Impact Factors Analysis on the Probability Characterized Effects of Time of Use Demand Response Tariffs Using Association Rule Mining Method

1

Kangping Li[1], Liming Liu[1], Fei Wang[1,2,3,*], Tieqiang Wang[4], Neven Duić[5], Miadreza Shafie-khah[6], João P. S. Catalão[7]

1. Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China
2. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China
3. Hebei Key Laboratory of Distributed Energy Storage and Micro-grid, North China Electric Power University, Baoding 071003, China
4. Dispatch and Control Center, State Grid Hebei Electric Power Supply Co., Ltd, Shijiazhuang 050021, China
5. Department of Energy, Power Engineering and Environment, University of Zagreb, Ivana Lučića 5, HR-10000 Zagreb, Croatia
6. School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland
7. Faculty of Engineering of University of Porto and INESC TEC, 4200-465 Porto, Portugal

Abstract—**Time of use (TOU) rate has been regarded as an effective strategy to associate utility companies to avoid peak time financial risks and make the most profit out of the market, while most programs are not effective as expected to reduce peak time demand of residents. Exploring the impact factors of peak demand reduction (PDR) can help policy makers find reasons that weaken effects of programs and corresponding measures can be carried out to maximize the benefits. However, averaging quantitative indicators for program assessment and incomplete impactor analysis method in existing research show limitations of revealing the complex reasons behind it. In this paper, an association rule mining based quantitative analysis framework is built to explore the impact of household characteristics on PDR under TOU price making up for the deficiencies in current research. Firstly, a probability distribution based customer PDR characterizing model is proposed, in which difference-in-difference model is adopted to quantify the effect of PDR and probability distribution fitting method is used to characterize the feature of PDR for households. Then a comprehensive association rule mining analysis using Apriori algorithm is presented to explore the impacts factors of PDR covering four categories of household characteristics including dwelling characteristics, socio-demographic, appliances and heating and attitudes towards energy. Finally, analysis results of a case study based on 2993 household records containing smart metering data and survey data illustrate that PDR level cannot be obtained simply based on the appliance's ownership and its usage habits. Socio-demographic information of households should be taken into consideration together; Internet connection and good house insulation contribute to the increase of PDR level. Moreover, the percentage of renewable generation for households also show a certain relationship with PDR. The proposed analysis framework and findings will associate retailer to improve the benefits of TOU programs and guide policy makers to design more efficient energy saving policies for residents.**

Keywords—Peak demand reduction; Household characteristics; Association rule mining; Demand response; Apriori algorithm

1. INTRODUCTION

1.1. Background and motivation

With fossil fuel depletion and environmental degradation, the clean and pollution-free renewable energy, such as wind power and solar power develop quickly all over the world[1]. However, due to the volatility and randomness characteristics, the increasing penetration of renewable sources in electricity generation brings challenges to the balance between supplies and demands[2], which is the fundamental of power system operation[3,4]. Although several technologies, including non-intermittent capacity, renewable energy forecasting that mainly includes solar energy forecasting[5] and wind power forecasting covering various time scales such as short term[6,7] as well as ultra-short term[8], microgrids, which are widely considered as an effective means to integrate distributed renewable generations into the main grid[9,10], and electricity storage[11,12], are introduced to mitigate the problem, these methods are not satisfactory after taking cost, accuracy and efficiency into consideration. Hence, as the proposal and maturity of the concept called demand response (DR)[13], it is regarded as a promising option for the integration of renewable energy[14].

DR programs aim to modify the demand patterns of electricity by encouraging consumers to reduce or shift their electricity consumption during peak time in response to price signals or financial incentives[15,16]. According to the two different forms of incentives, DR programs are usually divided into two categories[17], namely price-based programs[18] and incentive-based programs. Time-of-use (TOU) tariff, which is regarded as the most common price-based program (generally concludes in three forms: critical peak pricing, real time pricing and time-of-use)[19,20], gets extensive researches and applications due to its ease of

implementation[21]. For example, TOU tariffs play a significant role in the operation of microgrids through guiding customers reschedule electricity usage responding to price signals[22,23]. Moreover, the surging proportion of electric vehicles increases the burden of grid operation during peak time[24]. The implementation of TOU can excite the shiftable characteristics of electric vehicle loads realizing the reduction of costs while maintaining the safe and stable operation of power grid[25]. A large amount of TOU tests especially in the residential sector were conducted in many countries such as Ireland, Belgium, Finland, UK etc. It has proven to be an effective approach for reducing peak electricity demand in the residential sector around the world, especially in developed countries[26]. Evaluation of the peak demand reduction (PDR) effects of TOU programs is a significant part for the implementation of programs. It can provide authentic and effective information of executive conditions so as to direct the program designer modify the policy to be more cost effective. Different from industrial and commercial TOU programs, the electricity consumption behaviors of residents responding to the price signals may be different due to the diverse household characteristics of residential customers. And the household characteristics including the features of dwelling, home appliances, occupants and their behaviors can be regarded as the impact factors of residential PDR effects. Exploring the impact factors of PDR can give reasonable explanations on the contradictory results mentioned in reference[27] and help policy designers to deepen the understanding of electricity consumption behaviors. Based on the exploring results, the program designers can modify residential TOU programs according to the characteristics of residents in the area to maximize benefits. For example, utilizing the impact factors obtained through the research, more customized TOU programs may be designed and implemented, which are more likely to maximize the benefits and realize win-win goals for both residential customers and retailers[28]. Meanwhile, it can provide support for residential load forecasting especially probability predictions under the TOU mechanism. Generally speaking, it is of great significance to explore the impact factors of residential PDR.

1.2. Literature review

The analysis of the influence factors on the implementation effects of price-based DR programs usually consists of two steps: quantifying the program effects and exploring the impact factors[29]. Although there are many researches related to price-based program field experiments, most of them merely focus on the evaluation of the program effects. Ref. [30] estimated the results of a summer residential critical peak pricing experiment in Japan, which found that varied critical peak pricing priced by as much as 10 times the baseline rate and induced 6.5–8.8% additional maximum electricity-saving behaviors. Another large customer behavior experiment was carried out from 2009 by the Commission for Energy Regulation of Ireland[31]. The experiments illustrated that TOU tariffs did reduce electricity usage both in overall usage and peak time electricity consumption. A comprehensive review carried out by Ref. [32], the review summarized the impacts of three types of dynamic pricing pilots including 13 TOU pilot studies designed after 1997 and they concluded that basic TOU pricing programs could expect to see residents on peak demand change by −5%. Similar research was conducted by Ref. [33], the research covering 12 TOU pilot studies concluded that TOU pricing induced a −3% to −6% change in residential sector on peak demand. Among these programs, almost all the quantitative analyses to the effects of programs are conducted from the overall level to describe the whole reduction for all participants. However, households are not homogeneous due to their diverse characteristics. The existing overall evaluation results of the programs will cover up the differences among residential customers, influence the accuracy of assessment and make it hard for researchers to discover and explain the underlying causes of unexpected effects. Therefore, the methods suitable for the random behaviors of residential customers to measure the effects of program implementation should be studied more.

However, not all the programs show positive reduction in electricity consumption. The application of TOU tariffs to residential electricity consumers in Italy since 2010 is studied by Ref. [34]. The researchers found that TOU tariffs resulted in increases in electricity demand for substations at peak period. Some conflicting findings also show up in the existing researches. For example, ref. [35] found that a high fraction of residential households especially the richer ones did not respond to price signals. But the findings from Commission for Energy Regulation in Ireland illustrated that TOU tariffs did reduce electricity usage, and that households with higher consuming tended to show greater reductions[31]. This contradiction is caused by the lack of understanding of diverse electricity consumption behaviors, which are led by various characteristics of residents. Moreover, parts of researches merely focus on finding one or several impact factors that may be related to the responsiveness of participants. Ref. [36] conducted 44 interviews and home tours followed by a survey of households with children in Australia concluding that TOU tariffs were unlikely to effectively reduce peak period electricity consumption in households with children. The influence of weather on the effects of residential TOU tariff and the modeling method are presented in Ref. [37]. Considering the complexity of human behaviors and disparities of residential customers, comprehensive impact factors exploration work should be carried out.

Besides numbers of studies merely focusing on the evaluation of program effects, there are several papers explore the influence of feedback method on the effects. Ref. [29] attempted to reveal the underlying impacts of information feedback on electricity reductions and the findings illustrated that feedback methods do have influence on reducing and shifting demand because the feedbacks act as a reminder and motivator. Feedback methods can be regarded as the external impact factors of the implementation effects. Relatively speaking, household characteristics (HCs) of residential customers may show influences as internal factors. The electricity consumption reduction or shifting is the combination of all the factors. Hence, it is necessary to explore what are the key impact factors driving different effects and how the internal factors influence the effects especially the peak time demand reduction, which the power system operators most concern about. However, these researches are rarely included in the existing literature.

1.3. Contributions and paper structure

Facing the above issues, a probability distribution model is established to characterize the PDR of residential customers and an association rule mining (ARM) based quantitative analysis approach is proposed to find the key household characteristics and how the internal factors influence the peak time demand reduction. The main contributions of this paper can be summarized as follows:

(1) A PDR quantifying framework for residential customers under TOU programs is established

(2) A probability distribution based residential customer PDR characterizing model is proposed.

(3) Customer classification considering the features of peak time reduction behavior is presented. And an ARM method using Apriori algorithm is introduced to reveal the relations between HCs and customer types.

The proposed models and findings of the research serve multiple parties. From the perspectives of utilities, this research provides supports to customize electricity retail prices for various types of users, improve the load forecasting performance especially probability predictions in the circumstance of dynamic price rates widely used, increase the benefits of DR projects. Moreover, the findings of this study can help policy makers to explore more about electricity consumption habits of customers so as to enhance the effects of energy reduction and pave the way in the direction of the low-carbon future. As the research object of this study, customers can also benefit from customized and targeted services. More details will be discussed in 5.2.

The rest of the paper is structured as follows. Section 2 describes the dataset used in this paper. The method of PDR quantifying, probability characteristic description and association rules analysis are illustrated in section 3. Section 4 views the simulation results. In section 5, the findings and potential applications of this study are presented. Finally, the paper is concluded in Section 6.

2. DESCRIPTION OF DATA SET

The electricity consumption data and survey information used in this research are obtained from the Commission for Energy Regulation (CER) in Ireland[38]. CER began to design the TOU tariffs of Smart Metering Electricity Customer Behavior Trials (SMECBTs) as early as March in 2008 and the program was officially carried out during 2009 and 2010 for the sake of measuring the changes in residential consumer behaviors in terms of both reductions in peak demand and overall electricity use influenced by different TOU tariffs. Over than 5000 Irish residential customers took part in the program with smart metering installed to collect electricity usage data and they also completed comprehensively designed questionnaires, which include social and demographic questions, dwelling characteristics, household appliance information, energy consumption attitude of customers. The electricity consumption data associated with survey information will contribute to analyze the characteristics of residential electricity consumption behaviors.

