

Identification of charging behavior characteristic for large-scale heterogeneous electric vehicle fleet

Peng LIU¹, Jilai YU¹



Abstract This paper aims to accurately identify parameters of the natural charging behavior characteristic (NCBC) for plug-in electric vehicles (PEVs) without measuring any data regarding charging request information of PEVs. To this end, a data-mining method is first proposed to extract the data of natural aggregated charging load (ACL) from the big data of aggregated residential load. Then, a theoretical model of ACL is derived based on the linear convolution theory. The NCBC-parameters are identified by using the mined ACL data and theoretical ACL model via the derived identification model. The proposed methodology is cost-effective and will not expose the privacy of PEVs as it does not need to install sub-metering systems to gather charging request information of each PEV. It is promising in designing unidirectional smart charging schemes which are attractive to power utilities. Case studies verify the feasibility and effectiveness of the proposed methodology.

Keywords Plug-in electric vehicle, Natural charging behavior characteristic, Data-mining, Aggregated charging load, Theoretical model, Parameter identification, Heterogeneous

CrossCheck date: 11 July 2017

Received: 8 August 2016 / Accepted: 11 July 2017 / Published online: 2 December 2017

© The Author(s) 2017. This article is an open access publication

✉ Peng LIU
liupeng_hit@163.com

Jilai YU
yupwrs@hit.edu.cn

¹ Department of Electrical Engineering, Harbin Institute of Technology, Harbin 150001, China

1 Introduction

Plug-in electric vehicles (PEVs) are becoming increasingly popular due to their potential to enhance energy security as well as to address environmental issues [1, 2]. However, uncoordinated charging behaviors will not only impose negative influences on power grids as they may produce peak load strike but also degrade the eco-friendly advantages of PEVs as most of the uncoordinated charging energy is drained from traditional fossil plants [3, 4]. Therefore, shifting the aggregated charging load (ACL) from peak periods to off-peak periods will greatly mitigate the peak load strike on power systems and facilitate the integration of wind energy [5]. To this end, the smart charging schemes should be deployed to coordinate the charging behaviors of large-scale aggregated PEVs, which is defined as a PEV-fleet in this paper.

Designing the smart charging schemes needs to evaluate the demand response flexibility of the ACL [6], which first requires to identify two parameters for the natural charging behavior characteristic (NCBC) [7]. The first one is the charging probability (CP), i.e., the probability that a PEV is to be charged when it is parked. And the second one is the probability distribution function of the charging duration (CDPDF) when a PEV is charged. Note that the NCBC-parameters signify the statistical characteristics of natural charging behaviors. Thus, the effects of multiple charging factors (e.g., holidays) on natural charging behaviors can be inherently reflected in these two parameters. And NCBC-parameters under different charging factors will differ from each other. Consequently, the NCBC-parameters should stand out in comparison with other charging factors.

To date, the common methods to identify the NCBC-parameters generally fall into two categories, i.e., the



stochastic simulating methods [8–13] and the sub-metering methods [14–17]. For the stochastic simulating methods, travel patterns of internal combustion engine vehicles are used to simulate PEV charging behaviors and then to calculate NCBC-parameters [13]. For the sub-metering methods, a central authority or coordinator is required to gather and process the data regarding the charge request information of each PEV, such as the time when the PEV is plugged in, the duration of charging and the charging rate, etc. [14].

Most of the existing stochastic simulating methods, by and large, suffer from lack of accuracy. In practice, a PEV-fleet consists of heterogeneous vehicles and the charging parameters of them such as battery capacities, charging rates and the amounts of energy consumed per mile, etc. differ significantly [18]. Besides, the charging habits and usages (commuting or non-commuting) are also different among PEVs [4, 18]. Thus, the charging behaviours vary significantly among the vehicles in a PEV-fleet [18]. Most of the existing simulation methods may not accurately simulate charging behaviours of the PEV-fleet.

The sub-metering methods need to build expensive sub-metering systems to gather the data with regard to the charging request information of PEVs [6, 14]. These methods usually aim at designing demand management paradigms that are based on the real time pricing mechanisms. For many power utilities, the main impediment for such paradigms is that their residential customers may not be receptive to the true dynamic pricing schemes [3, 19]. For the power utilities whose residential customers prefer time-of-use price schemes, sub-metering systems are not necessities for the customers and the utilities are also unwilling to improve their measurement systems due to the high cost of investment, operation as well as maintenance [20]. Thus, the sub-metering methods are not suitable for these kinds of power utilities since the charging request information of PEVs are not available.

This paper aims to accurately identify the NCBC-parameters without measuring data regarding the charging request information of PEVs. If PEVs are popularized, the aggregated residential load will consist of three components, i.e., climate-sensitive load (CSL) [21], base load and ACL. The CSL is due to the utilization of air-conditioners and electric heaters during summer and winter, and the base load is owing to the daily usage of basic household appliances such as TVs, computers, water heaters, lighting lamps, washing machines, etc. For a large-scale PEV-fleet, the charging patterns of PEVs tend to be statistically steady and therefore, the ACLs during a specific time window will fall into a narrow band. The statistically steady charging patterns imply the predictability of NCBC-parameters which can be used to calculate ACL, leading to a remarkable regularity and predictability of time-varying

ACL. Our study is originally motivated by two basic ideas: ① The data of natural ACL can be mined from the big data of aggregated residential load of large-scale households. ② NCBC-parameters can be identified via the mined ACL data based on theories of linear convolution and parameter identification.

This paper proposes a novel methodology to identify the NCBC-parameters of the large-scale heterogeneous PEV-fleet. The contributions of this paper are summarized as follows:

- 1) A data mining method is originally proposed. By using the proposed method, the natural ACL data can be mined out from the available big data of the aggregated residential load. The mined ACL data includes the integrated charging characteristic information of a large-scale heterogeneous PEV-fleet. Thus, the proposed data mining method will help understand the integrated features of ACL from the system level's perspective.
- 2) A theoretical ACL model for the large-scale heterogeneous PEV-fleet is originally derived. Compared to existing models [22–25], the derived ACL model is able to calculate the actual ACL in an integrated way by using the actual NCBC-parameters. Thus, it does not need to divide the heterogeneous PEV-fleet into several homogenous sub-PEV-fleet and calculate the ACL for each homogeneous sub-PEV-cluster separately.
- 3) A mathematic model is originally build to identify the NCBC-parameters via the mined ACL data and the derived theoretical model. The identification model is formulated as a non-linear programming model that can be solved by the commonly applied interior point algorithm. The identified NCBC-parameters will help evaluate the demand response flexibility of natural ACL. Thus, the identification model is promising in designing unidirectional smart charging programs that rely heavily on the integrated information of ACL but slightly on individual information of each PEV, such as wind-to-vehicle pricing programs, etc.

The proposed methodology will acquire comparable results with the sub-metering methods, yet does not need to observe the individual information of each PEV. As a result, it does not need to build expensive sub-metering systems and will not expose the privacy of PEVs.

The rest of this paper is organized as follows. In Sect. 2, an overview of the proposed methodology is provided. In Sect. 3, a method of mining the natural ACL data is proposed. Section 4 derives a theoretical model for ACL calculation. Section 5 introduces the identification of NCBC-parameters based on the mined ACL data and the derived ACL model. Section 6 presents the case studies and Sect. 7 concludes this paper.

2 Overview of proposed methodology

The block diagram of the proposed methodology is shown in Fig. 1. The identification of NCBC-parameters consist of two modules. The first one is responsible for mining the big data of aggregated residential load to obtain the natural ACL data, and the second one is responsible for identifying the NCBC-parameters via the mined ACL data based on the derived ACL model. The procedures to identify the NCBC-parameters are summarized as follows:

- 1) Mine the big data of the aggregated residential load to obtain the data of the natural ACL. In this step, the natural ACL data are mined out from the aggregated residential load by using the proposed data-mining method.
- 2) Derive a theoretical ACL model for the large-scale heterogeneous PEV-fleet. In this step, a theoretical ACL model for the PEV-cluster with homogenous charging rate is first derived based on the linear convolution theory. Then, an improved model is proposed to calculate the ACL of the realistic (heterogeneous) PEV-fleet.
- 3) Estimate NCBC-parameters based on the mined ACL data and derived theoretical ACL model. In this step, the NCBC-parameters are identified by using the derived identification model.

