Forecasting Model Of Electricity Sale Market Based On User's Electricity Consumption Information

Shifa Gao[1], Ai Wang[2], Zhigang Wang[3], Zheng Wang[4], Xing Tong[5], Renhui Liu

[1]Planning and Review Center, Economic and Electrical Research Institute of Shanxi Electrical Power Company of SGCC, TaiYuan, ShanXi, 030002, China

[2]Planning and Review Center, Economic and Electrical Research Institute of Shanxi Electrical Power Company of SGCC, TaiYuan, ShanXi, 030002, China

[3]Planning and Review Center, Economic and Electrical Research Institute of Shanxi Electrical Power Company of SGCC, TaiYuan, ShanXi, 030002, China

[4]Planning and Review Center, Economic and Electrical Research Institute of Shanxi Electrical Power Company of SGCC, TaiYuan, ShanXi, 030002, China

[5]R & D department, Shenzhen Orange AI Technology Co., Ltd., ShenZhen, GuangDong, 518000, China

[6]R & D department, Shenzhen Orange AI Technology Co., Ltd., ShenZhen, GuangDong, 518000, China

[*]Corresponding author’s e-mail: 779656332@qq.com

Abstract. The prediction of electricity sale market is directly related to the balance of electricity supply and demand. The high-precision prediction method of electricity sale market will effectively avoid the imbalance of system supply and demand, and improve the security and economy of the system. In the past, it was difficult to obtain the information of users' electricity consumption, so the prediction methods of electricity sale market in the past were based on the overall electricity sale. The advantage of this method is that it can get the forecast value of the future electricity sales quickly and intuitively, while the disadvantage is that it fails to consider the electricity development trend of all walks of life that compose the electricity sales in detail.

1. Overview

The prediction of the electricity sale market is directly related to the balance of electricity supply and demand. The high-precision prediction method of electricity sale market will effectively avoid the imbalance of system supply and demand, and improve the security and economy of the system [1]. In the past, it was difficult to obtain the information about users' electricity consumption, so the prediction methods of electricity sale market in the past were based on the overall electricity sale. The advantage of this method is that it can get the forecast value of the future electricity sales quickly and intuitively, while the disadvantage is that it fails to consider the electricity development trend of all walks of life that compose the electricity sales in detail [2-5]. In the massive power consumption information environment, it is possible to start from the bottom-up, starting from the development law of each user's power consumption, pack and cluster users with similar power consumption characteristics, find the best prediction model for each type of user's power consumption, and finally get more accurate overall power sales prediction value[6].

Based on a large number of user electricity consumption data, this chapter will analyse the power consumption characteristics of users in different industries and different scales; design reasonable criteria to achieve clustering of users with similar power consumption characteristics; analyse the sensitivity of each type of user's power consumption to factors such as electricity price, national economic development, season and weather, and find out the correlation between the above factors and each type of user's power consumption. After a large number of user information resources are subdivided into market segments and formed into the smallest particles of prediction, the best prediction model is sought to realize the personalized prediction of different user groups, and the most appropriate prediction model for each type of user is developed.

2. Analysis of power consumption characteristics of users with different scales in different industries

Based on the collected power consumption data of users and related technologies used in power consumption data management system, the power consumption characteristics of users in different industries and scales can be analysed. As shown in figure 1, the load rate zoning map of 50 typical users is divided into four categories by monthly average load rate of 0.4, 0.6 and 0.8, and then further divided into seven categories by "standard deviation of load rate / monthly average load rate = 0.1".

In the figure 1, the horizontal axis is the monthly average load rate, the vertical axis is the standard deviation, and the blue circle represents each user. It can be seen that A / C / D / F has a large number of users in its categories, while B / E / G has a small number of users in its categories. The following describes the basic characteristics of various users:

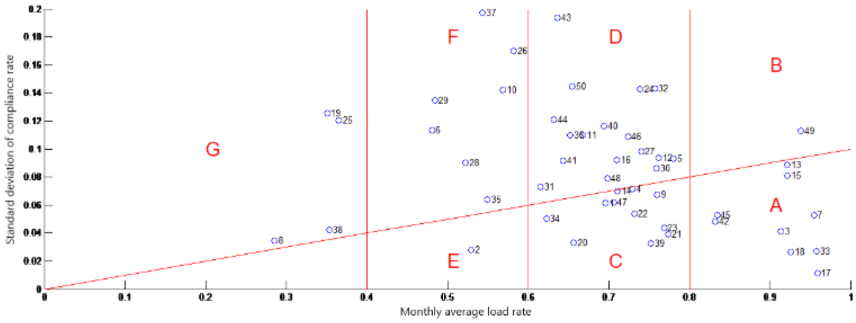

Figure 1. 50 typical user categories

A: This category is dominated by textile and high-tech industrial companies., with high load rate, stable load curve shape and small change in different days;

B: There is only one user in this category, which is Electronic Technology Co., Ltd. with high load rate, unstable load curve shape and large change in different days;

C: This kind of users are represented by steel / material / metallurgy, with medium load rate, stable load curve shape and small change in different days;

D: This kind of users are represented by cement / steel / metallurgy, with medium load rate, unstable shape of load curve and large change in different days;

E: There is only one such user, which is a special marine plate Co., Ltd. with low load rate, stable load curve shape and small change in different days;

F: This kind of users are represented by cement / steel / machinery manufacturing, with low load rate, unstable shape of load curve and large change in different days;

G: This kind of users are represented by steel / machinery manufacturing, with extremely low load rate, unstable load curve shape and large changes in different days;

From the above analysis, it can be seen that the load curve characteristics of users in textile and high-tech industries are stable, with little change in different days and high predictability; while the load curve of users in iron and steel / metallurgy / machinery manufacturing industries fluctuates greatly, and the difference between users in the same industry is also large, among which the load curve of some users (c) is relatively stable, and the daytime characteristics change little. The predictability is high, while the load curve of other users (d) is relatively stable, with large changes in daytime characteristics, which has certain predictability. There are also individual users (e.g. E / F / g) whose load curve fluctuates greatly, and is unstable between different days, which makes the prediction difficult.

The time delay analysis diagram of the above 7 types of typical users is drawn as follows:

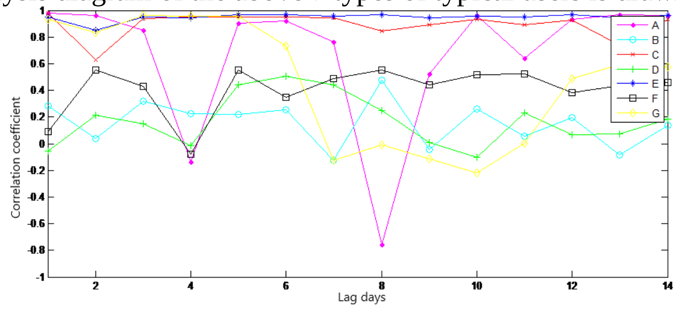

Figure 2. Delay analysis of typical users

3. Analysis of the relationship between different types of users' electricity consumption and related factors

It is not difficult to find from figure 3 that the daily electricity change characteristics of users which is belong to the class A have significant weekly fluctuation characteristics, in a week, it will occur the whole cycle of which electricity rising to the peak and then falling to the bottom. The annual electricity change consists of dozens of such fluctuation periods, including several holidays. During the holidays, the electricity of class A users will decline significantly, forming a step-by-step decline feature.

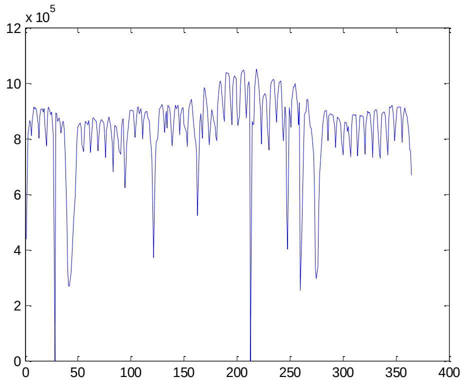

Figure 3. Daily electricity change curve of class A users

From figure 4, it can be seen that the weekly electricity change of class A users has been relatively stable, and its fluctuation is mainly affected by seasonal factors and holiday factors. For example, in summer, its electricity has increased significantly, and during the Spring Festival and national day, its electricity has decreased significantly.

Figure 5 shows the monthly electricity change curve of class A users. The monthly electricity change is relatively stable. In February (Spring Festival) and October (National Day), the electricity decreases, while in summer, the electricity increases.