2.1. Smart metering dataset

The smart metering dataset consists of the electricity consumption data of 4232 residential customers at 30 minutes interval over one and a half year. Only 3123 customers remained after excluding the households without intact electricity data. We use the electricity consumption data from July in 2009 to December in 2010, including the benchmark period (1st July to 31st December 2009) and the test period (1st January to 31st December 2010). In benchmark period, usage data was collected in order to establish a benchmark level of use. Participants were also allocated to a test or control group. During the test period, participants engaged in various time of use tariffs. Four various TOU tariffs named as tariff A to D and a weekend tariff named tariff E were executed in the SMECBTs. The details of tariff A to D in weekdays are presented in Fig. 1. The TOU structure (time bands including "Night", "Day" and "Peak") was based on the local system demand condition and the base TOU tariff would reflect the underlying cost of energy transmission, distribution generation and supply. It can be seen from Fig. 1 that from tariff A to D, the price rate during peak time increased significantly, while a small decline emerged in night and day time, which resulted in larger

peak-nonpeak price differences. We take the data of customers in tariff A for an example to carry out study in this paper.
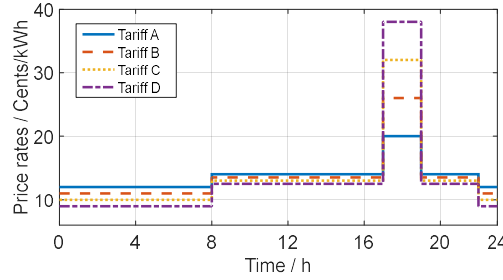

Fig. 1 Price rates of four tariffs. The time bands for TOU includes "Night", "Day" and "Peak". In weekdays, "Night" covers 23:00-08:00; "Day" covers 08:00-17:00 and 19:00-23:00; 17:00-19:00 is "Peak" period

2.2. Surveys dataset

The questionnaire in the survey is composed of 143 questions containing four main aspects, namely the dwelling characteristics, the socio-demographic data, the appliance's ownership and heating and the attitudes towards energy. The dwelling characteristics includes the type of dwellings, year and area of construction and so on. Questions like the age of householders, number of occupants are included in the socio-demographic data. The questions for appliance's ownership and heating embody detail information about the appliances in the house. And the attitudes towards energy exhibit the attitudes of customers, e.g. the willingness to reduce their consumption or their willingness to protect the environment. According to the ID of the customers, the survey information and the corresponding smart metering data can be linked. Customers with valid survey data are removed and the sample size ultimately decreases to 2993 including 818 customers in Tariff A for further analysis. Then the four kinds of survey questions without integrity and reasonableness are excluded. One hundred and three questions including 28 dwelling characteristic related questions, 18 socio-demographic related questions, 35 questions about the appliance's ownership and heating as well as 22 questions describing the attitudes towards energy remain and constitute HC set **R** .

3. METHODS

The proposed approach shown in Fig. 2 is divided into three steps. Firstly, the difference-in-difference model combined with weather-date based day matching and pattern matching principle based control group selection is established to calculate peak demand reduction for individuals. Secondly, using the quantified results of PDR for each resident, this paper adopts probability distribution fitting model combining Kolmogorov-Smirnov test to characterize the features of PDR. Based on the characterization results, customers with various PDR features are categorized into four clusters. Then, in order to facilitate the mining for impact factors of PDR, enhanced Apriori based ARM method is introduced to explore the relationships between PDR features and corresponding impact factors, namely HCs.
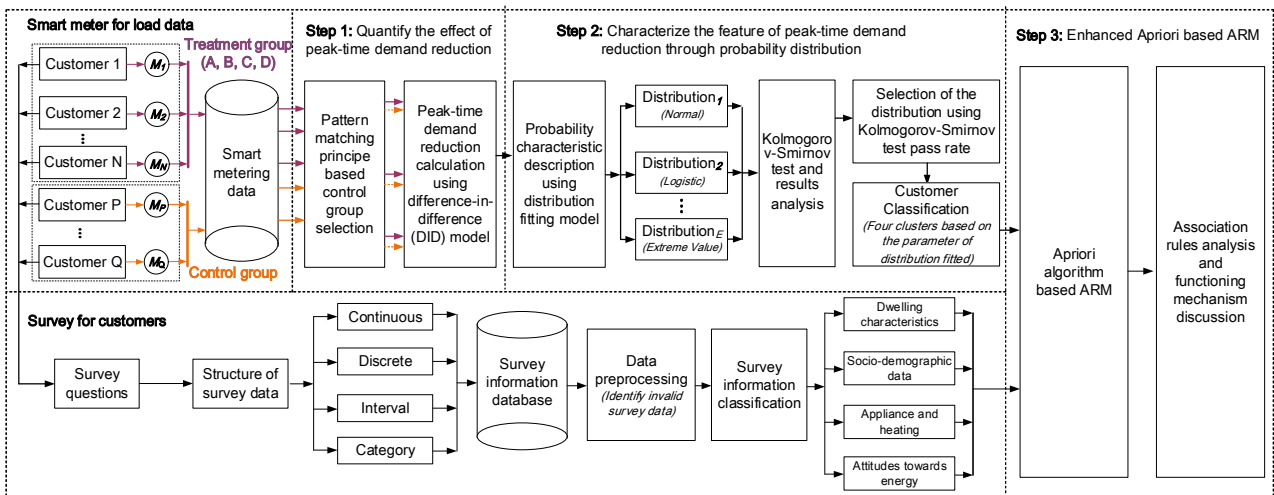

Fig. 2 Framework of the proposed approach

3.1. Quantify the effect of PDR for residents

With the wide implementing of TOU programs, more and more researches have focused on evaluating the effects of programs on changing the electricity consumption especially for the peak demand. Researchers usually take a group of residential customers in a particular area as a whole object to find out whether the TOU programs show significant PDR effects on users and how much the reduction will be. However, such average method may cover up the characteristics of each individual. For example, attending in a designed TOU tariff, residential customers in the group may show totally different responsive behaviors to the price signal. A large electricity usage reduction or shifting emerges in a part of customers and the others almost remain unchanged. But only a moderate effect will be obtained if the customers are regarded as a whole group. Designer cannot modify the policy effectively according to the result, which is detrimental for improving the benefits of the program. The deviation of the evaluation of executive performance is caused by their diverse characteristics and influenced by random electricity consumption behaviors of residential customers. In order to quantify the effect of the program more reasonably, details of PDR features at the individual level should be focused on, which takes full consideration of the characteristics of customers. Considering the randomness of electricity consumption behaviors of residential customers, probability models are usually adopted to characterize the behaviors[39]. In this paper, PDR probability distribution model is established based on the daily PDR data for each household to describe the PDR characteristics.

The simplest method to obtain daily PDR data is the "difference" model, which takes the difference of peak time demand in two days selected from pre-grogram and post-program respectively. Different from the characteristics of commercial and industrial customers, single residential customer shows much stronger randomness in electricity consumption. The variation in the peak time demand of the two selected days may be caused by occasional behaviors instead of impact from price signals. To solve this problem, weather-date based day matching method is proposed in 3.1.1.

In addition, during the TOU program, the electricity consumption behaviors of participants may be influenced by environmental factors such as climate changes, the development of regional economy [40,41]. The variations caused by these factors should also be excluded. So difference in difference model associated with load pattern matching principle based method is used to exclude the influences caused by environmental factors. Although the changes in the conditions of households (e.g., the changes in the composition of electrical appliances) may also affect the evaluation results of the program, we assume that the situations of the households remain unchanged in the study because these changes are hard to access in reality.

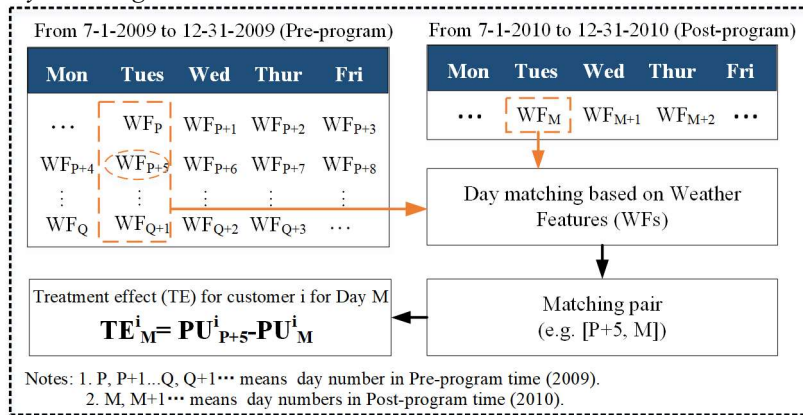*3.1.1. Weather-date based day matching method*

Fig. 3 Framework of weather-date based day matching method

To provide sufficient data for establishing the probability model, it is crucial for each day in the post-program period to select suitable corresponding days in pre-program period to constitute "day pairs". With regard to each day pair, it is considered that if there is no "treatment" (price signal) for the customer, the electricity consumption patterns of the customer in the selected two days may be very similar. It means that all the variations in the electricity usage during peak time are mainly considered as PDR. And the differences caused by occasional behaviors to some extent are lowered as much as possible. Therefore, weather-date based day matching method is used to constitute the day pairs. In this research, it is assumed that there are two main factors, weather feature and date, that may influence electricity usage patterns of customers. Weather features (WF) refer to the meteorological information of a day such as temperature, humidity. The WFs have a strong impact on electricity usage, especially the temperature. The atmospheric temperature usually affect the residential electricity usage by deciding the operation states of air conditionings and heaters, which are considered as main electricity consuming appliance[42]. In this paper, the daily temperature

features, namely maximum air temperature, minimum air temperature and average temperature, are taken as the WF. Moreover, the lifestyle of residential customers is another impact factor for the electricity usage pattern. For example, most people may go to work from Monday to Friday and stay at home on weekends. This kind of lifestyle results in various electricity consumption patterns on weekdays and weekends. Even among weekdays, the load pattern may present differences because of the various schedules of each day. So this paper maintains that the electricity usage of the two days that are in the same weekday with similar WFs is appropriate and the two days constitute the "day pair". The framework of the weather-date based day matching method is shown in Fig. 3.

In Fig 3, customer i is taken as an example to illustrate the process. Firstly, all the days in pre-program period are regrouped into five week sets namely Monday group to Friday group (The weekends are excluded from the pre-program period). Secondly, for day M ($M = \{1, 2 \ldots n\}$ means day numbers in post-program period), a corresponding week set K according to the "week" attribution of day M (i.e. Monday, Tuesday,…, Friday) is selected. Then, a global search is conducted in week set K to find the day with smallest Euclidean distance of WFs to day M and the daily PDR for day M can be obtained by equation (1).

$$TE_M^i = PU_P^i - PU_M^i \tag{1}$$

Where TE_M^i denotes the treatment effects of TOU program on customer for day M ; PU_P^i means the peak time usage of customer i in day P ; PU_M^i represents the peak time usage of customer in day M .

