Time-of-Use Period Partition Based on Improved Fuzzy C-Means and Abnormal Period Correction

Peng Wang, Yiwei Ma, Zhiqi Ling, Genhong Luo

***Abstract*—In time-of-use tariff period partition, clustering algorithms are commonly used. However, as load demands become more diverse in this big data era, large amount of non-linear data makes conventional clustering algorithms methods no longer be applicable in this field alone. Facing high-time-resolution daily load data with strong non-linearity, we propose a new method to partition periods. It consists of an improved fuzzy c-means clustering algorithm and a correction method for abnormal periods. Firstly, we propose modified fuzzy membership functions to improve the initialization of clustering for operation efficiency. Secondly, the method for calculating the fuzzy parameters based on the loss function is given. Thirdly, the initial period partition is obtained by the improved clustering. Next, the recognition model and fuzzy subsethood-based correction model for abnormal periods are structed, then the corrected period partition is confirmed. Finally, the effectiveness of the proposed methods is verified by two daily load data with a time resolution of 5 minutes.**

***Index Terms*—time-of-use tariff period partition, loss function, fuzzy membership functions, fuzzy c-means clustering algorithm, fuzzy subsethood, abnormal period correction.**

I. INTRODUCTION

WHIS the gradual deepening of electricity market reform, demand response based on price, an important interactive resource of the power system, can effectively guide customers to use electricity and improve system economy and reliability, and its role in the competitive market is becoming more and more obvious [1-4]. Time-of-use (TOU) tariff is a price-based demand response method which can solve the problems like rapid growth of electric load, supply that does not meet demand during the peak of electricity consumption and overgeneration during the valley period [5,6]. After the TOU implementation, the price differences between peak and valley periods make customers move some electricity consumption to the valley for reducing their electricity costs. Classical TOU based on the price elasticity matrix is studied in Literature [7,8], which provide many benefits, including promoting energy consumption, increasing load rate, and improving utilization efficiency [8]. So, it is necessary to make a reasonable period partition for customers to develop a feasible power consumption plan [9].

As a static electricity tariff mechanism, TOU is divided into various types [9]. For customers with time-of-day meters, the 24-hour day can be further divided into two prices for peak and off-peak [10-15], three prices for peak, flat, and valley [16-19]. Currently, various methods for partitioning TOU periods can be classified into two types: (i) load value-oriented, (ii) optimization objective-oriented.

(i) The load value-oriented methods can analyze the magnitude of the load, clustering algorithms like fuzzy c-means (FCM) and k-medoids are widely used. Fuzzy membership functions are used in Literature [20] to extract the relationship between the load magnitude at each moment to obtain the peak and valley periods. For practical implementation with minimal policy adjustment in Literature [21], the levels of TOU price remain unchanged, and the daily periods are divided into three clusters using FCM. Literature [22] also proposes a new FCM to partition daily periods. K-medoids is used in Literature [23] to build a peak-flat-valley (PFV) period partition model. In addition, the period partition model combined with probability distribution is also feasible in Literature [24].

(ii) On the other hand, the optimization objective-oriented methods allow for the allocation of periods to achieve the interests of policy maker. To reflect the connection between period partition and demand response, and maximize the utility of TOU, Literature [25] proposes the index of demand response to modify the period partition obtained only based on the membership functions. Literature [26] identifies valley periods by analyzing the redundancy of new energy generation, which can give full play to the demand response resources and promote renewable energy consumption. To fully consider the relationship between period partition and user behavior, and further expand the benefits, period partition is directly incorporated into the planning stage in Literature [27].

However, in some practical scenarios, we only have demand-side power data to support the period partition, without the support like historical period partition, prices of different periods and other related user information. Therefore, the role of the load value-oriented methods is not negligible.

This work was supported by the National Natural Science Foundation of China (No.61703068), and the Chongqing Postgraduate Research and Innovation Project (No. CYS22485). *(Corresponding author: Yiwei Ma.).*

Peng Wang is a postgraduate student of Department of Electrical Engineering, School of Automation and Industrial Internet, Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China (e-mail: 17623600676@163.com).

Yiwei Ma is an associate professor of Department of Electrical Engineering, School of Automation and Industrial Internet, Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China (e-mail: mayw@cqupt.edu.cn).

Zhiqi Ling is a postgraduate student of Department of Electrical Engineering, School of Automation and Industrial Internet, Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China (e-mail: lzq13535793794@163.com).

Genhong Luo is a postgraduate student of Department of Electrical Engineering, School of Automation and Industrial Internet, Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China (e-mail: 1468483560@qq.com).

The application of clustering algorithms in load values-based methods is extensive. But, the clustering effect is pretty sensitive to the initialization and iteration process. How to improve the algorithms has attracted widespread attention. Generally speaking, improvement methods can be divided into two classes: (i) improve the iterative approach, (ii) improve the initialization. (i): Literature [28] proposes an FCM with discriminative embedding to solve the problems including suboptimal results because of the influence of noises and redundant features. Literature [29] combines k-means and spectral clustering, then establishes a clustering method that can directly minimize the sum of the distances between points in the same cluster, the trouble of measuring the similarity between points is avoided. (ii): To deal with the random initialization, Literature [22] makes use of an improved particle swarm optimization based on support vector machine regression to find the initial cluster centers, makes it out of the local optimal, and thus classifies more accurate peak and valley periods. Literature [23] employs the maximum distance method to determine the initial center of each period. In this paper, we improve FCM based on initialization.

In terms of the time resolution of already researched power data, it is worth noting that all above discussed studies ranged from 15 minutes to 1 hour, as they make TOU policies based on different benefits or targets. However, we know that with the higher penetration of renewable energy sources and more diverse user demands, the load curve will become more complex and nonlinear. Studying more complex data can help us determine a more accurate distribution of PFV periods to improve overall benefits. The load data with different time resolutions are shown in Fig. 1.

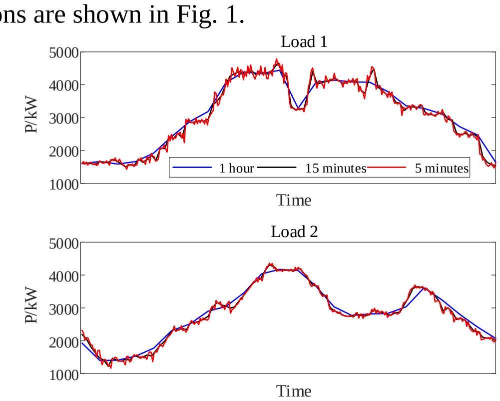

Fig. 1. Load with different time resolutions.

It is clear that lower time resolution means more information on demand response during any time interval will be blurred and thus difficult to be collected by the power supply side, which ultimately leads to lower user satisfaction as well as grid stability [11].

According to the above analysis, considering TOU with higher time resolution is meaningful. Therefore, we conduct a period partition study based on improved FCM (IFCM) and abnormal periods (AP) correction for load data with a time resolution of 5 minutes.

The contributions of this paper are described as follows:

(i) This paper proposes modified membership functions (MMF) which fully consider the characteristic of load curves. They can identify the period type of the load in specific intervals without the aid of clustering algorithms and prior knowledge. And the method for calculating fuzzy function parameters based on the loss function is given. This method ensures the maximum availability of MMF and is easy to explain and understand.

(ii) This paper establishes the AP correction method, which includes a recognition model and a correction model based on fuzzy subsethood (FSS). The correction is designed to address the problem that FCM clustering process cannot consider the scheduling rules in which the load period type remains unchanged over a longer period.

The remainder of the paper is organized as follows. Section II proposes the structure of MMF, and the calculation method of fuzzy parameters. Section III introduces the initial TOU period partition based on IFCM, and AP from clustering results. Section IV builds models for AP correction. Section V carries out the case study. The paper is summarized in Section VI.

II. Improved Initialization of FCM

Providing a reasonable initial membership degree in advance can overcome the randomness of FCM. Based on this, this paper aims to modify conventional membership functions (CMF) for improving the efficiency of clustering.

A. Conventional Membership Functions

Membership functions in general can be divided into three types: big-scale, small-scale, and middle-scale. In TOU period partition, CMF are usually used to describe the correlation between load and peak-flat-valley (PFV) periods [30]. The membership degree under CMF can make FCM clustering stable. The structure of CMF is shown in Fig. 2.

