PAPER • OPEN ACCESS

Analysis of Characteristics of Industrial Power Load in a Certain Region

Analysis of Characteristics of Industrial Power Load in a Certain Region

QIAN Tiantian, WANG Ke, SHI Fei, ZHU Kedong

China Electric Power Research Institute Co., Ltd., Nanjing, Jiangsu, China

qiantiantian@epri.sgcc.com.cn

Abstract. Cluster analysis was done on the industrial load in the same class industry, based on the industrial daily load data collected in a certain region. The results obtained indicate that even in the same industry, there are still large differences in load characteristics, and there is a certain similarity between loads in different industries. Compared with the traditional results of cluster analysis of the overall industrial load, cluster analysis by industry will help power companies to carry out more precise load management.

1. Introduction
The industrial load accounts for a large proportion of the existing power load, so it is important to study its load characteristics. The current industrial load analysis mostly clusters the daily load curves of all industries collected in a certain area [1]-[5]. This analysis method can cluster the industrial daily load curves of the region into several typical characteristic load curves, such as continuous production, peak production, midnight production, and reverse peak production [6]. Basically, a characteristic type load can represent an industry typical load. However, with the upgrade of the power company's measuring equipment, the number of users that can be collected in each industry has also increased exponentially. At this time, if cluster analysis is only performed on the overall load of the area, this is not precise enough. Characteristic analysis should also be carried out for certain industry loads with a large proportion of load in the region. In this way, the grid manager is provided with more detailed load information, which can help them to make decisions. For example, in the development plan, they should introduce which type of load in the area, to have a positive impact on the peak-filling and valley-loading and load management capabilities of the area.

2. Analysis
Taken the industrial load of a certain area as an example, according to the industrial user load data of a certain working day in 2018 provided by the regional power grid company, the cluster analysis is performed on the load of several industrial users with more data collection. First, according to the national economic industry classification, the industrial load in the region can be divided into casting and other metal products manufacturing, communication equipment manufacturing, and electronic device manufacturing. For the collected raw data, the number of users under each type is first counted, and then the characteristics of each type of power load are analyzed.

2.1. Detection and processing of data outliers
The data of each unit is collected locally by smart meter equipment and sent to the power company via the public communication network. Due to factors such as network conditions, the acquired values

usually contain missing values and abnormal points. The lost data at a certain moment can be replaced by the average of the simultaneous load of other similar days. The outliers are detected using the Grubbs test. The process of discrimination is as follows.

- Calculate the mean and standard deviation of a set of load data.

$$\mu = \sum_{i=1}^{n} x_i / n \tag{1}$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2} \tag{2}$$

- Calculating the Grubbs statistic.
- Determine the detection level α. Refer to GB4883 to obtain the corresponding Grubbs test threshold $D_{1-\alpha(n)}$.

- When $G_n > D_{1-\alpha(n)}$, judge x_i is an abnormal value, otherwise there is no abnormal value.

- Give the reject level $D_{1-\alpha'(n)}$. When $G_n > D_{1-\alpha'(n)}$, x_i is the highly outlier, eliminating the entire set of data.
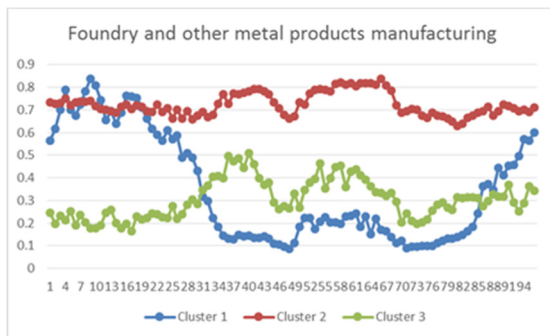
2.2. Normalization of data

Due to the different nature of the power users, the size of their electricity consumption data may vary greatly. In this way, in the clustering process, data with a large load value may affect the data with a small load value, and even annihilate it, so that the latter's role should not be reflected. To ensure that each individual has the same status for the entire data set in the analysis, the data is standardized. In this paper, the maximum normalization method is used, that is, each set of data is divided by the maximum value of the set of data, and converted into a number between 0-1.
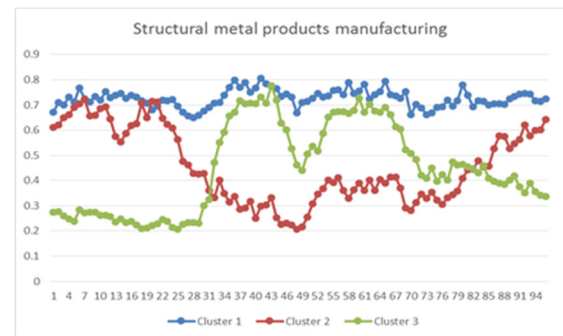
2.3. Cluster analysis

