1. Introduction

During the construction of smart grid, advanced metering equipment and intelligent terminal equipment are installed and used in a large scale, and the power consumption mode of residents diversifies. It is very important to gain thorough insights into the power consumption mode of residents, thus improving the accuracy of load forecasting and ensuring the normal operation of power systems, energy management and planning [1–3].

Short-term load forecasting methods can be divided into time series correlation forecasting method [4,5] and machine learning forecasting method [6–8]. However, these methods are difficult to meet the requirements of high-resolution and personalized short-term load forecasting nowadays. In order to build better user level load forecasting

model, how to effectively improve the accuracy of user power consumption clustering results and load forecasting values has become a topic issue for researchers in recent years.

Ref. [9] proposed a combined algorithm of cloud computing platform and K-means clustering. The clustering accuracy reaches 91.2%, which proves the effectiveness of K-means algorithm in clustering power customers. Ref. [10] analyzes the characteristics of users' electricity consumption behavior and conducts cluster load forecasting and direct load forecasting. Ref. [11] analyzes the difference of electricity consumption behavior among users. It takes users' historical load data as clustering samples, combines typical load forecasting methods with clustering methods, uses different combined forecasting models for different categories of users, and finally evaluates these models based on actual examples. In Ref. [12], K-means algorithm is combined with self-organizing map neural network to cluster users with different electricity consumption patterns. In Ref. [13], FCM algorithm is used to select similar period of time for users, combining with RBF neural network to conduct load forecasting. The results show that the method achieves ideal results in forecasting accuracy, and the periodicity of user load curve is verified. The combined model of wavelet de-noising and decision tree is used in Ref. [14] to analyze the personalized electricity consumption behavior of users and carry out load forecasting, but it ignores the influence of time sequence factor on load forecasting.

By analyzing relevant research, it can be found that the key to improving accuracy is to deeply understand the characteristics of consumers' electricity consumption behavior and cluster the users reasonably. Therefore, this study takes the user behavior characteristics as a starting point, analyzes the user's load curve and classifies the users with their consumption characteristics. By clustering users with K-means algorithm and filtering out local similar daily data with the help of improved FCM, the corresponding FCM–BP load forecasting model is built according to the correlation between historical load data and time series of different types of users.

2. User classification based on K-means clustering algorithm

Fig. 1 shows the load curves of five random users in an area of Nanjing in 2018. Different color load curves represent different users. In the figure, the vertical axis represents the load value, and the horizontal axis represents the moment when the load value is collected. Data is collected every 30 min, and 336 data points are collected from each user.

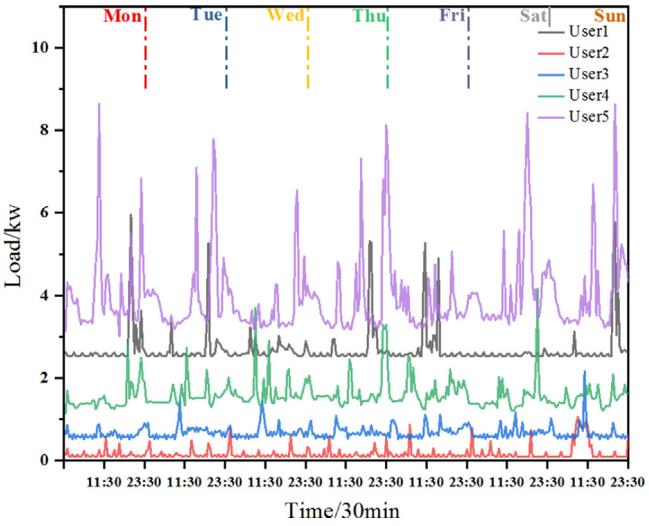

Fig. 1. A weekly load curve of 5 users. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

It can be seen from Fig. 1 that users 2 and 3 consume more electricity at night, and their load curves are cyclical and regular with very little fluctuation. The curves not only have similar load values at the same time on different days (daily correlation), but also have a gentle load curve in a short period of time (adjacent time correlation), which is common to residential users. The load curves of users 1, 4 and 5 have obvious peaks in mornings and afternoons

periodically. This means their load values are similar at the same time on different days (daily correlation), which is common among enterprise users. In this study, K-means algorithm is used for cluster analysis of users.