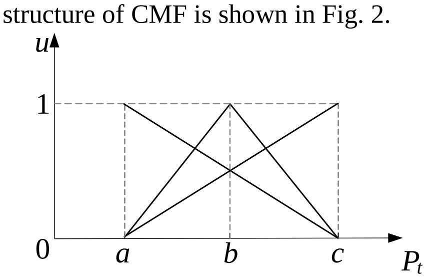

Fig. 2. CMF structure for PFV periods.

where, a, b and c are the minimum value, the mean value and the maximum value of the load, respectively.

Although CMF can calculate the PFV membership degree of any load curve easily and quickly. However, the uniform setting of their parameters ignores the variability of load curve features, resulting in CMF only stabilizing the clustering effect but not optimizing the clustering efficiency. And this is exactly what we are aiming at improving.

A. Modified Membership Functions

Unlike the structure and parameter setting of the CMF, modified membership functions (MMF) we proposed fully consider the characteristic of load curves.

This paper adopts the semi-trapezoid membership function to calculate the membership degree of peak period, as shown in Fig. 3 and (1). As the membership degree within $[a, b]$ is not high, the calculation of this part is not considered.

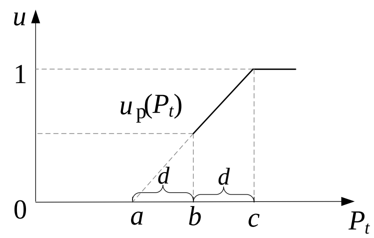

Fig. 3. MMF structure for peak period.

$$u_t^{\mathrm{p}} = \begin{cases} 0, & P_t < b \\ \dfrac{P_t - a}{c - a}, & b \le P_t < c \\ 1, & P_t \ge c \end{cases} \tag{1}$$

where, u_t^{p} is the membership degree of load point t to peak.

Similarly, MMF for valley period is shown in Fig. 4 and (2). The membership degree within [b , c] is not high, the calculation of this part is also not considered.

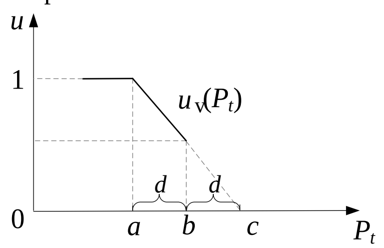

Fig. 4. MMF structure for valley period.

$$u_t^{\mathrm{v}} = \begin{cases} 1, & P_t < a \\ \dfrac{c - P_t}{c - a}, & a \le P_t < b \\ 0, & P_t \ge b \end{cases} \tag{2}$$

where, u_t^{v} is the membership degree of load point t to valley.

Here, MMF for flat period adopts a triangular shape which is shown in Fig. 5 and (3), and only the membership degree within [a , c] is considered.

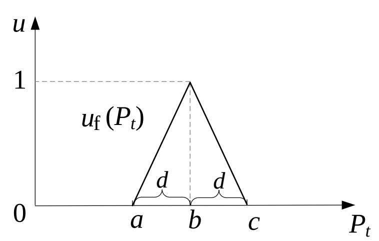

Fig. 5. MMF structure for flat period.

$$u_t^{\mathrm{f}} = \begin{cases} 0, & P_t < a \\ \dfrac{P_t - a}{b - a}, & a \le P_t < b \\ \dfrac{c - P_t}{c - b}, & b \le P_t < c \\ 0, & P_t \ge c \end{cases} \tag{3}$$

where, u_t^{f} is the membership degree of load point t to flat.

Compared to CMF, which fuzzes the entire load curve, some intervals of MMF are clear. In practical engineering, the fuzzy values of a and c can be confirmed based on the experience of TOU policy makers or a priori knowledge, but this still brings some risk. Therefore, MMF must be established with reasonable conservativeness. The fuzzy parameters: a, b and c are calculated by (4)-(6).

$$a = \min(\mathrm{Load}) + (1-\alpha)\left(\max(\mathrm{Load}) - \min(\mathrm{Load})\right) \tag{4}$$

$$c = \max(\mathrm{Load}) - (1-\alpha)\left(\max(\mathrm{Load}) - \min(\mathrm{Load})\right) \tag{5}$$

$$b = (a+c)/2 \tag{6}$$

where, $\alpha \in [0,1]$ is the level of conservativeness.

C. Confirmation of Fuzzy Parameters

Since CMF are applicable to any load curves, and the load points under these functions are correlated with each period (peak-flat-valley) at any time, we suppose that CMF are the most conservative. In this paper, the loss function (LF) based on mean square error is used to quantify the conservative difference between MMF and CMF. And α is confirmed by the variation of this difference to maximize the feasibility of MMF. The key calculation models are shown below.

$$\mathbf{u}_t^{\mathrm{mod}} = \left[u_{\mathrm{p},t}^{\mathrm{mod}}, u_{\mathrm{f},t}^{\mathrm{mod}}, u_{\mathrm{v},t}^{\mathrm{mod}}\right]^{\mathrm{T}} \tag{7}$$

$$\mathbf{u}_t^{\mathrm{con}} = \left[u_{\mathrm{p},t}^{\mathrm{con}}, u_{\mathrm{f},t}^{\mathrm{con}}, u_{\mathrm{v},t}^{\mathrm{con}}\right]^{\mathrm{T}} \tag{8}$$

$$\mathrm{LF}\left(\mathbf{u}_{\mathrm{mod}}, \mathbf{u}_{\mathrm{con}}\right) = \frac{1}{T}\sum_{t=1}^{T}\left\|\mathbf{u}_t^{\mathrm{mod}} - \mathbf{u}_t^{\mathrm{con}}\right\|^2 \tag{9}$$

$$\frac{d\mathrm{LF}}{d\alpha} = \lim_{\Delta\alpha \to 0}\frac{\Delta \mathrm{LF}}{\Delta\alpha} = \frac{\mathrm{LF}_{i+1} - \mathrm{LF}_i}{\alpha_{i+1} - \alpha_i} \tag{10}$$

where, $\mathbf{u}_t^{\mathrm{mod}}$ and $\mathbf{u}_t^{\mathrm{con}}$ are the membership degree vectors of the load under MMF and CMF, respectively.

The specific confirmation process of fuzzy parameters is shown in Fig. 6.

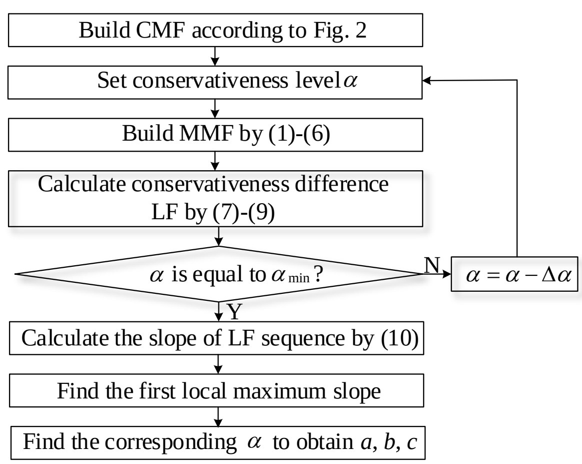

Fig. 6. Flow of fuzzy parameter confirmation.

Clearly, a lower α leads a larger value of LF. It is easy to understand that a larger value of LF means higher risk, the feasibility and conservativeness of MMF are more difficult to be guaranteed, and the selection of a and c is more unreasonable. So, the level of conservativeness α should be as high as possible. Besides, the local maximum slope means that the current variation of LF is the most noticeable in its neighborhood, indicating that the level α has a significant impact on the conservativeness of MMF. This is why we believe that the first local maximum slope is worth being chosen.

III. Initial TOU Period Partition

The initial membership degree from MMF help FCM to overcome the randomness, and optimize the rationality of initialization. Combining the above analysis, we then use FCM clustering to partition the initial TOU periods.