3.1.2. Difference in difference model

The difference in difference (DID) is a widely employed method, which can be used to explore the effects of public policies or program implementation. During the DID analysis, four information data sets are utilized to evaluate the impact of the program, including pre-program and post-program data for a treatment group and a control program respectively. Pre-program denotes prior to the treatment group in the program receiving the "treatment" and post-program denotes after the treatment group obtains the "treatment". DID model uses the measurements of Y (represents the output variable of interest and further denotes electricity usage during peak time in this study) for treatment group in pre-program and post-program time and evaluate the effects of treatments by calculating the differences between the values of Y in two periods. However, during the program, not only the "treatment" but also other external factors may influence the value of Y . For the sake of getting an accurate assessment of treatment effects, another difference is conducted between the pre-program and post-program values of Y for control group, the result of which is called "Common Trend". The modified treat effects β^{DID} can be obtained by the following equation (2) and the diagram shown in Fig. 4 visualizes the meaning of equation.

$$\beta^{DID} = (Y^{T-post} - Y^{T-pre}) - (Y^{C-post} - Y^{C-pre}) \tag{2}$$

Where Y^{T-pre} means the value of Y for treatment group sample before the "treatment"; Y^{T-post} denotes the value of Y after the "treatment". Similarly, let Y^{C-pre} and Y^{C-post} represent the values of Y for samples in control group before and after the "treatment". Usually there will be a process of regression followed to calculated the parameter β^{DID} , but the calculation process in this paper is conducted by the combination of the results from weather-date based day matching method and the pattern matching principle based method, which will be illustrated in detail in 3.1.3.
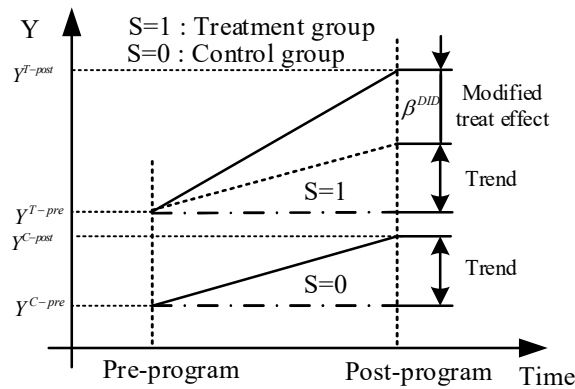

Fig. 4 Diagram of difference in difference model

Based on the discussion above , for each reduction calculation result, peak time electricity usage in a certain pre-program day is taken as Y^{T-pre} denoted by $PU_{day\text{-}pre}^{i}$. Similarly, Y^{T-post} can be replaced by $PU_{day\text{-}post}^{i}$. The DID model used in this paper can be described by equation (3).

$$PDR_{day-post}^{i}=(PU_{day-post}^{i}-PU_{day-pre}^{i})-(PU^{C-post}-PU^{C-pre}) \tag{3}$$

Daily PDR for each customer during the post-program period can be obtained through weather-date matching based method. And how to select the control objects is also necessary and the key point to the model. Therefore, the load pattern matching principle based method is proposed to address this problem.

3.1.3. Pattern matching principle based method

As equation (2) shows, the control group is introduced into the model to enhance the quantifying accuracy. Different from overall analysis to a group of residents, the selection of control objects for each individual should take the diverse characteristics of residents into consideration. The overall electricity usage of control group may not have synchronous changes as an individual does under the same impact from environmental factors. Therefore, it is crucial to select customers that exhibit similar changes as the objects in treatment group from the control group. Moreover, the electricity consumption patterns of residential customers exist a high correlation with HCs [43], which include characteristics covering dwelling, home appliances, demographic composition and behavior habits. Considering the homologous features, the residents with similar typical electricity consumption patterns are more likely to exhibit similar changes affected by the environmental factors. Hence pattern matching principle based method is proposed to select control objects for each resident in treatment group. The framework of the method is shown in Fig. 5.
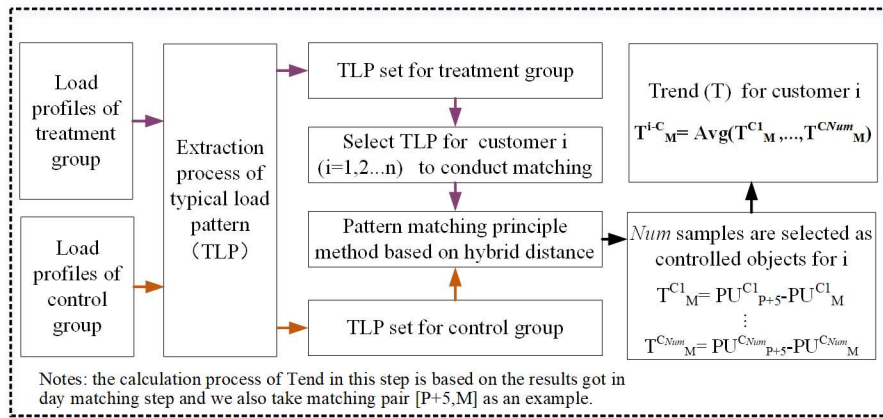

Fig. 5 Structure of the pattern matching principle based method

The pattern matching principle based method can be divided into two parts: typical load pattern (TPL) extraction and load pattern matching based hybrid distance. During the process of TPL extraction, average method is adopted. Average method conducts the extraction through averaging daily electricity load data for each customer and this method is wildly used in the previous research. Then, the hybrid distance, which is composed of Euclidean distance and Correlation distance, is introduced to measure the degree of similarity between two TLPs. The single Euclidean distance lack the measuring ability to the shape of load curves. The Correlation distance, which shows better preference in measuring trend or shape similarity, is introduced as a complement to improve the matching results. The arithmetic average method, that is the weights of two distances are same, is adopted to combine the two distance. Finally, according to the ranking results based on hybrid distance, the top *Num* users (the *Num* is set as 10 in this study) with the most similar load pattern curves to that of customer i are selected as control group. The "trend" of customer i can be described by equation (4).

$$T_{M}^{i-C}=\sum_{j=1}^{Num}TE_{M}^{Cj}\bigg/Num \tag{4}$$

Where TE_{M}^{Cj} represent the daily PDR for the $j\,th$ customers selected. Combining the results obtained from pattern matching principle based method and weather-date based day matching method, the modified PDR (MPDR) of customer i in day M can be calculated as equation (5) using the framework of DID. For each customer, the values of MPDR in post-program period

constitute the daily PDR set for subsequent probability modeling.

$$MPDR_M^i = TE_M^i - T_{M\_}^{i\ C}$$

(5)

3.2. Characterize the features of PDR through probability distribution

The behaviors with randomness in residential electricity consumption is complex to quantify. The approaches to the probabilistic characterization of residential customer electricity consumption behaviors, which are regarded as applicable methods to solve the problem, are extensively used. Distributions such as Weibull distribution, Gamma and Log-normal probability have exhibited good performance in modeling the electricity usage of households. However, different from the electricity usage, negative values show in the set of PDR. The distributions mentioned above are not suitable anymore. After comprehensive consideration, therefore, there widely used distributions, namely normal distribution, logistic distribution and extreme distribution, are adopted to characterize the PDR of residential customers.

3.2.1 Probability distribution

(1) Normal distribution

The normal distribution, which is also named Gaussian or Gauss distribution, is important in statistical field. It is often used in the natural and social sciences to represent real-valued random variables whose distributions are not known. The probability density function of normal distribution can be expressed as equation (6).

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(6)

Where x represents the random variable; μ denotes the mean, and σ is the standard deviation.

(2) Logistic distribution

The logistic distribution, which resembles the normal distribution in shape but has heavier tails (higher kurtosis), has an explicit closed form, so it has some advantages in practical applications[39]. The logistic distribution is a continuous probability distribution, whose cumulative distribution function is the logistic function appearing in logistic regression and neural networks. The probability density fiction of logistic distribution can be expressed as equation (7).

$$f(x \mid \mu, s) = \frac{e^{-\frac{x-\mu}{s}}}{s(1 + e^{-\frac{x-\mu}{s}})^2}$$

(7)

Where x represents the random variable; μ denotes the mean, and s is a scale parameter proportional to the standard deviation.

(3) Extreme value distribution

The extreme value distribution is appropriate for modeling the smallest value from a distribution whose tails decay exponentially fast. The probability density function for the extreme value distribution with location parameter μ and scale parameter σ is:

$$f(x \mid \mu, \sigma) = \sigma^{-1} \exp(\frac{x-\mu}{\sigma}) \exp(-\exp(\frac{x-\mu}{\sigma}))$$

(8)

Where x represents the random variable; μ denotes the mean, and σ is the scale parameter.

3.2.2 Parameter estimation and selection of distribution

In statistics, maximum likelihood estimation (MLE) is introduced to estimate the parameters of the distributions. In the estimation, given the observation data set $X = \{x_1, x_2, ..., x_N\}$ and an assumed probability density function f, maximum likelihood estimation attempts to find the unknown parameters μ and σ that maximize the likelihood function L_f [44]. L_f is shown as follows.

$$L_f(X \mid \mu, \sigma) = f(X \mid \mu, \sigma) \times f(X \mid \mu, \sigma) \times \cdots \times f(X \mid \mu, \sigma)$$
$$= \prod_{i=1}^{N} f(X_i \mid \mu, \sigma)$$

(9)

The goodness of fit tests measures the compatibility of a random sample with a theoretical probability distribution function, that is these tests show how well the distribution selected fits to the data. Several goodness of fit tests have shown in the literature and a widely used test is the two-sample Kolmogorov－Smirnov (KS) test. The test is used to decide if the original data set comes

from the proposed continuous distribution. It is based on the largest distance between empirical cumulative distribution functions of the two distribution. And tables with critical values are introduced to be the judgement of statistical significance[45]. The test statistic is given as follows:

$$D = \max_{x} | F_1(x) - F_2(x) | \tag{10}$$

Where D denotes the test statistic. $F_1(x)$ and $F_2(x)$ represent the two empirical distribution functions respectively.

In this paper, one sample is the set of PDR values, and the other sample is derived from the discrete output of cumulative distribution function fitted by the PDR values through maximum likelihood estimation. The test statistic D will be compared with the critical value from the table at 5%, and the p value is returned according to the comparison result. The null hypothesis of KS test is that the data of two samples are from the same continuous distribution; The alternative hypothesis is that two samples are from different continuous distributions. The sample is marked as "1" if the test rejects the null hypothesis at the 5% significance level (p<0.05), or marked as "0" if the test accepts the null hypothesis (p>0.05), which also represents the fit result passes the KS test in this paper.

3.3. Association rule mining

ARM is a wildly used data-mining method that can discover interesting knowledge from huge amounts of data within a given dataset. It is very effective in identifying novel, implicit and previously unknown relationships among items. During the process of rule mining, rule is defined as the form of $A \rightarrow B$ with two restrictions of $A, B \subset I$ and $A \cap B \neq \varnothing$, where I represents the set of items. A is a set of items called antecedent or the left hand side (LHS) and B denotes a set of items referred to as the consequent of the rule or the right hand side (RHS).

Given a set of data, the data mining method has the potential to generate enormous rules or patterns. However, only a small parts, which have to fulfill certain constraints,are actual interest for data analysis. There are several measures that are regarded as constraints to filter invalid rules and obtain the rules that provide useful information. Support and confidence are the most important indexes to complete the filtration. The support of an association rule shows the percentage of transactions containing the union of sets A and B , and it is taken to be the probability $P(A \cup B)$. Confidence is the proportion of the transactions with A that also contain the union of sets A and B . That is expressed as equation (11).

$$Conference(A \rightarrow B) = \frac{Sup(A \cup B)}{Sup(A)} \tag{11}$$

According to the formulation above, the ARM problem can be regarded as two sub problems, that is generating frequent itemsets, which means finding the itemsets with occurrences surpassing a given minimum support, and generating rules, which indicates discovering the itemsets with confidence over threshold from the frequent itemsets. The two sub problems are obviously straightforward while they are difficult to deal with in practice since global searching over all the possible items should be conducted, which need large computation ability. Therefore, Apriori algorithm is proposed to solve the problems. Apriori is a special method that adopts support based pruning for the sake of mitigating the impact of exponential growth of candidate itemsets[46]. This paper employs Apriori algorithm to conduct ARM.

Besides utilizing well established techniques, the preprocessing for the items is equally important. In Ref.[43], Chi-squared test of independence is introduced to find if there is a significant relationship that data mining is interested in between two items. Then, based the Chi-squared test results, items with significant relationships are selected for further analysis. This step can apparently reduce the dimension of the data so as to lower down the computation and enhance the performance of Apriori. In this paper, $\mathbf{HC} = \{hc_1, hc_2, ...hc_m\}$ denotes the HCs set presented above, and the goal is to find which HCs show influence on PDR and how they influence it. Thus, the HCs preliminary selection procedure is conducted first and then Apriori is carried out, with rules generated denoting as $\mathbf{R}_1$. Among the results, each association rule is a causality. Since we focus on the impact factors of PDR, which reflects through the user type, only the association rules with RHS containing PDR category variables (classification based on parameters of the PDR distribution fitted) remain and form set $\mathbf{R}_2$.