4. Forecasting model of electricity sale market based on user's electricity consumption information

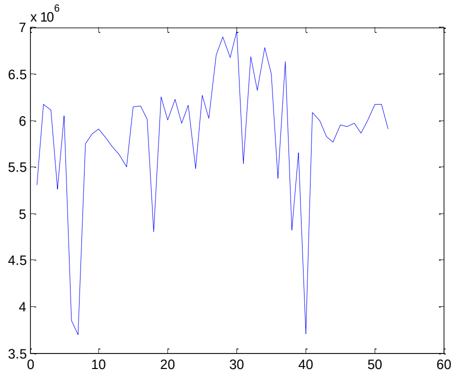

Figure 4. Weekly electricity change curve of class A users

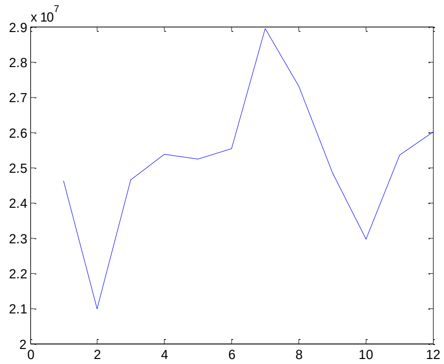

Figure 5. Monthly electricity change curve of class A users

According to the characteristics of large amount of electricity, different personalities and industry dominance of users, it is necessary to design reasonable criteria to realize the clustering of users with similar electricity consumption characteristics and the same industry [7-8]. After a large number of user information resources are subdivided into market segments and formed into the smallest particles of prediction, the best prediction model is sought to realize the individualization of prediction of different user groups, and each type of user is customized. Determine the most appropriate prediction model for electricity sales.

In this section, the idea of personalized user load forecasting based on decision tree is proposed. The core idea is to use decision tree method to build targeted load forecasting models and strategies for users with different load characteristics, so as to achieve high-precision user load forecasting. Specifically, for users with stable electricity consumption patterns such as electronics, shipping and textile, the typical electricity consumption patterns can be extracted directly through data mining, and then the prediction method can be selected according to the number of patterns; for users with large fluctuation in industries such as steel, metallurgy and non-ferrous metals, the load presents zigzag fluctuation characteristics, which needs to be firstly In order to improve the stability and regularity of the load, the frequency domain processing method is used to eliminate the fluctuation components, and the prediction is based on the remaining stable components.

4.1. Segmentation of user information resources based on decision of fluctuation component

Firstly, the user load fluctuation component is identified to determine whether the sawtooth fluctuation is significant. If it is significant, it will enter the wave component extraction phase, and then carry out

 1449 (2020) 012074 doi:10.1088/1742-6596/1449/1/012074

power consumption mode mining; if it is not significant, it will directly enter the power consumption mode mining phase.

The wave component identification method is wavelet packet decomposition method, through which the sawtooth wave can be adaptively eliminated. The removal effect is shown in figure 6: the blue curve is the original load curve of a certain day for a steel user, the orange curve is the smooth quantitation after the wavelet packet decomposition and the gray is the sawtooth wave component eliminated. The three are subject to the following equation:

Original smooth curve = Smooth component + Sawtooth wave component

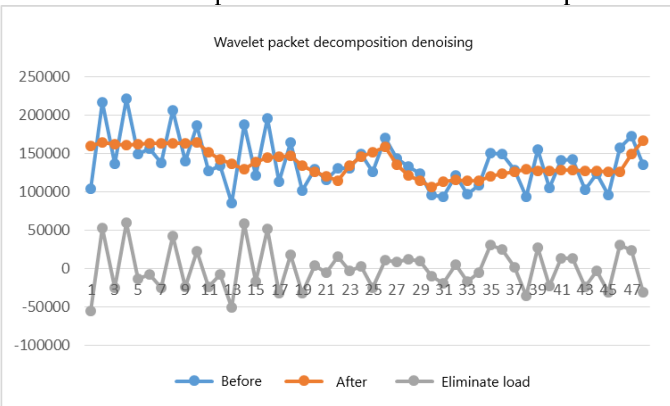

Figure 6. Denoising effect of wavelet packet decomposition

It can be seen that after wavelet packet decomposition, the load curve becomes very smooth and the predictability is greatly enhanced. The prediction method based on smooth component can significantly improve the prediction accuracy.