These steps will be discussed sequentially. Since most PEV owners prefer to charge their vehicles at home [4], only the charging behaviors occurring at home are studied in this paper. Moreover, each day is divided into 48 time-periods with each time-period lasting 0.5 h. The 48 time-periods are numbered from 1 to 48 with the index of 0:00–0:30 being 1 and 23:30–24:00 being 48.

Note that both the ACL and the NCBC-parameters are continuous variables that vary with time. To reduce the order of identification, we discretize these variables with each time-period (0.5 h). That is to say, the value of ACL in a time-period is actually the average value during that time-period. And the NCBC-parameters in a time-period are counted by using all of the PEVs that are plugged into the grid during that time-period. According to the law of large number, the ACL and NCBC-parameters of large-scale PEV-fleet during a time-period tend to be statistically steady.

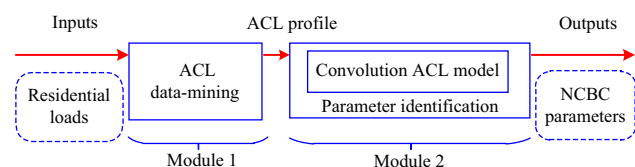


Fig. 1 Procedures for the identification of NCBC-parameter

3 Data-mining method

Travel patterns of a PEV during weeks seldom repeat, yet the integrated trip behaviors of large-scale PEVs tend to follow a repeatable pattern during weeks. Statistical distributions of variables used to describe the trip behaviors such as the time when a PEV arrives home, the daily trip distance, etc., repeat in different weeks [26]. Thus, the ACL and NCBC-parameters of a large-scale PEV-fleet will, by and large, recur in different weeks. Hence, a time window of 168 h (i.e., a week) will include all possible charging behaviors of PEVs’ owners [27]. Hence, the data of weekly ACLs should be used to identify the NCBC-parameters.

To demonstrate the proposed ACL data-mining method, the data of actual aggregated residential load of 6×10^5 households in a city of north China during the recent 8 years (2008–2015) are selected to form the original datasets. And the actual residential load data in 2008–2014 are gathered to form a dataset that is denoted as NRL. It is assumed that the load in NRL do not include ACL since the charging load of PEVs during these years can be neglected. As PEVs are not popularized to date, the data of weekly ACLs are generated by using the actual distributions of trip data via stochastically simulating the natural charging behaviors of 1×10^5 heterogeneous PEVs.

The distributions of trip data are obtained from the Beijing Transportation Research Centre [28]. The PEVs in the simulation include PHEVs, Mini EVs, Compact EVs and Medium EVs. PEVs of each type contain both commuting vehicles and non-commuting vehicles depending on their usages. A PEV-owner’s charging habit is quantified by his/her psychological buffer of the battery which is defined as a specific amount of battery energy [4], below which he/she will suffer the range anxiety. A parked PEV will be charged if the remaining battery energy is less than the sum of its owner’s psychological buffer and the battery energy to cover the next trip [4].

The required trip data and charging parameters are summarized in [18]. Noted that the trip data reflect the actual trip behaviors of the private cars in Beijing. The charging parameters are obtained by survey on the internet and thereby they are relatively accurate. Therefore, the generated ACL is reasonable. These ACL data are then added to the residential load data of 2015 to form another residential load dataset from which the ACL data should be mined out. This dataset is denoted as RL. The constructed dataset RL can be treated as a reasonable dataset that includes ACL. An overview of the stochastic simulation can be found in the Sect. 6. The details of the simulation method can be found in [18].



It should be noted that, the generated load data in RL are not the actual load data and they are only used for the demonstration of the data-mining method in this Section. After the PEVs are popularized, one can just use the load data of a year to form this dataset. In this section, we demonstrate that the ACL component of the residential load in RL can be obtained by using the residential load in NRL.

The weekly ACL curve can be easily obtained by sequentially connecting the daily ACL curves according to their days in the week. As a result, the core of the data-mining method is to extract all the daily ACLs from the residential loads in RL. The residential loads in RL on a regular day are shown in Fig. 2a. The green and red areas denote the base load component and CSL component respectively. And the blue area denotes the ACL component. Noted that, each load component shown in Fig. 2a is a time-sequence that consists of 48 load values in all time-periods of the day. For the residential loads on any given regular day in RL (see Fig. 2a), the procedures to mine the ACL component are shown in Fig. 3. They are summarized as 4 steps.

- 1) Generate two template pools for the per-unit base load profile and the per-unit CSL profile via the dataset NRL.
- 2) Identify per-unit load profiles for the components of base load (green area in Fig. 2a) and CSL (red area in Fig. 2a) by using the two template pools.
- 3) Identify actual load values in a specific time-period for the components of base load and CSL.
- 4) Calculate the base load and CSL components by using the identified per-unit load profiles and the actual load values in a specific time-period, and then subtract these two components from the daily residential loads to obtain the ACL component (blue area in Fig. 2a).

These 4 steps will be discussed in detail from Sects. 3.1 to 3.4.

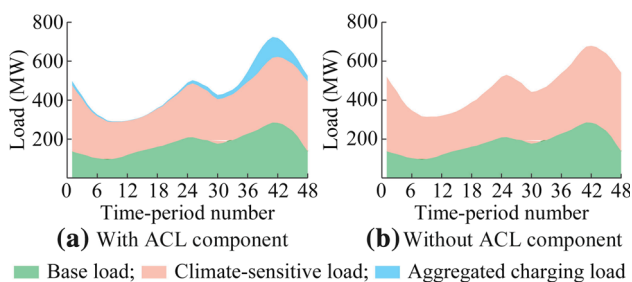


Fig. 2 Residential load curves on two regular days

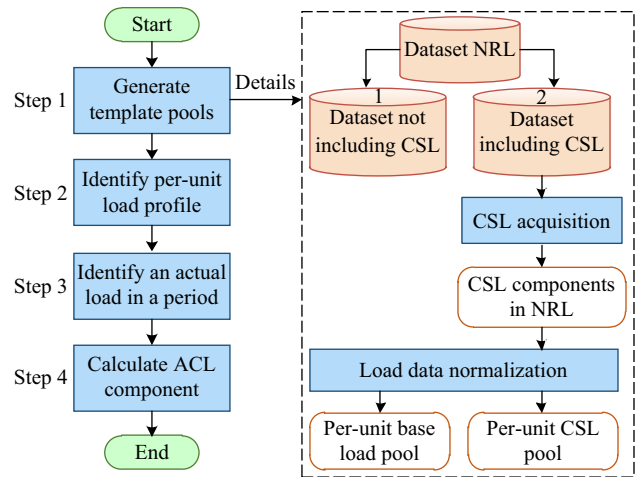


Fig. 3 Work flow of data mining method

3.1 Template pool generation

The generation of template pools for the per-unit base load profile and the per-unit CSL profile requires large amounts of base loads and CSL loads. The CSL heavily relies upon the temperature factors [29–33]. By using the dataset NRL, the linear correlation coefficients between the daily residential peak loads and the temperature factors are calculated. The results are shown in Table 1. In this table, T_{max} is the daily maximum temperature; T_{min} is the daily minimum temperature; T_2 , T_8 , T_{14} and T_{20} are temperatures at 2:00, 8:00, 14:00 and 20:00.