Most ARM algorithms adopt support-confidence framework, and the given thresholds of minimum support and confidence have excluded enormous useless rules. However, there is still a part of uninteresting rules generated. To tackle this problem, two measure indexes, life measure and improvement, are introduced. Lift measure is introduced to make up the shortcoming of confidence measure that ignores the baseline frequency of the consequent[43], given by equation (12).

$$Lift(A \to B) = \frac{Sup(A \cup B)}{Sup(A)Sup(B)} \tag{12}$$

If the resulting value of equation (12) equals to 1, then A and B are independent. If $Lift \leq 1$, the occurrence of A emerges negative correlation with the occurrence of B. The antecedents of the rules with $Lift \geq 1$ show a positive promoting effect on the consequent of the rules and these rules are what we are interested in. Then according to the values of $Lift$, rules with $Lift \geq 1$ are selected and constitute a new rules subset named $\mathbf{R}_3$. However, parts of rules in $\mathbf{R}_3$ with the same consequent but different antecedents probably imply nearly the same knowledge [43], which are redundant. Thus the improvement measurement is used to check the redundancy. The improvement measure of a rule exhibits the minimum difference between its confidence and the confidence of any proper sub-rule with the same consequent, which can be described by equation (13).

$$Imp(A \to B) = \min(\forall A' \subset A, Conf(A \to B) - Conf(A' \to B)) \tag{13}$$

Where A' represents the sub-set of A. The larger value of the improvement is, the larger predictive ability of the rule will be. In this paper, the rules in $\mathbf{R}_3$ whose *improvement* is less than 5% of their confidence indexes are removed to decrease the redundancy and ensure the conciseness.

4. CASE STUDY

In this section, the actual smart metering data and survey data for 2993 customers are used to present the preference of PDR quantifying model and explore the impact factors that influence the PDR effects of TOU programs through the Apriori algorithm based ARM model. All the case studies are performed on a standard PC with an Intel ® Core ™ i7-6700U CPU @ 2.60 GHz, and 8.0 GB RAM. The MATLAB (version 8.6.0.267246, R2015b) is adopted to perform the probability distribution based residential customer PDR characterizing model and the average running time of the modeling using data from each tariff is about 539.55 seconds. The enhanced Apriori algorithm is implemented on a widely used statistical program R (version 3.3.2) and the average running time is about 3.23 seconds

4.1. Results of quantifying the effects of PDR for residents

According to the method introduced in Section 3, we quantify the effects of PDR for residential customers who participated in the TOU programs of CER. In this part, we take the tariff "A" for an example due to its largest number of participants to illustrate the performance of the proposed model.

Fig. 6 shows a pair of matching results of the weather-date based day matching method. Prior to the program, the load pattern curve (the blue one) shows a significant increase during the peak period and the electricity consumption lower down after the peak time. As to the load pattern curve after the "treatment" (the orange one), there is also a peak but it moves to the day period before the peak time. This kind of load shift is led by the price signals and it is also called "Peak shifting". Although there are differences in the peak of the curves, it can be still found out that the two load curves exhibit the similar electricity consumption behaviors. Therefore, the diagram illustrates that the day pairs with similar load patterns can be identified accurately based on the proposed model.
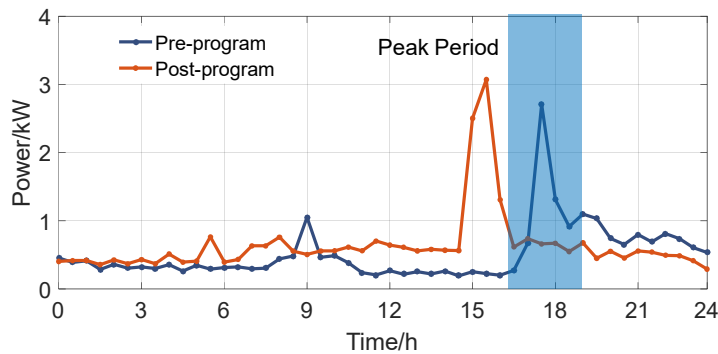

**Fig. 6** The results of the weather-date based day matching method

For each participant in treatment group (TG) "A", pattern matching principle based model using hybrid distance is adopted to select corresponding control objects. We take a customer (ID number 1280) in TG "A" as an example and results of selection are exhibited in Fig. 7. In Fig. 7, the black curve represents the TLP of customer 1280 during benchmark period and the red one

denotes the average of matching results selected from control group. It is obvious that the two curves are highly similar whatever in shape or amplitude because of the hybrid distance used.
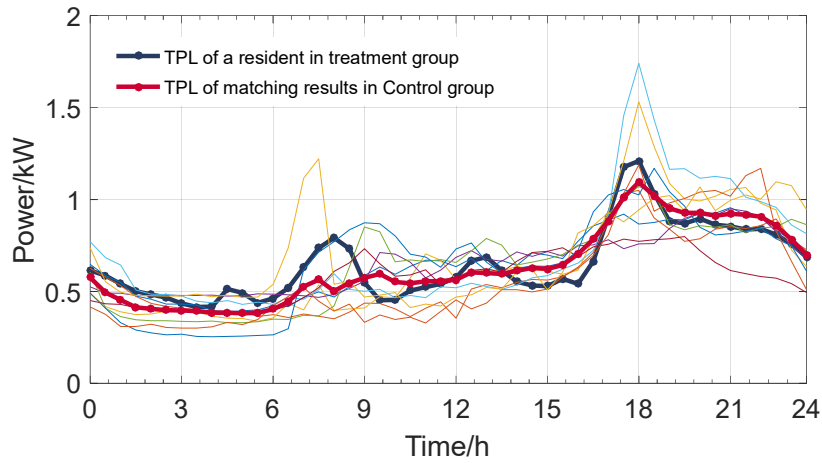
After the two steps above, each customer in TG "A" will obtain a corresponding PDR set, which is comprised of the calculation results of PDR for each day in post-test period. With regard to each PDR set, a frequency histogram can be printed. The diagram of customer 1280 is presented in Fig. 8 and the PDR represents the effects of the TOU policy on customer 1280. The frequency histogram emerges "Bell" shape and the value of PDRs in most days are positive. It means that the customer may lower their electricity consumption or shift the electricity load from peak time to other periods under the influence of the policy of TG "A". It is also obvious that there are several days in which the customer increases the load during period time. This phenomenon may be caused by three reasons: firstly, the calculation of PDR is conducted using daily electricity load data from an individual. The daily electricity consumption for a residential customer usually presents a strong randomness. Some occasional activities, such as visits from friends, holding a party in home may change the daily load curve temporarily through increasing the peak time usage. Secondly, during the process of weather-date based day matching, we hold the assumption that residents may own similar load pattern under similar weather condition in the same "week". However, it is imposed to say the matching results will be in full accordance with the assumption. Some matching deviations may lead to the negative values of PDR. Thirdly, whatever the first reason or the second one may also influence the calculation results of "trend" and then further impact the results of PDR. For a part of customers, the values of PDR in most days are small or negative. It is because these residents are barely influenced by the TOU program or several new household appliances are introduced during the period, which is not taken into consideration in this paper. In addition, not only parts of negative values but also some large values of PDR are created by the three reasons above. Although there are some deviations in calculation, the actual situation for reduction behaviors of customers will emerge after all the PDR calculation results get together just like the frequency of PDR shown in Fig. 8.
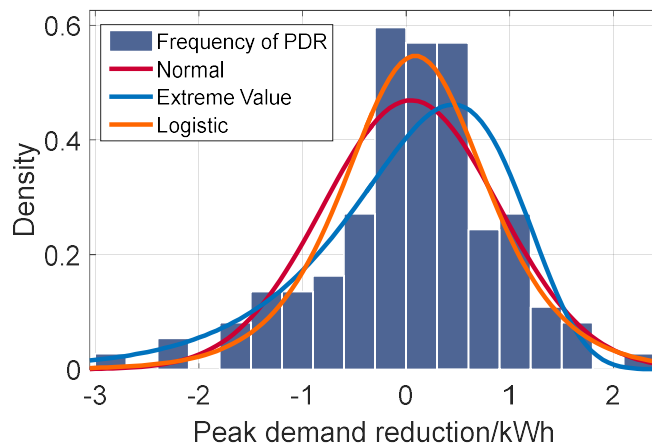
4.2. Results of characterizing the features of PDR

Based on the quantifying results calculated, three wildly used distributions are adopted to characterize the features of PDR for residents and the goodness of fit is measured by KS test in this part. Fig. 8 shows the frequency histogram and corresponding fitting results of three distributions for customer 1280. The PDR of customer 1280 is more concentrated in smaller numerical regions emerging the shape of a spike distribution. From the aspect of shape, the logistic distribution shows better performance than other distributions in describing the sharp peak of frequency histogram. Meanwhile, according to the results of probability distribution fitting and corresponding KS test shown in Tab. 1, it can be seen that the p values of the KS test results for three distributions are larger than 0.05, which means that the null hypothesis cannot be rejected i.e. the fitting results of three distributions all pass KS test. The results of parameter estimation through MLE illustrate that the logistic distribution exhibits the best fitting effects. Therefore, the logistic distribution with parameter μ=0.090 and $\sigma = 0.457$ is selected as the probability distribution to model the PDR for customer 1280.

Tab. 1 The results of probability distribution fitting for the PDR of 1280 and corresponding KS test

Distribution	Log likelihood	Parameter Estimate				KS test	
		μ	Std. Err.	σ	Std. Err.		
Normal	-154.088	0.048	0.077	0.850	0.055	Pass	p=0.797
Logistic	**-151.150**	0.090	**0.071**	0.457	**0.035**	**Pass**	**p=0.974**
Extreme Value	-157.560	0.453	0.076	0.797	0.051	Pass	p=0.797

Fig. 9 shows the fitting results of all customers in tariff A using three probability distributions respectively. For three distribution types, the goodness of each fitting process for a customer is performed by KS test. Then, the ratio of passed to the total number of customers is calculated. The percentage of pass for KS tests and the average of corresponding test statistic D for each distribution type are shown in Tab. 2. Higher passing rate and lower test statistic D indicate that logistic distribution shows better performance than the other distributions in characterizing the feature of PDR data. From the characteristics of the distribution curve, it can be seen that the logistic distribution and normal distribution show similar shape but the former can describe more about the sharp peak characteristic. Therefore, logistic is selected to model the PDR characteristics for customers and the analysis results are based on the fitting results of logistic distribution.
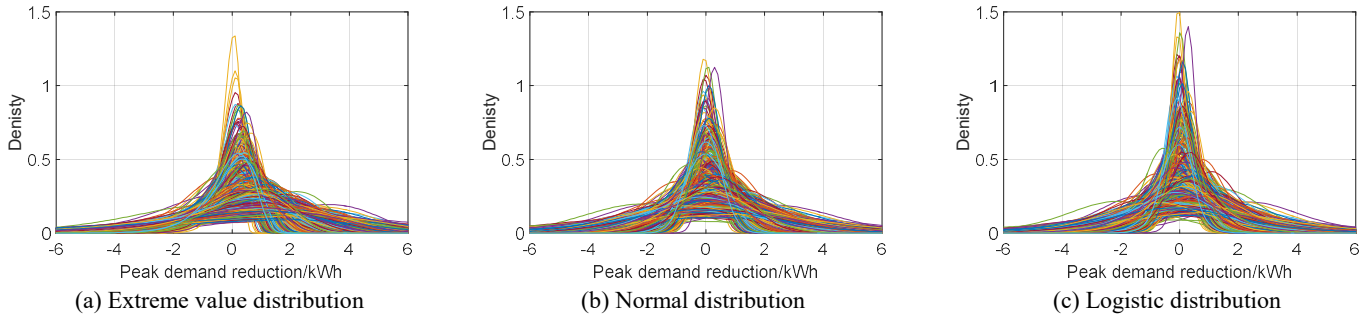

(a) Extreme value distribution (b) Normal distribution (c) Logistic distribution

Fig. 9 The fitting results of the PDR values for all customers in tariff A using three probability distributions.