4.2. Mining and forecasting method with electricity mode

The most direct and significant factor affecting the user load is the production and living plan of the user [9]. However, it is difficult for grid enterprises to get the specific information of production and living plan from the enterprise. The method of big data mining provides an effective way to solve this problem. We can carry out data mining on the user's historical load, so as to extract the user's power consumption mode, and establish a forecasting method based on the power consumption mode to achieve accurate user load forecasting.

4.3. Short term load forecasting method based on electricity consumption pattern mining

This method is suitable for user load with moderate number of modes, Statistics are made according to the order in which the historical power consumption modes appear, the state transition matrix of the model can be obtained by using the method of statistical decision. According to this matrix, the power consumption mode of users in the day to be predicted is determined. Then the load curve of the same day mode is exponentially smoothed to predict the load on the forecast day. The calculation steps of the algorithm are as follows:

1. Using hierarchical clustering (unsupervised clustering) to cluster historical load, generate a typical set of power consumption modes $A = \{1,2,\cdots,i,\cdots,N\}$, And get the production mode of each historical day.

2. According to the identification results of user's historical daily production mode, generating state transition matrix of M, The matrix records the condition that the reference daily power consumption mode which is fixed to i, The daily electricity consumption mode to be predicted is subject to the probability of M_{ij} which is from typical production mode $j(j = 1{\sim}N)$, Based on base date u_0 and production mode i, and it could determine the production mode $M_{ik}(M_{ik} = max(M_{i1}, M_{i2}, \cdots, M_{iN}))$ with the maximum transfer probability, That is, the most likely production mode on the day r_0 to be predicted.

3. Select the historical days of which all power consumption modes is k from the historical days$\{r_1, r_2, \cdots, r_n\}$, The smaller the subscript, the closer to the date r_0 to be predicted, the nearest is r_1, and the farthest is r_n.

4. Smooth weight of index is obtained according to the distance between historical daily load and forecast day r_0, The closer to the date to be predicted, the greater the weight, weight w_l is as follows:

$$W_l = \alpha(a-1)^{l-1} \quad l = 1,2, \ldots, n \tag{1}$$

5. The daily load curve P_{r_0} to be predicted is obtained by summing the historical daily load curve r_l with weighting W_l.

$$\boldsymbol{P}_{r_0} = \sum w_l \boldsymbol{P}_{r_l} \tag{2}$$

P_{r_l} is the load curve of r_l.

5. A case study of user pattern mining prediction model

In order to analyse the advantages and disadvantages of the prediction method of electricity pattern mining, we compare it with the latest day prediction method, the latest three days prediction method and SVM prediction method. The basic principles of these methods are briefly described as follows:

Last day forecast method: the inertia of user load is small, and the daily load is usually the most similar to the load of the previous day, so the load of the previous day can be used as the forecast value directly.

Last three days forecast method: this method is an extension of the last three days forecast method. The load of the last three days is smoothed to get the forecast value.

SVM prediction method: SVM, also known as support vector machine, has very strong nonlinear modelling ability and adaptive prediction ability. The user's historical load is taken as the input data, and the daily load to be predicted is taken as the output data to form the "input-output" data pair. A large number of input-output data pairs are used to train SVM model, which can be predicted after training.

Using the above four methods, the load curve of each day of a month is predicted continuously for 50 typical user loads which sampled from A Province. As shown in figure 7, the horizontal axis is the user ID, the vertical axis is the prediction error value CMAPE, the Yellow curve is the prediction error of the pattern mining prediction method, and the blue, orange and gray are the error of the latest day prediction method, the latest three days prediction method and SVM prediction method respectively.

