This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

Synchronous Pattern Matching Principle Based Residential Demand Response Baseline Estimation: Mechanism Analysis and Approach Description

Fei Wang, *Senior Member, IEEE*, Kangping Li, *Student Member, IEEE,* Chun Liu, Zengqiang Mi, Miadreza Shafie-khah, *Senior Member, IEEE* and João P. S. Catalão, *Senior Member, IEEE*

***Abstract*—Most current customer baseline load (CBL) estimation methods for incentive-based demand response (DR) rely heavily on historical data and are unable to adapt to the cases when the load patterns (LPs) in the DR event day are not similar enough to those in non-DR days. After the error generation mechanism of current methods is revealed, a synchronous pattern matching (SPM) principle based residential CBL estimation approach without historical data requirement is proposed. All customers are split into DR and CONTROL group, including DR participants and non-DR customers, respectively. First, all CONTROL group customers are clustered into several non-overlapping clusters according to LPs similarity in the DR event day. Second, each DR participant is matched to the most similar cluster in the CONTROL group according to the similarity between its load curve segments (LCS) in DR event day, excluding DR part and cluster centroids. Third, the CBL of each DR participant is estimated with an optimized weight combination method using the load data within the DR event period of all the customers in the very matching cluster in the CONTROL group. A comparison with five well-known CBL estimation methods using a dataset of 736 residential customers indicates that the proposed approach has better overall performance than other current CBL estimation methods.**

***Index Terms*—Incentive-based demand response; Customer baseline load; Synchronous pattern matching; Optimized weight combination**

This work was supported in part by the National Natural Science Foundation of China (51577067), the National Key Research and Development Program of China (2017YFF0208106), the Beijing Natural Science Foundation of China (3162033), the Beijing Science and Technology Program of China (Z161100002616039), the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (LAPS18008), the Science and Technology Project of State Grid Corporation of China (SGCC), the China Scholarship Council, the Open Fund of State Key Laboratory of Operation and Control of Renewable Energy & Storage Systems (China Electric Power Research Institute) (5242001600FB). J.P.S. Catalão acknowledges the support by FEDER funds through COMPETE 2020 and by Portuguese funds through FCT, under Projects SAICT-PAC/0004/2015 - POCI-01-0145-FEDER-016434, POCI-01-0145-FEDER-006961, UID/EEA/50014/2013, UID/CEC/50021/2013, UID/EMS/00151/2013, and 02/SAICT/2017 - POCI-01-0145-FEDER-029803, and also funding from the EU 7th Framework Programme FP7/2007-2013 under GA no. 309048. (Corresponding author: Fei Wang)

F. Wang, K. Li and Z. Mi are with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Source (North China Electric Power University), Baoding 071003, China. F. Wang is also with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA (e-mail: feiwang@ncepu.edu.cn).

C. Liu is with the State Key Laboratory of Operation and Control of Renewable Energy & Storage Systems, China Electric Power Research Institute, Beijing 100192, China. (e-mail: liuchun@epri.sgcc.com.cn).

M. Shafie-khah is with C-MAST, University of Beira Interior, Covilhã 6201-001, Portugal (e-mail: miadreza@gmail.com)

J.P.S. Catalão is with INESC TEC and the Faculty of Engineering of the University of Porto, Porto 4200-465, Portugal, also with C-MAST, University of Beira Interior, Covilhã 6201-001, Portugal, and also with INESC-ID, Instituto Superior Técnico, University of Lisbon, Lisbon 1049-001, Portugal (e-mail: catalao@ubi.pt).

I. Introduction

In recent years, the reliability and flexibility of power system operation face serious challenges due to the growing uncertainties from the supply side because of the fast increasing penetration of renewable energy sources [1,2]. In addition to the measures focusing on the supply side itself such as power forecasting technologies of wind [3] and solar [4], the measures leveraging the resources from the demand side, especially demand response (DR), are widely applied as effective alternative means to maintain the reliability and improve the flexibility of power system operation more economically and environment-friendly [5].

To better utilize the massive demand side resources, many attempts have been conducted all over the world to design effective DR programs for industrial, commercial and residential customers [6-8]. Generally, DR programs can be classified into two categories: price-based and incentive-based [9]. For price-based DR programs, participants change their electricity consumption behaviors from their normal patterns in response to the dynamic changes of electricity price over time. The electricity prices increase when the system operates near its peak load, which would encourage end users to reduce their electricity usage demand or shift it to a non-peak time period. For incentive-based DR programs, multiple DR aggregators compete to sell DR products to the system operator in the electricity market by providing financial compensation to end users in order to modify their electricity consumption patterns.

Although DR can provide attractive solutions in the market to help operator running the power system in a more efficient way, there are many challenges facing in its practice [10]. One of the challenges in the implementation of incentive-based DR is the financial settlement of DR participation compensations. DR aggregators usually give financial compensations to the participants according to their load reduction amount (LRA) during the DR event. The customer baseline load (CBL) is introduced to calculate the LRA, which refers to the amount of electricity that would have been consumed by the participants in the absence of the DR event [11]. The difference between CBL and actual load is regarded as the LRA.

The accurate estimation of CBL is critical to the incentive-based DR programs because it involves the interests of both DR aggregators and the participants. If the CBL is underestimated, the participants will probably feel their efforts for load reduction being not fully rewarded and therefore end up reducing their desires to participate or the response deepness. If the CBL is overestimated, DR aggregators will have to pay more compensation that would reduce their own benefits. Therefore, DR aggregators need to figure out the accurate method to avoid the under- or over-estimations of CBL.

It can be found that most current CBL estimation methods all rely on historical data [11-15] according to the literature review in section II, which causes these methods to be unable to adapt to the cases when the load patterns (LPs) in DR event day are not similar enough to those in non-DR event days. This dissimilarity probably is caused by different weather conditions in DR and non-DR event days [16]. Moreover, most current methods are originally designed for industrial or commercial customers, the methods tailored for residential customers are relatively rarely investigated. However, residential customers exhibit more volatile LPs in contrast to industrial and commercial customers due to their random electricity consumption behaviors, which makes current CBL estimation methods unreliable. Therefore, this paper aims to develop a more accurate CBL estimation approach particularly for residential customers.

The contributions can be summarized as follows:

(1) The error generation mechanism of current CBL estimation methods involving historical data is revealed and summarized as "non-synchronous matching". For averaging and regression methods, it refers to the non-synchronous matching from the perspective of time frame between the input (historical data) and output (CBL). For CONTROL group method, it refers to the non-synchronous matching from the perspective of mapping relations of DR participant and its corresponding similar cluster in CONTROL group between DR and non-DR event days. This mechanism explains the essential reason causing the main drawback of current methods involving historical data that large errors will occur while the LPs in DR event days are not similar enough to those in non-DR event days.

(2) A synchronous pattern matching (SPM) principle based residential CBL estimation approach is proposed to address the non-synchronous matching issues. The idea is that only the concurrent data in DR event day instead of the historical data before DR event day should be used (to match DR participant to the corresponding similar cluster in CONTROL group and estimate CBL), which means the input data of SPM all belong to the DR event day within which the CBL to be estimated is supposed to exist. Therefore the non-synchronous matching issues in terms of time frame and mapping relations for current methods both can be resolved by the proposed SPM approach.

(3) The specific steps of SPM are presented subsequently. First, clustering is utilized to match the DR participant whose CBL to be estimated to the cluster (in CONTROL group) showing the most similar LPs in the same DR event day. Second, an optimized weight combination method is developed to form CBL in order to make full use of the information contained in the load data of the customers belonging to the cluster (in CONTROL group) matched in the first step sharing the most similar LPs to DR participant. An additional benefit of SPM is no historical data but only the concurrent load data in DR event day is required, which is especially useful for the new customers who just sign up DR programs but have not yet accumulated enough available historical load data.

The paper is organized as follows. Section II reveals and summarizes the error generation mechanism of most current CBL estimation methods after the literature review. The proposed approach is illustrated in Section III. In Section IV, the comparison between the proposed approach and five well-known methods under different scenarios are presented. In Section V, the impact of three factors on the performance of the proposed approach is analyzed and discussed. Section VI highlights the conclusions and future works.

II. Literature Review

CBL estimation methods proposed in the literature can be classified into three categories: Averaging, Regression and CONTROL group methods. In this section, these estimation methods are reviewed and discussed as follows.

A. Averaging methods

Averaging methods use the average load of X days in the past Y non-DR event days prior to the DR event days to estimate the CBL. According to the difference of data selection criteria, these methods can be further divided into several different sub-categories including HighXofY [11], LowXofY [12] and MidXofY [13]. Generally, averaging methods are easy to understand and implement, but these methods typically generate large estimation errors because the information of the DR event day (e.g. weather condition, the day of week) is not taken into consideration in these methods. In other words, these methods cannot adapt to the change of conditions for the DR event day.

B. Regression methods

The regression methods try to fit a linear/non-linear function to describe the relationship between the load and explanatory variables including historical load and weather data (such as temperature, humidity and wind speed) and then the CBL can be estimated by this function [11-13]. However, DR event days most likely correspond to those extreme weather days (i.e., the temperature is very high or low), thus it is difficult to find past non-DR event days in which the consumption patterns are similar to those in DR event days. Another issue is that the regression methods require a large amount of historical data to estimate suitable coefficients, which probably is unavailable prior to the first DR event day.