Tab. 2 KS test percentage of pass and the average of corresponding test statistic D

Distribution	KS pass (%)	Average of test statistic D
Normal	88.26	0.276
Logistic	91.81	0.243
Extreme Value	84.60	0.301

As the most important parameters for distributions, μ and σ can describe the major features of distributions. For the fitting results of PDR, μ represents the average of PDR, which embodies the influence of price signal on the electricity consumption behaviors of customers during peak time. σ is a measure that is adopted to quantify the dispersion of the data. The variation of PDR can reflect the regularity of the customer reposing to the incentives. Some customers show more stable electricity consumption behaviors reduction while the reduction behaviors of other people emerge large variations in different days, which results in a big value of σ.

In the existing research, different models are designed for DR to simulate the electricity consumption behaviors of customers. For incentive-based DR programs, researchers in Ref. [47] model the DR program through adopting the step-wise bid form by DR aggregators, which makes a possibility for the involvement of small loads. With regard to price-based DR programs, especially TOU programs, the DR models reflect the characteristics of load shifting or reduction under the incentives of price signals[48]. One of the most widely used models is the price elasticity model, which utilizes self-elasticity coefficient and mutual-elasticity coefficient to describe the load reducing and load shifting behaviors of users respectively[49]. The price elasticity model and the quantifying model proposed in the paper are all DR models to simulate the responsive behaviors of customers. The former focuses on describing the overall load changes covering all time periods (peak time, day time, and night time), while the latter performs better in describing the characteristics of the responsive behaviors during peak period, particularly the residential customer behaviors with randomness. Moreover, characterizing the PDR features of each residential customer is the key point to the exploration of impact factors, which can be achieved by the proposed model. But it is hard for the price elasticity model to obtain characteristics for each customer due to the acquisition method of the self-elasticity coefficient and mutual-elasticity coefficient. Therefore, considering the emphasis of this study, using the probability distribution based residential customer PDR characterizing model is more effective than other DR models.
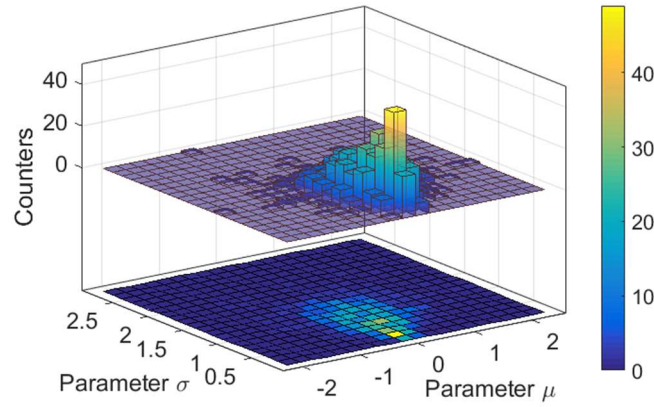
Fig. 10 Distribution of customers in tariff A.
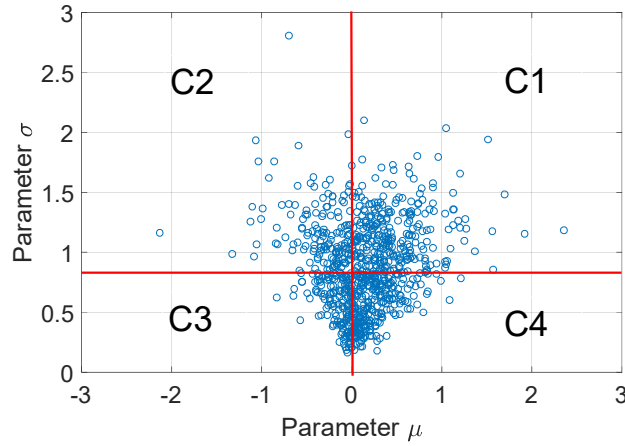
Fig. 11 The distribution of the classification results of the four PDR categories

According to the values of μ and σ, the distribution of customers in tariff A is presented in Fig. 10. Fig. 10 shows that most customers emerge stable peak demand reduction behaviors and the customer with a large amount of PDR usually show a big variation in the electricity consumption behaviors. Most of customers reduce their household demand during peak time under the incentives of price signal. However, a small part of customers shows less reduction in peak demand even negative cut. The differences in reduction level as well as the regularity of behaviors is led by the diversity of HCs. The details will be discussed in section 4.3. For the convenience of further analysis, customers are divided into four categories based on the PDR level (the values

of μ) and the regularity of their corresponding reduction behaviors (the values of σ). The median for σ and "0" for μ are selected as the classification thresholds. Then, four PDR categories named Cluster 1 to Cluster 4 (C1 to C4) are obtained. C1 contains customers with high PDR and large behavior changes. People in C2 show similarly irregular cutting behaviors as C1 but they are much less responsive to the price signals. C3 and C4 show totally opposite characteristics as C1 and C2. The division results are presented in Fig. 11. Using the same approach, similar analysis is conducted to tariff B, C, and D while results are not presented in detail.

4.3. Association rules analysis

In order to conduct the mining of association rules, the cluster information for customers in Cluster 1 to 4 marked as peak demand reduction type 1 to 4 (PDRT 1 to 4) respectively is added to corresponding data set of HCs and constitutes an item set. The combination process is implemented on classification results of the four TOU programs and then the Apriori algorithm is applied to produce holistic knowledge about significant relationships between the HCs and the PDRT of customers. Relative minimum support, which can be obtained through the proportion of each type of users in population, is adopted in the rule mining process to overcome the problem caused by a too small number of rules related to a certain PDRT [43]. Chi-square test is applied to check if the association rules picked exhibit statistical significance at 95% confidence level. Taking tariff A for an example, the rules obtained are summarized in Tab. 3 and sorted by the value of lift.

Tab. 3 Summary of the rules obtained by Apriori algorithm for Tariff A

Rules	LHS	RHS	Sup(%)	Conf(%)	Lift
		Rules related to Cluster1			
1	{Internet_access=Y, Electric cooker=Y, Games consoles=Y}	{T1}**	12.35%	54.01%	1.67
2	{Dishwasher=Y, Electric cooker=Y, Games consoles=Y}	{T1}**	11.61%	53.98%	1.67
3	{Others_use_internet=Y, Electric cooker=Y, Games consoles=Y}	{T1}**	11.25%	53.49%	1.65
4	{Others_use_internet=Y, House =own_with_mortgage, Home_insulated=Y}	{T1}*	9.90%	53.29%	1.64
5	{Tumble dryer=Y, Electric cooker=Y, Games consoles=Y}	{T1}**	11.25%	53.18%	1.64
6	{Dishwasher=Y, Games consoles=Y, Home_insulated=Y}	{T1}*	10.15%	52.53%	1.62
7	{Internet_access=Y, Dishwasher=Y, Games consoles=Y}	{T1}**	13.33%	52.15%	1.61
8	{Others_use_internet=Y, Dishwasher=Y, Games consoles=Y}	{T1}**	12.71%	52.00%	1.61
9	{Heat_Electric_immersion=Y, Tumble dryer=Y}	{T1}*	10.15%	51.88%	1.60
10	{Dishwasher=Y, Lap-top computers=Y, Games consoles=Y}	{T1}*	10.39%	51.83%	1.60
		Rules related to Cluster 2			
1	{Live_state=under_15, Cook=Electric cooker, Tumble dryer=Y}	{T2}**	5.50%	35.16%	2.00
2	{Internet_access=Y, Live_state=under_15, House=own_with_mortgage}	{T2}*	5.38%	33.08%	1.88
3	{Others_use_internet=Y, House =own_with_mortgage}	{T2}*	5.38%	32.59%	1.85
4	{Others_use_internet=Y, Live_state=under_15, Electric cooker=Y}	{T2}*	5.50%	32.37%	1.84
5	{House =own_with_mortgage, Electric cooker=Y}	{T2}*	5.38%	32.35%	1.84
6	{Others_use_internet=Y, House=own_with_mortgage, Cook=Electric cooker}	{T2}*	5.87%	32.00%	1.82
7	{House =own_with_mortgage, Dishwasher=Y}	{T2}*	5.75%	31.97%	1.82
8	{Others_use_internet=Y, Cook=Electric cooker, Games consoles=Y}	{T2}*	5.99%	31.82%	1.81
9	{Cook=Electric cooker, Dishwasher=Y, Washing machine_Fre=1}	{T2}*	5.38%	30.99%	1.76
10	{Live_state=under_15, Lap-top computers=Y}	{T2}*	5.38%	30.77%	1.75
		Rules related to Cluster 3			
1	{Live_state=live_alone, Desk-top computers=0}	{T3}**	5.75%	37.90%	2.08
2	{Others_use_internet=N, Heat_Electric_immersion=N, Desk-top computers=0}	{T3}*	5.50%	36.29%	1.99
3	{Others_use_internet=N, Dishwasher=N,Desk-top computers=0}	{T3}**	6.11%	36.23%	1.99
4	{Heat_Electric_immersion=N,Desk-top computers=0,Washing machine_Fre=<1}	{T3}*	5.87%	36.09%	1.98
5	{Others_use_internet=N, Heat_Electric_immersion=N, Washing machine_Fre=<1}	{T3}*	5.99%	36.03%	1.98
6	{Others_use_internet=N, Dishwasher=N, Washing machine_Fre=<1}	{T3}*	5.99%	35.77%	1.96
7	{Dishwasher=N, Desk-top computers=0, Washing machine_Fre=<1}	{T3}*	6.23%	35.66%	1.96
8	{Dishwasher=N, Desk-top computers=N, Games consoles=N}	{T3}**	6.60%	35.29%	1.94
9	{Internet_access=N,TV>21=1}	{T3}*	5.99%	35.25%	1.94
10	{Internet_access=N,TV>21=Y, Washing machine _Fre=<1}	{T3}*	6.48%	35.10%	1.93

	Rules related to Cluster 4					
1	{SOCIAL CLASS=DE, Done a lot to reduce=2,Washing machine_Fre=<1}	{T4}**	9.66%	50.97%	1.60
2	{SOCIAL CLASS=DE, House=own_not_mortgaged, Washing machine_Fre=<1}	{T4}*	10.15%	50.61%	1.59
3	{SOCIAL CLASS=DE, Games consoless=0,Washing machine_Fre=<1}	{T4}*	10.88%	50.00%	1.57
4	{Done a lot to reduce =2, Lap-top computers=0, Washing machine_Fre=<1}	{T4}**	11.37%	49.73%	1.56
5	{SOCIAL CLASS=DE, Lap-top computers=0, Games consoless=0}	{T4}*	10.76%	49.44%	1.56
6	{SOCIAL CLASS=DE, House =own_not_mortgaged, Lap-top computers=0}	{T4}*	9.78%	49.38%	1.55
7	{House =own_not_mortgaged, Lap-top computers=0,Washing machine_ Fre=<1}	{T4}*	11.49%	48.45%	1.52
8	{Others_use_internet=N, Tumble dryer=N}	{T4}*	9.66%	48.17%	1.52
9	{Retired,Lap-top computers=N}	{T4}*	9.90%	45.76%	1.44
10	{Retired,Washing machine_usage_Fre=<1}	{T4}*	10.76%	45.60%	1.43

** P-value<1% * P-value<5%

4.3.1 *Mining results analysis for Cluster 1 to 4*

Customers in Cluster 1 exhibit high demand reduction during peak time, while the electricity consumption behaviors changes responding to the price signal is not stable. There are 39 rules related to Cluster 1 and 10 rules with the largest lift value are presented in Tab. 3. It is not surprising that customers with a variety of electrical household appliances tend to be type 1. There are 6 appliances, namely electricity cooker, games console, dishwasher, tumble dryer, electric immersion and computer presenting, in the ten rules and the results are consistent to the conclusions in previous research[53]. Among the five rules with largest lift value, the combination of "Games consoles=Y" and "Electricity cooker=Y" appears 4 times, which illustrates that households with game consoles and electric cookers are more likely to be type 1. "Dishwasher=Y" and the utilization of internet, including "Internet_access=Y" and "Others_use_internet=Y", both appear 5 times among 10 rules showing a strong relationship with type 1. Regarding the dwelling characteristics, households with "Home_insulated=Y" tend to be Cluster 1 as presented in Rule 4 and 6.