A. Flow of Period Partition Based on IFCM

From MMF, we can get PFV membership degree matrix **U**. It is expressed as follows.

$$\mathbf{U}=\begin{bmatrix} u_1^{\mathrm{p}} & u_2^{\mathrm{p}} & \cdots & u_T^{\mathrm{p}} \\ u_1^{\mathrm{f}} & u_2^{\mathrm{f}} & \cdots & u_T^{\mathrm{f}} \\ u_1^{\mathrm{v}} & u_2^{\mathrm{v}} & \cdots & u_T^{\mathrm{v}} \end{bmatrix} \tag{11}$$

The membership degree must meet constraint (12).

$$\text{s.t.} \sum_{k=1}^{c} u_t^k = 1, \quad 0 \le u_t^k \le 1 \tag{12}$$

So, **U** must be normalized to be applicable in clustering.

During the algorithm operation, the cluster centers and the membership degree update each other iteratively and finally converge. The calculation methods are shown below.

$$v_k = \frac{\sum_{t=1}^{T} (u_t^k)^2 P_t}{\sum_{t=1}^{T} (u_t^k)^2} \tag{13}$$

$$u_t^j = \frac{1}{\sum_{k=1}^{c} \left(\frac{|P_t - v_j|}{|P_t - v_k|} \right)^2} \tag{14}$$

When objective function (15) is satisfied, we stop clustering

$$\min J = \sum_{k=1}^{c} \sum_{t=1}^{T} (u_t^k)^2 |P_t - v_k|^2 \tag{15}$$

The flow chart is shown in Fig. 7.

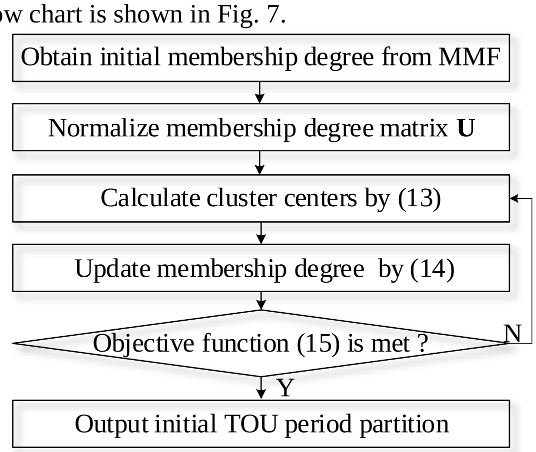

Fig. 7. Flow of initial period partition.

B. Abnormal Period Description

Based on experience and practical engineering, we need to ensure that the period type does not change frequently within a short period of time to support the formulation of TOU. However, from initial TOU period partition above, we can see that high-time-resolution load data with strong non-linearity leads excessively rapid changes of period type, which are not conducive to the decisions of actual scheduling plans.

For convenience, we name the rapidly changing periods as abnormal periods (AP) and the remaining periods as normal periods (NP). We select typical period partition to analyze the structure of AP and NP in Fig. 8.

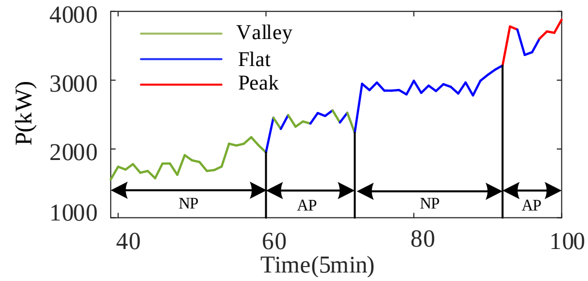

Fig. 8. Typical structure of AP and NP.

It can be seen that in a short period, the valley-flat period types in the first AP changes multiple times, while the peak-flat period types in the second AP also changes. This feature also exists in other periods. Extract this structure and obtain the distribution diagram as shown in Fig. 9.

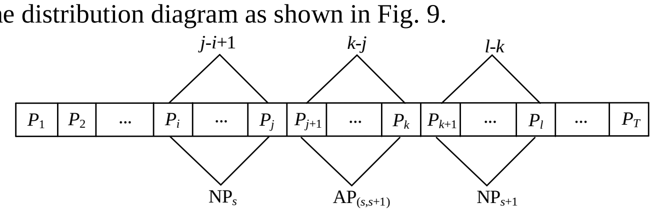

Fig. 9. AP and NP abstract distribution diagram.

The interval lengths (time lengths) of the first NP and the second NP are *j*-*i*+1 and *l*-*k*, respectively, and there is only one period type within them. The AP (length is *k*-*j*), contains at least two period types within it.

IV. Corrected TOU Period Partition

To correct AP, this paper firstly proposes the recognition model to filter the clustering results. After obtaining AP, the correction model is then constructed to optimize the clustering results and make them usable.

A. AP Recognition Model

The clustering results show that AP is distributed between different NP. Due to the complex period types within AP, it is difficult to build a universal model to directly find all different AP. But NP only has three types: peak, flat, and valley, therefore, filtering AP based on NP recognition is feasible.

We first provide some symbols to describe the model.

The clustering results of IFCM are represented as **X**.

$$\mathbf{X} = [v_1, v_2 \ldots f_i, p_{i+1} \ldots v_T] \tag{16}$$

Then, perform attribute value integer processing on **X**. We set peak-flat-valley as 3-2-1, then obtain the integer period partition **Y**.

$$\mathbf{Y} = [1,1 \ldots 2,3 \ldots 1] \tag{17}$$

If **Y** has n subsequences and T numbers, then they are represented as $\mathbf{Y}_{s1}, \mathbf{Y}_{s2}, \ldots, \mathbf{Y}_{sn}$ and $y_1, y_2, \ldots, y_T$, respectively; Len{.} is the length of a sequence; First{.} represents the first element of a sequence.

The flow of AP recognition model is shown in Fig. 10.

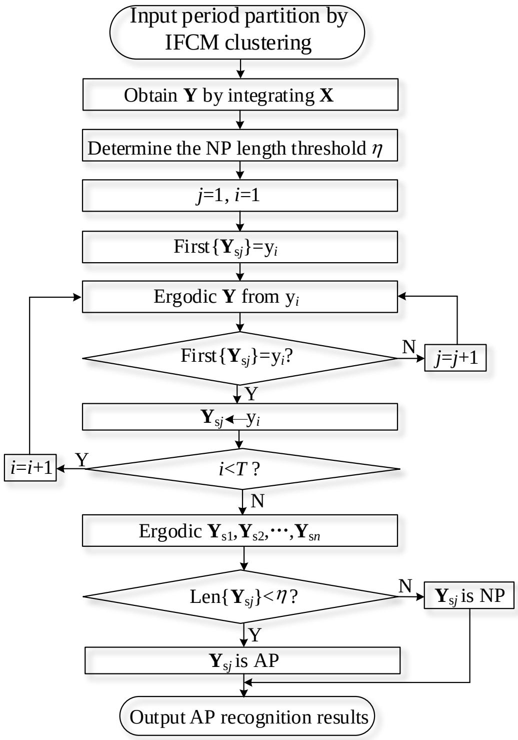

Fig. 10. The flow of AP recognition model.

B. AP Correction Rules

We clarify that there is only one period type in any NP, while there are at least two types in AP, which come from adjacent NP. This is related to the load data structure and algorithm iteration process, the reasons for this phenomenon are not be elaborated here. We view each AP and its adjacent NP as one single correction unit, and correct AP one by one. So, constraint (18) should be met.

$$\mathrm{AP}_{\{s,s+1\}} \in \mathrm{NP}_w, w = \{s, s+1\} \tag{18}$$

It is worth mentioning that (18) can effectively ensure the maximum continuity of period types, preventing valley periods from being corrected to peak periods, or peak periods to valley periods.

The correction rules for AP are shown in Table I.

TABLE I
CORRECTION RULES FOR AP

Correction rules	Adjacent NP types	AP correction
Rule 1	$\mathrm{NP}_s = \mathrm{NP}_{s+1}$	Partitioned as type of NP_w
Rule 2	$\mathrm{NP}_s = \mathrm{peak}$ $\mathrm{NP}_{s+1} = \mathrm{flat}$	AP is partitioned as
Rule 3	$\mathrm{NP}_s = \mathrm{flat}$ $\mathrm{NP}_{s+1} = \mathrm{valley}$	type of NP_s or NP_{s+1}
Rule 4	$\mathrm{NP}_s = \mathrm{flat}$ $\mathrm{NP}_{s+1} = \mathrm{peak}$	by correction model
Rule 5	$\mathrm{NP}_s = \mathrm{valley}$ $\mathrm{NP}_{s+1} = \mathrm{flat}$	

C. AP Correction Model

From the correction rules, the correction model needs to be designed to choose the optimal period type for AP based on the distribution of load data. So, we propose the model based on fuzzy subsethood (FSS).

