

Research on Ultra-Short-Term Load Forecasting of Distribution Network Based on Fuzzy Clustering and RBF Neural Network

Ma Guozhen¹, Hu Po¹, Wang Yunjia¹, Wang Yongli², Cai Chengcong^{2,*}, Sun Yaling², Zhang Xinya²

¹Economic and Technological Research Institute of State Grid Hebei Electric Power Co., Ltd, No. 27, Fuqiang Street, Yuhua District, Shijiazhuang, Hebei Province, 050022, China

²North China Electric Power University, Beijing, 102206, China

Abstract. In order to solve the diversification of the load characteristics of the distribution network due to the difference in the electric structure and the electricity consumption habits of users, the calculation accuracy of the forecast model is difficult to meet the actual demand. In this paper, through in-depth study of the characteristics of ultra-short-term load, an ultra-short-term load forecasting model based on fuzzy clustering and RBF neural network (FCM-RBF) is constructed. The model not only considers the historical load characteristics of locally similar days, but also considers the current load characteristics of the forecast days. The load on a locally similar day can well reflect the overall trend of the predicted load; the current load on the forecast day can well reflect the changing law of real-time data during the forecast period and some random factors in the forecast period. Finally, a power grid load in a certain area of southwestern China is selected as an example to verify the effectiveness and accuracy of the proposed method.

1. Introduction

At present, there are many methods of load forecasting. Compared with traditional methods, there are regression analysis methods, time series method, line extrapolation method, gray forecasting, etc. These algorithms are simple, fast, and widely used. However, with the rapid growth of the social economy, industrial electricity, commercial electricity, and residential electricity consumption in the distribution network show rapid growth and diversified load characteristics [1]. Load characteristics between different users are different due to the production process and the peak and valley periods of electricity consumption between different industries. Simultaneously, power users' load characteristics also change with the seasons, weather changes, typical days, and power consumption areas. Part of the load fluctuates sharply in a short time, and there is a tremendous peak-to-valley difference. The current load forecasting technology has a new development trend, and some intelligent algorithms led by neural networks have sprung up like mushrooms. Ref. [2] first applied a neural network to power load forecasting, and thus load forecasting entered the era of intelligent algorithms. Ref. [3] uses a radial basis function neural network (RBF) for short-term load forecasting research. The results show that in load forecasting, the RBF network has higher accuracy than the BP network. Ref. [4] uses a traditional neural network (CNN) and the hierarchical neural network (HNN) to study ultra-short-term load forecasting. The results obtained by the two reach the same level, but the

hierarchical neural network has more advantages in dealing with dynamic load characteristics. Good robustness, but its network training characteristics are relatively poor. Ref. [5] uses an adaptive neuro-fuzzy system to study ultra-short-term load forecasting, and this method has achieved ideal results in forecasting accuracy. Ref. [6] proposed an improved grey model algorithm to study ultra-short-term load forecasting, which improved the forecasting accuracy for the load inflection point.

Due to the rapid development of the current power industry, the distribution network's spatial load forecasting plays an important role. The internal structure of the power system is involved. The entire power industry has gradually increased requirements for low-consumption and high-efficiency load forecasting, achieving the load forecast results more accurately. Therefore, this paper uses the method of combining fuzzy clustering and RBF neural network to perform ultra-short-term load forecasting on the distribution network and verifies that the method can effectively improve the load forecasting accuracy through an example.

2. Load forecasting model based on fuzzy clustering and RBF neural network

2.1. Locally similar day model

* Corresponding author: 18830251935@163.com

Firstly, we can see from the load characteristics that historical load can be divided into general working day load, rest day (Saturday and Sunday) load, general holiday load, November and Spring Festival holiday load, etc. According to the data type because both National Day and the Spring Festival are both. It is a significant holiday, and the loads on the November and Spring Festival are usually classified into one category. Then we can divide the historical load into four categories. Because the fuzzy clustering algorithm is unsupervised, it often does not require too many historical samples in the classification process. Therefore, we often only need to use the data of the previous month of the prediction date as historical samples for fuzzy classification, and the most recent month after the historical data is classified by fuzzy classification, the fuzzy membership degree of the category of the predicted daily data can often be mapped well. When considering the selection of the clustering parameter C , this article considers two schemes. Option 1: When the selected historical load sample does not include the load during the National Day or Spring Festival, we can take the fuzzy clustering parameter $C=3$; when the selected load sample includes the load during the National Day, the Spring Festival, we take the parameter $C=4$.