C. CONTROL group methods

The CONTROL group method shown in Fig. 1 is to estimate the CBL by using the load data of the non-DR customers who exhibit the most similar LPs to the DR participants. Yi Zhang et al. [14] propose a cluster-based CBL estimation method. First, typical load pattern (TLP) of each customer is generated by averaging the historical load curves (LCs) of all available non-DR event days. Second, all customers are grouped into several clusters according to the similarity of their TLPs through K-means algorithm. Third, the CBL is estimated by averaging the actual load within the DR event period of all non-DR customers in the same cluster. The results show that this method is more accurate than traditional methods. Nevertheless, we note that the clusters derived by clustering are not adjusted over time in that work, thus this method is most possibly unable to adapt to the ever-changing consumption patterns of customers.

In order to find the most similar CONTROL group for each customer in DR group, Leslie et al. [15] develop a new CONTROL group selection method based on the individual LCs. The authors try to select a suitable CONTROL group by minimizing the distance between the LCs of CONTROL group and DR group in historical non-DR event days. The results show that this method can significantly improve the accuracy of CBL estimation. However, due to the volatile LPs of residential customers, there is no guarantee that the LCs in DR event days would certainly similar enough to those in non-DR event days. Namely, the optimal match between the DR participant and its corresponding the most similar cluster in CONTROL group found in non-DR event days does not ensure that it can be achieved in future DR event days.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

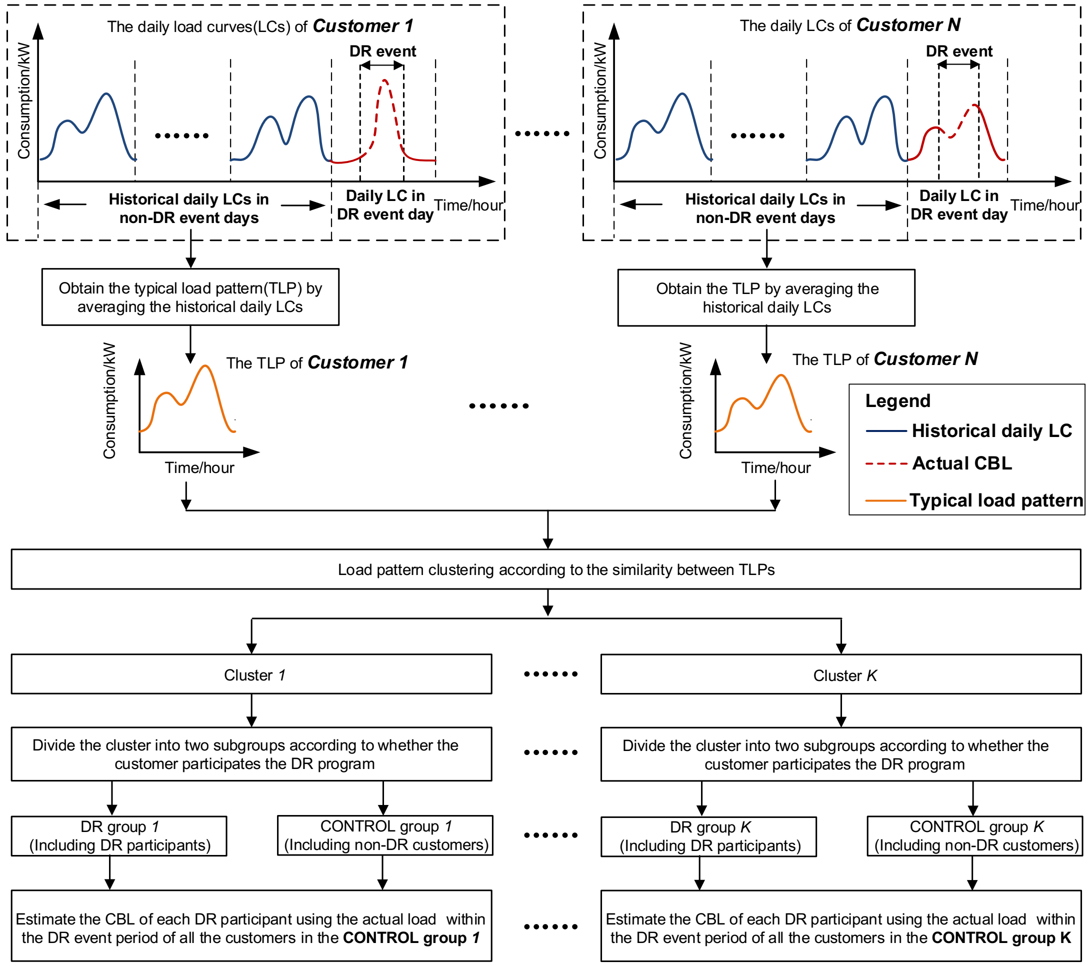


Fig. 1. The flow chart of CONTROL group method.

D. The error generation mechanism of most current methods

Both averaging and regression methods need historical data to estimate CBL. These two methods can work when the LPs in DR event days are similar to those in non-DR event days. However, usually the DR event days are somehow different, which probably results in large errors for these two methods. The error generation mechanism of averaging and regression CBL estimation methods is demonstrated in Fig. 2.

As shown in Fig. 2, it is possible that the LPs of DR participants in DR event day are not similar enough to those in historical non-DR days due to weather or behavioral variations. Then the CBL estimation results of averaging and regression methods would greatly differ from the real CBL, so these two methods cannot provide accurate CBL estimation in the above cases. This is caused by the non-synchronous matching from the perspective of time frame between the input (historical data) and output (CBL) of the methods. CONTROL group method can provide more accurate estimation results because it uses concurrent load data of the non-DR customers in CONTROL group during DR event period to estimate CBL. The key to this method is to match the DR participant with an appropriate cluster in CONTROL group that exhibits the most similar LPs to the DR participant. Currently, the match performed between DR participant and the cluster in CONTROL group only relies on historical data, which could lead to some disadvantages. First, it needs to extract a TLP for each customer to reflect the typical consumption pattern. Sufficient historical data is always required to generate a reasonable TLP. However, these data prior to the first DR event day is probably not available. Second, also more importantly, even though the reasonable TLPs are generated, there is no guarantee that the participants' LPs in DR event days must be similar to their own TLPs, since DR event days usually correspond to some untypical days (e.g. days with unusual temperatures). In other words, although the current CONTROL group method uses the load data of the customers in CONTROL group within DR event days to estimate CBL, the mapping relations between each DR participant and its corresponding similar cluster in CONTROL group are still conducted by using historical data. The mapping relations found in historical data will not be applicable in DR event days in the above cases. The DR participant will be matched to an inappropriate cluster in CONTROL group that exhibits dissimilar LP and thereby large CBL estimation errors will consequentially occur. Therefore, the current CONTROL group method is essentially a non-synchronous matching method just like averaging and regression from the perspective of conducting mapping relations relying on historical data.

In summary, the main drawback of most current methods involving historical data that large errors will occur while the LPs in DR event days are not similar enough to those in non-DR event days, is caused by the non-synchronous matching issues in terms of time frame and mapping relations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

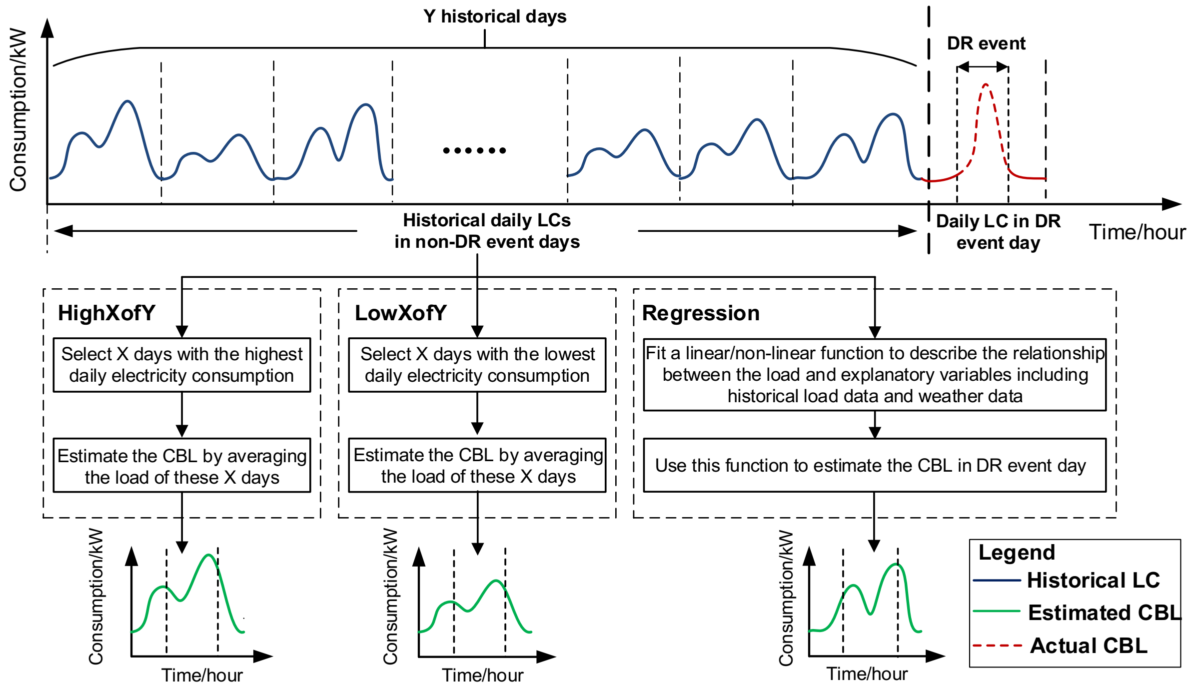


Fig. 2. The error generation mechanism of averaging and regression CBL estimation methods.