After the above data processing, then the daily load curve analysis of each industry is carried out. The most widely used clustering algorithm is k-means. First select k objects from all data objects as the initial cluster center. For the remaining objects, they are assigned to the cluster closest to them based on their distance from these cluster centers. Then recalculate the average of each new cluster as a new clustering center. This process is repeated until the criterion function converges. Usually the sum of the mean squared deviations of all data is used as the criterion function.
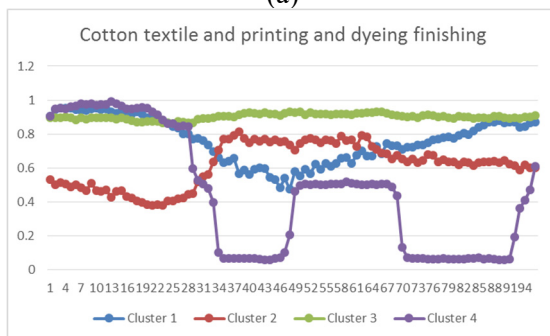
In this paper, a total of 1,440 power user data were screened, and the data collection density was one point every 15 minutes, a total of 96 points a day. Each of the data represents a working day load curve of a power user. A total of 188 industrial sector categories. Among them, the number of loads is more in the manufacture of foundry and other metal products, structural metal products manufacturing, cotton textile and printing and dyeing finishing, metal processing machinery manufacturing, and cement manufacturing, Manufacture of general-purpose components, manufacture of electronic components and electronic materials, manufacture of electronic components, manufacture of ships and related equipment, manufacture of transportation equipment and other transportation equipment. The load clustering results for each of the above industries are shown as Figure 1.
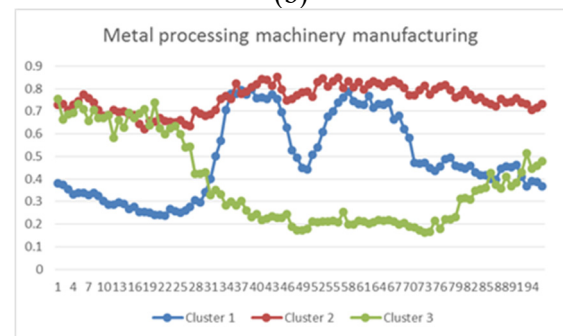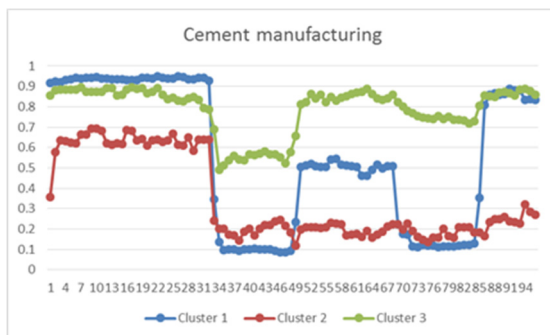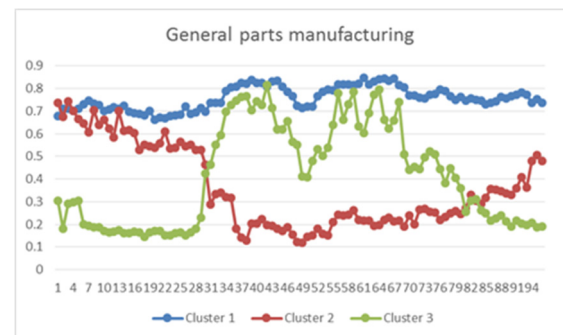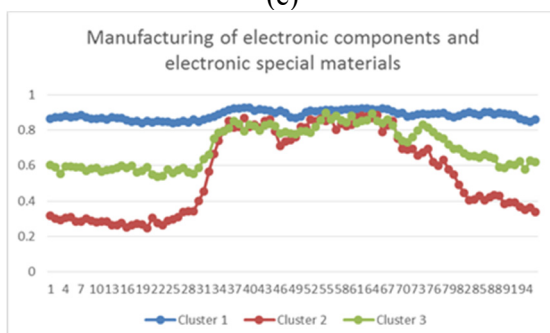
(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

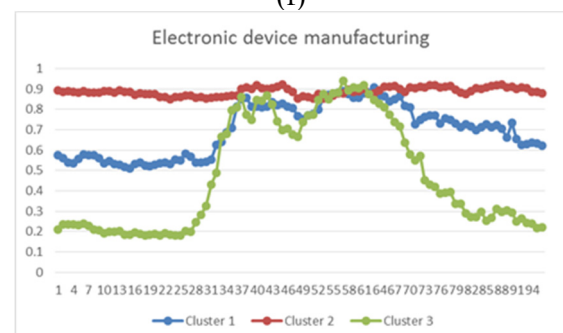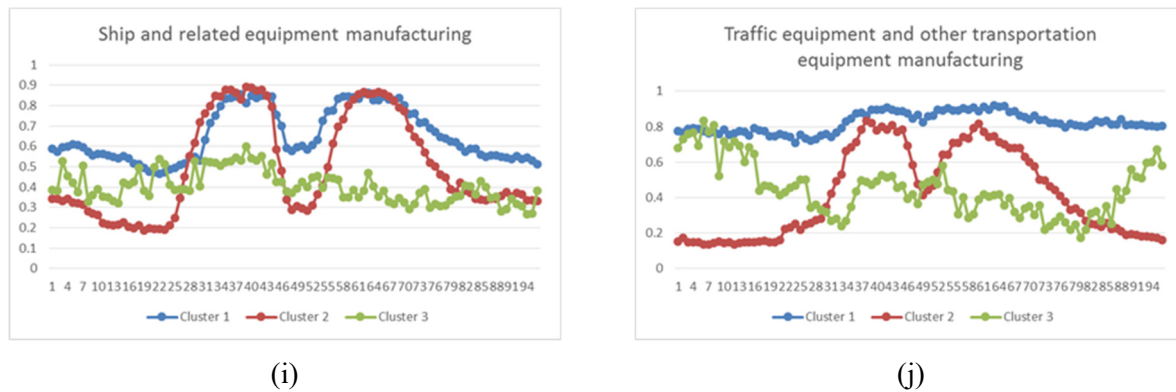(i)   (j)