FSS values represent the degree to which a fuzzy set is a subset of another fuzzy set [31-33]. For example, for two fuzzy sets A and B, the FSS value of A to B can be denoted as S(A, B).

$$S(\mathrm{A},\mathrm{B}) = \frac{1}{n}\sum_{i=1}^{n}\frac{u_{\mathrm{A}\cap\mathrm{B}}(\chi_i)}{u_{\mathrm{A}}(\chi_i)} = \frac{1}{n}\sum_{i=1}^{n}\frac{\min\{u_{\mathrm{A}}(\chi_i), u_{\mathrm{B}}(\chi_i)\}}{u_{\mathrm{A}}(\chi_i)} \tag{19}$$

where, $S(\mathrm{A},\mathrm{B}) \in [0,1]$, $u_A(\chi_i)$ and $u_B(\chi_i)$ are the membership degree of χ_i to A and B, respectively.

Here, we give two possible situations in this paper [34]:

SIT 1 if S(A, B)=1 and S(A, B)=0. It is evident that $\mathrm{A} \subset \mathrm{B}$, while B is not a subset of A.

SIT 2 if $0<S(\mathrm{A}, \mathrm{B})<S(\mathrm{B}, \mathrm{A})\le 1$. Neither $\mathrm{A} \subset \mathrm{B}$ nor $\mathrm{B} \subset \mathrm{A}$, but FSS indicates that B is more a subset of A than conversely.

Doing so is good for AP correction, on the one hand, dividing $\mathrm{AP}_{(s,s+1)}$ into NP_s or NP_{s+1} can be equivalent to comparing the FSS of two NP. On the other hand, each AP having two possible correction results is helpful for the construction of FSS.

Specifically, the correction model is as follows: calculate $S(u_s^{\mathrm{NP}}, u_{s+1}^{\mathrm{NP}})$ and $S(u_{s+1}^{\mathrm{NP}}, u_s^{\mathrm{NP}})$ by (20)-(21), then correct the period type of AP based on their values.

$$u_s^{\mathrm{NP}} = f_s\left[\mathrm{AP}_{(s,s+1)}\right] \tag{20}$$

$$u_{s+1}^{\mathrm{NP}} = f_{s+1}\left[\mathrm{AP}_{(s,s+1)}\right] \tag{21}$$

where, u_s^{NP} is the fuzzy set mapped by the load points in $\mathrm{AP}_{(s,s+1)}$ according to f_s. As only two different fuzzy subsets need to be considered here, f_s can applied as (22)-(23).

$$f_s = \begin{cases} \varpi, & P_t < \varsigma \\ \dfrac{(P_t - \varsigma) - \varpi(2P_t - \varsigma - \xi)}{\xi - \varsigma}, & \varsigma \le P_t < \xi \\ 1 - \varpi, & P_t \ge \xi \end{cases} \tag{22}$$

$$f_{s+1} = 1 - f_s \tag{23}$$

where, $P_t \in \mathrm{AP}_{(s,s+1)}$, ϖ is the function switching coefficient. In rules 2 and 3, the period type level of NP_s is higher than that of NP_{s+1}, so we set $\varpi = 0$. Similarly, in rules 4 and 5, the type level of NP_s is lower than that of NP_{s+1}, we set $\varpi = 1$. ξ and ς are fuzzy parameters, which are confirmed by the period types of NP_s and NP_{s+1}. The two values should reflect the load feature of NP. We set (24)-(25) and (26)-(27) to calculate ξ and ς, respectively.

$$\left.\begin{array}{l} \text{Rule 2}\begin{cases} \min|\xi - c| \\ \min|\varsigma - b| \end{cases} \\ \text{Rule 3}\begin{cases} \min|\xi - b| \\ \min|\varsigma - a| \end{cases} \end{array}\right\} \Rightarrow \varpi = 0 \tag{24}$$

$$\text{s.t. } \begin{array}{l} \xi \in \mathrm{NP}_s \\ \varsigma \in \mathrm{NP}_{s+1} \end{array} \tag{25}$$

$$\left.\begin{array}{l} \text{Rule 4}\begin{cases} \min|\varsigma - b| \\ \min|\xi - c| \end{cases} \\ \text{Rule 5}\begin{cases} \min|\varsigma - a| \\ \min|\xi - b| \end{cases} \end{array}\right\} \Rightarrow \varpi = 1 \tag{26}$$

$$\text{s.t. } \begin{array}{l} \varsigma \in \mathrm{NP}_s \\ \xi \in \mathrm{NP}_{s+1} \end{array} \tag{27}$$

The flow of the correction model is shown as follows.

Algorithm 1 AP Correction Model

Input: AP recognition results
Output: Period partition without AP

```
Ergodic each AP_(s,s+1)
    if AP_(s,s+1) is in accordance with rule 1 do
        NP_w ← AP_(s,s+1)
    else if AP_(s,s+1) is in accordance with rule 2 or rule 3 do
        Calculate ξ and ς by (24)-(25) then
        Set ϖ=0
    else if AP_(s,s+1) is in accordance with schemes 4 or 5 do
        Calculate ξ and ς by (26)-(27) then
        Set ϖ=1
    end if
    Calculate S(u^s_NP, u^{s+1}_NP) and S(u^{s+1}_NP, u^s_NP) by (19)-(23)
    if S(u^NP_s, u^NP_{s+1}) > S(u^NP_{s+1}, u^NP_s) then
        NP_{s+1} ← AP_(s,s+1)
    else if S(u^NP_s, u^NP_{s+1}) < S(u^NP_{s+1}, u^NP_s) then
        NP_s ← AP_(s,s+1)
    end if
```

V. Case Study

In order to highlight the superiority of modified membership functions (MMF) and compare the clustering effect before and after the abnormal periods (AP) correction, the following methods are used to evaluate the results.

Partition coefficient (PC) and partition entropy (PE) based on membership degree can be used as feasible evaluation indexes in this paper [35-37]. PC and PE can be expressed as

$$\mathrm{PC} = \frac{1}{n}\sum_{k=1}^{c}\sum_{i=1}^{n}\left(u_i^k\right)^2 \tag{28}$$

$$\mathrm{PE} = -\frac{1}{n}\sum_{k=1}^{c}\sum_{i=1}^{n} u_i^k \log u_i^k \tag{29}$$

where, the largest PC and the smallest PE mean the best effect.

Silhouette coefficient (SC) is employed for clustering effect evaluation in this paper. Assume that the average distance from χ_i to other data samples in the same cluster is noted as $a(i)$. The average distance from χ_i to all samples in different clusters is $b_{k,i}$, and $b(i)=\min\{b_{1,i},\ldots,b_{k,i}\}$.

$$\mathrm{SC} = \frac{\sum_{i=1}^{n} \dfrac{b(i)-a(i)}{\max\{a(i),b(i)\}}}{n} \tag{30}$$

where, the bigger the SC is, the better the clustering quality is.

To avoid the contingency of algorithm to data set, and enhance the effectiveness and reproducibility of the method proposed in this paper. We select two daily load curves in Fig. 1 to carry out study case. In parameter settings, the NP length threshold η is set as 0.5 hours, and step length $\Delta\alpha$ is 0.01.

A. Fuzzy Parameter Confirmation

According to (1)-(10), we can obtain the relationship between the risk and conservativeness level α . They are given in Fig. 11.

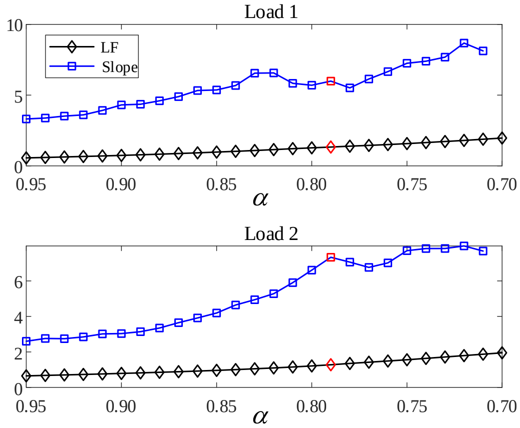

Fig. 11. LF, slope-α .