III. Methodology

A. Basic idea of SPM principle based approach

A SPM principle based CBL estimation approach performing the match between DR participants and the corresponding cluster in CONTROL group synchronously in each DR event day is proposed to address the above issues.

There are two steps included in CONTROL group method: the first step is the mapping between each DR participant and its corresponding the most similar cluster in CONTROL group, the second step is the estimation using concurrent load data.

Using concurrent load data of the customer in CONTROL group within DR event day to estimate CBL is an essential change to lead non-synchronous to half-synchronous estimation from the perspective of time frame for input data resources, which can eliminate the misleading influences coming from the historical data for those averaging and regression methods.

However, the estimation still probably won't be accurate if inappropriate mapping relations are obtained in the first step while the LPs in DR event day are not similar enough to those in non-DR event days because they are conducted by using historical data.

Two basic ideas to address the above issues are proposed.

1. SPM is proposed that only the concurrent load data in the DR event day, instead of the historical data before the DR event day, should be used to estimate CBL, which means that the input data used in SPM all belong to the same day of the CBL.

2. An optimized weight combination method is developed to form CBL in order to make full use of the information contained in the load data of those customers extracted by SPM in CONTROL group sharing the most similar LPs to DR participant in DR event day.

B. Framework of the proposed SPM approach

Assume $\boldsymbol{D} = \{d \mid d = 1,2,...,D\}$ is the set of DR event days and $\boldsymbol{T} = \{t \mid t = 1,2,...,T\}$ is the set of timeslots for a DR event day.

For a given DR event day $d \in \boldsymbol{D}$, the residential customer set can be divided into two subgroups according to whether it participates the DR program or not:

1) *DR group*, $\boldsymbol{N} = \{n \mid n = 1,2,...,N\}$, which consists of N DR participants;

2) *CONTROL group*, $\boldsymbol{M} = \{m \mid m = 1,2,...,M\}$, which is composed of M non-DR customers (also known as CONTROL customers).

The flow chart of the SPM principle based approach is shown in Fig. 3.

C. LP clustering

LP clustering refers to grouping customers into several clusters such that the customers in the same cluster will share similar electricity consumption pattern while customers in different clusters exhibit distinct LPs. K-means, one of the most widely used clustering algorithms, is used to perform LP clustering due to its advantages such as fast computation speed, effective clustering results and simplicity in the input parameter [16]. K-means separates M CONTROL customers into K clusters based on their LCs in DR event day through an iterative process. K is the number of clusters, which is given by user before clustering. The centroid of each cluster can be obtained by calculating the average of all the data points in the cluster.

For the DR event day d, the aim of K-means is to minimize the sum of squared error between CONTROL customers' LCs and cluster centroids over all K clusters, which is expressed in formula (1).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

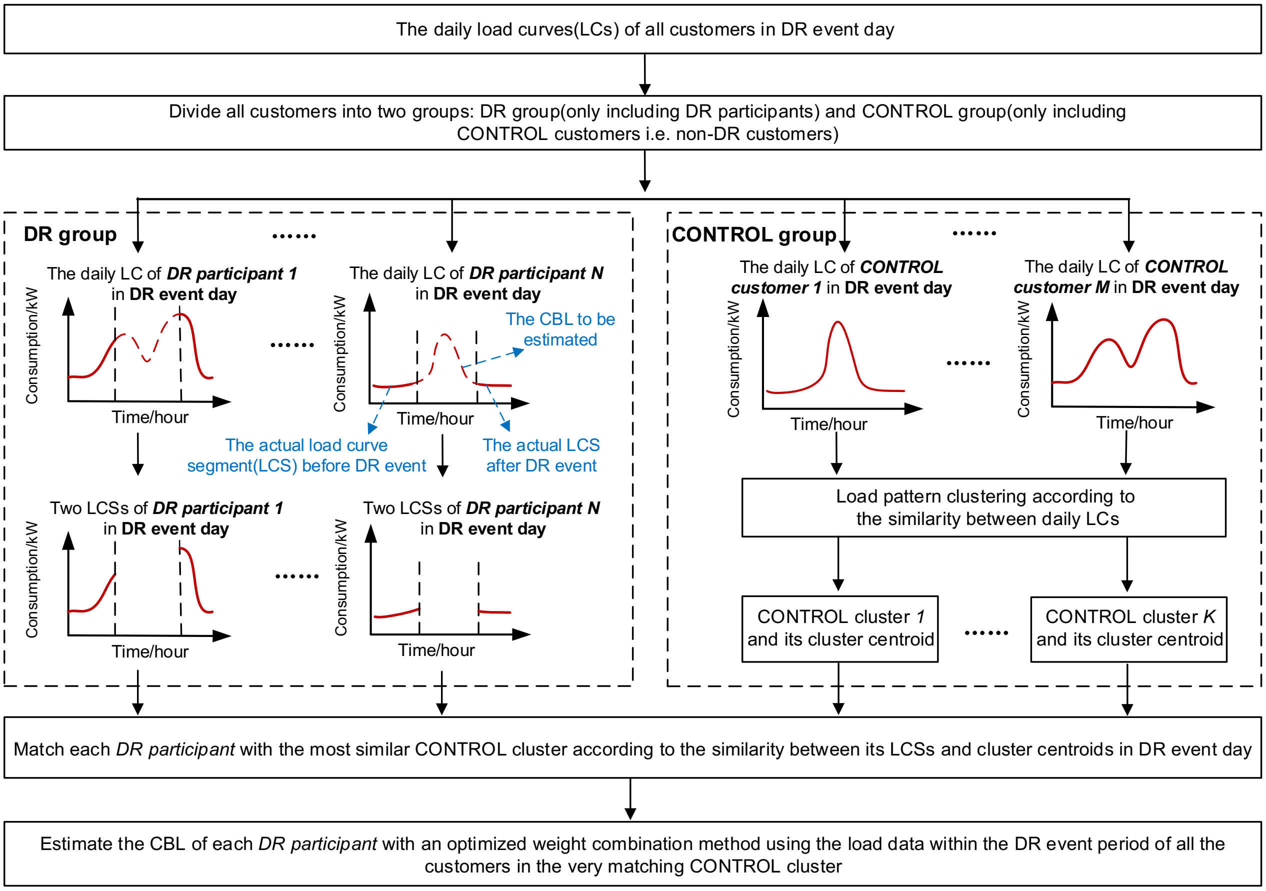


Fig. 3. The flow chart of the proposed SPM principle based approach.

$$\min f = \sum_{k=1}^{K} \sum_{m=1, m\in C_k}^{M} \left\| \boldsymbol{L}_{m,d} - \boldsymbol{C}_{k,d} \right\| \tag{1}$$

where $\boldsymbol{L}_{m,d} = [l_{m,d}^1, l_{m,d}^2, ..., l_{m,d}^T]$ is the actual LC of CONTROL customer m in DR event day d. $\boldsymbol{C}_{k,d} = [c_{k,d}^1, c_{k,d}^2, ..., c_{k,d}^T]$ is the cluster centroid $k, k = 1, 2, ..., K$.

Two indexes for clustering performance evaluation applied in many works [17,18], Davies-Bouldin index (DBI) and Ratio of within Cluster Sum of Squares to Between Cluster Variation (WCBCR), are chosen to determine the suitable value of K for the clustering in this paper.

Even the smaller values of these two indexes generally indicate the better performance of clustering, but the directive property of DBI and WCBCR are not always accordant over the same variation tendency of the clustering number, which means different indexes represent different aspects of clustering performance. Hence, the optimal number of clusters is determined under the balance between the performance requirement and the corresponding complexity of clustering.

The clustering will be performed every time for each DR event day to accommodate the ever-changing consumption patterns of customers. Note that the raw LCs without any normalized processing are used as the input of the clustering because the shape and magnitude of the LC are equally important for CBL estimation. That is to say, only the customers with both similar shape and magnitude of the LC could be clustered together into the same cluster.

D. *SPM based on similarity metric*

For each DR event day, once the K clusters in CONTROL group are obtained, each DR participant should be matched to one of the K clusters through SPM. Here, "SPM" refers to the principle that the data used for matching between DR participant and CONTROL group all belong to the very DR event days. Namely, the matching is performed without any requirements of the historical data out of the DR event day.

The time period of the DR event can be represented by a tuple $\boldsymbol{\delta}$, $\boldsymbol{\delta} = \{\delta_s, \delta_s + 1, ..., \delta_e\}$, where δ_s is the start time and $\delta_e (\delta_e < T)$ is the end time. For each DR day $d \in \boldsymbol{D}$, the actual load data of DR participant n before and after DR event can be utilized to perform the SPM based CBL estimation, which is defined as load curve segments (LCSs), denoted by $\boldsymbol{LCS}_{n,d}^{before} = [l_{n,d}^1, l_{n,d}^2, ..., l_{n,d}^{\delta_s - 1}]$ and $\boldsymbol{LCS}_{n,d}^{after} = [l_{n,d}^{\delta_e+1}, l_{n,d}^{\delta_e+2}, ..., l_{n,d}^T]$, respectively.

Similarly, the load data of the cluster centroid k before and after DR event are defined as cluster centroid segments (CCSs), denoted by $\boldsymbol{CCS}_{k,d}^{before} = [c_{k,d}^1, c_{k,d}^2, ..., c_{k,d}^{\delta_s-1}]$ and $\boldsymbol{CCS}_{k,d}^{after} = [c_{k,d}^{\delta_e+1}, c_{k,d}^{\delta_e+2}, ..., c_{k,d}^T]$, respectively.