Secondly, the local similarity day model is used to select the appropriate sample data as the training sample of the neural network. This paper selects m similar simultaneous segment loads from the most recent n historical days as the selection of local similar day sequence objects $L_{i,t}$, among which,

$$L_{o,t} = \{l_{o,t-d-1}, l_{o,t-d-2}, \dots, l_{o,t}\} \quad (1)$$

$$L_{i,t} = \{l_{i,t-d+1}, l_{i,t-d+2}, \dots, l_{i,t}\}, i = 1, 2, \dots, m \quad (2)$$

Where, $l_{o,t-d+1}$ represents the load value at time $t-d+1$ on the forecast day; $l_{i,t-d+1}$ represents the load value at time $t-d+1$ on the i day.

Take the first d points of the predicted time $t+1$ from the latest n historical daily data to form a sequence group $L_{k,t} (k = 1, 2, \dots, n)$, and then send the sequence group $L_{k,t} (k = 1, 2, \dots, n)$ and the target sequence $L_{o,t}$ to the FCM algorithm for fuzzy classification, using the local similarity model equation (3) to Determine the similarity difference between each group sequence in the sequence group $L_{k,t} (k = 1, 2, \dots, n)$ and the target sequence.

Definition: the "similar difference" between vector x_k and target vector x_0 is:

$$D_k = \sum_{i=1}^C \mu_{i,o} * |\mu_{i,k} - \mu_{i,o}| \quad (3)$$

Among them, D_k represents the similarity difference between the vector x_k and the target vector x_0 ; $\mu_{i,o}$ represents the membership degree of the target vector belonging to the i cluster prototype; $\mu_{i,k}$ represents the membership degree of the vector x_k belonging to the i cluster prototype. The smaller the similarity difference in this model, the higher the trend similarity between the vector x_k and the target vector x_0 .

Finally, by comparing the size of the similarity difference, select m of the n historical days with smaller similarity differences as the local similarity days.

2.2. Super short-term load forecasting models

RBF neural network has strong nonlinear mapping ability and good convergence characteristics, and can theoretically approximate any nonlinear curve. Therefore, when building a load forecasting model through an RBF network, historical load changes are often considered to predict future load changes. This article proposes the following prediction models:

$$L_{k,t+1} = (1 + \tilde{\delta}_{k,t+1})L_{k,t} \quad (4)$$

Where, $L_{k,t+1}$ is the predicted load at time $t+1$ on the predicted day; $L_{k,t}$ is the real-time load at time t on the predicted day, and is the known load; $\tilde{\delta}_{k,t+1}$ is the rate of change of the load at time $t+1$ on the predicted day.

$$\tilde{\delta}_{k,t+1} = f(\delta_{j,t+1}, \delta_{j-1,t+1}, \dots, \delta_{j-n,t+1}) \quad (5)$$

Where, $\tilde{\delta}_{k,t+1}$ is obtained through neural network prediction and the load change rate at time $t+1$ on the predicted day; $\delta_{j,t+1}$ is the load change rate at time $t+1$ on the j day of the similar day. The similar daily load change rate can be obtained smoothly by formula (6).

$$\delta_{j,t+1} = \frac{L_{j,t+1} - L_{j,t}}{L_{j,t}} \quad (6)$$

The specific process of load forecasting of RBF neural network combined with cluster analysis is as follows:

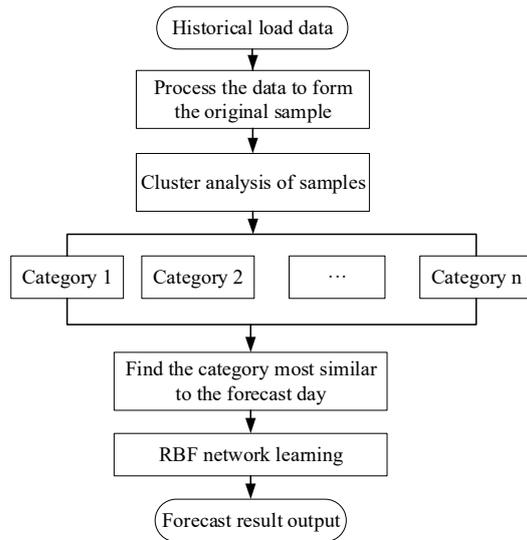


Fig. 1. Flow chart of load forecasting

3. Case analysis

3.1. Spatial load data processing

This paper needs to use a large amount of historical data as the basis when constructing the distribution network load forecasting neural network, but when obtaining the corresponding parameter information, it needs to be carried out through the power collector and other equipment, so the data may be affected by the equipment itself. Error, so before constructing, we must first process the acquired load data, and choose different processing methods for different data types. Data types can be divided into error data, missing data, and series error data. The error data is based on a certain value range, using effective filtering means. For example, for a certain load data, first assume that its actual value is Q and $Q > 0$, and the corresponding temperature parameter is set to $[-25^{\circ}\text{C}, 35^{\circ}\text{C}]$, the corresponding humidity, rainfall and water pressure are all numbers greater than zero. According to these conditions, select the qualified value to input into the neural network [7]. The source of the missing data is extracted from the original parameters and the corresponding filtering links. If the actual missing data are less, it can be filled by linear interpolation. The filling formula is

$$T_{n+j} = \frac{T_{n+i} - T_n}{i} \times j \quad (0 < j < i) \quad (7)$$

Where, T_{n+j} is expressed as an inferred intermediate value; T_n, T_{n+i} expressed as a specific temperature parameter at a known time $n, n+i$.

Series error data are adjusted according to the mean value of the sum of the parameter data before and after,

and the evolution parameters of the previous week are integrated to obtain new data.

Finally, in order to prevent the phenomenon of neuron saturation in the constructed neural network, corresponding processing operations should be carried out for the load data before the actual training, so as to solve other negative effects caused by the difference of the original data.

3.2. Forecast simulation results and analysis

Aiming at the ultra-short-term load forecasting method of distribution network, this paper uses pure RBF algorithm, grey relational analysis-RBF method (GRA-RBF) method and fuzzy clustering-RBF (FMC-RBF) method proposed in this paper to simulate. The gray correlation analysis-RBF method uses gray correlation analysis to extract similar local days. It then uses the ultra-short-term load forecasting model mentioned in this article to predict. In the load curve of a day, the peak and trough periods of the load are the most difficult to predict because the load curve changes in these two periods are complicated, and the curve fluctuation range is large, and it is difficult to track the trend of the curve. The load curve has an obvious upward or downward trend in the flat peak period, and the load curve is relatively flat and easy to predict. To prove the algorithm's feasibility, this article takes the grid load in a particular area of southwestern China in March 2019 as an example. In March, a load of this network area is between $[800, 1800]$ MW, with 12 data per hour. The data of a trough period on a particular day in March are predicted. The specific simulation results and analysis are as follows:

1) Pure RBF algorithm

Table 1. RBF ultra-short-term prediction results table

Time	True load value / $1.0\text{e}+003\text{*MW}$	RBF forecast load value / $1.0\text{e}+003\text{*MW}$	Relative error /%
15:00	1.3746	1.3715	0.36
15:05	1.3741	1.388	1.01
15:10	1.3814	1.3895	0.58
15:15	1.3757	1.3909	1.11
15:20	1.377	1.3924	1.12
15:25	1.3674	1.3939	1.94
15:30	1.3764	1.3953	1.38
15:35	1.3737	1.3968	1.68
15:40	1.3764	1.3982	1.59
15:45	1.3733	1.3917	1.34
15:50	1.3761	1.3931	1.24
15:55	1.3755	1.3964	1.39
Average relative error /%		1.23	

It can be seen from Table 1 that the RBF network is used to predict the ultra-short-term load sequence directly, the prediction result is not very ideal, and the average relative error is 1.23%. Besides, because the RBF network is used to directly predict the load value instead of predicting the load change rate, the relative

error at the time point is mostly more than 1%, so the forecast curve cannot accurately match the target curve in the trend.

2) GRA-RBF Method

Table 2. GPR-RBF ultra-short-term prediction results table

Time	True load value / 1.0e+003*MW	GRA-RBF forecast load value / 1.0e+003*MW	Relative error /%
15:00	1.3764	1.3711	0.38
15:05	1.3741	1.3721	0.15
15:10	1.3814	1.3732	0.6
15:15	1.3757	1.3742	0.11
15:20	1.377	1.3753	0.12
15:25	1.3674	1.3764	0.66
15:30	1.3764	1.3773	0.07
15:35	1.3737	1.3784	0.34
15:40	1.3764	1.3795	0.22
15:45	1.3733	1.383	0.71
15:50	1.3761	1.3841	0.58
15:55	1.3755	1.3852	0.7
Average relative error /%		0.39	

Using grey relational analysis to select locally similar sequences and then using the prediction model of this paper. The prediction effect obtained is correct. However, there is little difference between the predicted curve and the actual curve in amplitude, and the average relative error of the forecast through period is only 0.39%. From Table 2, it can be seen that the average relative error of each data is within 1%. The GRA-RBF algorithm uses the prediction model constructed in this paper, which shows that the prediction model proposed in this paper is quite feasible.