It is clear that lower conservativeness levels mean larger risk in MMF construction. From the above figure, we can see that the value of LF increases monotonically, for power dispatchers, selecting the appropriate α requires sufficient work experience. When $\alpha = 0.21$, the slope reaches its first local maximum value in both daily load curves, indicating that the conservativeness change is the most obvious within the confidence level neighborhood. The values of fuzzy parameters: *a*, *b* and *c* of two daily load curves are shown in Table II.

TABLE II
FUZZY PARAMETERS OF MMF

Load curves	Fuzzy parameters	Values(kW)
Load 1	*a*	4086.62
	b	3110.53
	c	2134.44
Load 2	*a*	3688.95
	b	2762.67
	c	1836.39

B. Initialization Based on MMF

It is necessary to evaluate the membership degree to prove the effectiveness of MMF. The PFV membership degree of two daily load curves after normalization is given in Fig. 12 and Fig. 13.

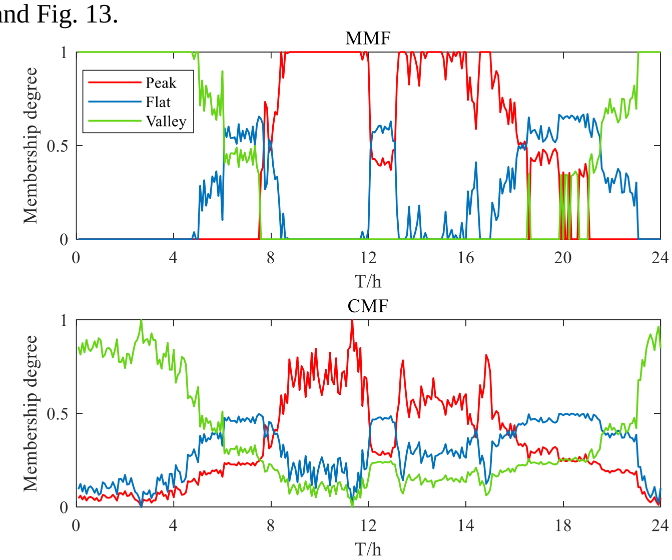

Fig. 12. PFV membership degree of Load 1.

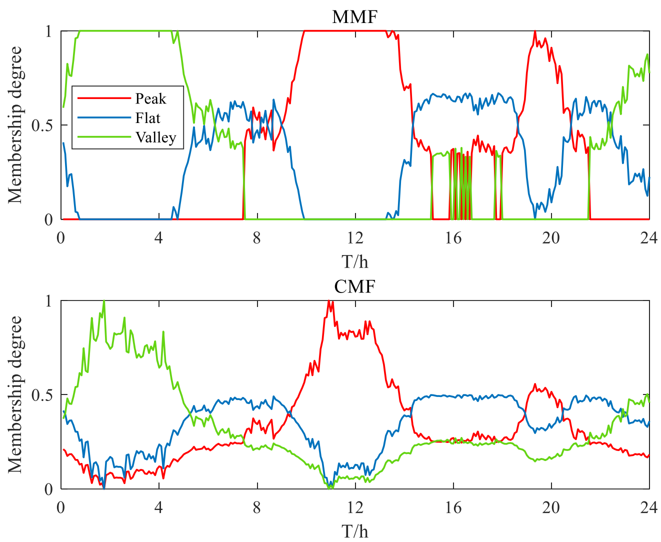

Fig. 13. PFV membership degree of Load 2.

We can see that the membership degree under MMF is able to maintain 0 or 1 in a continuous interval, which is clear, not fuzzy. And it is worth noting that only two period types (flat and valley) need to be considered in early morning and late night (about 23:00~04:00). This surely follows the work and rest patterns of people. While, for CMF, we need to calculate the possibility of three different period types, which undoubtedly makes FCM harder.

PC and PE evaluation results for two membership functions are shown in Table III.

TABLE III

EVALUATION FOR MMF AND CMF

Load curves	Functions	PC	PE
Load 1	MMF	0.7904	0.1315
	CMF	0.4900	0.3773
Load 2	MMF	0.7172	0.1770
	CMF	0.4519	0.4013

As can be seen from Table III, for Load 1 and Load 2, PC of MMF are both higher than those of CMF, and PE of MMF are both lower. The evaluation results show that MMF can more accurately describe the PFV membership degree.

In addition, according to (13), the initial membership degree is able to generate initial cluster centers, which are also applicable to the initialization of other clustering algorithms like k-means clustering algorithm. They give us another view to verify MMF. To expand the use of MMF, we do the case study again using k-means and put the results in Fig. 20 and Fig. 21 in Appendix.

C. Evaluation of Clustering Process

After the initialization, the initial periods can be partitioned, and the iteration process of J is given in Fig. 14.

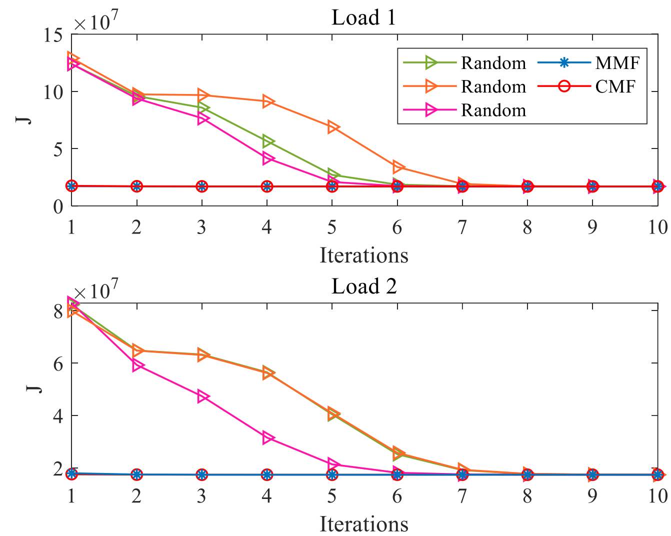

Fig. 14. The iterations of FCM based on random initialization, MMF and CMF.

Multiple runs of FCM with randomly initializing membership degree clearly show that they have a worse operation performance. The convergence speed is slower while the instability of their iterations is also noticeable, which can greatly affect the partition efficiency and effect.

The specific comparison of iteration processes initialized by MMF and CMF is shown in Fig. 15.

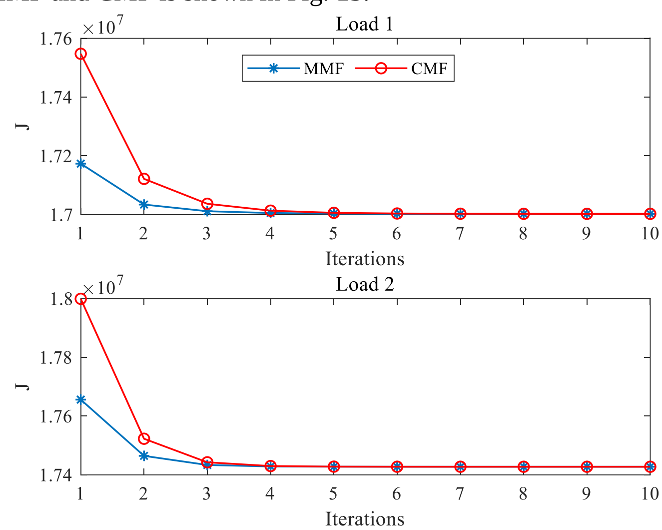

Fig. 15. Iteration processes initialized by MMF and CMF.

As seen above, despite the approximate convergence speed, the clustering process initialized by MMF is more efficient, the mean value and variance of J are shown in Table IV.

TABLE IV

EVALUATION OF ITERATIONS

Load curves	Functions	Mean values(kW)	Variance(kW)
Load 1	MMF	17024234.10	53340.64
	CMF	17074093.63	170384.50
Load 2	MMF	17454938.62	71620.08
	CMF	17496146.57	179206.94

The iterations of cluster centers are given in Fig. 16 and Fig. 17.