Table 1 shows that, there is no strong correlation between the temperature factors and the daily residential peak load on days in spring and autumn. As a result, residential loads during spring and autumn can usually be treated as base loads which are rarely influenced by temperature factors [29–32]. Table 1 also shows that, the residential peak loads are most correlated with the daily maximum temperature on days in summer and winter. This indicates that the daily maximum temperature has the heaviest influence on the CSL among the temperature

Table 1 Linear correlation coefficients between daily peak load and temperature factors

| Temperature | Linear correlation coefficients | | | |
|-------------|---------------------------------|--------|--------|--------|
| | Spring | Summer | Autumn | Winter |
| T_{max} | 0.21 | 0.92 | 0.13 | -0.91 |
| T_{min} | 0.12 | 0.68 | 0.22 | -0.35 |
| T_2 | 0.16 | 0.78 | 0.16 | -0.34 |
| T_8 | 0.23 | 0.61 | 0.18 | -0.36 |
| T_{14} | 0.18 | 0.81 | 0.21 | -0.69 |
| T_{20} | 0.22 | 0.79 | 0.19 | -0.59 |

factors shown in Table 1. Therefore, the daily maximum temperature is more suitable than the other temperature factors to signify whether there is CSL or not on days in summer and winter [29–33].

For further demonstration, the temperature sensitivity of the daily residential peak load is analyzed by using the dataset NRL. Figure 4 shows the analyzing results. For easy understanding, the daily peak load in a year is normalized to the average peak load at 20 °C within the same year. Due to the operation of air conditioners and electric heaters, the daily peak load starts to increase once the maximum temperature is above 27 °C or below 14 °C. Figure 4 indicates that air-conditioners and electric heaters barely work on days with the maximum temperature between 14 °C and 27 °C. Thus, residential loads in the dataset NRL (for the case in the paper) can be treated as base loads on days with maximum temperatures between 14 °C and 27 °C. And most of the CSLs exist on days with the maximum temperature being above 27 °C or below 14 °C.

For the used dataset NRL in this paper, we only select the residential loads on days with the maximum temperature being in [14, 27 °C] as the actual base loads. This will exclude the effects of CSLs and obtain the relatively accurate base loads. It is worthy to point out that the range of temperature ([14, 27 °C] for the case in the paper) varies among cities [33]. Even though, similar results are also obtained in the studies [29–31]. Once the dataset NRL of a city is obtained, one can easily identify this temperature range.

The template pools can be generated via the dataset NRL. The generation procedures can be found Fig. 3. The dataset NRL is first divided into 2 sub-datasets. The residential load data on days with the maximum temperature being in [14, 27 °C] are gathered to form the 1st sub-dataset and the other load data in NRL are collected to form the 2nd sub-dataset. As indicated in Fig. 4, the load in the 1st sub-dataset can be treated as the base load. And the load in the 2nd sub-dataset is the summation of base load and CSL. Then, the 48 load values of each day in the 1st sub-dataset are normalized to their maximum load value to obtain the corresponding per-unit base load profile. Per-unit load

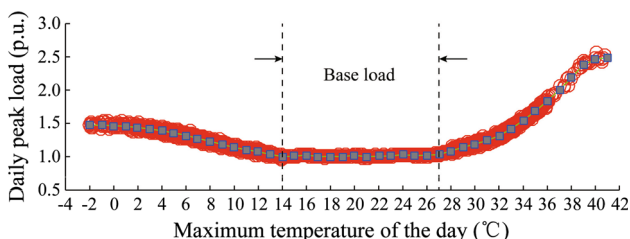


Fig. 4 Influence of maximum temperature of the day on daily residential peak load

profiles of all days are then gathered to form the per-unit base load pool.

To generate the per-unit CSL pool, the CSL components of the daily residential loads in the 2nd sub-dataset are first obtained. The residential loads on a regular day in the 2nd sub-dataset are shown in Fig. 2b. The CSL component shown in Fig. 2b (red area) can be obtained by subtracting the base load component (green area) from the overall residential loads.

Note that the base load is owing to the daily usage of basic household appliances such as TVs, computers, water heaters, washing machines, etc. The number of major basic appliances per 100 urban households is shown in Fig. 5 (see [34]). It can be seen that, the basic household appliances have become saturated for developed cities such that the number of them increases slightly in successive years [35]. Accordingly, the basic load increases slightly in successive years with the increment in one year being little. It is true that some particular factors (e.g., sudden outage, etc.) may results in the variation of base loads. Yet, the base loads for large-scale households on most regular days possess of remarkable and steady regularity at the system-level [36]. That is to say, the base loads on regular days within a year differ little from each other. As a result, the average (typical) base load curve is approximately treated as the actual base load curve in this paper. Figure 6a, b show the average (typical) base load curve of each year which are obtained by using the base loads in 2008–2014 (i.e., the loads in the 1st sub-dataset).

Based on the above analysis, the base load component of the residential loads on a regular day during the n^{th} year in the 2nd sub-dataset (shown in Fig. 2b) can be represented by the average base loads in the same year shown in Fig. 6a or b. Then, the corresponding CSL component can be easily calculated by subtracting the base load component from the residential loads. Once all the CSL components of the residential loads in the 2nd sub-dataset are obtained, one can generate the per-unit CSL profiles by normalizing the 48 load values of each component to their maximum load value. These per-unit load profiles are then collected to form the per-unit CSL pool.

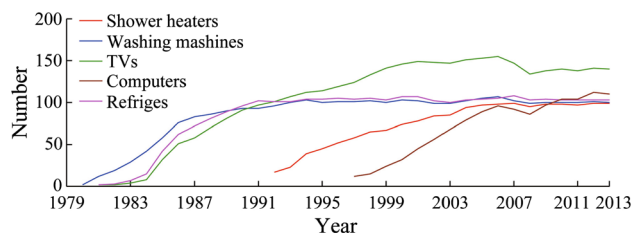


Fig. 5 Number of major basic appliances per 100 urban households



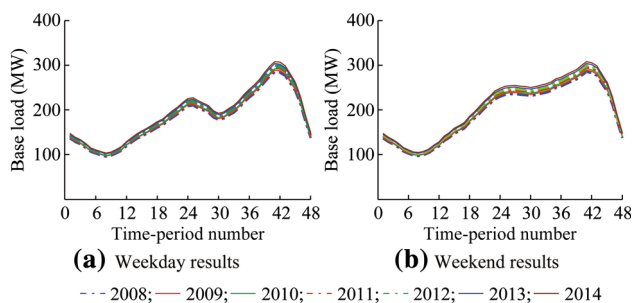


Fig. 6 Typical base load curves in 2008–2014

3.2 Per-unit load profiles identification for the base load and CSL components

The generated template pools can be used to obtain the per-unit load profiles for the base load and CSL components of the residential loads in RL.

Note that the integrated usage patterns of basic appliances for large-scale households in developed cities possess steady regularity with little randomness on regular days [36]. Thus, the per-unit load profile for the base load component on a regular weekday (weekend) in RL can be picked out from the per-unit base load pool. Typical per-unit base load curves, chosen from the base load template pool, are shown in Fig. 7a, b.

The per-unit CSL profile on a summer (winter) day can signify the integrated usage pattern of air-conditioners (electric heaters) that is majorly impacted by the maximum temperature of the day [29–33]. For large-scale households, the integrated usage patterns of air-conditioners (electric heaters) keep relatively steady on regular days with identical maximum temperature [29–31, 33]. Thus, the per-unit CSL load profile for the CSL component on a regular weekday (or weekend) in RL can be picked out from the per-unit CSL pool according to the maximum temperature of that day. Typical per-unit CSL profiles, chosen from the per-unit CSL template pool, under various maximum temperatures during summer are shown in Fig. 8.