3) FCM-RBF method

Table 3. FCM-RBF ultra-short-term prediction results table

Time	True load value / 1.0e+003*MW	GRA-RBF forecast load value /1.0e+003*MW	Relative error /%
15:00	1.3764	1.3709	0.4
15:05	1.3741	1.3717	0.18
15:10	1.3814	1.3725	0.65
15:15	1.3757	1.3733	0.18
15:20	1.377	1.3741	0.21
15:25	1.3674	1.3749	0.55
15:30	1.3764	1.3757	0.05
15:35	1.3737	1.3765	0.2
15:40	1.3764	1.3773	0.07
15:45	1.3733	1.3808	0.55
15:50	1.3761	1.3844	0.6
15:55	1.3755	1.3852	0.7
Average relative error /%		0.36	

The FCM-RBF method proposed in this paper is used for prediction, and the result is basically ideal. The amplitude tracking can be seen from Table 3. All relative

errors of the method proposed in this paper are less than 1%, and the average relative error is only 0.36%, which is largely due to the selection and model of locally similar sequences forecast load change rate. Selecting relatively similar sequences from the historical load as the training samples of the neural network can better map the input space of the neural network. Using the most similar sequence as the test sample can ensure that the output of the neural network is closer to the real data. This is essential for ultra-short-term load forecasting.

4. Conclusion

This paper proposes an ultra-short-term load forecasting method for distribution network based on fuzzy clustering and RBF neural network. Combined with fuzzy cluster analysis, the RBF neural network method is used to simulate and predict the load of the distribution network. By analyzing and comparing the simulation forecast results of the three distribution network load forecasting methods, the errors of the forecast data during the peak and trough periods of a week are summarized. It proves the effectiveness and accuracy of the described ultra-short-term load forecasting method of distribution network based on fuzzy clustering and RBF neural network, and can provide reference for the research of ultra-short-term load forecasting of distribution network.

Acknowledgments

This article is funded by the Science and Technology Research Project of the Economic and Technical Research Institute of State Grid Hebei Electric Power Co., Ltd.: Research on seasonal power grid load characteristics and development trends based on data mining (SGHEJY00GHJS1900055).

References

1. Zhao T., Wang L.T., Zhang Y., et al. (2016) Relation factor identification of electricity consumption behavior of users and electricity demand forecasting based on mutual information and random forests. *J. Proceedings of the CSEE*, 36(3):604-614.
2. Park D.C., El-Sharkawi M.A. (1991) Electric loads forecasting using an artificial neural network. *IEEE Transaction on Power Systems*, 6(2):442-449.
3. Luo Y., Shao Z.C., Zhang L., et al. (2018) Robust economic dispatch and backup configuration considering wind power uncertainty and gas grid operation constraints. *J. Journal of Electrotechnical Technology*, 33(11): 2456-2467.
4. Liu M., Shi Y., Fang E.A. (2012) new operational strategy for CCHP systems with hybrid chillers. *J. Applied energy*, 95: 164-173.
5. Wu Z., Gu W., Wang R., et al. (2011) Economic optimal schedule of CIP microgrid system using chance constrained programming and particle

- swarm optimization. C. In: Power and Energy Society General Meeting. pp.1-11.
6. Li J.H., Sang C.C. (2105) Energy com-prehensive system optimization planning and operational framework. *J. Electric Power Construction*, 36(8): 41-48.
 7. Cui P.C., Shi J.Y., Wen F.H., etal. (2017) Optimal configuration of energy hub considering comprehensive demand side response. *J. Electric Power Automation Equipment*, 37(6): 101-109.
 8. Moradi M. H., Hajinazari M., Jamasb S., etal. (2013) An energy management system(EMS) strategy for combined heat and power(CIP) systems based on hybrid optimization method employing fuzzy programming. *J. Energy*,49: 86-101.