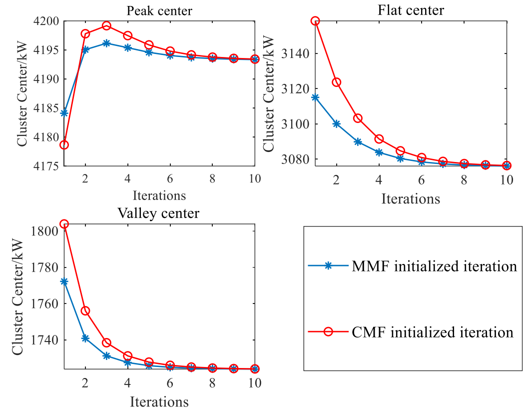

Fig. 16. Iteration of cluster centers of Load 1.

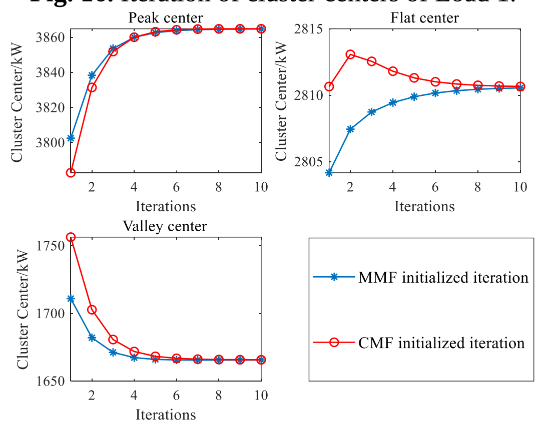

Fig. 17. Iteration of cluster centers of Load 2.

It can be seen that the cluster centers converge to the same value for all period types regardless of which membership functions are used. However, just like the iteration process of objective function, the iterations of cluster centers are faster and more stable.

D. Correction of Initial TOU Period Partition

Since objective function value J and cluster centers converge to the same value, the final clustering results are same. So, period partitions no longer distinguish between CMF and MMF here. The initial period partitions and their correction of two daily load curves are shown in Fig. 18 and Fig. 19.

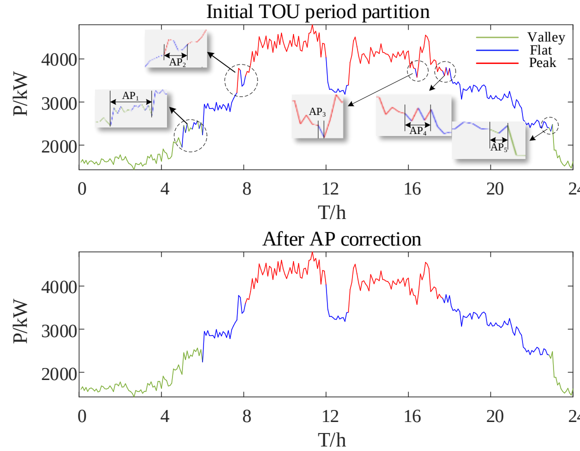

Fig. 18. Initial period partition and correction of Load 1.

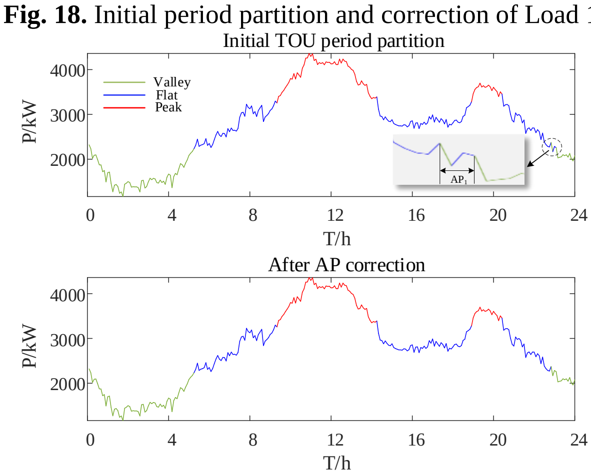

Fig. 19. Initial period partition and correction of Load 2.

Compared to Load 1, the Load 2 curve is much smoother, there is only one AP that needs to be corrected here. The specific time and period types of each AP in Load 1 and Load 2 are shown in Table V and Table VI, respectively.

TABLE V
THE SPECIFIC PERIOD PARTITION OF LOAD 1

	Initial TOU period partition			Corrected TOU period partition	
Period types	NP time nodes	AP time nodes	Total time	NP time nodes	Total time
Peak	08:05~12:00 13:05~16:20 16:25~17:40	07:40~07:50 17:45~17:50 17:55~18:00	8h 45min	08:05~12:00 13:05~17:40	8h 30min
Flat	06:00~07:40 12:00~13:05 18:00~22:50	05:00~05:05 05:10~05:15 05:30~05:45 05:50~05:55 07:50~08:05 16:20~16:25 17:40~17:45 17:50~17:55 22:55~23:00	8h 40min	06:00~08:05 12:00~13:05 17:40~22:50	8h 20min
Valley	00:00~05:00 23:00~24:00	05:05~05:10 05:15~05:30 05:45~05:50 05:55~06:00 22:50~22:55	6h 35min	00:00~06:00 22:50~24:00	7h 10min

TABLE VI
THE SPECIFIC PERIOD PARTITION OF LOAD 2

	Initial TOU period partition			Corrected TOU period partition	
Period types	NP time nodes	AP time nodes	Total time	NP time nodes	Total time
Peak	09:20~14:15 18:55~20:25	/	6h 25min	09:20~14:15 18:55~20:25	6h 25min
Flat	05:15~09:20 14:15~18:55 20:25~22:50	22:55~23:05	11h 20min	05:15~09:20 14:15~18:55 20:25~22:50	11h 10min
Valley	00:00~05:15 23:05~24:00	22:50~22:55	6h 15min	00:00~05:15 22:50~24:00	6h 25min

From the above tables, it can be seen that in Load 1, the total time of peak and valley becomes less, while the time of valley period increases. The change of Load 2 is not obvious, the time of peak period does not change, while the time of flat period is slightly shortened. The correction for each AP is shown in Table VII.

TABLE VII
PERIOD TYPES INCLUDED IN AP

Load curves	AP location	Before correction	After correction
Load 1	AP_1	flat-valley	valley
	AP_2	peak-flat	flat
	AP_3	peak-flat	peak
	AP_4	peak-flat	flat
	AP_5	flat-valley	valley
Load 2	AP_1	flat-valley	valley

The effectiveness of the correction model can be proved by SC index evaluating for the quality of the period partition. It is shown in Table VIII.

TABLE VIII
SC INDEX EVALUATION

Load curves	Period partition	SC
Load 1	Initial	0.7248
	Corrected	0.7284
Load 2	Initial	0.6486
	Corrected	0.6379

The evaluating results change slightly after the AP correction (0.4967% for Load 1 and 1.649% for Load 2, respectively). This is acceptable, because the correction model changes the initial clustering results, which necessarily leads to better or worse clustering quality, but this change is within tolerable limits. Generally speaking, the slight clustering effect loss ensures the feasibility of corrected TOU period partition and proves the effectiveness of the correction strategy.

VI. CONCLUSION

In the field of time-of-use (TOU) tariff period partition, to deal with high-time-resolution load data with strong non-linearity, this paper proposes an improved FCM (IFCM) and the abnormal periods (AP) correction method to form a new partition method.

The modified membership functions (MMF) can effectively describe the correlation of load with each period. Based on the membership degree initialized by MMF, the clustering efficiency is obviously improved.

The selection of fuzzy parameters for MMF is risky. It is feasible to quantify conservativeness difference using loss function, and the parameters chosen by finding the first local maximum slope can make MMF applicable mostly.

The correction method can effectively correct AP from clustering, producing feasible TOU period partition. The additional study in the Appendix shows that the cluster centers initialized by MMF are also applicable to k-means.

Based on the study above, replacing FCM and using k-medoids for clustering is worth investigating. In addition, the AP correction method can be developed based on maximizing the clustering quality.