The *similarity* between two vectors is denoted by $S(\boldsymbol{x}, \boldsymbol{y})$ and calculated by formula (2).

$$S(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{dis(\boldsymbol{x}, \boldsymbol{y})} \tag{2}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

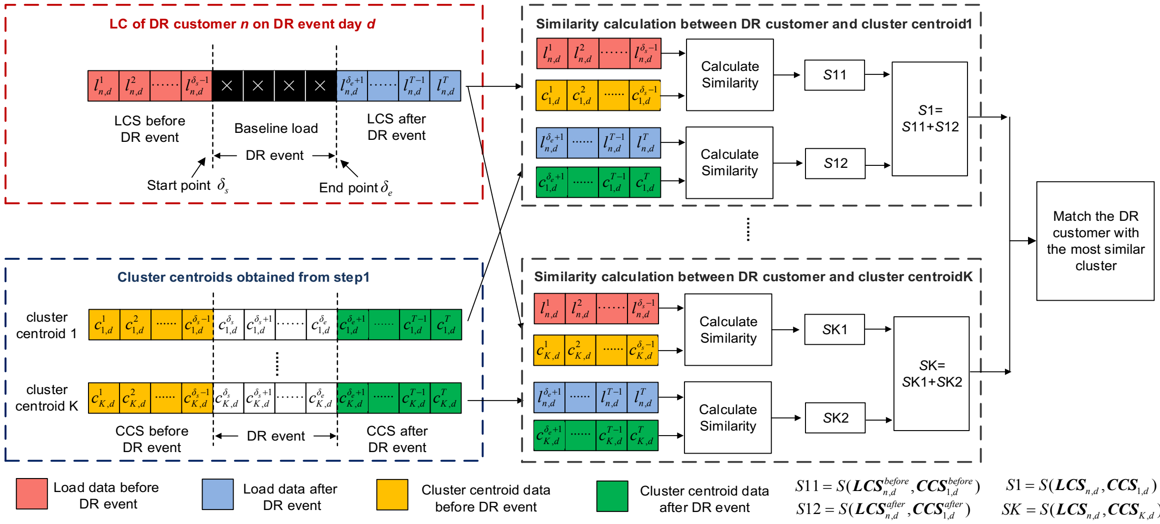


Fig. 4. Illustration of SPM between DR participant and each cluster centroid in CONTROL group.

where $dis(\boldsymbol{x},\boldsymbol{y})$ is the *distance* between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. The *distance* can be a common distance metric, such as Euclidean distance. The larger the *similarity* is, the more similar these two vectors are. The SPM is performed according to the similarity between LCSs and CCSs, which is shown in Fig. 4 and illustrated as follows.

For each DR participant n, calculate the *similarity* between $\boldsymbol{LCS}_{n,d}^{before}$ and each $\boldsymbol{CCS}_{k,d}^{before}, k=1,2...,K$, denoted by $S(\boldsymbol{LCS}_{n,d}^{before},\boldsymbol{CCS}_{k,d}^{before})$. Meanwhile, calculate the *similarity* between $\boldsymbol{LCS}_{n,d}^{after}$ and each $\boldsymbol{CCS}_{k,d}^{after}, k=1,2...,K$, denoted by $S(\boldsymbol{LCS}_{n,d}^{after},\boldsymbol{CCS}_{k,d}^{after})$. The *similarity* between DR customer and each cluster centroid $\boldsymbol{C}_k, k=1,2,...,K$ can be expressed as the sum of $S(\boldsymbol{LCS}_{n,d}^{before},\boldsymbol{CCS}_{k,d}^{before})$ and $S(\boldsymbol{LCS}_{n,d}^{after},\boldsymbol{CCS}_{k,d}^{after})$. Then the DR participant will be matched to the most similar cluster that shows the maximum *similarity* with the DR participant.

E. CBL estimation using optimized weight combination

After matching all DR participants with the corresponding clusters, the next step is to estimate the CBL for each DR participant. For DR participant n, assume she is assigned to the cluster k, find all of the CONTROL customers belonging to the same cluster, indexed by $\boldsymbol{I}_k=\{i\,|\,i=1,2,...,M_k\}$, where M_k is the number of the CONTROL customers in cluster k. These CONTROL customers share the similar LPs to the DR participant n in the DR event day, thus the load data of each CONTROL customer i during DR event can be seen as an independent single estimation for CBL of DR participant n.

Inspired by the basic idea of combination forecasting [19] that properly combining multiple results of several different individual forecast models can effectively improve the forecasting accuracy, a combination estimation model is established to estimate the CBL of DR participant by combining the load data information of all CONTROL customers in the same cluster, given by formula (3).

$$b_{n,d}^t = \underset{i\in \boldsymbol{I}_k}{f}(l_{i,d}^t) \quad ,\forall n\in \mathbf{C}_k, t=\delta_s,...,\delta_e \tag{3}$$

where $b_{n,d}^t$ is the estimated CBL and $f(\;)$ represents a function mapping the load data of CONTROL customers to the CBL of DR participant.

A linear combination is used for the mapping function in this paper, expressed by formula (4).

$$b_{n,d}^t = \sum_{i=1}^{M_k} w_i l_{i,d}^t, t=\delta_s,...,\delta_e \tag{4}$$

where w_i is the weight to the ith individual estimation model (corresponds to the ith CONTROL customer in cluster k).

How to find a set of non-negative weights $\boldsymbol{W}=[w_1,w_2,...,w_{M_k}]$ to make the estimated CBL as close as possible to the actual CBL is the key to this problem.

Since the actual baseline load is unknown in practice, thus we can only use the load data outside the DR event duration to determine the weights. Let us denote the time period without DR event in each DR event day by $\boldsymbol{\varepsilon}=\{1,2,...,\delta_s-1\}\cup\{\delta_e+1,\delta_e+2,...,T\}$ and $\boldsymbol{T}=\boldsymbol{\delta}\cup\boldsymbol{\varepsilon}$. The error of the ith individual estimation model at timeslot t is denoted by e_{it}, $e_{it}=l_{n,d}^t-l_{i,d}^t, t\in\boldsymbol{\varepsilon}$.

All of the errors for individual estimation model i during the time period $\boldsymbol{\varepsilon}$ can form a vector, denoted by the estimation error vector $\boldsymbol{e}_i=[e_{i1},e_{i2},...,e_{i|\varepsilon|}]^T$, where $|\varepsilon|$ is the number of timeslots during the time period $\boldsymbol{\varepsilon}$. The error of the combined estimation model at the time slot t can be calculated by formula (5).

$$e_t = b_{n,d}^{t\;*} - b_{n,d}^t = l_{n,d}^t - \sum_{i=1}^{M_k} w_i l_{i,d}^t = \sum_{i=1}^{M_k} w_i e_{it} \tag{5}$$

where $b_{n,d}^{t\;*}$ is the actual CBL, which is equal to the actual load $l_{n,d}^t$ in the time period without DR event.

The linear combination can be formulated as an optimization model to find a set of optimal weights to minimize the sum of squared errors, which is described in formula (6) as follows.

$$\min J=\sum_{t=1}^{|\varepsilon|} e_{it}^2=\sum_{t=1}^{|\varepsilon|}\sum_{i=1}^{M_k}\sum_{j=1}^{M_k} w_i e_{it} w_j e_{jt}$$
$$s.t.\begin{cases}\sum_{i=1}^{M_k} w_i=1\\ w_i\geq 0, i=1,2,\dots,M_k\end{cases} \tag{6}$$

To present the above model in a matrix form, a square matrix named error information matrix with the size of $M_k\times M_k$ is introduced and denoted as $\boldsymbol{E}_{(M_k)}=(E_{ij})_{M_k\times M_k}$. The element in the matrix $\boldsymbol{E}$ is given by formula (7).

$$\begin{cases}E_{ii}=\boldsymbol{e}_i^T\boldsymbol{e}_i=\sum_{t=1}^{|\delta|} e_{it}^2\\ E_{ij}=\boldsymbol{e}_i^T\boldsymbol{e}_j=\sum_{t=1}^{|\delta|} e_{it}e_{jt}\end{cases} \tag{7}$$

Define a vector whose element values are all equal to 1, denoted by $\boldsymbol{R}=[1,1,\dots,1]^T$. The above optimization problem can be written as formula (8).

$$\min J=\boldsymbol{W}^T\boldsymbol{E}\boldsymbol{W}$$
$$s.t.\begin{cases}\boldsymbol{R}^T\boldsymbol{W}=1\\ \boldsymbol{W}\geq 0\end{cases} \tag{8}$$

This problem can be solved by Lagrange multiplier method, and the optimal weights can be calculated by formula (9).

$$\boldsymbol{W}=\frac{\boldsymbol{E}^{-1}\boldsymbol{R}}{\boldsymbol{R}^T\boldsymbol{E}^{-1}\boldsymbol{R}} \tag{9}$$

F. Error performance metrics

1) Single evaluation indexes

The actual baseline is always unknown when there is a DR event, thus CBL estimation methods are usually tested and evaluated in those days without DR events. Hence, the actual load during DR event duration is regarded as the actual baseline in these test days. Three single indexes including *accuracy, bias* and *variability* are chosen to evaluate the performance of a CBL estimation method in this paper [12]. *Accuracy* can be measured using the average of the baseline's mean absolute error (MAE), which is given by formula (10).

$$MAE=\frac{\sum_{n=1}^{N}\sum_{d=1}^{D}\sum_{t=\delta_s}^{\delta_e}\left|b_{n,d}^t-l_{n,d}^t\right|}{N\cdot D\cdot|\delta|} \tag{10}$$

where $|\delta|$ is the number of timeslots for a DR event.