By contrast, Cluster 2 shows low PDR as well as relatively unstable reduction behavior under the tariff A. After mining the association rules, 89 rules related to Cluster 2 emerged and the top 10 rules with largest lift values are listed in Tab. 3. As Rule 1, 2 and 10 illustrate, customers live with children under 15 years old are more like to be Cluster 2. Among all the rules, "House =own_with_mortgage" appears 5 times showing a significant association with Cluster 2. Regarding the appliance's ownership and heating, the mining results show that households with electricity cooker, game console, dishwasher, tumble dryer, and lap-top are more likely to belong to Cluster 2, which is similar with Cluster 1. It is interesting that customers in Cluster 2 have similar categories of electrical appliances as Cluster 1, which mean semblable demand reduction potentials, but the actual PDR of customers in Cluster 2 is low. To further explore the reasons for the differences of PDR between two type users, another ARM is performed and results are analyzed in 4.3.2.

For households in Cluster 3, they are hardly responsive to the price signal. As Rule 1 shown, people live along without desk-table computer tend to be Cluster 3. Different from Cluster 1 and Cluster 2, customers in Cluster 3 rarely have any electricity appliances expect a TV>21 inch. Moreover, almost no network connection is another feature for Cluster 3.

Cluster 4 shows a stable and remarkable demand reduction during peak time. Only 20 rules are found for Cluster 4. "SOCIAL CLASS=DE" exhibits a significant association with Cluster 4. It appears 6 times among the 10 rules. In terms of electrical appliances, cooking without electric cookers seems to be a feature for Cluster 4. Households without game consoles, laptop as well as tumble dryer are more likely to be Cluster 4. It should be noticed that the attitude towards energy items first emerge in results. As illustrated in Rule 1 and 4, people who believe that they have done a lot to reduce the amount of electricity usage tend to be this cluster. Moreover, people retired with less electricity appliance usage tend to be Cluster 4 according to Rule 9 and 10.

4.3.2 *ARM results of subitem from Cluster 1 and Cluster 2*

It is obvious that Cluster 1 and Cluster 2 show many common LHS in the rules, especially the items "Electric cooker=Y", "Internet_access=Y" and "Others_use_internet=Y". To further explore how the HCs influence the PDR of customers in Cluster 1 and Cluster 2, households with the three common HCs are selected from Cluster 1 and Cluster 2 respectively constituting a sub-item set, which includes 232 customers. Then, Apriori algorithm is conducted on the sub-item set. Sixteen interesting rules with highest lift values severally associated with Cluster 1 and Cluster 2 are listed in Tab. 4.

It can be seen that several new rules emerge after further investigation. With respect to socio-demographic factors, "SOCIAL CLASS=C1" appear 3 times among the 8 rules related to Cluster 2 and the combination of "SOCIAL CLASS=C1" and "Live_state=under_15" appear 2 times illustrating a strong association with Cluster 2. By contrast, the social class shows less association with Cluster 2 and only one rule shows people with "SOCIAL CLASS=DE, Electric shower=Y" tend to be Cluster 1. Regarding electrical appliances, households in both two clusters show abundant types of electrical appliances and similar usage.

But it should be noticed that Rule 3 related to PDRT1 and Rule1, 2 associated with Cluster 2 demonstrate that the percentage of usage being generated from renewable sources show relationship with the type of customers. High proportion tends to be Cluster 2 while households with lower percentage are more like to be Cluster 1. In terms of dwelling characteristics, house with insulation tend to be the feature of customers in Cluster 1 illustrated by Rules 1, 5, 8 related to PDRT1. Most interestingly, Rule 6, 8 related to PDRT2 show people in Cluster 2 own the attitudes that they have done a lot to reduce electricity consumption and they agree with the point that they can reduce the electricity bill by changing the way the people they work with and they use electricity.

Tab. 4 Interesting rules associated with PDRT1 and PDRT 2 discovered from sub-item set

Rules	LHS	RHS	Sup(%)	Conf(%)	Lift
		Rules related to PDRT1			
1	{Done a lot=2, Home_insulated=Y,Chief income earner=E}	{T2}**	13.79%	84.21%	1.30
2	{Heat_Electric_immersion=Y,Desk-top computers=Y, Lap-top computers=N}	{T2}*	13.36%	83.78%	1.30
3	{Generated from renewable sources=4, Overall cost of electricity=4}	{T2}*	13.36%	79.49%	1.23
4	{Electric shower=Y,Desk-top computers=1,overall cost of electricity=4}	{T2}*	14.66%	79.07%	1.22
5	{Age_of_interview=36_45,Electric shower=Y, Home_insulated=Y}	{T2}*	13.79%	78.05%	1.21
6	{SOCIAL CLASS=DE, Electric shower=Y}	{T2}*	13.36%	77.50%	1.20
7	{Heat_Electric_immersion=Y, Games consoless=1, Learn how to reduce=N}	{T2}*	13.36%	77.50%	1.20
8	{Age_of_interview=36_45,Electric shower=Y, Home_insulated=Y}	{T2}*	17.24%	76.92%	1.19
		Rules related to PDRT2			
1	{Immersion=Y,Lap-top computers=Y, Generated from renewable sources=3}	{T1}*	8.19%	59.38%	1.68
2	{Cook=Electric cooker,Lap-top computers=Y, Generated from renewable sources=3}	{T1}*	8.62%	58.82%	1.66
3	{House_age=>40,Lap-top computers=Y,TV>21=1}	{T1}*	8.19%	55.88%	1.58
4	{SOCIAL CLASS=C1,Live_state=under_15,Heat_Electric_immersion=Y}	{T1}*	7.33%	54.84%	1.55
5	{SOCIAL CLASS=C1,Live_state=under_15,Electric shower=Y}	{T1}*	7.33%	54.84%	1.55
6	{Done a lot=1,Lap-top computers=Y,Games consoles=Y}	{T1}*	7.33%	51.52%	1.46
7	{SOCIAL CLASS=C1,done a lot=2,Games consoles=Y}	{T1}*	7.76%	51.43%	1.46
8	{Can_change=2,Tumble dryer=Y,Lap-top computers=Y}	{T1}*	7.33%	50.00%	1.41

** P-value<1% * P-value<5%

4.3.3 *ARM results for tariff B, C, and D*

(a) Results for tariff B   (b) Results for tariff C
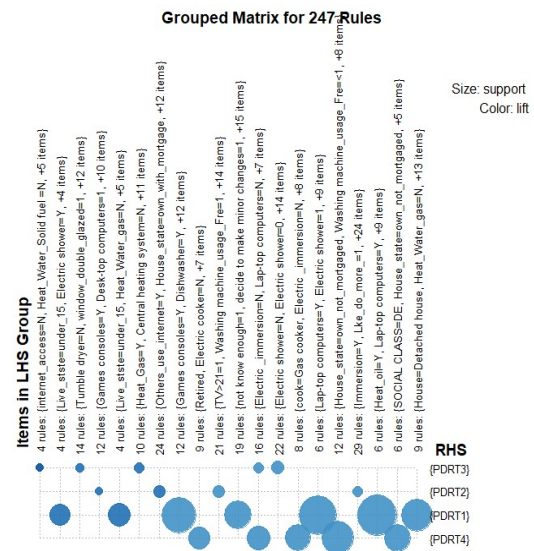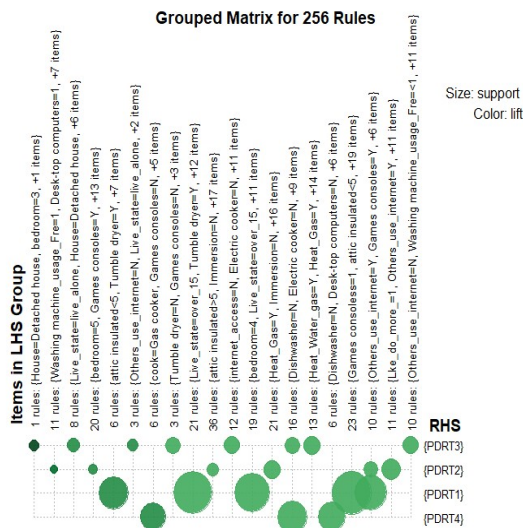
Grouped Matrix for 402 Rules
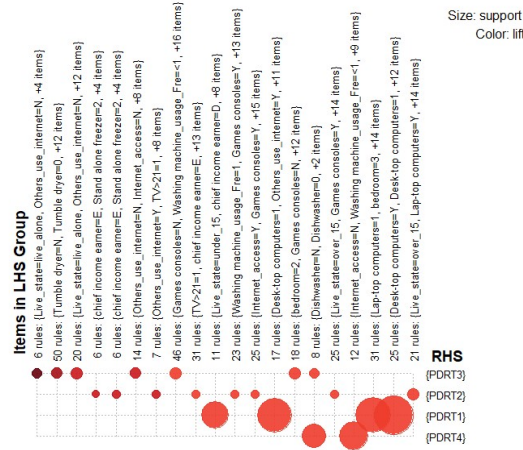
(c) Results for tariff C

Fig. 12 The ARM results for tariff B, C, and D presenting by balloon plot; A balloon plot with antecedent groups as columns and consequents as rows was used. The number of antecedents and the most important (frequent) items in the group are displayed as the labels for the columns. The size of each balloon shows the support value and bigger size of balloon means larger support value. The dark for balloons represents the lift and the darker the color of balloons are, the larger value for the aggregated lift in the group will be.

The ARM results for tariff B, C, and D are presented in Fig. 12 respectively through grouped matrix-based visualization method[51]. We adopt R-extension package ArulesViz[52] to perform the grouped matrix using balloon plot. Consequents items are shown as rows and antecedent groups as columns. Taking the results in tariff D for example, the group of most interesting rules with the largest values of Lift measure are presented in the top-left corner of Fig. 12. Six rules contain "Live_state=live_alone" and "Others_use_internet=N" as well as up to 4 other items as antecedent and "PDRT3" as consequent, which is consistent with the results from the analysis of tariff A.

5. DISCUSSIONS

Through the ARM analysis presented in the section 4, some interesting association rules emerged. In this section, all the results obtained in last part are summarized and some possible mechanisms of the results are discussed as well. Moreover, the applications of the study are also illustrated in this part from customers, retailers and policy makers three aspects respectively.