The K-means algorithm is a typical clustering algorithm. The main idea of the algorithm is to divide the samples into different clusters by continuous iteration until the objective function reaches its optimal value.

The specific procedure is as follows:

(1) Choose K appropriate points as the initial clustering centers.

(2) Calculate the value of d:

$$d = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{1}$$

In formula (1), d represents the distance between the other points and the center points, x_i represents the value of the ith variable in sample x, and y_i represents the value of the ith variable of the clustering center y.

(3) Classify all the points according their nearest center points, thus dividing all the sample points into K clusters.

(4) Calculate the center of mass of the K clusters and update them as new clustering centers.

(5) Repeat step (2), (3) and (4), keep iterating until the clustering centers stop shifting.

This study used the K-means algorithm to cluster the users into clusters A and B: the electricity consumption behavior of class A users has both daily correlation and adjacent time correlation while that of class B users only has daily correlation.

3. User side load forecast based on FCM– BP

3.1. FCM algorithm

This study used the FCM (Fuzzy C-Means) to select local similar days. The procedure of FCM is as follows:

(1) Select c appropriate clusters and a weighted index m. Generate an initial cluster center matrix V^0 randomly. The number of iteration is $0(l = 0)$.

(2) Calculate each cluster center V_i^{l+1}

$$V_i^{l+1} = \frac{\sum_{k=1}^{n}(u_{ik}^l)^m x_k}{\sum_{k=1}^{n}(u_{ik}^l)^m} \quad (i = 1, 2, \ldots, c) \tag{2}$$

In formula (2), x_k represents the kth element, u_{ik} represents the kth element's membership towards the ith cluster.

(3) Update the membership matrix and calculate the value of the object function.

$$u_{ik}^{l+1} = \left[\sum_{j=1}^{c}\left(\frac{d_{ik}^l}{d_{jk}^l}\right)^{\frac{2}{m-1}}\right]^{-1} \quad (i = 1, 2, \ldots, c; k = 1, 2, \ldots, n) \tag{3}$$

$$J^l(U^l, V^l) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^l (d_{ik}^l)^2 \tag{4}$$

In the formula above, U represents the membership matrix, d_{ik} represents the distance between the kth sample point and the ith cluster center.

(4) The given threshold $\varepsilon > 0$, if the object function satisfies the condition $\left|J^l - J^{l+1}\right| \leq \varepsilon$, then the iteration is stopped. Otherwise, let $l = l + 1$, iterate again and return to step (2).

(5) When the iteration is finished, observe the membership of all the samples of matrix U and classify them into the cluster with the maximal membership.

3.2. Selection of local similar days based on FCM

In order to select similar days for users, it is necessary to classify users' historical load curves first. According to the load features, the load can be divided into four types: ordinary workday load, ordinary weekend load, ordinary holiday load and major holiday load (National Day or Chinese Spring Festival). Since the FCM algorithm cannot determine the similarity between the target vector and each sample vector, in order to find the local similar days

of the load, the model is proposed in this study, which can reasonably determine the similarity between the target vector and each sample vector by amplitude and trend. The specific principle is as follows:

The "similarity difference" between the target vector x_0 and the sample vector x_k is defined as:

$$D_k = \sum_{i=1}^{c} u_{i,0} * \left| u_{i,k} - u_{i,0} \right| \tag{5}$$

In formula (5), D_k is the similarity difference between the target vector x_0 and the sample vector x_k; $\mu_{i,0}$ indicates the membership degree to which target vector x_0 belongs to the ith clustering prototype; and $\mu_{i,k}$ indicates the membership degree to which target vector x_k belongs to the ith clustering prototype. The smaller the D_k, the higher the trend similarity between the target vector x_0 and the sample vector x_k.

The value of the target sequence is the first d load points of the (t+1) times to be predicted. This study select m historical loads with high similarity from recent n historical days as a local similar sequence $L_{i,t}$, where:

$L_{0,t} = \left\{ l_{0,t-d+1}, l_{0,t-d+2}, \ldots, l_{0,t} \right\}$, $l_{0,t-d+1}$ represents the historical load at (t−d+1) times on the unpredicted day.