Accuracy represents the absolute value of the difference between the baseline and the actual load. Lower MAE values indicate more accurate estimation results. *Bias* is measured using the mean of the average error between the estimated baseline and the actual load, which given by formula (11).

$$bias=\frac{\sum_{n=1}^{N}\sum_{d=1}^{D}\sum_{t=\delta_s}^{\delta_e}(b_{n,d}^t-l_{n,d}^t)}{N\cdot D\cdot|\delta|} \tag{11}$$

A positive bias indicates over-estimation while a negative bias indicates under-estimation. When bias is close to zero and MAE is larger than zero, it indicates that the CBL method sometimes over-estimates and sometimes under-estimates the baseline, but overall, the over- and under-estimations balance each other out.

As a matter of fact, bias is more relevant than accuracy on determining the compensation given to DR participants. So, generally the closer the bias is to zero, the better the CBL estimation method is.

Variability is used to evaluate the robustness of a method under various conditions (e.g. different DR event days, different customers) [13]. The index chosen for measuring variability is the relative error ratio (RER), which is defined as the standard deviation of the baseline's prediction errors expressed as a fraction of average load during the event period of time. The variability of DR participant n can be calculated by formula (12).

$$RER_n=\frac{\sum_{d\in\mathbf{D}} std(\boldsymbol{B}_{n,d}(\boldsymbol{\delta})-\boldsymbol{L}_{n,d}(\boldsymbol{\delta})/avg(\boldsymbol{L}_{n,d}(\boldsymbol{\delta}))}{D} \tag{12}$$

where $\boldsymbol{B}_{n,d}(\boldsymbol{\delta})=[b_{n,d}^{\delta_s},b_{n,d}^{\delta_s+1},\dots,b_{n,d}^{\delta_e}]$, $\boldsymbol{L}_{n,d}(\boldsymbol{\delta})=[l_{n,d}^{\delta_s},l_{n,d}^{\delta_s+1},\dots,l_{n,d}^{\delta_e}]$. $std(\ )$ represents the calculation of standard deviation and $avg(\ )$ is the calculation of average value. The smaller the RER_n, the more stable a baseline's error is for participant n. Furthermore, the average value of RER denoted by RER^{avg} can be used to measure the stability of the CBL estimation method.

2) Overall evaluation index

The above three indexes evaluate the performance of a CBL estimation method from different points of view. In order to evaluate the overall error performance of a CBL estimation method, a new performance metric named overall performance index (OPI) is proposed, which is defined as the weighted sum of the absolute normalized value of *accuracy*, *bias* and *variability*. Assume there is a set of CBL estimation methods to be compared indexed by $\boldsymbol{O}=\{o\,|\,o=1,2,\dots,O\}$, the OPI of the *o*th method can be calculated by formula (13).

$$OPI_o=\lambda_1\frac{MAE_o}{\max\limits_{o\in\boldsymbol{O}}(MAE_o)}+\lambda_2\frac{|bias_o|}{\max\limits_{o\in\boldsymbol{O}}(|bias_o|)}+\lambda_3\frac{RER_o^{avg}}{\max\limits_{o\in\boldsymbol{O}}(RER^{avg})} \tag{13}$$

where $MAE_o, bias_o, RER_o^{avg}$ are the values of the corresponding indexes for the *o*th method. $MAE_{\max}$, $|bias|_{\max}$ and $RER_{\max}$ are the maximum values of the corresponding indexes among the O CBL estimation methods to be compared. $\lambda_1,\lambda_2,\lambda_3$ are the weight coefficients and are set to be equal in this paper, i.e. $\lambda_1=\lambda_2=\lambda_3=1$, since these three indexes are equally important in assessing the performance. A lower OPI value indicates a better overall performance.

IV. CASE STUDY

A. Dataset

The data used in this research is obtained from the Commission for Energy Regulation (CER) in Ireland [20]. CER carried out the Smart Metering Electricity Customer Behavior Trials (CBTs) during 2009 and 2010 for the purpose of assessing the impact of smart meters on consumer's electricity consumption to inform the cost-benefit analysis for a national rollout. Over 4,000 Irish residential customers participated this trial. In order to consider the seasonal effect

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

on load, we chose a full year load data with 1-hour interval from Jan.1st to Dec.31st of 2010, and only the load data of customers from "control group" (note here the "control group" is different from the aforementioned one) are used. These customers' consumption behaviors are not affected by the TOU tariff.

The data set is trimmed by removing the customers with missing load data, and finally 763 customers with a full year load data are obtained for the further analysis.

B. Experimental Settings

1) Selection of event-like days

Because there is no DR event in the dataset, we ran the CBL methods including the proposed approach on several event-like days in order to test and evaluate their performance. Considering DR events are usually conducted on those extreme weather days (i.e., the temperature is very high or low), therefore, we selected 5 hottest days in summer (from Jun. to Aug.) and 5 coldest days in winter (from Dec. 01 to Dec.31 combine with Jan. 01 to Feb. 28) as the DR event-like days to test the CBL estimation methods. The temperature data is obtained from a professional weather website named Wunder Ground [21]. The DR event period is chosen as 4 hours from 4:00 pm to 8:00 pm for these DR event-like days.

2) CBL estimation methods used for comparison

The proposed approach is compared with other five well-known CBL estimation methods: three averaging methods, the regression method and the CONTROL group method named "TLP-cluster" proposed in Ref [14]. Table I presents the overview of these six methods.

3) Scenarios setting

Multiple different scenarios are set to test the performance so as to better illustrate the differences between the other methods and the proposed approach, in which the lengths of available historical data vary from three months to three days. MAE, bias and variability are calculated for every method under each scenario.

TABLE I. OVERVIEW OF SIX CBL METHODS FOR COMPARISON.

CBL estimation method	Categories	Description
High5of10	Average	Adopted by NYISO [13]
Mid4of6	Average	Recommend by PJM [13]
Low5of10	Average	Proposed in Ref [12]
Regression	Regression	Adopted by ERCOT [13]
TLP-cluster	CONTROL	Proposed in Ref [14]
SPM	CONTROL	Proposed in this paper

C. Results and analysis

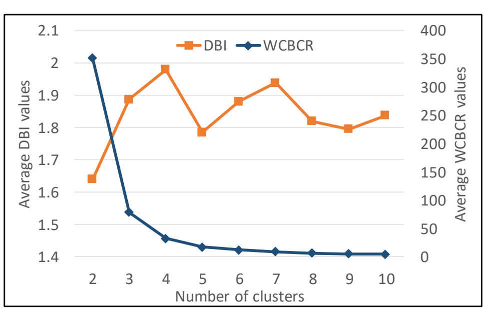
1) Clustering results

K-means was implemented separately for each DR event day through MATLAB R2012b. Euclidean distance was selected as the similarity metric for clustering and matching. Over 100 rounds with the number of clusters ranging from 2 to 10 were performed for each DR event day *d* in order to find an optimal number of clusters. A new set of initial centroids was chosen randomly in each round. DBI and WCBCR were calculated for each round as well. The average values of these two clustering validity indexes for the first DR event day are shown in Fig. 5 (a). It was noted the average DBI value was lowest for two clusters. However, two clusters are too rough to distinguish all LPs. So more than two clusters were sought. In detail, two extremum points can be observed at 5 and 9 for DBI. Hence, the optimal number of clusters can be chosen as 5 or 9. Regarding the WCBCR, it presented decreasing tendency as the number of clusters increased and showed a convergence trend when the number of clusters was larger than 5. In summary, considering the balance between clustering quality and complexity, we finally chose 5 as the number of clusters. The same process was used for the other event days as well. The five derived cluster centroids are shown in Fig.5 (b). It could be observed that the derived five clusters differ from each other in terms of the load shape and amplitude. Cluster2 was comprised of customers with low electricity consumption and flat LPs. Cluster5 corresponded to customers with high electricity usage levels.

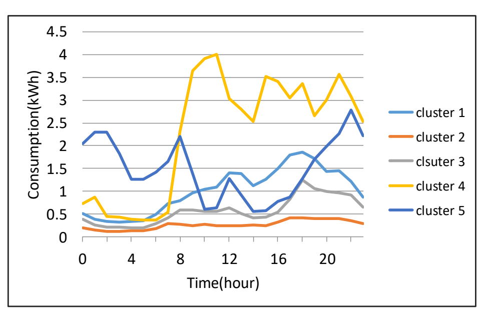
2) CBL estimation results

To make the comparison results more reliable, the procedure was run 100 rounds. In each round, 100 customers were randomly selected as DR participants and the remaining 663 customers are selected as CONTROL customers.

Table II presents the comparison results (mean performance) for six CBL methods under different scenarios. T-test was used to determine if the performance of the proposed approach was statistically significantly better than other CBL estimation methods. The results of T-test are shown in Table III. In general, the proposed approach showed statistically significantly better overall performance than the other five CBL estimation methods. CONTROL group based methods such as TLP-cluster and SPM showed better overall performance than averaging and regression methods. Additionally, Mid4of6 outperformed the other two averaging and regression methods.



(a)



(b)

Fig. 5. Clustering results for the first event day. (a).Two clustering evaluation indexes for k-means. (b).The derived five cluster centroids.