5.1. Summary

The relationships between four kinds of HCs and user types, which reflect the responsive behaviors of customers incentivized by price signals during peak time, are illustrated in Section 4 through Apriori based ARM methods. Among the discoveries, HCs related to the appliance's ownership and usage habits show the largest impacts on the PDR. Several common electrical appliances show an obvious relationship with the PDR for households. Besides the electricity appliances composition, socio-demographic feature and dwelling characteristics as another two main characteristics also show strong relationship with PDR, especially in accounting for the differences of responsive behaviors of customers in type 1 and type 2. However, the attitudes towards energy usage of households do not show the expected impacts, which may be caused by the deficiencies in the cognition of their self-behaviors. The details for each category are presented as follows.

As the main characteristic influencing the PDR, electrical appliances, such as electric cooker, games console, dishwasher, tumble dryer, electric immersion and computers emerge an obvious relationship with the PDR for households. Some similar results that households with flexible appliances, such as dishwasher, electric cooker show more demand response potential are illustrated in existing research[50,53,54]. However, households in type 1 and type 2, which own similar electricity appliances composition, show various responsive behaviors during the peak time. The difference is the result of multi-factor action including socio-demographic data, the attitudes towards energy as well as dwelling characteristics. So it is inaccurate to model the responsive behavior solely based on the appliance's ownership and heating characteristics. In addition, it can be found that network connection has impacts on the PDR for people. Because the internet is a significant source for residents to access knowledge about the price policy and how to reduce electricity consumption.

Regarding the dwelling characteristics, several association rules are found through the Apriori based analysis. Residential customers with insulated house tend to be type 1, which means more demand reduction in peak time responding to the price signal. Better insulation effect can improve the thermal inertia of house and it contributes to enhance the ability to shift thermal load based on the price signals. Similarly, the installation time of attic insulation is another characteristic that impact the PDR for customers and this association rule was found in the analysis of customers in tariff B. Moreover, another fascinating result emerges during the analysis to tariff A illustrating that the percentage of electricity being generated from renewable sources may influence the PDR level of customers. Households with higher percentage are more likely to be type 2, which means less response amount to the signal price and vice versa.

Socio-demographic data is another significant impact factor for PDR. The population composition of a house may influence the response level of households. The results show that living with children under 15 may influence the responsive behaviors because of the irregular lifestyle and some special need such as the demand of dryer for infants, which low down the flexible ability of some appliances. Customers who live alone tend to have low PDR responding to the price and that is led by a smaller number of appliances. Social class, which is defined as the occupation of chief income earner according to the respondents, have obvious impact on the responsive behaviors to the tariff. For example, customers in type 3 and type 4 both have few electrical appliances. However, households in type 4 prefer to low down their electricity usage and react to the peak price, even the responsive behaviors may decrease comfort level. Moreover, the employment state of retirement for people in type 4 provides more flexible time to adjust their load and response to the price signal.

The attitudes towards energy usage of households can usually reflect their willingness to respond to the price signal. However, based on the ARM results related to type 2, it can be found that the low PDR levels of households in type 2 do not conform to their attitudes "They have done a lot to reduce electricity consumption" and "They agree with the point that they can reduce the electricity bill by changing the people they work with". These differences may be caused by lacking the understanding of household appliances, poor knowledge about how to reduce their consumption or deficiencies in the cognition of self-behavior.

5.2. Applications of the study

The proposed PDR qualifying model and Apriori based ARM model show widespread applications in three perspectives. The details are presented from three aspects, namely customers, retailers and policy makers.

For customers, the results obtained from the ARM model may help them learn more about what influence the electricity usage demand and how they can respond to price signal better during the peak time. Moreover, utilizing the accurate qualifying results and the impact factors obtained through the model, more customized TOU programs may be given by retailers, which are more likely to maximize the benefits of customers.

From the perspective of retailer, the PDR qualifying model enables retailer to have the ability to assess the effect of peak load reduction more accurately and particularly rather than merely through an overall reduction index to evaluate the program. Refined probability characterized qualifying results can provide more information about the responsive behaviors of customers, which can be used to develop customized programs. Several customized TOU programs are proposed in the existing research[55, 56], but the various responsive behavior characteristics and meteorological impact factors[57] such as EV charging behavior[58], daily weather classifications[59], temperature[60] and solar irradiance[61] are not taken into full consideration in these methods. Therefore, the finding in this study will contribute to develop more effective customer targeted program. Moreover, the competition in the electricity retail market becomes increasingly intense with the deepening of electricity market reform in part of countries (e.g. China). The discoveries of the study can also assist electricity retailers in exploring new business models to enhance their competitiveness in the retail market[62].

In addition, the most significant HCs affecting the peak demand responsive behaviors can guide the retailers on how to improve the PDR effect. For example, customers in type 2 show low response level to the financial incentive so as to reduce the overall effect of the program and new methods that utilize social pressure to help customer decrease the electricity usage can be introduced to improve the peak demand reduction effect[63]. Moreover, more feedback information about their electricity consumption and energy saving tips can be sent to them in order to get more demand reduction in peak time.

In order to protect the consumers from the inevitable consequences of a monopoly industry, all the TOU price portfolios should be checked by state or provincial regulation departments, and some equity issues raise concerns with the promotion of TOU programs[64]. Some researchers believe that high electricity prices would present disproportionate influences on low income and high-income households. Because low income customers do not have the capacity to shift or reduce consumption and avoid paying high peak prices. Utilizing the qualifying model, more details about customers with various characteristics can be obtained, which

may help to modify the portfolios and mitigate the equity issues at the most extent. Moreover, the impact factors of PDR explored in this research may also provide the policy makers with information about what influence the electricity consumption behaviors responding to external signal and associate policy makers to design more efficient energy saving policies for residents.

6. CONCLUSIONS

In this paper, a probability distribution based residential customer PDR characterizing model is established and an association rule mining (ARM) based quantitative analysis approach of HCs impact on residential PDR is proposed to comprehensively explore the impact factors. The main findings are presented as follows:

(1) The amount of PDR responding to price signals cannot be obtained simply based on the appliance's ownership as well as its usage habits. The socio-demographic information of households should be taken into consideration together.

(2) To a certain extent, Internet connection contributes to the increase of PDR level of households by providing feedback information and energy saving tips.

(3) The dwelling characteristics are associated with PDR levels of households. Under the similar conditions, customers with insulated house tend to show higher PDR levels.

(4) The percentage of electricity being generated from renewable sources for households has influence on their PDR. Households with larger percentage of renewable generation tend to show lower PDR levels.

Influenced by the diverse household characteristics, parts of the impact factors of PDR found in this paper may alter in different regions or under various price signals. Moreover, it should be noticed that due to the limitation of dataset, the resolution of data used in this study is 30-min time resolution. Higher time resolution of electricity consumption data (15-min or 1-min) from different dataset may lead to various conclusions. Nevertheless, the PDR qualification model and ARM based quantitative analysis approach can be regarded as a general data-driven framework mining impact factors based on different datasets. In the future, cross-comparison of multiple datasets analysis will be conducted to explore the varieties of impact factors

REFERENCES

[1] Wang F, Zhen Z, Mi Z, Sun H, Su S, Yang G. Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. Energy Build 2015;86:427-38. https://doi.org/10.1016/j.enbuild.2014.10.002.

[2] Li G, Li G, Zhou M. Model and application of renewable energy accommodation capacity calculation considering utilization level of inter-provincial tie-line. Prot Control Mod Power Syst 2019;4(1):1. https://doi.org/10.1186/s41601-019-0115-7.

[3] Talari S, Shafie-Khah M, Wang F, Aghaei J, Catalão JPS. Optimal Scheduling of Demand Response in Pre-Emptive Markets Based on Stochastic Bilevel Programming Method. IEEE Trans Ind Electron 2019;66(2):1453-64. https://doi.org/10.1109/TIE.2017.2786288.

[4] Wang F, Xu H, Xu T, Li K, Shafie-khah M, Catalão JPS. The values of market-based demand response on improving power system reliability under extreme circumstances. Appl Energy 2017;193:220-31. https://doi.org/10.1016/j.apenergy.2017.01.103.

[5] Zhen Z, Pang S, Wang F, Li K, Li Z, Ren H, Shafie-khah M, Catalão JPS. Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting. IEEE Ind Appl 2019; 55(4): 3331-3342. https://doi.org/ 10.1109/TIA.2019.2904927.

[6] Wang F, Zhang Z, Liu C, Yu Y, Pang S, Duić N, Catalão JPS. Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. Energy Convers Manag 2019;181:443-62. https://doi.org/10.1016/j.enconman.2018.11.074.

[7] Wang F, Mi Z, Su S, Zhao H. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies 2012;5:1355-70. https://doi.org/10.3390/en5051355.

[8] Wang F, Zhen Z, Liu C, Mi Z, Hodge BM, Shafie-khah M, Catalão JPS. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. Energy Convers Manag 2018;157:123–35. https://doi.org/10.1016/j.enconman.2017.11.080.

[9] Lai J, Lu X, Yu，Xinghuo. Stochastic Distributed Frequency and Load Sharing Control for Microgrids with Communication Delays. IEEE Syst J 2019, Early Access. https://doi.org/10.1109/JSYST.2019.2901711.

[10]   Lai J, Lu X, Wang F, Dehghanian P, Tang R. Broadcast Gossip Algorithms for Distributed Peer-to-Peer Control in AC Microgrids. IEEE Trans Ind Appl 2019, 55(3):2241-2251. https://doi.org/10.1109/TIA.2019.2898367.

[11]   Nojavan S, Zare K, Mohammadi-Ivatloo B. Application of fuel cell and electrolyzer as hydrogen energy storage system in energy management of electricity energy retailer in the presence of the renewable energy sources and plug-in electric vehicles. Energy Convers Manag 2017;136:404-17. https://doi.org/10.1016/j.enconman.2017.01.017.

[12]   Zhang D, Li J, Hui D. Coordinated control for voltage regulation of distribution network voltage regulation by distributed energy storage systems. Prot Control Mod Power Syst 2018;3:35–42. https://doi.org/ 10.1186/s41601-018-0077-1

[13]   Motalleb M, Thornton M, Reihani E, Ghorbani R. Providing frequency regulation reserve services using demand response scheduling. Energy Convers Manag 2016;124:439-52. https://doi.org/10.1016/j.enconman.2016.07.049.

[14]   Redondo Gil C, Esquibel LA, Alonso Sánchez AM. European strategic energy technology plan. Renew Energy Power Qual J 2009;1(7):512-21. https://doi.org/10.24084/repqj07.403.

[15]   Wang F, Li K, Zhou L, Ren H, Contreras J, Shafie-khah M, Catalão JPS. Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting. Int J Electr Power Energy Syst 2019;105:529-40. https://doi.org/10.1016/j.ijepes.2018.08.039.

[16]   Chen Q, Wang F, Hodge BM, Zhang J, Li Z, Shafie-Khah M, Catalão JPS. Dynamic Price Vector Formation Model-Based Automatic Demand Response Strategy for PV-Assisted EV Charging Stations. IEEE Trans Smart Grid 2017;8(6):2903-15. https://doi.org/10.1109/TSG.2017.2693121.

[17]   Majidi M, Nojavan S, Zare K. Optimal stochastic short-term thermal and electrical operation of fuel cell/photovoltaic/battery/grid hybrid energy system in the presence of demand response program. Energy Convers Manag 2017;144:132-42. https://doi.org/10.1016/j.enconman.2017.04.051.

[18]   Nojavan S, Majidi M, Najafi-Ghalelou A, Ghahramani M, Zare K. A cost-emission model for fuel cell/PV/battery hybrid energy system in the presence of demand response program: ε-constraint method and fuzzy satisfying approach. Energy Convers Manag 2017;138:383-92. https://doi.org/10.1016/j.enconman.2017.02.003.