APPENDIX

The initial cluster centers based on MMF and CMF are applied to k-means clustering algorithms. Then, the clustering results are corrected. The following iteration processes are obtained to prove the effectiveness of MMF and AP correction method. Here, we simply describe the principle of k-means as follows.

Step 1: Generate initial cluster centers.

Step 2: For each sample, its distance to all initial cluster centers is calculated respectively. Then they are divided into the cluster with the shortest distance.

Step 3: The mean value of cluster is taken as a new center.

Step 4: Repeat ***Step 2-Step 3*** until the cluster centers no longer change.

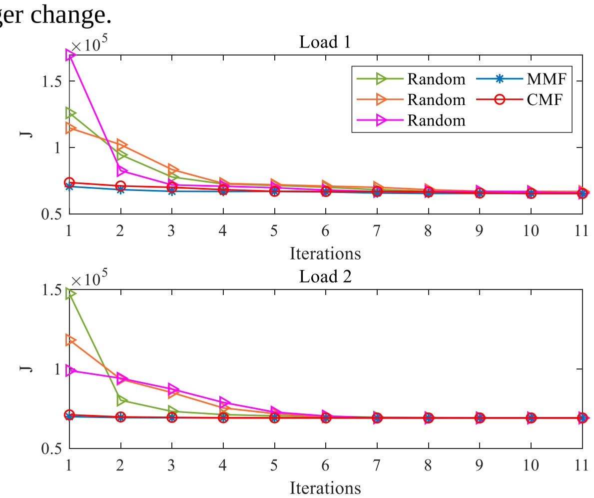

Fig. 20. The iterations of k-means based on random

initialization, MMF and CMF.

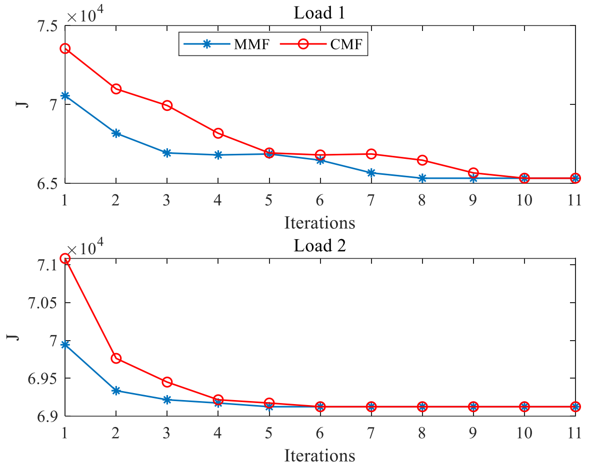

Fig. 21. The iterations of k-means based on MMF and CMF.

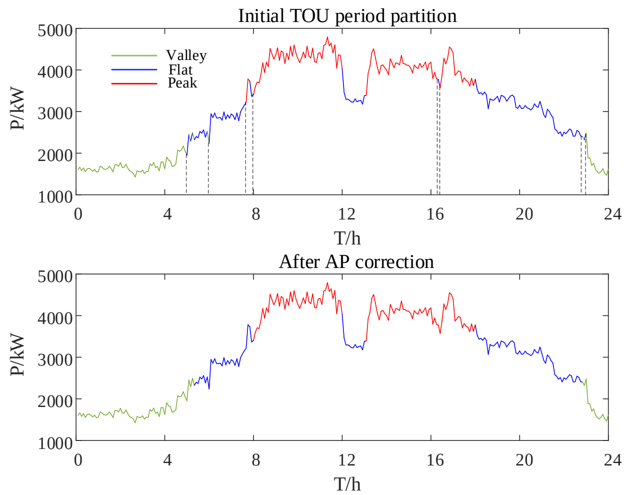

Fig. 22. Initial period partition and correction of Load 1.

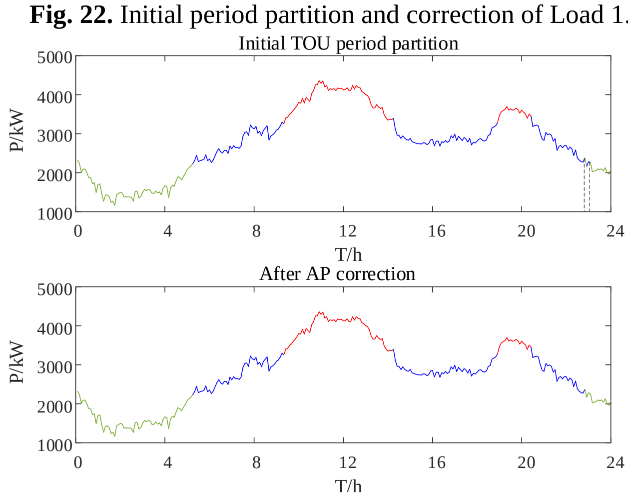

Fig. 23. Initial period partition and correction of Load 2.

TABLE IX

SC INDEX EVALUATION

Load	Correction	SC
Load 1	Initial	0.7184
	Corrected	0.7079
Load 2	Initial	0.6486
	Corrected	0.6379

REFERENCES

[1] Cortes-Arcos T , Bernal-Agustin J L, Dufo-Lopez R , et al. “Multi-objective demand response to real-time prices (RTP) using a task scheduling methodology,” Energy, July.2017, vol.138, pp.19-31, doi: 10.1016/j.energy.2017.07.056.
[2] Wang J , Zhong H, Lai X, et al. “Distributed real-time demand response based on Lagrangian multiplier optimal selection approach,” Applied Energy, Dec.2017, vol.190, pp.949-959, doi: 10.1016/j.apenergy.2016.12.147.
[3] L. Jia and L. Tong, “Dynamic Pricing and Distributed Energy Management for Demand Response,” IEEE Transactions on Smart Grid, March 2016, vol. 7, no. 2, pp. 1128-1136, doi: 10.1109/TSG.2016.2515641.
[4] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing and T. Başar, “Demand Response Management in the Smart Grid in a Large Population Regime” IEEE Transactions on Smart Grid, Jan. 2016, vol. 7, no. 1, pp. 189-199, doi: 10.1109/TSG.2015.2431324.
[5] J. C. Galvis and A. Costa, “Demand Side Management Using Time of Use and Elasticity Price,” IEEE Latin America Transactions, Oct. 2016vol. 14, no. 10, pp. 4267-4274, doi: 10.1109/TLA.2016.7786304.
[6] C. M. Flath, “An optimization approach for the design of time-of-use rates,” IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society, Vienna, Austria, 2013, pp. 4727-4732, doi: 10.1109/IECON.2013.6699899.
[7] K. Men, J. Wu and S. Zhou, “Regulation Power Evaluation Method of Demand Response by Electricity Price-Demand Elasticity Matrix Considering Equipment Fault Rate,” 2019 IEEE 2nd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 2019, pp. 66-70, doi: 10.1109/AUTEEE48671.2019.9033403.
[8] P. Yang, G. Tang and A. Nehorai, “A game-theoretic approach for optimal time-of-use electricity pricing,” IEEE Transactions on Power Systems, May 2013, vol. 28, no. 2, pp. 884-892, doi: 10.1109/TPWRS.2012.2207134.
[9] D. Zhao, H. Wang, J. Huang and X. Lin, “Time-of-Use Pricing for Energy Storage Investment,” IEEE Transactions on Smart Grid, March 2022, vol. 13, no. 2, pp. 1165-1177, doi: 10.1109/TSG.2021.3136650.
[10]S. Faddel and O. A. Mohammed, “Automated Distributed Electric Vehicle Controller for Residential Demand Side Management,” IEEE Transactions on Industry Applications, Jan.-Feb. 2019, vol. 55, no. 1, pp. 16-25, doi: 10.1109/TIA.2018.2866255.
[11]H. Yang, Y. Gao, Y. Ma and D. Zhang, “Optimal Modification of Peak-Valley Period Under Multiple Time-of-Use Schemes Based on Dynamic Load Point Method Considering Reliability,” IEEE Transactions on Power Systems, Sept. 2022, vol. 37, no. 5, pp. 3889-3901, doi: 10.1109/TPWRS.2021.3131519.
[12]M. Z. Oskouei, S. Zeinal-Kheiri, B. Mohammadi-Ivatloo, M. Abapour and H. Mehrjerdi, “Optimal Scheduling of Demand Response Aggregators in Industrial Parks Based on Load Disaggregation Algorithm,” IEEE Systems Journal, March 2022, vol. 16, no. 1, pp. 945-953, doi: 10.1109/JSYST.2021.3074308.
[13]L. Zhao, Y. Zhou, F. L. Quilumba and W. -J. Lee, “Potential of the Commercial Sector to Participate in the Demand Side Management Program,” IEEE Transactions on Industry Applications, Nov.-Dec. 2019, vol. 55, no. 6, pp. 7261-7269, doi: 10.1109/TIA.2019.2937503.
[14]C. Feng, Y. Wang, K. Zheng and Q. Chen, “Smart Meter Data-Driven Customizing Price Design for Retailers,” IEEE Transactions on Smart Grid, May 2020, vol. 11, no. 3, pp. 2043-2054, doi: 10.1109/TSG.2019.2946341.
[15]W. Zhong, K. Xie, Y. Liu, C. Yang and S. Xie, “Multi-Resource Allocation of Shared Energy Storage: A Distributed Combinatorial Auction Approach,” IEEE Transactions on Smart Grid, Sept. 2020, vol. 11, no. 5, pp. 4105-4115, doi: 10.1109/TSG.2020.2986468.
[16]H. Jin, Z. Li, H. Sun, Q. Guo and B. Wang, “Coordination on Industrial Load Control and Climate Control in Manufacturing Industry Under TOU Prices,” IEEE Transactions on Smart Grid, Jan. 2019, vol. 10, no. 1, pp. 139-152, doi: 10.1109/TSG.2017.2733578..
[17]S. Al-Rubaye, A. Al-Dulaimi, S. Mumtaz and J. Rodriguez, “Dynamic Pricing Mechanism in Smart Grid Communications Is Shaping Up,” IEEE Communications Letters, July 2018, vol. 22, no. 7, pp. 1350-1353, doi: 10.1109/LCOMM.2018.2822798.
[18]H. Yang, L. Wang, Y. Zhang, H. -M. Tai, Y. Ma and M. Zhou, “Reliability Evaluation of Power System Considering Time of Use