TABLE II. COMPARISON RESULTS OF MAE, BIAS, RER AND OPI FOR SIX CBL ESTIMATION METHODS UNDER DIFFERENT SCENARIOS.

methods	Three-month historical data				Three-week historical data				Three-day historical data			
	MAE	Bias	RER^{avg}	OPI	MAE	Bias	RER^{avg}	OPI	MAE	Bias	RER^{avg}	OPI
High5of10	0.4973	0.1135	0.8065	2.2768	0.4973	0.1140	0.8029	2.2866	—[a]	—	—	—
Mid4of6	0.4499	-0.0651	0.7053	1.8614	0.4500	-0.0646	0.7031	1.8632	—	—	—	—
Low5of10	**0.4415**[b]	-0.2187	0.5953	2.3996	**0.4415**	-0.2182	0.5941	2.4183	—	—	—	—
Regression	0.6469	0.2234	0.5856	2.7261	0.6508	0.1891	0.6002	2.6142	0.6666	0.1617	0.6103	3.0000
TLP-cluster	0.5137	-0.0102	0.5808	1.5599	0.5204	-0.0113	0.5967	1.5946	0.5293	-0.0119	0.6102	1.8675
SPM	0.4874	**-0.0091**	**0.5644**	**1.4940**	0.4882	**-0.0065**	**0.5662**	**1.4851**	**0.4884**	**-0.0082**	**0.5665**	**1.7116**

a: Fail to provide CBL estimation results due to lack of sufficient historical data;
b: The best result for each column is shown in bold.

TABLE III. RESULTS OF T-TEST.

Methods	Three-month historical data				Three-week historical data				Three-day historical data			
	p-value				p-value				p-value			
	MAE	Bias	RER^{avg}	OPI	MAE	Bias	RER^{avg}	OPI	MAE	Bias	RER^{avg}	OPI
SPM-High5of10	**	**	**	**	**	**	**	**	—	—	—	—
SPM-Mid4of6	**	**	**	**	**	**	**	**	—	—	—	—
SPM-Low5of10	**	**	**	**	**	**	**	**	—	—	—	—
SPM-Regression	**	**	**	**	**	**	**	**	**	**	**	**
SPM-TLP-cluster	**	*	**	**	**	*	**	**	**	*	**	**

*: p-value>0.05
**: p-value<0.05

Specifically, the proposed method presented the best performance over the six CBL methods in terms of *bias* and *variability* and also demonstrated competitive results on *accuracy*.

Mid4of6 and TLP-cluster showed overall average performance on all single evaluation indexes but did not show overall best results on any index. Low5of10 showed the best performance in terms of *accuracy* but the worst results for *bias*. Specifically, it was not surprising that Low5of10 had more negative bias than other methods, while High5of10 showed much more positive bias than the other methods. Due to the lack of refined weather data (such as humidity, wind speed and other weather information), the regression method presented the overall worst estimation results in terms of *accuracy* and relative poor *bias* and *variability*.

In terms of the applicability of the scenarios with different length of historical data, the proposed approach can well accommodate to multiple scenarios even with limited historical data. This was because no historical data was needed for the proposed approach. Both the regression and TLP-cluster methods were sensitive to the amount of the available historical data. All single evaluation indexes of these two methods became worse with the decrease of available historical data. Regression needs a long historical data for the model training and insufficient samples will make the model under-fitting. Sufficient historical data is also required for TLP-cluster method to generate a reasonable TLP for clustering.

Regarding averaging methods, the CBL estimation results were not affected by the length of historical data, nevertheless they all failed to provide valid estimation results when there was no sufficient historical data (i.e. less than 10 days).

Taking the participant#1 as an example, the daily LCs of DR participant#1 in the last ten weekdays prior to December 17th, 2010 (i.e. one of the DR event days) are shown in Fig. 6.

It can be seen that the LP on December 17th, 2010 was apparently different from those LPs in the last ten weekdays in terms of both magnitude and shape. The above six methods were used to estimate the CBL and the results are shown in Fig. 7. It can be observed that the estimated baseline obtained by the proposed approach was very close to the actual baseline. By contrast, the other five non-synchronous CBL estimation methods presented large positive bias and MAE.

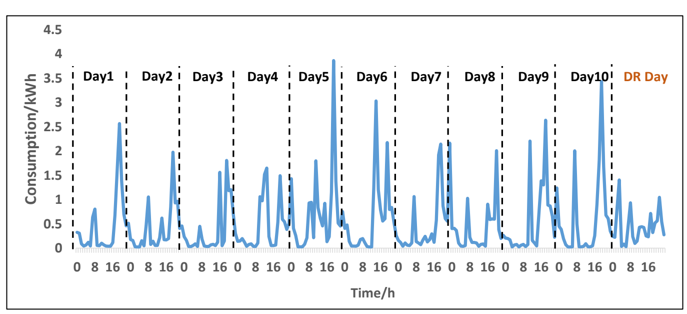


Fig. 6. The daily LCs of DR participant#1 in the last ten weekdays prior to December 17th, 2010.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

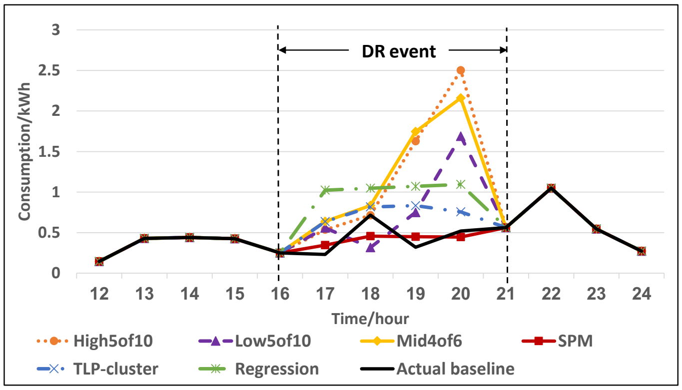


Fig. 7. Comparison of the CBL estimation results obtained by six methods for DR participant#1 on December 17th, 2010.

V. Discussions

The key to the proposed approach is whether the DR participant with incomplete load data (because the CBL is to be estimated) can be properly matched to the correct CONTROL cluster (i.e. the cluster show similar LP to DR participant). In this section, the impact of three factors on the matching results are investigated and discussed.

A. The impact of DR event duration

Various DR event durations from 1 hour (19:00-20:00) up to 8 hours (12:00-20:00) were set to explore its impact on the performance of the proposed approach. After 100 rounds tests for the 8 different DR event durations, the MAE and Bias of the proposed approach were calculated in each round. The distribution of the estimation results of the 100 rounds is illustrated in Fig. 8.

Interestingly, we found that the value of MAE first increased with the increase of the DR duration, but then presented a decreasing tendency when the duration was over 3 hours. The reasons can be considered as follows.

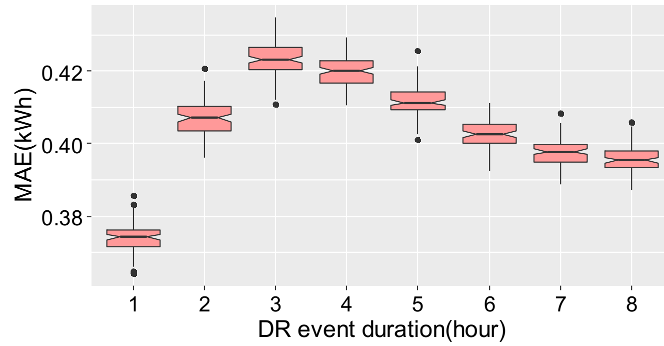
On the one hand, the DR event duration affects the length of loss load data and thereby affects the accuracy of LP matching. On the other hand, the difficulty of CBL estimation varies in different time periods. These two factors have combined effect on the final estimation results. We tried to isolate these two factors to figure out the specific function mechanism of each of them.

For the first aspect, we used the whole LC of each DR participant to perform the LP matching and the obtained matching results were treated as the perfect matching results considered as the ground truth. Then the accuracy of LP matching can be obtained by comparing the LP matching result derived by LCS to the ground truth. The results are shown in Fig. 9 (a). It is not surprising that the accuracy of matching decreased with the increase of DR event duration, because the available information used to perform LP matching reduces with the increase of DR duration.

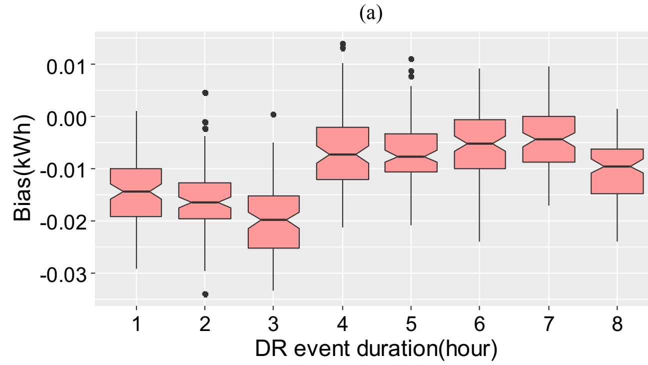
For the second aspect, in order to explore the estimation difficulty for different time periods, we still use the whole LC of each DR participant to perform the LP matching and utilize the perfect matching results to estimate the CBL.

The MAE values for the estimation results were calculated for different DR durations and illustrated in Fig. 9 (b).

The MAE values showed the similar tendency to the results in Fig. 8 (a), which indicated that the time period from 4:00 pm to 8:00 pm was most difficult to estimate. It was probably because that the customers have diversified electricity demand in this period.