[19]   Zhu Q, Luo X, Zhang B, Chen Y. Mathematical modelling and optimization of a large-scale combined cooling, heat, and power system that incorporates unit changeover and time-of-use electricity price. Energy Convers Manag 2017;133:385-98. https://doi.org/10.1016/j.enconman.2016.10.056.

[20]   Jannati J, Nazarpour D. Optimal energy management of the smart parking lot under demand response program in the presence of the electrolyser and fuel cell as hydrogen storage system. Energy Convers Manag 2017;138:659-69. https://doi.org/10.1016/j.enconman.2017.02.030.

[21]   Mazidi M, Monsef H, Siano P. Incorporating price-responsive customers in day-ahead scheduling of smart distribution networks. Energy Convers Manag 2016;115:103-116. https://doi.org/10.1016/j.enconman.2016.02.030.

[22]   Aghajani GR, Shayanfar HA, Shayeghi H. Presenting a multi-objective generation scheduling model for pricing demand response rate in micro-grid energy management. Energy Convers Manag 2015;106:308-21. https://doi.org/10.1016/j.enconman.2015.08.059.

[23]   Ma T, Wu J, Hao L. Energy flow modeling and optimal operation analysis of the micro energy grid based on energy hub. Energy Convers Manag 2017;133:292-306. https://doi.org/10.1016/j.enconman.2016.12.011.

[24]   Bayati M, Abedi M, Gharehpetian GB, Farahmandrad M. Short-term interaction between electric vehicles and microgrid in decentralized vehicle-to-grid control methods. Protection and Control of Modern Power Systems. Prot Control Mod Power Syst 2019;4:42–52. https://doi.org/ 10.1186/s41601-019-0118-4

[25]   Yu L, Li YP, Huang GH, An CJ. A robust flexible-probabilistic programming method for planning municipal energy system with considering peak-electricity price and electric vehicle. Energy Convers Manag 2017;137:97-112. https://doi.org/10.1016/j.enconman.2017.01.028.

[26]   Rahman MM, Hettiwatte S, Shafiullah GM, Arefi A. An analysis of the time of use electricity price in the residential sector of Bangladesh. Energy Strateg Rev 2017;18:183-98. https://doi.org/10.1016/j.esr.2017.09.017.

[27]   Srivastava A, Van Passel S, Laes E. Assessing the success of electricity demand response programs: A meta-analysis. Energy Res Soc Sci 2018;40:110-17. https://doi.org/10.1016/j.erss.2017.12.005.

[28]   Yang J, Zhao J, Wen F, Dong ZY. A Framework of Customizing Electricity Retail Prices. IEEE Trans Power Syst 2018;33(3):2415-28. https://doi.org/10.1109/TPWRS.2017.2751043.

[29]   Carroll J, Lyons S, Denny E. Reducing household electricity demand through smart metering: The role of improved information about energy saving. Energy Econ 2014;45:234-43. https://doi.org/10.1016/j.eneco.2014.07.007.

[30]   Mizutani F, Tanaka T, Nakamura E. The effect of demand response on electricity consumption under the existence of the reference price effect: Evidence from a dynamic pricing experiment in Japan. Electr J 2018;31(1):16-22. https://doi.org/10.1016/j.tej.2018.01.004.

[31]   Commission for Energy Regulation. Electricity Smart Metering Technology Customer Behaviour Trials (CBT) Findings Report. [Online]Available: http://www.cer.ie/GetAttachment.aspx?id=fa5f2b08-9774-4878-9617-01709eda5bfc

[32]   Newsham GR, Bowker BG. The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. Energy Policy 2010;38(7):3289-96. https://doi.org/10.1016/j.enpol.2010.01.027.

[33]   Faruqui A, Sergici S. Household response to dynamic pricing of electricity: A survey of 15 experiments. J Regul Econ 2010;38(2):193-225. https://doi.org/10.1007/s11149-010-9127-y.

[34] Torriti J. Price-based demand side management: Assessing the impacts of time-of-use tariffs on residential electricity demand and peak shifting in Northern Italy. Energy 2012;44(1):576-83. https://doi.org/10.1016/j.energy.2012.05.043.

[35] Gyamfi S, Krumdieck S, Urmee T. Residential peak electricity demand response - Highlights of some behavioural issues. Renew Sustain Energy Rev 2013;25:71-7. https://doi.org/10.1016/j.rser.2013.04.006.

[36] Nicholls L, Strengers Y. Peak demand and the "family peak" period in Australia: Understanding practice (in)flexibility in households with children. Energy Res Soc Sci 2015;9:116-24. https://doi.org/10.1016/j.erss.2015.08.018.

[37] Miller R, Golab L, Rosenberg C. Modelling weather effects for impact analysis of residential time-of-use electricity pricing. Energy Policy 2017;105:534-46. https://doi.org/10.1016/j.enpol.2017.03.015.

[38] Irish Social Science Data Archive. Data from the Commission for Energy Regulation (CER)-smart metering project. [Online]. Avaliable: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/, accessed Dec.10, 2017.

[39] Munkhammar J, Rydén J, Widén J. Characterizing probability density distributions for household electricity load profiles from high-resolution electricity use data. Appl Energy 2014;135:382-90. https://doi.org/10.1016/j.apenergy.2014.08.093.

[40] Mukherjee S, Vineeth CR, Nateghi R. Evaluating regional climate-electricity demand nexus: A composite Bayesian predictive framework. Appl Energy 2019;235:1561-82. https://doi.org/10.1016/j.apenergy.2018.10.119.

[41] Zhang H. Antinomic policy-making under the fragmented authoritarianism: Regulating China's electricity sector through the energy-climate-environment dimension. Energy Policy 2019;128:162-9. https://doi.org/10.1016/j.enpol.2019.01.003.

[42] Wang F, Li K, Liu C, Mi Z, Shafie-Khah M, Catalão JPS. Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description. IEEE Trans Smart Grid 2018;9(6):6972–85. https://doi.org/10.1109/TSG.2018.2824842.

[43] Wang F, Li K, Duić N, Mi Z, Hodge BM, Shafie-khah M, Catalão JPS. Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns. Energy Convers Manag 2018;171:839–54. https://doi.org/10.1016/j.enconman.2018.06.017.

[44] Morris P, Pawitan Y. In All Likelihood: Statistical Modelling and Inference Using Likelihood. Math Gaz 2008. https://doi.org/10.2307/3621915.

[45] Carpaneto E, Chicco G. Probability distributions of the aggregated residential load. 2006 9th Int. Conf. Probabilistic Methods Appl. to Power Syst. PMAPS, 2006. https://doi.org/10.1109/PMAPS.2006.360235.

[46] Agrawal R, Srikant R. Fast Algorithms for Mining Association Rules. He 20th Int Conf Very Large Data Bases 1994. https://doi.org/ 10.1016/S0045-6535(98)00288-4.

[47] Ghahramani, M., Nazari-Heris, M., Zare, K., Mohammadi-Ivatloo, B. Energy and Reserve Management of a Smart Distribution System by Incorporating Responsive-Loads/Battery/Wind Turbines Considering Uncertain Parameters. Energy 2019;183:205-19. https://doi.org/10.1016/j.energy.2019.06.085.

[48] Nazari-Heris M, Abapour S, Mohammadi-Ivatloo B. Optimal economic dispatch of FC-CHP based heat and power micro-grids. Applied Thermal Engineering 2017:114:756-69. https://doi.org/10.1016/j.applthermaleng.2016.12.016.

[49] Sadeghian, O., Nazari-Heris, M., Abapour, M., Taheri, S. S., Zare, K. Improving reliability of distribution networks using plug-in electric vehicles and demand response. Journal of Modern Power Systems and Clean Energy, 2019: 1-11. https://doi.org/10.1007/s40565-019-0523-8.

[50] Zhao L, Yang Z, Lee WJ. The Impact of Time-of-Use (TOU) Rate Structure on Consumption Patterns of the Residential Customers. IEEE Trans Ind Appl 2017;53(6):5130-38. https://doi.org/10.1109/TIA.2017.2734039.

[51] Hahsler M, Karpienko R. Visualizing association rules in hierarchical groups. J Bus Econ 2017;87(3):317-35. https://doi.org/10.1007/s11573-016-0822-8.

[52] Hahsler M, Chelluboina S. Visualizing Association Rules: Introduction to the R-extension Package arulesViz. R Proj Modul 2011. https://doi.org/ 10.3109/17453679809000941

[53] Chen H, Peng Y hang, Wang Y ling. Thermodynamic analysis of hybrid cooling system integrated with waste heat reusing and peak load shifting for data center. Energy Convers Manag 2019;183:427-39. https://doi.org/10.1016/j.enconman.2018.12.117.

[54] Nojavan S, Majidi M, Zare K. Optimal scheduling of heating and power hubs under economic and environment issues in the presence of peak load management. Energy Convers Manag 2018;156:34-44. https://doi.org/10.1016/j.enconman.2017.11.007.

[55] Huang W, Zhang N, Kang C, Li M, Huo M. From demand response to integrated demand response: review and prospect of research and application. Protection and Control of Mod. Power Syst., 2019;4:148-150. https://doi.org/10.1186/s41601-019-0126-4.

[56] Yang J, Zhao J, Wen F, Dong Z. A Model of Customizing Electricity Retail Prices Based on Load Profile Clustering Analysis. IEEE Trans Smart Grid 2018;10(3):3374-86. https://doi.org/10.1109/TSG.2018.2825335.

[57] Sun Y, Wang F, Wang B, Chen Q, Engerer N. A, Mi Z. Correlation feature selection and mutual information theory based quantitative research on meteorological impact factors of module temperature for solar photovoltaic systems. Energies 2017;10:1-20. https://doi: 10.3390/en10010007.

[58] Bayati M, Abedi M, Gharehpetian G. B., Farahmandrad M. Short-term interaction between electric vehicles and microgrid in decentralized vehicle-to-grid control methods. Protection and Control of Mod. Power Syst., 2019;4:42-52. https://doi.org/10.1186/s41601-019-0118-4

[59]   Wang F, Zhen Z, Wang B, Mi Z. Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV power forecasting. Appl Sci 2018;8:1-23. https://doi: 10.3390/app8010028.

[60]   Sun Y, Wang F, Zhen Z, Mi Z, Liu C, Wang B, Lu J. Research on short-term module temperature prediction model based on BP neural network for photovoltaic power forecasting. 2015 IEEE Power & Energy Society General Meeting, PES GM, 2015. https://doi.org/ 10.1109/PESGM.2015.7286350.

[61]   Wang Z, Wang F, Su S. Solar irradiance short-term prediction model based on BP neural network. Energy Procedia 2011;12: 488-494. https://doi.org/10.1016/j.egypro.2011.10.065.

[62]   Li K, Mu Q, Wang F, Gao Y, Li G, Shafie-Khah M, Catalão JPS, Y Y, Ren J. A Business Model Incorporating Harmonic Control as a Value-added Service for Utility-owned Electricity Retailers. IEEE Trans Ind Appl 2019; Early Access. https://doi.org/ 10.1109/TIA.2019.2922927.

[63]   Wang F, Liu L, Yu Y, Li G, Li J, Shafie-Khah M, Catalão JPS. Impact analysis of customized feedback interventions on residential electricity load consumption behavior for demand response. Energies 2018;11(4):770-. https://doi.org/10.3390/en11040770.

[64]   Yan X, Ozturk Y, Hu Z, Song Y. A review on price-driven residential demand response. Renew Sustain Energy Rev 2018;96:411-19. https://doi.org/10.1016/j.rser.2018.08.003.