Peak-valley time division model based on net load curve

Junhong Guo¹, Xiaqiang Xue¹, Ling Li¹, Yiping Cheng²,³ and Yajie Li²

¹ State Grid Jibei Electric Power Company Limited, Beijing, 100053;
² School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, 102206
³ Email: 13264243308@163.com

Abstract. As China has been quickly promoting its power market reform, the degree of thermal power units participating in market transaction has been gradually deepened. Aiming at different situations of load, thermal power units need to adopt bidding strategy accordingly. In power system with high proportion of renewable energy, the competition curve that thermal power units actually face is no longer total load curve. Proper division of peak-valley time intervals will guide thermal power units to adjust bidding strategy positively and effectively in the competition. A net-load-curve-based peak-valley time division model is built in this paper with the help of K-means algorithm. And then the paper contrasts and analyses the distributive characteristics of total and net load curves through a calculation example to verify the positive effect this model has on accurate bid for thermal power units.

1. Introduction
In recent years, China has been speeding up the process of electricity marketization, especially that of spot market reform. The first eight pilot provinces of electricity spot market have started up the trial thoroughly [1-3]. The proportion of bidding electricity quantity increases year after year because of thermal power units’ full participation in electricity-marketized competition [4]. It is crucial for thermal power units to bid flexibly according to different situations of load.

Over the late years, the proportion of renewable energy in our power system has been rapidly augmented. What’s more, our country is energetically taking several measures to guarantee the renewable-energy-first consumption [5-7]. Thus, instead of total load curve, the actual electricity quantity that thermal power units put into competition should follow net load curve, where the contribution of renewable energy has been deducted. This character particularly stands out in power system with high-proportion renewable energy. Therefore, this paper will research peak-valley time division based on net load curve.

At the moment, some experts and scholars have made some relevant analysis about peak-valley time division. In documents [8-9], the method of high-dimension of data is used to build up a data-sample set, which is covered with all information of load within a longer time period. Then, on this foundation, the peak-valley time division model is created. Documents [10-11] optimize the threshold with the help of threshold-optimization function to produce the single-day peak-valley time division model. Next, further efforts to dig the data help to build yearly division model. Although some improvements do have been made in the documents above aiming at the division of high-dimension and long-period time intervals, they are still based on total load curve essentially. Limitations will spring up when we apply them in power system with high-proportion renewable energy.
A peak-valley time division model based on net load curve is put forward in this paper. First, the paper analyses the reason why we choose net load curve to divide time intervals in power system with high-proportion renewable energy. Additionally, setting typical examples of California in the United State and Bai City, Jilin Province in China, it furtherly confirms the significance of net-load-curve-based time division. On this basis, the seasonal peak-valley time division model based on K-means algorithm is successfully set up. Finally, the calculation example contrasts the distributive characteristics of total and net load curves, authenticating the positive effect on accurate bid for thermal power units if we take advantage of the latter to divide time intervals.

2. Peak-valley time division based on net load curve

2.1. Reason analysis
In previous power system where thermal power units constitute majority parts, peak-valley time intervals were divided simply according to total load curve. However, along with the continuous development of the construction of our high-proportion renewable-energy system, there are some certain limitations in this traditional method.

System’s total load minus the contribution of renewable energy turns to be the remaining load, which is defined as net load. At present, the government is taking several measures to wholeheartedly guarantee the prior consumption of wind power and photovoltaic power, which two are the main types of renewable energy in our country. In this case, we can approximately regard the difference between total load and the contribution of wind and photovoltaic power as net load. Net load curve encompasses dual information of both sources and loads, not only representing actual load demand that thermal power units face, but also reflecting the responding competition curve of thermal power units in spot market. The peak-valley time division based on net load curve is more applicable for power system with high-proportion renewable energy.