$L_{i,t} = \left\{ l_{i,t-d+1}, l_{i,t-d+2}, \ldots, l_{i,t} \right\}, i = 1, 2, \ldots, m, l_{i,t-d+1}$ represents the historical load at (t−d+1) times on Day i.

The selection of local similar days is divided into the following four steps:

Step 1: the first d points of (t+1) time on n historical days before the unpredicted day are selected as historical load sequence set $L_{k,t}$ (k = 1, 2, … n), and the first d points of (t+1) time on unpredicted day are selected as the target sequence $L_{0,t}$.

Step 2: the historical load sequence set $L_{k,t}$ (k = 1, 2, … n) and the target sequence $L_{0,t}$ are selected as samples, and the FCM Algorithm is used to implement fuzzy clustering.

Step 3: the similarity difference between the target sequence $L_{0,t}$ and various historical load sequences $L_{k,t}$ of the same cluster can be obtained by using the Formula (5).

Step 4: get the value of the similarity difference, and select m historical days having the smallest similarity difference as local similar days.

3.3. FCM– BP load prediction model

Generally speaking, BP Neural Network uses a three-layer network, including Input layer, Hidden layer and Output layer. Its structure is shown in Fig. 2.

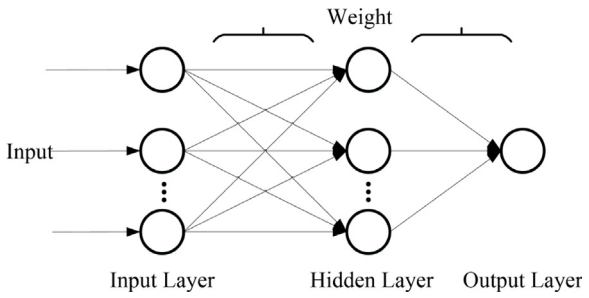

Fig. 2. Structure of BP neural network.

The input layer obtains the input vector, and then neurons will receive the information to generate corresponding weights and transmit them to the hidden layer. After that, the hidden layer neurons also generate weights and transmit them to the output layer. Finally, the output layer will generate the output vector. The generated output vector will compare deviation with the expected vector every time, and constantly revise itself, and finally the ideal neural network is obtained [15].

Since the load curves has the characteristics of daily correlation and adjacent time correlation, K-means algorithm can be used to classify users into two clusters —- A and B. The load curves of Cluster A users contains both daily correlation and adjacent time correlation, while the load curves of Cluster B users has only daily correlation

characteristics. FCM algorithm is used to select similar days. For Cluster A users, the local similar sequence of (d+1) times of the unpredicted day can be calculated by using the method described in Section 3.2. The selected local similar sequence and the previous d points of $L_{0,t}$ historical load value are taken as the user information set. For Cluster B users, FCM algorithm is used to calculate the local similar sequence with the minimum similarity difference as the user information set. Finally, the sets are used as input data for BP Neural Network, which is utilized to forecast the short-term load.

The specific process is shown in Fig. 3.

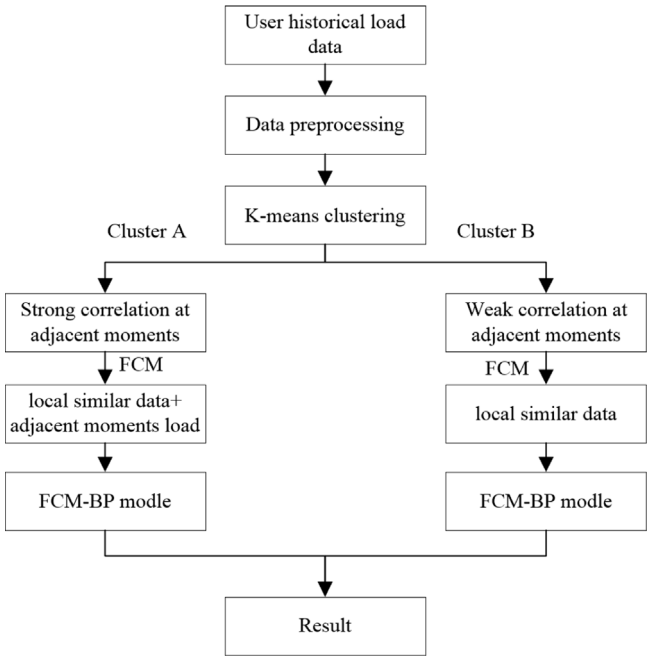

Fig. 3. Flow chart of short term load forecasting model based on K-means and FCM–BP..