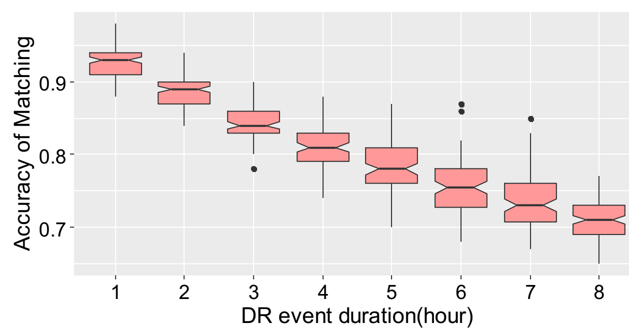


(a)

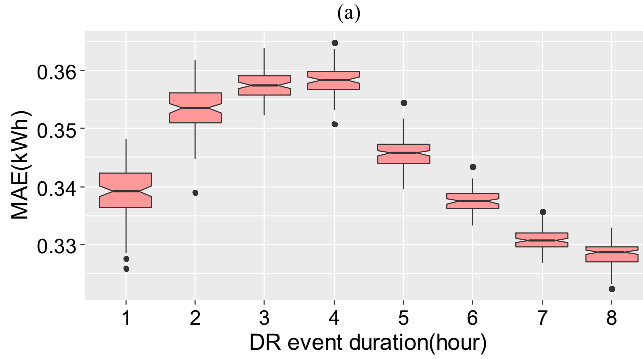


(b)

Fig. 8. Distribution of the CBL estimation results for different DR event durations. (a). MAE and (b). Bias.



(a)



(b)

Fig. 9 (a). Accuracy of LP matching for different DR event durations. (b). Distribution of MAE of the estimation results derived by perfect matching for different DR durations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

B. The impact of CONTROL group size

Various size ratios (SRs) of CONTROL group to DR group from 0.5 to 5 were set to investigate its impact on the performance of the proposed approach. The number of DR participants was fixed as 100 and the number of CONTROL customers changed from 50 to 500. The proposed approach was run 100 rounds for each SR and different CONTROL customers were randomly selected from the whole CONTROL group for each round. The distribution of the CBL estimation results is illustrated in Fig. 10.

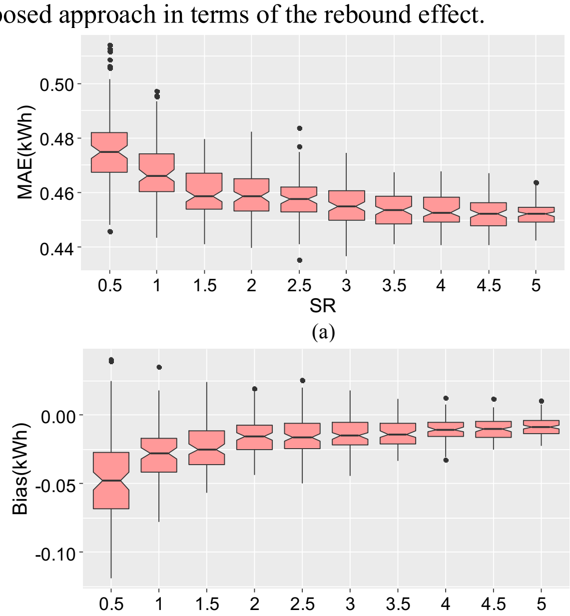
Tukey test, a single-step multiple comparison procedure, was used to determine if these results corresponding to different SR options were statistically significantly different from each other. The results of Tukey test are shown in Tables IV and V.

Both accuracy and bias became better with the increase of SR. However, in terms of MAE, there was no significant difference between these estimation results when the SR was larger than 2. In terms of Bias, a ratio of 3.5 or higher provided statistically similar performance. This was because the diversity of LPs for CONTROL group increased when the SR became larger. There was a higher possibility that each DR participant can be matched to a suitable CONTROL cluster.

C. The impact of rebound effect

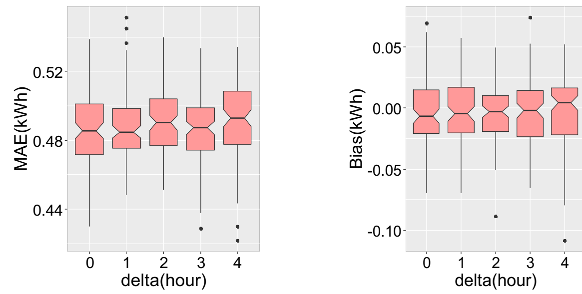
Once the DR event ends, the demand of DR participants will typically exceed the baseline load for a period of time known as rebound, which may affect the clustering and thereby have an impact on the performance of the proposed approach.

Only the data after a period of time following the end point of DR event (end+*delta*) was used in addition to the data before DR event to avoid the influence caused by the rebound on the proposed approach as lower as possible. The *delta* time after DR event was set from 0 to 4 hours with the interval of one hour (0, 1h, 2h, 3h, 4h) to test the robustness of the proposed approach in terms of the rebound effect.



(b)

Fig. 10. Distribution of the CBL estimation results for different SRs between CONTROL group and DR group. (a). MAE and (b). Bias.



(a) (b)

Fig. 11. Distribution of the CBL estimation results for different values of *delta*. (a). MAE and (b). Bias.

The DR event time was still set to be 4:00 pm-8:00 pm. For example, only the data before DR event was used in the clustering and estimation process when the delta time is 4 hours. The distribution of the CBL estimation results for different values of *delta* is shown in Fig. 11. Tukey test was also used to determine if the obtained estimation results were statistically significantly different from each other. The p-values were 0.322 and 0.922, which indicated that there was no significant difference between these estimation results in terms of both MAE and Bias. In other words, the rebound effect had a negligible impact on the estimation results.

VI. CONCLUSIONS

In order to improve the accuracy of CBL estimation for incentive-based DR programs, a SPM principle based residential CBL estimation approach is proposed. The SPM idea is feasible since CBL estimation isn't necessary a real-time computation problem.

Compared to those current non-synchronous estimation methods, the proposed approach presented statistically significantly better overall performance. Moreover, it also can well adapt to the scenario, with very limited historical data. Namely, the proposed approach not only overcomes the shortcoming of current methods in terms of the over reliance on historical data, but also obtains higher accuracy for those DR event days in which the LPs are not similar enough to those in non-DR event days. The applicability and overall performance of CBL estimation have been improved by the proposed approach. This work is valuable for DR aggregators to estimate the CBL for residential customers, especially useful for new customers who just signed up the DR program but have not yet accumulated enough available historical load data.

The future works of this research are listed as follows:

(1) Test the proposed approach on more datasets to further verify its effectiveness.

(2) More and more customers have installed distributed photovoltaic (PV) systems, which could significantly affect the net load profiles and thereby have an impact on the SPM based approach. We will study the impacts of the presence of different penetration rates of distributed PV systems on the proposed approach and figure out the robustness of our approach under these cases.

(3) One single customer may participate in multiple DR programs under different scenarios in the future [22-24]. How the proposed approach will be impacted by the presence of multiple DR programs will also be investigated.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2018.2824842, IEEE Transactions on Smart Grid

TABLE IV RESULTS OF TUKEY TEST ON ALL SR OPTIONS IN TERMS OF MAE

p-value		SR									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
SR	0.5	0	0	0	0	0	0	0	0	0	0
	1.0		0	0.016	0	0	0	0	0	0	0
	1.5			0	0.077	0	0	0	0	0	0
	2.0				0	0.061	0.005	0.002	0	0	0
	2.5					0	1.000	1.000	0.952	0.140	0.106
	3.0						0	1.000	1.000	0.493	0.420
	3.5							0	1.000	0.603	0.530
	4.0								0	1.000	1.000
	4.5									0	1.000
	5.0										0

TABLE V RESULTS OF TUKEY TEST ON ALL SR OPTIONS IN TERMS OF BIAS

p-value		SR									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
SR	0.5	0	0	0	0	0	0	0	0	0	0
	1.0		0	0.509	0.002	0	0	0	0	0	0
	1.5			0	0.856	0.071	0.001	0	0	0	0
	2.0				0	1.000	0.716	0.007	0	0	0
	2.5					0	1.000	0.115	0.058	0.002	0
	3.0						0	0.990	0.958	0.327	0.001
	3.5							0	1.000	1.000	0.254
	4.0								0	1.000	0.217
	4.5									0	0.777
	5.0										0

REFERENCES

[1] F. Wang, H. Xu, T. Xu et al, "The values of market-based demand response on improving power system reliability under extreme circumstances", *Appl Energy*, vol. 193, pp.220-231, 2017.

[2] P. Samadi, V. W. S. Wong, R. Schober, "Load scheduling and power trading in systems with high penetration of renewable energy resources," *IEEE Trans Smart Grid*, vol.7, pp.1802-1812, 2016.

[3] J. Shi, Z. Ding, W. J. Lee et al. "Hybrid Forecasting Model for Very-Short Term Wind Power Forecasting Based on Grey Relational Analysis and Wind Speed Distribution Features," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 521-526, Jan. 2014.

[4] F. Wang, Z. Mi, S. Su et al, "Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters," *Energies*, vol.5, pp.1355-1370, 2012.

[5] Q. Chen, F. Wang, B.M. Hodge et al, "Dynamic price vector formation model based automatic demand response strategy for PV-assisted EV charging station," *IEEE Trans. Smart Grid*, pp.1-1,2017.

[6] Mohagheghi S, Raji N, "Managing industrial energy intelligently: demand response scheme," *IEEE Ind Appl Magazine*, vol.20, pp.53-62, 2014.