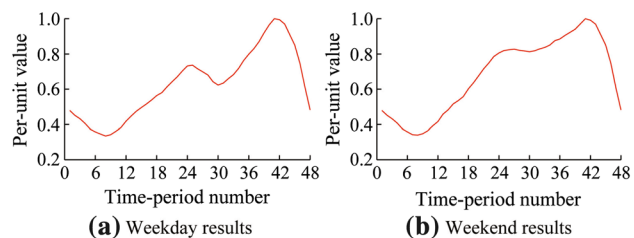


Fig. 7 Typical per-unit base load profiles

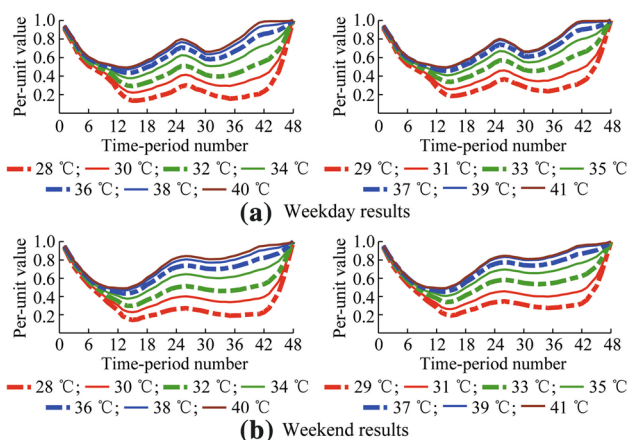


Fig. 8 Typical per-unit CSL profiles during summer

3.3 Actual load values identification in a specific time-period for the base load and CSL components

The per-unit load profile only determines the shape of the load pattern. One of the actual load values in a specific time-period is still required to obtain the complete load component.

If charging behaviors are natural (uncoordinated), it is likely that PEVs will be fully charged during evening and early night [8–13, 18]. Thus, the ratio of ACL to residential load in time-period 13 (6:00–6:30 on the day) is quite close to zero (which can also be verified in Fig. 2a) [15]. Thus, residential loads in RL in time-period 13 only contain base loads (and CSLs). Accordingly, these load values can be used to obtain the actual base load and CSL that are included in the residential load in RL during time-period 13.

Note that residential loads in RL during time-period 13 on regular weekdays (weekends) with their maximum temperatures being in [14, 27 °C] can be treated as base loads as there barely exist CSL and ACL. The average of them can be used as the actual base load that is included in the residential load in RL during time-period 13, as the base loads in the same time-period on regular weekdays (weekends) within a year are almost identical (which is discussed in Sect. 3.1). For the residential load in time-period 13 on a regular day with the maximum temperature being above 27 °C or below 14 °C (see Fig. 2a), it can be treated as the summation of the CSL and base load. Subtract the identified (average) base load from this residential load, then one can obtain the CSL included in the residential load in RL during time-period 13.

3.4 ACL component identification

For the base load and CSL components that are included in the daily residential loads in RL, if the per-unit load profile and the actual load value in time-period i_0 (here, $i_0 = 13$) are denoted as $\{L^{pu}(i)\}(i = 1, 2, \dots, 48)$ and $L(i_0)$ respectively, then the actual load in a given time-period i ($i = 1, 2, \dots, 48$) can be calculated as:

$$L(i) = L^{pu}(i)L(i_0)/L^{pu}(i_0) \tag{1}$$

Once the base load and CSL components of the daily residential loads in RL are identified, one can easily calculate the corresponding ACL component. Note that the residential loads in RL on days with their maximum temperature between 14 and 27 °C only consist of base loads and ACLs. To mine these ACLs, only base load components need to be identified. Once all ACL components are obtained, the weekly ACL curve of each season can be acquired conveniently.

4 Theoretical ACL model

In this section, the theoretical ACL model for a PEV-cluster with identical charging rate is first proposed based on the linear convolution theory. Then, this model is improved to calculate the ACL of realistic PEV-fleet with heterogeneous charging rates.

The instantaneous charging rate of a PEV-cluster can be considered to be constant during the entire charging process [6–18]. Assume there are S types of charging rates in the realistic PEV-fleet, and then the PEV-fleet can be categorized into S PEV-clusters according to their corresponding charging rates. For the cluster s ($s = 1, 2, \dots, S$), its charging rate is denoted as P_{ch}^s .

4.1 Theoretical ACL model for the PEV-cluster s

For the PEVs that are plugged into the grid in a time-period, the charging loads posed by these vehicles can be analogous to the responses that are excited by them. Also, the ACLs of the PEV-cluster s can be analogous to the responses excited by this PEV-cluster. Here, the number of time-periods required to charge batteries from fully discharged to fully charged with the charging rate of P_{ch}^s is denoted as T^s . According to the linear convolution theory, the ACL in time-period j is the linear superposition of charging loads enforced in time-period j by PEVs that are plugged into the grid from time-period $j - T^s + 1$ to j . If the number of PEVs plugged into the grid during time-period i is denoted as N_i^s , and the charging load of these PEVs in time-period j is denoted as $P_{i,j}^s$, then the ACL of the PEV-cluster s in time-period j can be expressed as:

$$P_j^s = \sum_{i=j-T^s+1}^j P_{i,j}^s \tag{2}$$

Calculating P_j^s with the NCBC-parameters (i.e. CPs and CDPDFs) of PEV-cluster s needs a model to describe the quantitative relationship between the charging load $P_{i,j}^s$ and the NCBC-parameters in time-period i . This in turn needs to analyze the evolution of charging process for the N_i^s PEVs that are plugged into the grid in time-period i . Noted that, these PEVs will be gradually disconnected from the grid during the following time-periods. Accordingly, the ACL of these PEVs will gradually decrease in a specific pattern from the initial pulse (the ACL excited by these PEVs in time-period i). The decreasing pattern depends on the CDPDF of these N_i^s PEVs.

The CDPDF for the N_i^s PEVs that are plugged into the grid during time-period i is denoted as $H_i^s = \{h_{i,t}^s\}(t = 1, 2, \dots, T^s)$. $h_{i,t}^s$ is the probability that a PEV will be charged for t time-periods. Obviously, $\sum_{t=1}^{T^s} h_{i,t}^s = 1$. As $r - 1$ time-periods elapsed, the number of

PEVs that have left the grid is $N_i^s \sum_{t=1}^{r-1} h_{i,t}^s$. Hence, the number of PEVs still being connected with the grid in time-period $i + r - 1$ is $N_i^s \left(1 - \sum_{t=1}^{r-1} h_{i,t}^s\right)$. Multiplying the number of remaining PEVs in each following time-period by P_{ch}^s , one can obtain the ACL-time-sequence of these N_i^s PEVs.

Table 2 summarizes the charging process of the N_i^s PEVs that are plugged into the grid in time-period i . The ACL gradually decreases from the initial value $P_{ch}^s N_i^s$ and after $r - 1$ time-periods, it decreases to a value of

$P_{ch}^s N_i^s \mu_{i,r}^s$. Noted that, $\mu_{i,r}^s = 1 - \sum_{t=1}^{r-1} h_{i,t}^s$ is the decreasing coefficient, which reflects the decreasing pattern of the ACL excited by these N_i^s PEVs. Hence, the ACL that is posed by the N_i^s PEVs in time-period j is:

$$P_{i,j}^s = P_{ch}^s N_i^s \mu_{i,j-i+1}^s \quad (j \geq i) \tag{3}$$

For the PEV cluster s , the probability that a vehicle arrives home in time-period i and does not leave again on that day is denoted as g_i^s , and the CP in time-period i (i.e., the probability that a vehicle is to be charged in time-period i when it is parked) is denoted as f_i^s . Let N^s denote the number of PEVs in cluster s , then the number of PEVs that are plugged into the grid in time-period i can be written as:

$$N_i^s = N^s g_i^s f_i^s \tag{4}$$

Substituting (4) into (3) yields:



