convergence factor. During the convergence process of the algorithm, the weights of w_1, w_2 and w_3 dynamically adapt to the algorithm environment, constantly changing the learning weights of the offspring grey wolves to the first three levels of grey wolves, and update the position of the grey wolves.

3 Experiment and analysis

3.1 Data set

The experimental data by the 9th Electrician Mathematical Contest in Modeling, which includes electrical data and meteorological data. Carry out daily and weekly power load forecasting. Model input variable setting: in the power load data, enter the same time every week, a total of seven, two fixed in a day; in the meteorological data, enter the highest temperature of the previous day, minimum temperature, average temperature, relative humidity and precipitation. The output variable is electrical load value in the future.

The model includes 14 input variables and 1 output power load data variable. In this paper, the data of July and August are used for daily forecast and weekly forecast.

3.2 Experiment model analysis

3.2.1 Data normalization

Select the appropriate power load data and preprocess it. According to the characteristics of the filtered power load data, in the power load data, the logarithmic method is used for normalization. The calculation formula is as follows:

$$L^{'} = log_{10} L \tag{5}$$

In meteorological data, the normalization of temperature, humidity and rainfall is to scale the value of the feature to between 0 and 1. The formula is as follows:

$$x^i_{norm} = \frac{x^i - x_{min}}{x_{max} - x_{min}} \tag{6}$$

3.2.2 Forecast method

Divide the data set into training set and test set, use GWO to optimize the regularization coefficient of KELM and RBF kernel parameters, obtain I-GWO-KELM, and verify the test set after model training.

Use ELM, KLEM and GWO-KELM to conduct comparative experiments to verify the optimized performance of GWO on KELM. Use I1-GWO-KELM and I2-GWO-KELM and I3-GWO-KELM models, these are based on the improvement of the convergence factor in the GWO-KELM model.

Based on the proportion of the influence of and on the wolf in the hierarchy, they are given different position update weights. I1-GWO-KELM is an improved method of weighted average ([10]), I2-GWO-KELM is an improved method based on the proportional weight of fitness value ([10]), I3-GWO-KELM is the method proposed in this paper, the vector sum is weighted, and weighted sum of the best position is used to update the position.

3.2.3 Model evaluation indicator

In power load forecasting, the evaluation indicator usually selects the average absolute percentage error. The smaller the value, the higher the prediction accuracy of the model. The MAPE formula is as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{7}$$

3.3 Experimental results

3.3.1 Data normalization verification experiment

The necessity and effectiveness of data normalization are verified through experiments. Table 1 is normalized daily predicted MAPE indicators, and Table 2 is the comparison of normalized weekly predicted MAPE indicators.

Table 1. Data normalization daily forecast.

Date \ Model		ELM	KELM	GWO -KELM	I1-GWO -KELM	I2-GWO -KELM	I3-GWO -KELM
Jul.	MAPE	4.113	2.931	1.952	1.885	1.637	1.172
	Time/s	1.8	1.9	2.2	2.2	2.4	2.3
Aug.	MAPE	11.484	8.554	3.894	3.442	1.885	1.192
	Time/s	1.8	1.8	2.1	2.1	2.3	2.1

Table 2. Data normalization weekly forecast.

Date \ Model		ELM	KELM	GWO -KELM	I1-GWO -KELM	I2-GWO -KELM	I3-GWO -KELM
Jul.	MAPE	5.159	3.699	2.779	2.232	1.928	1.758
	Time/s	1.9	1.9	2.2	2.1	2.2	2.2
Aug.	MAPE	11.724	10.947	2.499	2.387	1.981	1.748
	Time/s	2	1.9	2.1	2.2	2.5	2.2

It can be seen from the above four tables that after normalizing the data, the MAPE index is significantly reduced, thus verifying the effectiveness and necessity of the experiment to normalize the data. It can be seen from the MAPE index that I3-GWO-KELM has better results than the unimproved model and other improved models. The following table models all use abbreviations.

3.3.2 Short-term power load forecast graphics

The following is a comparison chart of power load forecast and actual power load values in July and August. The better the fitting effect between the curve in the experimental graph and the actual value, the better the performance of the prediction model.

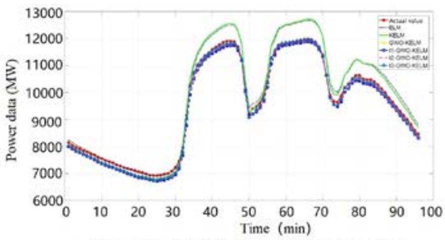

Fig.1. Daily forecast based on July data

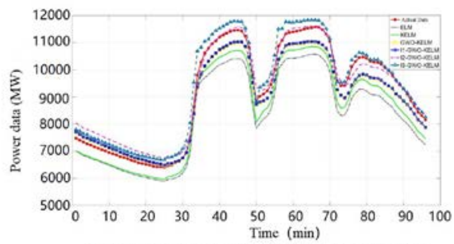

Fig.2. Daily forecast based on August data

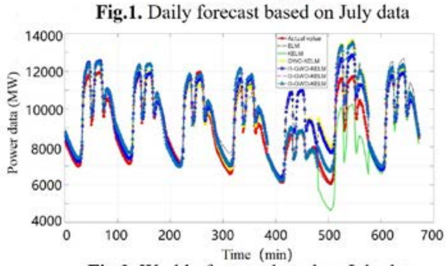

Fig.3. Weekly forecast based on July data

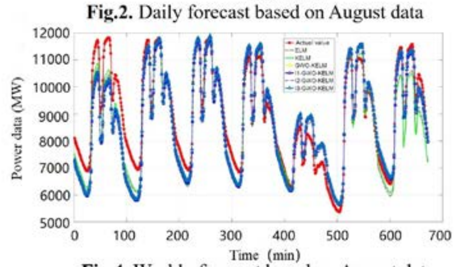

Fig.4. Weekly forecast based on August data

In this paper, I3-GWO-KELM is used for short-term load forecasting. According to experimental analysis, it is concluded that in the daily and weekly forecasts based on July and August, the predicted value of the I3-GWO-KELM model is the closest to the actual value. Experiments verify that the improved GWO model I3-GWO-KELM has the best effect, thus proving that the improvement of GWO is effective and feasible.

References

1. Veeramsetty V, Deshmukh R. SN Applied Sciences, 2(5) (2020)
2. Zhongcheng Y. Qinghai Electric Power. (2019)
3. Jia P, Miao Y. Journal of Physics: Conference Series, 1693(1):012116 (7pp) (2020)
4. Hu R, Wen S, Zeng Z, et al. Neurocomputing, 221: 24-31 (2017)
5. Zhang X, Wang J, Zhang K. Electric Power Systems Research, 146:270-285 (2017)
6. Chen X, Tuo J, Wang Y. 2019 Chinese Control Conference (CCC). 3000-3005 (2019)
7. Chen X, Wang Y, Tuo J. International Federation of Automatic Control (to be published)
8. ChenWei, Deng, GuangBin. Science China Information Sciences, 58(2):1-16 (2015)
9. Joshi H, Arora S. International Journal of Swarm Intelligence, 3(2/3):126 (2017)
10. Chen X. China National Knowledge Infrastructure. (2020)