The load rate-of-change of each group can be obtained by using Formula (6) to process the local similar days' load sample data.

$$\delta_{j,t+1} = \frac{L_{j,t+1} - L_{j,t}}{L_{j,t}} \tag{6}$$

For Cluster A users, the load rate-of-change and the load value at adjacent times are taken as BP training samples; while for Cluster B users, only load rate-of-change is taken as BP training samples. Finally, the load data at the unpredicted time can be calculated by Formula (7).

$$L_{k,t+1} = (1 + \widetilde{\delta}_{k,t+1})L_{k,t} \tag{7}$$

where h is the load rate-of-change at (t+1) times of the unpredicted day, which is calculated by Formula (8):

$$\widetilde{\delta}_{k,t+1} = f(\delta_{j,t+1}, \delta_{j-1,t+1}, \ldots \delta_{j-n,t+1}) \tag{8}$$

4. Example simulation

4.1. Data description and pre-processing

The data set used in this study is the load data of 200 users from an area of Nanjing in 2018. One data point is sampled every 30 min, and 48 load point data can be obtained daily. This study pre-processes the collected samples first. For the distorted or missing abnormal data, this study adopts the horizontal smoothing method to process, that

is, to replace the abnormal value with the average value of the first 16 adjacent data points of the abnormal value. After processing the abnormal data, normalize all load data. The normalization formula is as follows:

$$x_i' = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}} \tag{9}$$

In formula (9), x_i represents the sample data, x_i' represents the normalized sample data, x_{max} represents the maximum value in the sample data, and x_{min} represents the minimum value in the sample data.

4.2. Evaluation index

In order to effectively evaluate the accuracy of the model and compare it with other algorithms, MAPE (Mean Absolute Percentage Error) is adopted as the model evaluation index. The formula is as follows:

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\left(\left|\frac{\hat{y}_i - y_i}{y_i}\right| * 100\%\right) \tag{10}$$

In the formula, n represents the number of points to be predicted, y_i represents the real load data of the ith point to be predicted, and $\hat{y}_i$ represents the predicted load data of the ith point.

4.3. Analysis of prediction results

In this study, K-means algorithm is used to cluster 200 users into cluster A and cluster B. Cluster A users have both daily correlation and adjacent time correlation; Cluster B users have only daily correlation characteristics, and 15 typical users are selected for display. The classification results are shown in Table 1.

Table 1. User clustering results.

User Name	Cluster	User Name	Cluster
User 1	A	User 9	A
User 2	B	User 10	B
User 3	A	User 11	A
User 4	B	User 12	A
User 5	A	User 13	A
User 6	B	User 14	A
User 7	B	User 15	A
User 8	B		

Extract the classified user information set as the training set. The training set uses the data from January 1, 2018 to August 1, 2018; the test set uses the data from August 2, 2018 to August 31, 2018 data. The parameters of the BP neural network in this article are: 128 hidden layer neurons, the training accuracy is 0.0001, the maximum number of iterations is 15 000, the learning rate is 0.1, and the maximum number of failures is 256. Calculate the average MAPE of this model, RBF, GRNN, BP respectively, and the results are shown in Table 2.

Table 2. Comparison of MAPE of different algorithms (%).

Cluster	BP	RBF	GRNN	Proposed Model
A	2.4733	2.5034	2.4220	2.1066
B	2.4990	2.4862	2.2282	2.1887

By observing Table 2, it is not difficult to find that the model is more accurate than the three other models. From Fig. 4, it can be more intuitively seen that the load curve approximation degree of the model is better than that of RBF, GRNN and BP model. In summary, the model proposed can effectively improve the accuracy of user-level load forecasting. The reasons are as follows:

(1) The model can screen out the historical sequence with the highest similarity in the sample as the BP training sample. In this way, the input space of BP can be mapped more reasonably, and the historical load sequence with the highest similarity as the input vector can ensure that the output of BP is closer to the true value.