2.2. Function
California and Bai City of Jilin are respectively typical areas with high proportion of grid-connected photovoltaic power and wind power. Figure 1 and 2 illustrate the situations of power output and load demand on a typical day in these two areas in different seasons.

![Figure 1](image1.png) Net load curve of US California on a typical day in spring.

![Figure 2](image2.png) Net load curve of Heilongjiang Harbin on a typical day in winter.
According to the figures above, net load curve of California appears to be duck-curve-shaped, utterly different from total load curve [12]. Compared with local total load curve, net load curve of Bai City fluctuates more severely and has a more obvious valley trend at night. By contrast, net load curve can present more correctly the actual peak-valley situations thermal power units face. Therefore, it is not only an important condition to ensure the safe and steady operation of power system, but also provides more practical price-leading signal for thermal power units participating in spot market competition to divide time intervals according to the distributive characteristics of net load curve.

3. Peak-valley time division based on K-means

3.1. Principle of K-means clustering algorithm

The levels of output differ dramatically in different seasons. In practical application, we should take typical-day net load curves of some area as study objects in spring, summer, autumn and winter respectively. Dividing time intervals separately over each season helps to meet the demand of system dispatching operation more flexibly and dynamically.

K-means is one of the most commonly-used clustering algorithm. It can effectively deal with a collection of larger-scale and high-dimension data and can also classify large data-sets efficiently. Thus, it is widely used in peak-valley time division of power system. The algorithm firstly divides data into several groups by the classification standard given in definition. Next, it calculates the distances from each object to the sub cluster center by iteration. Then, in turn, allocates each object to its nearest cluster center. Each group formed by such a method is called a cluster. This algorithm is a division scheme searching n clusters by iteration to make the global error minimal when all kinds of samples are replaced by these n mean values [13-14].

K-means has higher efficiency, while it can only deal with numerical data. In fact, the peak-valley time division problem is exactly to cluster and classify data according to load numerical values at each time point. Therefore, it just falls into the application category of the algorithm.

3.2. Steps of K-means clustering algorithm

Peak-valley time division based on net load curve mainly contains the following steps:

(1) Define boundary moments in studied section as time points so the net-load-value set corresponding to 24 time points is \( Q = \{ q_1, q_2, \ldots, q_{24} \} \).

(2) Randomly select three values from the net-load-sample-data set \( Q \) as peak, flat, and valley cluster centers \( C_j(I) \) \((j = 1, 2, 3)\). Meanwhile, set initial \( I = 1 \). For each sample datum, calculate its Euclidean Distance \( D \) to each cluster center:

\[
D = \sqrt{(q_i - C_j(I))^2}
\]

(3) Allocate all sample data to their nearest cluster center and define the data sets, belonging to three cluster centers, as \( W_j \). According to \( q_i \in W_j \), write down as \( q_i^{(j)} \). On this basis, regard the mean values of each set as new cluster centers \( C_j(I) \):

\[
C_j(I) = \frac{1}{n_j} \sum_{k=1}^{n_j} q_i^{(j)}
\]

(4) According to new cluster centers, recalculate the Euclidean Distances and reclassify the data. In order to set the terminating condition of the clustering process, we need to define criterion function \( J_e \) as followed of the sum of squared errors:

\[
J_e(I) = \sum_{j=1}^{3} \sum_{k=1}^{n_j} ||q_i^{(j)} - C_j(I)||^2
\]
Where, $n_j$ is the number of data in cluster set $j$. When two contiguous values of error come to the given accuracy, the clustering process ends. The corresponding result is the division scheme of the peak, flat and valley time intervals in the season.

4. Calculation example analysis

4.1. Net load curves in different seasons
Take area A for example, where the proportion of installed renewable energy is nearly up to 40%. It’s a typical grid-connected area with high-proportion renewable energy. Figure 3 demonstrates the power output and load demand curves on a typical day in area A separately in spring, summer, autumn and winter.