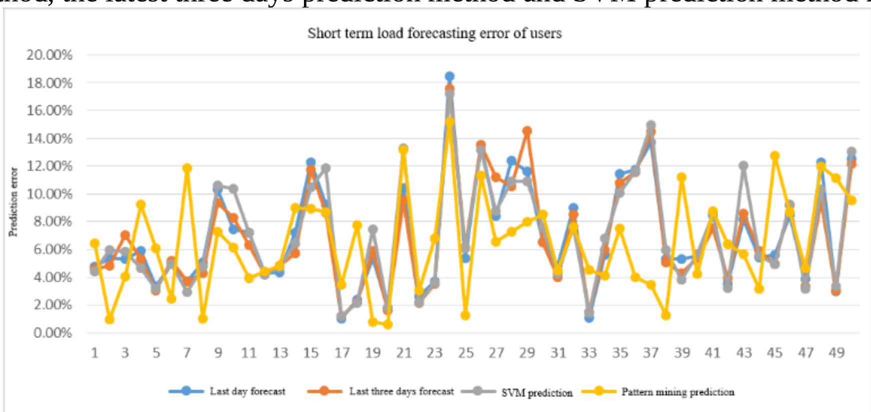

Figure 7. Short term load forecasting error of users

For each user, the prediction accuracy of the four methods is sorted to find the best prediction method. The proportion of each prediction method is the optimal prediction, as shown in Table 1. It is not difficult to find that the proportion of pattern mining prediction method is as high as 52.00%, which shows that for more than half of users, the effect of pattern mining prediction method is the best.

Table 1. Proportion of optimal prediction

Method	Best forecast share
Last day forecast	10.00%
Last three days forecast	20.00%
SVM prediction	18.00%
Pattern mining prediction	52.00%

The average accuracy of the four methods is calculated. As shown in Table 2, the accuracy of the model mining prediction method is 93.51%, which is 0.49% higher than that of other methods on average.

Table 2. Comparison of average prediction accuracy of methods

Method	1-CMAPE
Last day forecast	93.01%
Last three days forecast	93.14%
SVM prediction	92.91%
Pattern mining prediction	93.51%

6. conclusion

Based on a large number of user electricity data, this paper analyses the power consumption characteristics of users in different industries and scales. Through the design of reasonable user electricity behaviour criterion, the clustering of users with similar electricity consumption characteristics can be realized. On this basis, the sensitivity of each type of user's electricity consumption to electricity price, national economic development, season, weather and other factors is analysed, and the relationship between the above factors and each type of user's electricity consumption is found. After a large number of user information resources are subdivided into market segments and clustering of users with similar electricity consumption characteristics is realized, an appropriate electricity sales forecasting model is designed for different types of users.

Acknowledgement

Science and technology support project of State Grid Corporation of China, Contract number: SGSXJY00PSJS190003

References

[1] Liu WD, Yu DZ, Qiu HD, et al., (2019)Reform of Retail Electricity Marketization and Its Comprehensive Assessment System, Modern Electric Power, vol. 36(1), pp. 22-28.
[2] Lin NG, (2018)Regional Electricity Sales Forcasting Based on Empirical Mode Decomposition and ARIMA Model, Low Voltage Apparatus, vol.7, pp.73-77.
[3] Li W, Xu XY, Xia K, et al., (2018)Short-term Load Forecast Based on Radial Basis Function Extreme Learning Machine, Journal of Wuhan University of Technology(Information & Management Engineering), vol. 40(5), pp. 563-567.
[4] Ji B, Tan JC, Zeng XT, (2018)Energy storage optimal control strategy for renewable energy based industrial park under electricity market, Electrical Engineering, vol.19(8), pp.22-29.
[5] Liu XJ, Su XL, Yan XX, et al., (2017)Load Forecast of Active Distribution System Based on Active Response and Electricity Market, Proceedings of the CSU-EPSA, vol. 29(2), pp. 121-128.
[6] GARULLI A, PAOLETTI S, VICINO A. (2015)Models and Techniques for Electric Load Forecasting in the Presence of Demand Response. IEEE transactions on control systems technology: A publication of the IEEE Control Systems Society,3(3):1087-1097.

[7] CHEN Y-i bo, ZHENG Ling, YAO Jian-gang.(2016)Improved multiple regression load forecasting method based on rough set theory and D-S evidence theory[J].Power System Protection and Control, 44(6):62-67.
[8] YAN Wei, CHENG Chao, XUE Bin, et al. (2016)Forecasting for monthly electricity consumption using X12 multiplication method and ARIMA model[J].Proceedings of the CSU-EPSA, 28(5):74-80.
[9] WANG He, HU Zhi jian, CHEN Zhen, et al. (2013)A hybrid model for wind power forecasting based on ensemble empirical mode decomposition and wavelet neural networks [J].Transactions of China Electrotechnical Society, 28(9):137-144.