Table 2 ACL sequence imposed by PEVs plugged to the grid in time-period i

| Present time-period | Elapsed time-period | Decreasing coefficient | Remaining PEV number | ACL |
|---------------------|---------------------|------------------------------------|-----------------------|--------------------------------|
| i | 0 | 1 | $N_i^s \mu_{i,1}^s$ | $P_{ch}^s N_i^s \mu_{i,1}^s$ |
| $i + 1$ | 1 | $1 - h_{i,1}^s$ | $N_i^s \mu_{i,2}^s$ | $P_{ch}^s N_i^s \mu_{i,2}^s$ |
| \vdots | \vdots | \vdots | \vdots | \vdots |
| $i + r - 1$ | $r - 1$ | $1 - \sum_{t=1}^{r-1} h_{i,t}^s$ | $N_i^s \mu_{i,r}^s$ | $P_{ch}^s N_i^s \mu_{i,r}^s$ |
| \vdots | \vdots | \vdots | \vdots | \vdots |
| $i + T^s - 1$ | $T^s - 1$ | $1 - \sum_{t=1}^{T^s-1} h_{i,t}^s$ | $N_i^s \mu_{i,T^s}^s$ | $P_{ch}^s N_i^s \mu_{i,T^s}^s$ |
| $i + T^s$ | T^s | 0 | 0 | 0 |

$$P_{ij}^s = P_{ch}^s N^s g_i^s f_i^s \mu_{i,j-i+1}^s \quad (j \geq i) \tag{5}$$

Then, substitute (5) into (2), one will obtain:

$$P_j^s = P_{ch}^s N^s \sum_{i=j-T^s+1}^j g_i^s f_i^s \mu_{i,j-i+1}^s \tag{6}$$

Equation (6) is the theoretical model that can only calculate the ACL of a PEV-cluster with identical charging rate. It is noteworthy that PEVs will arrive home with different initial values of state of charge (SOC). Therefore, the ACL is actually produced by PEVs with different initial values of SOC. PEVs with different initial values of SOC require different charging times to fully charge their batteries. Hence, the parameter CDPDF can be used to deal with the effect of initial SOC on the calculation of ACL. Considering that the parameter CDPDF reflects the distribution of initial SOC, the derived theoretical ACL model inherently concerns the issues of initial SOC.

4.2 Theoretical ACL model for realistic PEV-fleet

In this part, the NCBC-parameters of each PEV-cluster with homogenous charging rate are combined together to obtain the equivalent NCBC-parameters of the realistic PEV-fleet. The equivalent NCBC-parameters are thereafter used to derive the theoretical model of the realistic ACL.

Since the realistic PEV-fleet contains S PEV-clusters and the numbers of PEVs for the cluster s is N^s , the total number of PEVs in the realistic fleet can be calculated as

$$N = \sum_{s=1}^S N^s. \text{ For the realistic PEV-fleet, let } P_{ch}^{\min} \text{ denote the}$$

minimum charging rate within the realistic PEV-fleet. Besides, the probability that a PEV arrives home in time-period i and does not leave again on that day is denoted as g_i . Then, the following equivalence principle is used to obtain the equivalent NCBC-parameters for the realistic PEV-fleet.

Equivalence principle: charging one PEV with the charging rate P_{ch}^s in some time-periods is equivalent to charging P_{ch}^s/P_{ch}^{\min} PEVs with the charging rate P_{ch}^{\min} in the same time-periods from the system-level's perspective.

For the realistic PEV-fleet, the equivalent number of PEVs plugged into the grid in time-period i is denoted as N_i . According to the Equivalence principle, N_i can be calculated as:

$$N_i = \sum_{s=1}^S (P_{ch}^s/P_{ch}^{\min}) N_i^s \tag{7}$$

Substituting (4) into (7) yields:

$$N_i = \sum_{s=1}^S (P_{ch}^s/P_{ch}^{\min}) N^s g_i^s f_i^s \tag{8}$$

The distribution of time when a PEV arrives home relies on the living and working habits of PEV owners. These habits tend to be independent of PEVs' charging rate P_{ch}^s . Thus, we have $g_i = g_i^s$. Since $g_i = g_i^s$, the equivalent CP of the realistic PEV-fleet in time-period i , denoted as f_i , can be calculated as:

$$f_i = N_i/(N g_i) = \sum_{s=1}^S (P_{ch}^s/P_{ch}^{\min}) N^s f_i^s / N \tag{9}$$

According to the Equivalence principle, the equivalent probability that a PEV will be charged for t time-periods can be calculated as:

$$h_{i,t} = \sum_{s=1}^S h_{i,t}^s N_i^s (P_{ch}^s/P_{ch}^{\min}) / N_i \tag{10}$$

Since $N_i = N g_i f_i$ and $g_i = g_i^s$, substituting (4) into (10) yields:

$$h_{i,t} = \sum_{s=1}^S h_{i,t}^s N^s f_i^s (P_{ch}^s / P_{ch}^{min}) / (N f_i) \tag{11}$$

By using (9) and (11), equivalent NCBC-parameters of the realistic PEV-fleet can be obtained. The ACLs of the realistic PEV-fleet can be calculated by these equivalent NCBC-parameters in an integrated way. The ACL of the realistic PEV-fleet in time-period j , P_j , can be calculated as:

$$P_j = P_{ch}^{min} N \sum_{i=j-T+1}^j g_i f_i \mu_{i-j+1} \tag{12}$$

where $\mu_{i-j+1} = 1 - \sum_{k=1}^{j-i} h_{i-k}$ is the time-decreasing coefficient of ACL and T is the maximum number of time-periods that is required to charge the PEVs in the realistic PEV-fleet. $T = \max\{T^s\}$ ($s = 1, 2, \dots, S$). The proof of (12) can be found in the ‘‘Appendix A’’.

Equation (12) is the theoretical ACL model for the realistic PEV-fleet. In (12), P_{ch}^{min} and g_i can be easily obtained. Note that PEVs should be licensed in the traffic departments with a unique license number. Thus, it will be easy to get the information of PEVs’ total number (N) from the traffic departments. In addition, PEVs tend to be registered in the power utilities for the acquisition of charging services such as the installations and maintenances of charging devices, the inquiries of charging tariffs as well as the payments of charging bills, etc. As a result, the information of PEVs’ total number (N) can also be obtained from the power utilities. As a result, only the CPs and CDPDFs in (12) need to be identified.

5 NCBC-parameters identification

A non-linear programming problem is formulated to identify NCBC-parameters. For easy demonstration, the weekly ACL data used to construct the dataset RL and the distribution of time when vehicles arrive home in week are shown in Fig. 9 [18, 28]. Note that one week is divided into 336 time-periods with each time-period lasting 0.5 h. These time-periods are further grouped into 7 time-sections that are indexed from 1 to 7, wherein each time-section includes 48 time-periods. The d^{th} ($d = 1, 2, \dots, 7$) time-section begins from the 13th time-period (6:00–6:30) of the d^{th} day and ends at the 12th time-period (5:30–6:00) of the $(d+1)^{th}$ day. Notice that Monday is the 1st day in week.

For the d^{th} time-section, PEVs arrive home from the 13th time-period of the d^{th} day to the 6th time-period (2:30–3:00) of the $(d + 1)^{th}$ day [28]. And most of these PEVs will be fully charged in time-section d . Thus, the ACLs of these vehicles usually have large values in time-section d whereas they are extremely small at the beginning

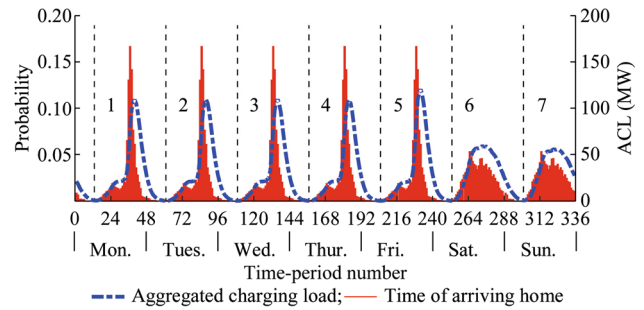


Fig. 9 ACL and distribution of time when vehicles arrive home in the week (two consecutive time-sections are separated with black dot lines)

of time-section $(d + 1)$ [13, 18]. Accordingly, ACLs in the d^{th} time-section can be used to identify NCBC-parameters corresponding to this time-section. Thus, weekly NCBC-parameters can be divided into 7 groups with the parameters of each group being identified separately via ACLs in the corresponding time-section. For easy demonstration, the 48 time-periods of each time-section are renumbered from 13 to 60 with the index of 6:00–6:30 being 13 and 5:30–6:00 of the next day being 60. For each time-section, there is no need to identify the NCBC-parameters in time-periods with the index ranging from 55 to 60 (3:00–6:00), as there is usually few PEVs arriving home from 3:00 to 6:00 of each day [28].