[7] H. T. Roh and J. W. Lee, "Residential Demand Response Scheduling With Multiclass Appliances in the Smart Grid," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 94-104, Jan. 2016.

[8] FERC.<https://www.ferc.gov/industries/electric/indus-act/demand-response/dr-potential.asp>.

[9] N. G. Paterakis, O. Erdinc, J. P. S. Catalão, "An overview of demand response: key-elements and international experience," *Renew Sust Energ Rev*, vol.69, pp.871-891, 2017.

[10] S. Nolan, M. O'Malley, "Challenges and barriers to demand response deployment and evaluation," *Appl Energy*, vol.152, pp.1-10, 2015.

[11] S. Mohajeryami, M. Doostan, A. Asadinejad, P. Schwarz. Error Analysis of Customer Baseline Load (CBL) Calculation Methods for Residential Customers. *IEEE Trans. Ind Appl*, vol.53, pp.5-14, 2017.

[12] T.K. Wijaya, M. Vasirani, K. Aberer, "When bias matters: an economic assessment of demand response baselines for residential customers," *IEEE Trans. Smart Grid*, vol.5, pp.*1755*-1763, 2014.

[13] PJM Empirical Analysis of Demand Response Baseline Methods, PJM Load Manage. Task Force, KEMA Inc., Clark Lake, MI, USA, Apr. 2011.

[14] Y. Zhang, W. Chen, R. Xu et al, "A cluster-based method for calculating baselines for residential loads," *IEEE Trans. Smart Grid*, vol.7, pp.2368-2377, 2016.

[15] L. Hatton, P. Charpentier and E. Matzner-Løber, "Statistical estimation of the residential baseline," *IEEE Trans. Power Syst*, vol.31, pp.1752-1759, 2016.

[16] F. Wang, Z. Zhen, Z. Mi, H. Sun, S. Su, and G. Yang, "Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting," *Energy and Buildings*, vol. 86, pp. 427-438, Jan. 2015.

[17] J. Kwac, J. Flora and R. Rajagopal, "Household Energy Consumption Segmentation Using Hourly Data," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 420-430, Jan. 2014.

[18] G. Chicco, R. Napoli and F. Piglione, "Comparisons among clustering techniques for electricity customer classification," *IEEE Trans. Power Syst*, vol.21, no.2, pp.933-940, May. 2006.

[19] J. M. Bates and C. W. J. Granger, "The combination of forecast," *Operational Research Society*, vol.20, pp.451-468, 1969.

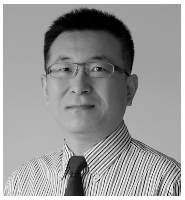
[20] Irish Social Science Data Archive. Data from the Commission for Energy Regulation (CER)-smart metering project. http://www.ucd.ie/issda/data/commissionforenergyregulationcer/; 2012 [accessed 15-04-2016].

[21] https://www.wunderground.com/history/

[22] T. Lu, Z. Wang, J. Wang, Q. Ai and C. Wang, "A Data-Driven Stackelberg Market Strategy for Demand Response-Enabled Distribution Systems," *IEEE Trans. Smart Grid*, available online.

[23] F. Wang, L. Zhou, H. Ren, X. Liu, S. Talari, M. Shafie-khah and J. P. S. Catalão, "Multi-Objective Optimization Model of Source-Load-Storage Synergetic Dispatch for a Building Energy Management System Based on TOU Price Demand Response," *IEEE Trans. Ind. Appl*, vol. 54, no. 2, pp. 1017-1028, Mar/Apr. 2018.

[24] Z. Xu, T. Deng, Z. Hu, Y. Song and J. Wang, "Data-Driven Pricing Strategy for Demand-Side Resource Aggregators," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 57-66, Jan. 2018.



Fei Wang (M'09-SM'17) received the B.S. degree from Hebei University, Baoding, China in 1993, the M.S. and Ph.D. degrees in electrical engineering from North China Electric Power University (NCEPU), Baoding, China, in 2005 and 2013, respectively. Currently, he is a Professor with the Department of Electrical Engineering at NCEPU and the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, Baoding, Beijing, China. He is the Director of Smart Energy Network Integrated Operation Research (SENIOR) Center at NCEPU. He was a visiting Professor with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA, from 2016 to 2017. He was a Post-Doctoral Fellow with the Department of Electrical Engineering at Tsinghua University, Beijing, China, from 2014 to 2016. Since March 2018, he is a Guest Editor for the Special Issue on "Demand Side Management and Market Design for Renewable Energy Support and Integration" of the *IET Renewable Power Generation*. His current research interests include integrated energy system modeling and optimization; electricity price, electricity load and renewable energy power forecasting; demand response and electricity market; smart grid and micro-grid. He was the recipient of the 2014 Natural Sciences Academic Innovation Achievement Award of Hebei Province and the 2014 Outstanding Doctoral Dissertation Award of North China Electric Power University.

Kangping Li (S'15) received the B.S. degree in electrical engineering from North China Electric Power University, Baoding, China, in 2015. Currently, he is pursuing the Ph.D. degree in the Department of Electrical Engineering at North China Electric Power University, Baoding, China. His research interests include demand response, electricity market and power system optimization.

Chun Liu is a Professor at China Electric Power Research Institute (CEPRI), Beijing, China, with an appointment as the Vice Director of Renewable Energy Research Center, CEPRI. His research interests include renewable energy (wind and solar PV) power forecasting, power system dispatch operation and optimization, smart grid and renewable energy integration technology.

Zengqiang Mi received the B.S. and M.S. degree in electrical engineering from North China Electric Power University (NCEPU), Baoding, China, in 1983 and 1986, respectively. He is currently a Professor with the Department of Electrical Engineering at NCEPU and the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, Baoding, Beijing, China. He is the Director of Power System Automation Research Institute at NCEPU. His research interests include power system operation, mechanical elastic energy storage, energy policy and smart grid technology.

Miadreza Shafie-khah (M'13-SM'17) received the M.Sc. and Ph.D. degrees in electrical engineering from Tarbiat Modares University, Tehran, Iran, in 2008 and 2012, respectively. He received his first postdoc from the University of Beira Interior (UBI), Covilha, Portugal in 2015, while working on the 5.2-million-euro FP7 project SiNGULAR ("Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration"). He received his second postdoc from the University of Salerno, Salerno, Italy in 2016. He is currently an Assistant Professor eq. and Senior Researcher at CMAST/UBI, where he has a major role of coordinating a WP in the 2.1-million-euro national project ESGRIDS ("Enhancing Smart GRIDS for Sustainability"), while co-supervising four PhD students and two post-doctoral fellows. He was considered one of the Outstanding Reviewers of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, in 2014 and 2017, one of the Best Reviewers of the IEEE TRANSACTIONS ON SMART GRID, in 2016 and 2017, and one of the Outstanding Reviewers of the IEEE TRANSACTIONS ON POWER SYSTEMS, in 2017. His research interests include power market simulation, market power monitoring, power system optimization, demand response, electric vehicles, price forecasting and smart grids.

João P. S. Catalão (M'04-SM'12) received the M.Sc. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. degree and Habilitation for Full Professor ("Agregação") from the University of Beira Interior (UBI), Covilha, Portugal, in 2007 and 2013, respectively.

Currently, he is a Professor at the Faculty of Engineering of the University of Porto (FEUP), Porto, Portugal, and Researcher at INESC TEC, INESC-ID/IST-UL, and C-MAST/UBI. He was the Primary Coordinator of the EU-funded FP7 project SiNGULAR ("Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration"), a 5.2-million-euro project involving 11 industry partners. He has authored or coauthored more than 625 publications, including 225 journal papers (more than 65 IEEE Transactions/Journal papers), 350 conference proceedings papers, 2 books, 34 book chapters, and 14 technical reports, with an *h*-index of 38, an *i10*-index of 145, and over 5950 citations (according to Google Scholar), having supervised more than 50 post-docs, Ph.D. and M.Sc. students. He is the Editor of the books entitled *Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling* and *Smart and Sustainable Power Systems: Operations, Planning and Economics of Insular Electricity Grids* (Boca Raton, FL, USA: CRC Press, 2012 and 2015, respectively). His research interests include power system operations and planning, hydro and thermal scheduling, wind and price forecasting, distributed renewable generation, demand response and smart grids.

Prof. Catalão is an Editor of the IEEE TRANSACTIONS ON SMART GRID, an Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, and a Subject Editor of the *IET Renewable Power Generation*. From 2011 till 2018 (seven years) he was an Editor of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY and an Associate Editor of the *IET Renewable Power Generation*. He was the Guest Editor-in-Chief for the Special Section on "Real-Time Demand Response" of the IEEE TRANSACTIONS ON SMART GRID, published in December 2012, and the Guest Editor-in-Chief for the Special Section on "Reserve and Flexibility for Handling Variability and Uncertainty of Renewable Generation" of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, published in April 2016. Since May 2017, he is the Corresponding Guest Editor for the Special Section on "Industrial and Commercial Demand Response" of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS. Since March 2018, he is the Lead Guest Editor for the Special Issue on "Demand Side Management and Market Design for Renewable Energy Support and Integration" of the *IET Renewable Power Generation*. He was the recipient of the 2011 Scientific Merit Award UBI-FE/Santander Universities, the 2012 Scientific Award UTL/Santander Totta, the 2016 FEUP Diploma of Scientific Recognition, and the Best INESC-ID Researcher 2017 Award, in addition to an Honorable Mention in the 2017 Scientific Awards ULisboa/Santander Universities. Moreover, he has won 4 Best Paper Awards at IEEE Conferences.