Figure 1. Industrial load analysis in different industries

As can be seen from Figure 1, although the same is the industrial load, the load curve is also different between different industries. But there are some similar load curves, such as cluster 2 of graph (a), cluster 1 of graph (b), cluster 2 of graph (d), cluster 1 of graph (f), and cluster1 of graph (j). The shape of the load curve is similar, but it is still slightly different. Therefore, it is necessary to use quantitative indicators to make a difference. In this paper, the load rate is used as a quantitative indicator, and the calculation method is as follows. The calculation results of typical load curve load rates in various industries are shown in Table 1.

$$\beta = P_{av} / P_{max} \qquad\qquad (3)$$

Table 1. Average value and load rate of each load curve

Figure	Cluster	Average value	Load factor	Figure	Cluster	Average value	Load factor
(a)	1	0.35	41.93%	(f)	1	0.76	89.53%
	2	0.72	86.54%		2	0.36	48.72%
	3	0.30	59.93%		3	0.41	50.69%
(b)	1	0.73	90.07%	(g)	1	0.89	95.62%
	2	0.45	62.98%		2	0.56	63.08%
	3	0.45	57.62%		3	0.70	78.54%
(c)	1	0.76	80.14%	(h)	1	0.70	77.25%
	2	0.62	76.00%		2	0.89	96.18%
	3	0.90	96.75%		3	0.48	50.97%
	4	0.46	46.12%				
(d)	1	0.49	61.46%	(i)	1	0.66	76.04%
	2	0.76	88.45%		2	0.50	56.00%
	3	0.38	50.65%		3	0.40	67.43%
(e)	1	0.56	58.64%	(j)	1	0.82	89.28%
	2	0.34	49.68%		2	0.42	50.11%
	3	0.79	88.21%		3	0.45	53.84%

In Figure (a), at 9:00-21:00, the load of cluster1 is very low, which should be not put into the main productive load. At 0:00-7:00 and 22:00-24:00, the load is at a high level, which is a typical peak production load. Cluster2 has always maintained a high load level. At around 12:00, there is a low load, which is consistent with the general work schedule of workers. Cluster3 has the same trend as cluster1, but the average load and load rate are different. Figures (b), (d), (f) and (j) are similar to the case of Figure (a). In Figure (c), cluster4 is a typical inverse peak production type, but a high level of load occurs between 12:00 and 18:00. The load type of Figure (e) is quite different from that of other industries, and it is mainly the reverse peak production type. The load curves of figure (g), (h) and (i) are similar, and can be divided into continuous production type and daytime production type, in which the load rate of daytime production type is quite different.

3. Conclusions

This paper selects the industrial load of several different industries in a certain area. Cluster analysis was done on these industry loads. Several typical electrical load curves for their load within each industry are derived. Quantitative analysis was done with the load rate. It can provide theoretical analysis basis for the future load management of the power grid and the establishment of peak and valley electricity prices.

Acknowledgment

References

[1]   Zhong S, Tam K S 2015 Hierarchical classification of load profiles based on their characteristic attributes in frequency domain *IEEE Trans. Power Systems*, **30** 2434–41

[2]   Zhao Wenqing, Gong Yaqiang 2016 Load curve clustering based on Kernel K-means *Electric Power Automation Equipment* **36** 203–7

[3]   ZHU Xiaoqing 2012 Classification of power load and its application *Ph.D. dissertation* South China University of Technology

[4]   LIU Yongguang, SUN Chaoliang, NIU Zhenzhen 2014 Research on the improved fuzzy C-means clustering algorithm based power load characteristic classification technology *Electrical Measurement & Instrumentation* **51** 5–9

[5]   LI Zhiyong, WU Jingying, WU Weilin 2008 Power customers load profile clustering using the SOM neural network *Automation of Electric Power Systems* **32** 66–70

[6]   FENG Xiaopu, ZHANG Tiefeng 2010 Research on electricity users classification technology based on actual load curve *Electric Power Science and Engineering* **26** 18–22