Let $\{P_j^m\}$ ($j = 13 \dots 60$) denote the ACL data mined in a given time-section. And the period within this time-section when the power system experiences peak ACL is denoted as $[j^{pb}, j^{pc}]$. Then, the problem to identify the NCBC-parameters in this time-section can be formulated as a non-linear programming problem as follows:

$$\min Q = \sum_{j=j^{pb}}^{j^{pc}} (1 - P_j / P_j^m)^2 \tag{13}$$

s.t.

$$\begin{cases} f_i > 0 \\ h_{i,t} \geq 0 \end{cases}, \forall i \in \{13 \dots 54\}, \forall t \in \{1 \dots T\} \tag{14}$$

$$|f_i - f_{i+1}| \leq \delta_i, \forall i \in \{13 \dots 53\} \tag{15}$$

$$h_{i,t} = h_{i+1,t}, \forall i \in \{13 \dots 53\} \tag{16}$$

$$\sum_{t=1}^T h_{i,t} = 1, \forall i \in \{13 \dots 54\} \tag{17}$$

$$|P_j / P_j^m - 1| \leq \xi_j, \forall j \in \{13 \dots 60\} \tag{18}$$

where P_j is calculated via (12) by using the identified NCBC-parameters, and $(1 - P_j / P_j^m)^2$ is the quadratic function of the relative difference between the mined ACL and the calculated ACL. The objective function is to minimize the sum of quadratic functions of relative



differences during the peak ACL period in the given time-section. In (15), δ_i is the threshold of difference of CPs between the i^{th} and $(i+1)^{\text{th}}$ time-periods. In (18), ξ_j is the threshold of the relative difference between P_j and P_j^m . Details about setting δ_i and ξ_j will be discussed later.

The constraint described in (14) ensures that the values of NCBC-parameters are greater than zero. Equation (16) indicates that the CDPDFs in all time-periods within the same time-section can be considered to be identical, which is verified through our case studies. Equation (17) guarantees the sum of all elements of CDPDF equals to one. The non-linear programming problem described above can be solved via the widely used interior point algorithm [37]. Note that the interior point algorithm has the polynomial computation complexity when it is used to solve the non-linear programming problems [38].

As mentioned previously, the NCBC-parameters can be used to evaluate the demand response flexibility of ACL [6, 7]. For most of PEVs, the times they parked at home are much longer than the times to fully charge their batteries and thus the ACL can be shifted to the off-peak periods. In this paper, the demand response flexibility of ACL is defined as the maximum amount of time that the ACL can be shifted to later hours. By using the identified NCBC-parameters, the number of PEVs that are plugged into the grid in the i^{th} time-period can be calculated as $N_i = Ng_i f_i$. The number of PEVs that need to be charged for t time-periods can be expressed as $N_t^i = N_i h_{i,t} = Ng_i f_i h_{i,t}$. And the ACL posed by these PEVs is $P_{\text{ch}}^{\min} N_t^i = P_{\text{ch}}^{\min} Ng_i f_i h_{i,t}$. Let time-period k denote the charging deadline before which the PEV have to be fully charged. Then, the charging of the N_t^i PEVs can be at most delayed for $k-i-t$ time-periods. In other words, the demand response flexibility for the ACL that is posed by the N_t^i PEVs is $k-i-t$.

6 Case studies

In this section, the methods of mining ACL data and identifying NCBC-parameters are tested. As PEVs are not popularized to date, this paper generates the ACLs and NCBC-parameters by stochastically simulating charging behaviors of 1×10^5 PEVs [18]. They are treated as the sub-metering results (benchmarks) to verify the feasibility of the proposed methodology. We would like to point out that, the generated ACL and NCBC-parameters are not the actual ACL and NCBC-parameters. They are only used to verify the feasibility of the proposed data-mining method and the parameter-identification model. After the PEVs are popularized, one can just mine the data of actual ACL from the data of residential loads and identify the actual NCBC-parameters by using the actual ACL data.

PEVs are first divided into several subgroups according to the vehicle types, usages and charging habits. Then, the stochastic simulation is performed on each subgroup in consecutive 13 weeks [18]. The average weekly ACLs of each subgroup are summed up. And during each time-period in week, NCBC-parameters of each subgroup are combined together via the Equivalence Principle described in Sect. 4. The details of the simulation method can be found in [18].

The benchmark of the weekly ACLs are shown in Fig. 9 (blue dot line). The CDPDFs during all time-periods in week are shown in Fig. 10 wherein the CDPDFs in the same time-section are displayed with the same color. Notice that the CDPDFs within the same time-section exhibit some differences, especially for those at the beginnings and endings of the time-section when there are few PEVs arriving home. However, treating the CDPDFs in the same time-section to be identical (16) will bring little error in the calculation of ACLs, which will be verified in Sect. 6.2.

6.1 Tests for data-mining method

The ACL data included in the residential loads of RL are extracted to verify the feasibility of the proposed data-mining method. The ratios of ACL to the peak load on weekdays and weekends are 14% and 7%, respectively.

Let $\{P_j^b\}$ ($j = 13 \dots 60$) denote the benchmark ACL data in a given time-section. Then, the relative error (absolute value) between the benchmark ACL data and mined ACL data in time-period j can be calculated as $\left| \frac{P_j^m}{P_j^b} - 1 \right|$. Figure 11a, b show the benchmark ACL data and the relative errors of mined ACL data during the 1st and 7th time-sections in week. Note that Fig. 11a, b do not display the relative errors that are greater than 0.16 at the beginnings and endings of the time-section.

It can be seen that, the relative errors of mined ACL data are generally greater than 0.08 during [6:00, 9:30] of the first day and [3:30, 6:00] of the following day. However, it is worthy to point out that the relative errors are quite low (mostly less than 0.04) during the periods when the ACLs are notably high. This indicates that the ACL data-mining

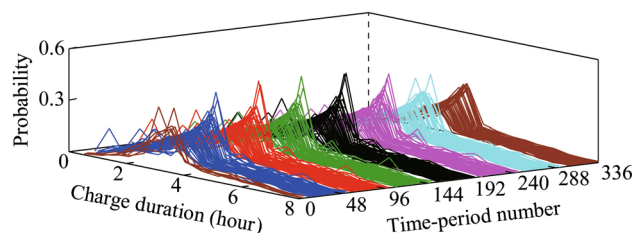


Fig. 10 CDPDF of each time-period in the week

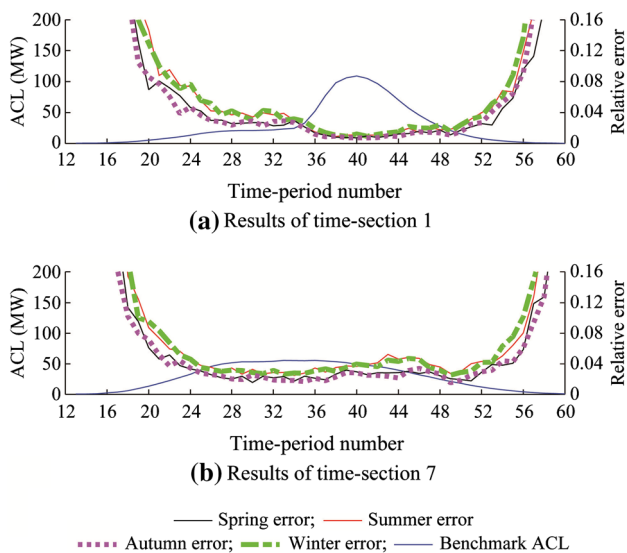


Fig. 11 Benchmark ACL and relative error of mined ACL in time-section 1 and time-section 7

method can obtain accurate data of ACLs in the period when there are notable amount of PEVs arriving home. Similar results can be obtained by analyzing the relative errors in other time-sections. Since the underlying NCBC-information mainly depends on the ACLs that are notably high, the minded ACL data can be used to identify the corresponding NCBC-parameters.