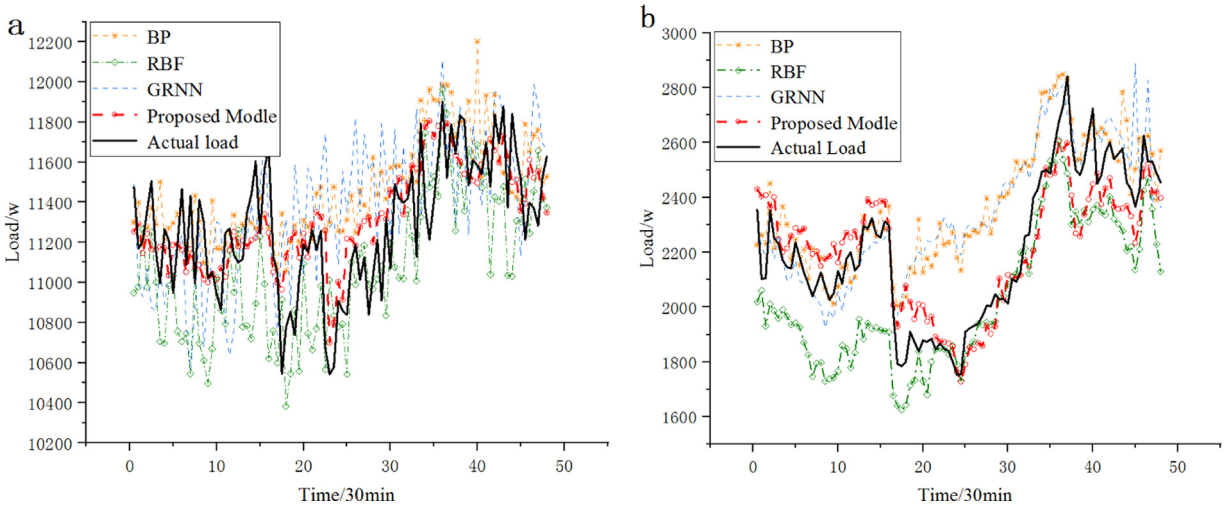

Fig. 4. (a) Comparison of A-type user load curve prediction (b) Comparison of B-type user load curve prediction.

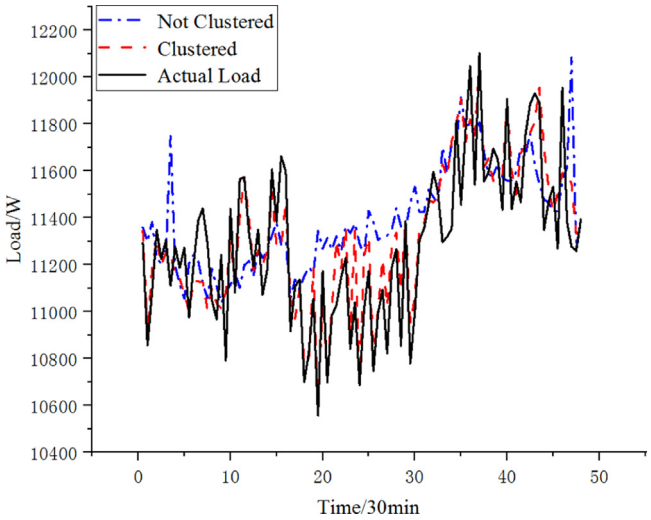

Fig. 5. Comparison of clustering effect.

(2) The model forecasts the load value by calculating the load rate-of-change at the moment before it is predicted without direct prediction. The load value actually includes the climatic factors, system operation factors, and other random factors at the time to be predicted. Making full use of the external information will help improve the accuracy of short-term load forecasting.

In order to test the effectiveness of user clustering, the model in this study is compared with the unclustered FCM–BP model, and the prediction curve is shown in Fig. 5.

In Fig. 5, the FCM–BP prediction curve fits better after clustering than without clustering. This fully demonstrates that the K-means clustering algorithm can distinguish users with different power consumption characteristics.