From the pictures we can see, total and net load curves of area A differ markedly in each season, which may lead to greater discrepancies in peak-valley time division results according to different curves. Net load curve reflects the real situation of generating units participating in competition, so in power system with high-proportion renewable energy, it will be more scientifically reasonable to divide peak-valley time intervals based on net load curve.

![Power output and load demand curves of area A in four seasons.](image)

Figure 3. Power output and load demand curves of area A in four seasons.

4.2. Outcomes of peak-valley time division in different seasons
Respectively studying the net load curves in four seasons above and utilizing K-means algorithm, we can reach the outcomes of peak-flat-valley time division, as shown in Figure 4.

Contrasting Figure 3 and 4, we will find that since the shapes of net load curves in different seasons differ remarkably, the results of peak-valley time division also vary dramatically with season.
5. Conclusions

(1) Along with the increasement of renewable-energy proportion in power system, there exist larger distinctions between total and net load curves. Thermal power units take part in market competition depending on net load curve, which is supposed to be divided based on its own distributive characteristics.

(2) Comparing the mean net load curves in different seasons in the calculation example, because photovoltaic power stations generate heavily at noon and wind power stations have a higher output-level at nights, net load is generally low at noon and nights. At the same time, compared with load curve, net load curve fluctuates more severely after a great deal of renewable energy join in power system.

References

[1] The National Development and Reform Commission, The National Energy Administration 2019
How to put a further approach to Chinese electricity spot market[J] Public Electricity Consumption 34(09) 4-5

[2] Anonymity 2019 The characteristics of the transactions of the first pilot units in electricity spot market[J] Public Electricity Consumption 34(09) 5

[3] The National Development and Reform Commission,The National Energy Administration
Notice on carrying out the pilot work of electric power spot market construction (Development and Reform Office Energy [2017]1453) [EB/OL] [2017-08-28].

[4] Li Ting, Miao Zengqiang 2017 A new model for generation scheduling distribution under the background of electric power system reformation[J] Guangxi Electric Power 40(02) 1-4+32

[5] Zhang Tao, GU Jie 2018 Markov short-Term load forecasting method for power system with high proportion renewable energy[J] Power System Technology 42(04) 1071-1078

[6] Zhu Jizhong, Xie Pingping, Xuan Peizheng, et al. Renewable energy consumption technology under energy internet environment [C] // The Institution of Engineering and Technology.2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Nov 26-28, 2017, Beijing, China.

[7] Cui Demin, Zhao Haibing, Fang Yanqiong, et al. 2018 Renewable energy production simulation...
considering the coordination between local consumption and transmission[J] Power System Protection and Control 46(16) 112-118

[8] Cheng Yu, Zhai Nana 2012 Electricity price peak and valley periods division based on customer response [J] Automation of Electric Power Systems 36(09) 42-46+53

[9] Liu Shuyong, Li Na, Fu Jingshuai 2018 Peak-valley time division model based on high-dimensional norm and SGHSA algorithm [J] China Electric Power 51(01) 179-184

[10] Yu Xinhua, Shu Yan, Zhao Huiru, et al. 2017 Peak-valley time-period partitioning model of social electricity consumption based on optimal threshold[J] Shanxi Electric Power 41(21) 9-16

[11] Liu Ming 2018 Pricing strategy research of residents’peak-valley time-of-use price based on data dig[D] Guizhou University

[12] Shaker H, Zareipouri H, Wood D, et al. 2016 Impacts of large-scale wind and solar power integration on California's net electrical load [J] Renewable & Sustainable Energy Reviews 58 761-774

[13] Yang Junchuang, Zhao Chao 2019 Overviews of K-means clustering algorithm research[J/OL] Computer Engineering and its Application 1-11 [2019-11-15]

[14] Li Qiuyan, Wang Yan, Sun Yujun Xiao Yong Zhang Chaoxin 2017 Applications of improved K-means algorithm in power users’ clustering recognition[J] Information Technology 10 108-112+117