It should be noted that, there are some related works on getting the accurate data of ACLs (see [15, 16, 39]). In these works, the data of charging load for each vehicle are submitted to a central authority by using the expensive sub-metering system. The central authority then aggregates charging load of each vehicle to form the ACL. Compared with these methods, the proposed method will obtain comparable results of ACL yet does not need to install the expensive sub-metering systems. Unlike the studies aiming to get the accurate ACL, the proposed method targets to identify the NCBC-parameters which can be used to evaluate the demand response flexibility of the ACL. As a result, the proposed methodology will be more suitable to help design the pricing incentive programs for charging coordination.

6.2 Tests for identification model

In this part, the performances of the identification model is tested. First, the feasibility of constraint (16) is verified. Then, the NCBC-parameters are identified via the ACL data that are mined from the load data of each season in RL.

To verify the feasibility of (16), the 48 CDPDFs within each time-section are first combined to obtain one CDPDF.

And then these 7 combined CDPDFs along with the CPs of all time-periods in the week are substituted into (12) to calculate the weekly ACLs which are denoted as $\{P_j^{com}\}$ ($j = 1, 2, \dots, 336$). The relative errors (absolute values) between the benchmark ACL and $\{P_j^{com}\}$ are shown in Fig. 12. It can be seen that, the relative errors are less than 0.05 in most of the time except at the endings of each time-section. Since the ACLs at endings of each time-section are quite low, the relatively large errors barely affect the identification of NCBC-parameters.

It should be noted that, commuting PEVs usually arrive home simultaneously during [16:00, 22:00] in the first 5 time-sections [28]. For PEVs arriving home in these periods, the ratio of commuting ones to non-commuting ones in each time-period varies considerably. As trip distances of commuting and non-commuting PEVs are somewhat different, CPs in each time-period of [16:00, 22:00] vary significantly. Thus, δ_i s with $i \in [33, 44]$ are set to 0.06. Besides, PEVs arriving home during [6:00, 16:00] \cup [22:00, 3:00] mainly consist of non-commuting vehicles. Considering that the distributions of trip distances vary smoothly for non-commuting PEVs that arrive home in adjacent time-periods, the differences of corresponding CPs of them are small. Here, δ_i s with $i \in [13, 32] \cup [45, 54]$ are set to 0.03. For the last two time-sections, δ_i s are set to 0.03 as PEVs arriving home during time-periods in these two time-sections are usually non-commuting vehicles.

The underlying NCBC-information corresponding to the mined ACL data during [6:00, 9:30] of one day and [3:30, 6:00] of the following day are inaccurate due to the large errors in the mined ACL data. As a result, ξ_j s with $j \in [13, 19] \cup [56, 60]$ are set to ∞ . Besides, the errors in mined ACL data during [9:30, 12:00] of one day and [0:00, 3:30] of the following day are somewhat large due to the relatively low ratios of ACL to the residential load. Thus, ξ_j s with $j \in [20, 24] \cup [49, 55]$ are set to 0.07. Here, ξ_j s in other time-periods of each time-section are set to 0.03. Initial values for NCBC-parameters are set to 0.5. The total PEV number (N) and the minimum charging rate of PEVs in the realistic PEV-fleet (P_{ch}^{min}) are equal to 1×10^5 and 3.5 kW

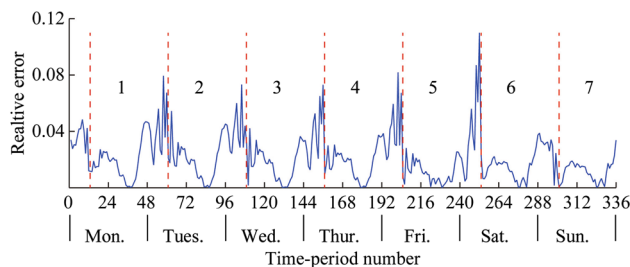


Fig. 12 Relative error between P_j^b and P_j^{com} in the week (two consecutive time-sections are separated with red dot lines)



respectively. The maximum number of time-periods that is needed to charge the PEVs in the realistic PEV-fleet (T) is 16 (8 h). The time distribution when PEVs arrive home in each time-section can be found in Fig. 9. According to the mined ACL data, the peak ACL period [j^{pb} , j^{pe}] is set as [36, 45] in the first 5 time-sections and [25, 44] in the last two time-sections.

Figure 13a, b show the CDPDFs that are identified via the mined ACL data during the 1st and 7th time-sections. Figure 14a shows the CPs of all time-periods in the 1st time-section and Fig. 14b displays those in the 7th time-section. In Figs. 13 and 14, the blue lines represent the benchmark data with other lines being marked with legends. Figures 13 and 14 show that the identified NCBC-parameters are consistent with the benchmark. As the benchmark ACL data used to construct the dataset RL is the same among seasons, NCBC-parameters identified by using the mined ACL data of each season are almost identical. One can obtain same conclusions by analyzing the corresponding results in other time-sections.

To show the identified results more clearly, the values of expectation and standard deviation of charge durations in each time-section are shown in Fig. 15a, b. And the RMS (root mean square) values of relative errors between the