5. Conclusion

In order to predict the user-level load more accurately, this study analyzed the characteristics of user electricity consumption, and then used the K-means algorithm to cluster users into two clusters. After clustering, FCM is used to extract the user's characteristic sequence. Finally the user information set is input and the BP neural network

is applied to forecast. By analyzing the experimental results, the prediction accuracy of the model in this study is higher than that of RBF, GRNN, BP neural network, and unclustered FCM–BP model, which proves that the method proposed can effectively improve the prediction accuracy of the model. Since the current model only considers the time-series correlation of the load and involves few external factors, it is difficult to improve the model accuracy any further. In the future, subsequent work includes mining the massive user load data in the smart grid, and make full use of it to improve the clustering effect on power users, optimizing the clustering method, and further analyzing the influence of economic and climatic factors on the accuracy of the load forecasting model, building a user-level high-resolution and high-precision load forecasting model.

References

[1] Jiye Wang, Zhixiang Ji, Mengjie Shi, Fupeng Huang, Chaoyang Zhu, Dongxia Zhang. Demand analysis and application research of intelligent distribution electricity big data. Chin J Electr Eng 2015;35(8):1829–36.

[2] Dongxia Zhang, Xin Miao, Liping Liu, Yan Zhang, Keyan Liu. Research on the development of smart grid big data technology. Chin J Electr Eng 2015;35(1):2–12.

[3] Ran Hao, Qian Ai, Fei Xiao. Research on power consumption analysis framework based on multivariate big data platform. Power Autom Equip 2017;37(8):20–7.

[4] Al-Hamadi HM, Soliman SA. Fuzzy short-term electric load forecasting using Kalman filter. IEE Proc Gener Transm Distrib 2006;153(2):217–27.

[5] Xiaojing Li, Chuntao Li, Lanmei Cong, Ziyi Ren, Hongliang Luo, Yuwen Wang, Hui Yuan, Hao Qiu. Short-term load forecasting based on dynamic weight similarity day selection algorithm. Power Syst Prot Control 2017;45(6):1–8.

[6] Zhiyou Cheng, Baihong Ding, Guoxiao Yu. Short-term load forecasting method based on IPSO-LSVM. New Technol Electr Eng Energy 2020;39(5):41–8.

[7] Dandan Zhang, Gang Hu, Jing Lu, Xiaodong Yin, Qiwen Ren. Short-term load forecasting based on GAD-BP neural network. Electron Measur Technol 2019;42(24):143–7.

[8] Jianhuan Zhang, Ying Ji, Lidong Chen. Application of deep learning in power load forecasting. Autom Instrum 2019;40(8):8–12, +17.

[9] Suxiang Zhang, Jianming Liu, Bingzhen Zhao, Jinping Cao. Research on the analysis model of residential electricity consumption based on cloud computing. Power Grid Technol 2013;37(6):1542–6.

[10] Beckel Christian, Sadamori Leyna, Staake Thorsten, Santini Silvia. Revealing household characteristics from smart meter data. Elsevier Ltd.; 2014, p. 78.

[11] Hang HH, Lin LS, Chen N, et al. Particle swarm optimization based non-intrusive demand monitoring and load identification in smart meters. In: Industry applications society meeting. 2012, p. 1–8.

[12] Bingyu Zhou, Bo Liu, Dan Wang, Yu Lan, Xiran Ma, Dongdong Sun, Qiuyi Huo. Cluster analysis of user interaction electricity consumption behavior based on self-organizing center K-means algorithm. Electr Power Constr 2019;40(1):68–76.

[13] Peng Xu. Research on load forecasting method based on fuzzy clustering and RBF neural network. Guangxi University; 2012.

[14] Mingfeng Yuan, Tao Liu, Xianwu Shan, Yichen Xu. Load characteristics analysis and forecasting based on industry clustering. Electr Autom 2019;41(5):77–9, +88.

[15] Dandan Zhang, Gang Hu, Jing Lu, Xiaodong Yin, Qiwen Ren. Short-term load forecasting based on GAD-BP neural network. Electron Measur Technol 2019;42(24):143–7.