Electricity Pricing," IEEE Transactions on Power Systems, May 2019, vol. 34, no. 3, pp. 1991-2002, doi: 10.1109/TPWRS.2018.2879953.
[19]B. Wang, Y. Li, W. Ming and S. Wang, "Deep Reinforcement Learning Method for Demand Response Management of Interruptible Load," IEEE Transactions on Smart Grid, July 2020, vol. 11, no. 4, pp. 3146-3155, doi: 10.1109/TSG.2020.2967430.
[20]S. Wang, S. Zhang, K. Wang, et al. "Multi-objective optimal operation of distributed energy storage considering user demand response," Electric Power Automation Equipment, Nov.2020, vol.40, no.1, pp.125-132, doi:10.16081/j.epae.201911029.
[21]M. Ding, X. Wang, J. Wang, Z. Yang, H. Zhong and J. Yang, "A dynamic period partition method for time-of-use pricing with high-penetration renewable energy," 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 2017, pp. 1-6, doi: 10.1109/EI2.2017.8245547.
[22]X. Zhang and X. Li, "Research on Peak and Valley Periods Partition and Distributed Energy Storage Optimal Allocation Considering Load Characteristics of Industrial Park," 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Mauritius, Mauritius, 2021, pp. 1-6, doi: 10.1109/ICECCME52200.2021.9591133.
[23]W. Huang et al., "Peak and Valley Periods Partitioning Based on Improved K-medoids Algorithm," 2020 IEEE Sustainable Power and Energy Conference (iSPEC), Chengdu, China, 2020, pp. 1860-1866, doi: 10.1109/iSPEC50848.2020.9351236.
[24]Q. Hui-ting and L. Yi-jie, "Research of peak and valley period partition approach on statistics," 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), Weihai, China, 2011, pp. 1774-1779, doi: 10.1109/DRPT.2011.5994186.
[25]Y. Cheng, N. Zhai. "Electricity price peak and valley periods division based on customer response," Automation of Electric Power Systems, May. 2012, vol. 36, no.9, pp.42-46+53, doi: 10.3969/j.issn.1000-1026.2012.09.008.
[26]Y. Wang, J. FU, J. Shao, et al. "Development and assessment of time-of-use electricity pricing policy for integrating wind power curtailment by peak load regulation," Jilin Electric Power, Nov.2018, vol.46, no.257, pp.1-5, doi:10.16109/j.cnki.jldl.2018.04.001.
[27]Y. Wu, L. LÜ, L. XU, et al. "Optimal planning of micro-grid considering time division of TOU," Electrical Measurement &Instrumentation, Nov.2020, vol.57, no.11, pp.55-62, doi:10.19753/j.issn1001-1390.2020.11.009.
[28]F. Nie, X. Zhao, R. Wang, X. Li and Z. Li, "Fuzzy K-Means Clustering With Discriminative Embedding," IEEE Transactions on Knowledge and Data Engineering, March 2022, vol. 34, no. 3, pp. 1221-1230, doi: 10.1109/TKDE.2020.2995748.
[29]S. Pei, H. Chen, F. Nie, R. Wang and X. Li, "Centerless Clustering," IEEE Transactions on Pattern Analysis and Machine Intelligence, Jan. 2023, vol. 45, no. 1, pp. 167-181, 1 doi: 10.1109/TPAMI.2022.3150981.
[30]T. T. Teo et al., "Optimization of Fuzzy Energy-Management System for Grid-Connected Microgrid Using NSGA-II," IEEE Transactions on Cybernetics, Nov. 2021, vol. 51, no. 11, pp. 5375-5386, doi: 10.1109/TCYB.2020.3031109.
[31]Bart, Kosko. "Fuzzy entropy and conditioning," Information Sciences, Dec.1986, vol.40, no.2, pp.165-174, doi: 10.1016/0020-0255(86)90006-X.
[32]Wang C C , Hon-Son D . "A modified measure for fuzzy subsethood," Information Sciences, Mar.1994, vol.79, no.3-4, pp.223-232, doi: 10.1016/0020-0255(94)90121-X.
[33]K. A. Rasmani and Q. Shen, "Modifying weighted fuzzy subsethood-based rule models with fuzzy quantifiers," IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542), Budapest, Hungary, July.2004, vol.3, pp. 1679-1684, doi: 10.1109/FUZZY.2004.1375433.
[34]P. Grzegorzewski and E. Mrowka, "Subsethood measure for intuitionistic fuzzy sets," IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542), Budapest, Hungary, July.2004, vol.1, pp. 139-142, doi: 10.1109/FUZZY.2004.1375704.
[35]P. Wang. "Pattern Recognition with Fuzzy Objective Function Algorithms (James C. Bezdek)," SIAM Review, Dec.1983, vol.25, no.3, pp.442, doi: 10.1137/1025116.
[36]Bezdek, James C. "Cluster Validity with Fuzzy Sets." Journal of Cybernetics, Feb.1973, vol.3, no.3, pp.58-73, doi:10.1080/01969727308546047.
[37]M. R. Rezaee, P. M. J. van der Zwet, B. P. E. Lelieveldt, R. J. van der Geest and J. H. C. Reiber, "A multiresolution image segmentation technique based on pyramidal segmentation and fuzzy clustering," IEEE Transactions on Image Processing, July 2000, vol. 9, no. 7, pp. 1238-1248, doi: 10.1109/83.847836.

Peng Wang, received the B.Sc. degree in electrical engineering from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2021. He is currently working toward the M.S. degree at School of Automation and Industrial Internet, Chongqing University of Posts and Telecommunications. His major research interests include power system dispatch and power markets.

Yiwei Ma, received the M.Sc. degree in control engineering (2007) and the Ph.D. degree in electrical engineering (2015) from South China University of Technology in China, and the B.Sc. degree (2003) from Yanshan University. In 2015, She joined Chongqing University of Posts and Telecommunications, where she is currently an Associate Professor of Electrical Engineering, and conduct more research on microgrid, smart grid, V2G and related technologies