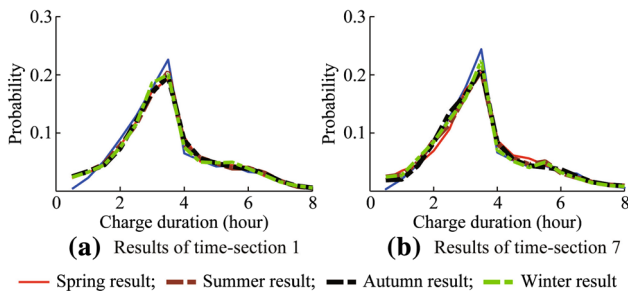


Fig. 13 CDPDFs in time-section 1 and time-section 7

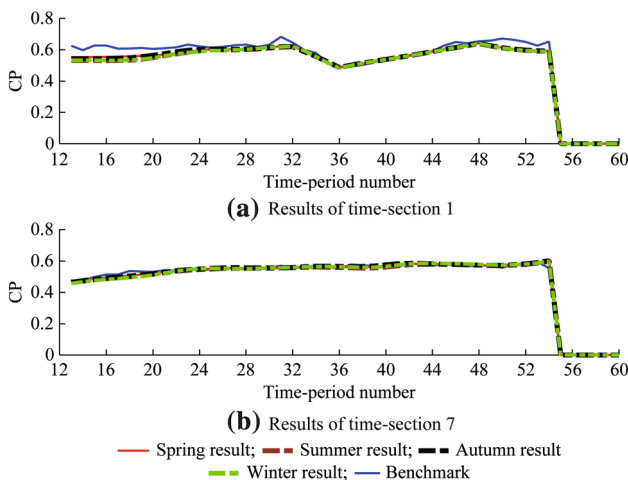


Fig. 14 CPs in time-section 1 and time-section 7

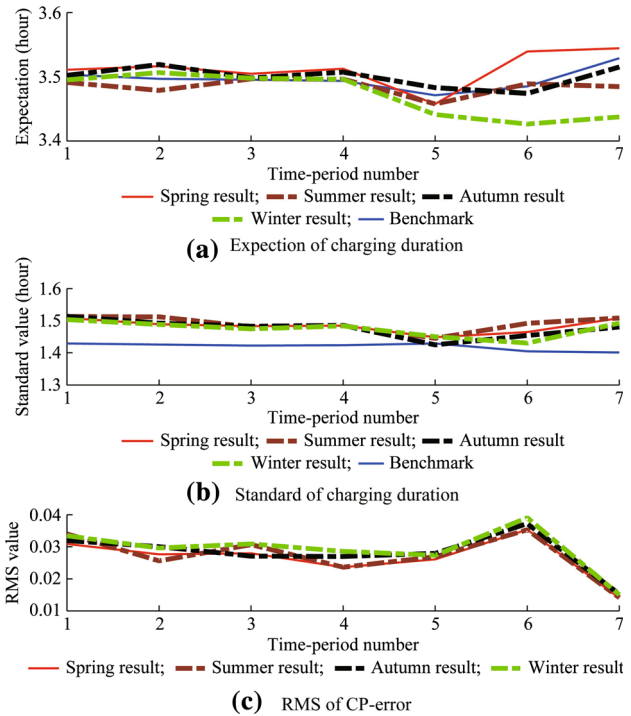


Fig. 15 Statistical results of the identified NCBC-parameters

benchmark CPs and the identified CPs during [12:00, 24:00] of each time-section are shown in Fig. 15c.

Figure 15a shows that the mean values of charge durations calculated by using the identified CDPDFs are quite close to those obtained by using the benchmark CDPDFs. The maximum difference is only 0.09 h which occurs in the 7th time-section during winter when the identification value and the benchmark value are 3.44h and 3.53h respectively. Figure 15b shows that the standard deviations of charge durations for the identification values and the benchmark values are quite close. The maximum difference is only 0.1 h which happens in the 7th time-section during summer when the identification value and the benchmark value are 1.51h and 1.41h respectively. Fig. 15c shows that the RMS values for the relative errors between the benchmark CPs and the identified CPs in each time-section are less than 0.04. It indicates that the values of identified CPs are almost equal to those of benchmark CPs during [12:00, 24:00] of each time-section in the week. The results shown in Figs. 13, 14 and 15 verify that the identification model is feasible.

7 Conclusion and future work

This paper proposed a methodology to identify the NCBC-parameters for large-scale heterogeneous PEV-fleets.

The proposed methodology is well suitable for the power utilities to understand the integrated features of ACL from the system level’s perspective. Therefore, it will help the power utilities to solve some problems that rely heavily on the integrated information of ACL but slightly on the PEV’s individual information. For example, the evaluation of demand response flexibility, the design of pricing incentive programs, etc. Case studies show that the proposed methodology will obtain comparable results with the sub-metering methods. However, it does not need to install the expensive sub-metering systems to gather and process the PEV’s individual information. As a result, the proposed methodology significantly reduces the cost of investment, operation as well as maintenance and will not pose the privacy of PEVs.

We would like to point out that, the extraction of ACL data requires huge amounts of historical residential loads. Due to lack of historical load data, the data of ACL on statutory holidays (e.g., the Spring Festivals, the National Day, etc.) cannot yet be accurately mined out. Thus, the proposed methodology does not work well for the scenarios on statutory holidays. It is desirable to improve the methodology to identify the NCBC-parameters on statutory holidays in the future.

This paper only studies the identification of the natural charging behaviors. It is also desirable to analyze the scenarios where PEVs have pricing incentives in off-peak periods. In future work, the proposed methodology will be extended to mine the data of coordinated ACL and identify the parameters of coordinated charging behaviors. It will facilitate some studies that relate with the identification of PEVs’ actual response behaviors. For example, the variation of ACL at dimensions of both time and space (region) can be easily analyzed by comparing the natural ACL with the coordinated ACL. And the effect of pricing incentives on the charging behaviors can be conveniently obtained by analyzing the parameters of coordinated charging behaviors. On one hand, these studies will help the power utilities improve the pricing incentive programs in the power markets. On the other hand, they will also help the power utilities use the direct load control techniques to shave the ACLs that remain in the peak periods under emergency operating scenarios.

Acknowledgements This work was jointly supported by the NSFC-RCUK_EPSRC under Grant 51361130153 and the National Natural Science Foundation of China under Grant 51377035. The authors would also like to thank the Jinan Power Supply Company for providing us the residential load data.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted

use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A

Proof of (12)

Let

$$A = P_{ch}^{min} N \sum_{i=j-T+1}^j g_i f_i \mu_{i,j-i+1} \tag{A1}$$

According to the definition of P_j , one can obtain:

$$P_j = \sum_{s=1}^S P_j^s \tag{A2}$$

Hence, (12) is equivalent to:

$$A = \sum_{s=1}^S P_j^s \tag{A3}$$

Substituting $N_i = Ng_i f_i$ and $\mu_{i,j-i+1} = 1 - \sum_{t=1}^{j-i} h_{i,t}$ into (A1) yields:

$$A = P_{ch}^{min} \sum_{i=j-T+1}^j N_i \left(1 - \sum_{t=1}^{j-i} h_{i,t} \right) \tag{A4}$$

Substituting (10) into (A4) yields:

$$\begin{aligned} A &= P_{ch}^{min} \sum_{i=j-T+1}^j N_i \left(1 - \sum_{t=1}^{j-i} \sum_{s=1}^S \frac{N_i^s P_{ch}^s}{N_i P_{ch}^{min}} h_{i,t}^s \right) \\ &= P_{ch}^{min} \sum_{i=j-T+1}^j N_i \left(1 - \sum_{s=1}^S \frac{N_i^s P_{ch}^s}{N_i P_{ch}^{min}} \sum_{t=1}^{j-i} h_{i,t}^s \right) \\ &= P_{ch}^{min} \sum_{i=j-T+1}^j \left(N_i - \sum_{s=1}^S N_i^s \frac{P_{ch}^s}{P_{ch}^{min}} \sum_{t=1}^{j-i} h_{i,t}^s \right) \end{aligned} \tag{A5}$$

Substitute (7) into (A5), we will obtain:

$$\begin{aligned} A &= P_{ch}^{min} \sum_{i=j-T+1}^j \sum_{s=1}^S \frac{P_{ch}^s}{P_{ch}^{min}} N_i^s \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \\ &= \sum_{i=j-T+1}^j \sum_{s=1}^S P_{ch}^s N_i^s \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \end{aligned} \tag{A6}$$

Substitute (4) into (A6), one can get:



$$\begin{aligned}
 A &= \sum_{i=j-T+1}^j \sum_{s=1}^S P_{\text{ch}}^s N^s g_{ij}^{sf} \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \\
 &= \sum_{s=1}^S P_{\text{ch}}^s N^s \sum_{i=j-T+1}^j g_{ij}^{sf} \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right)
 \end{aligned} \tag{A7}$$

$\forall s = 1, 2, \dots, S$, we denote:

$$B^s = P_{\text{ch}}^s N^s \sum_{i=j-T+1}^j g_{ij}^{sf} \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \tag{A8}$$

$\forall s = 1, 2, \dots, S$, considering that $T = \max\{T^s\}$, then $T^s \leq T$. For the PEV-clusters that satisfy $T^s = T$, substituting (6) into (A8) yields:

$$B^s = P_j^s \forall s \in S_{T^s=T} \tag{A9}$$

where $S_{T^s=T}$ is the set of indexes for the PEV-clusters that satisfy $T^s = T$. For the PEV-clusters that satisfy $T^s \leq T - 1$, deriving (A8) yields:

$$\begin{aligned}
 B^s &= P_{\text{ch}}^s N^s \sum_{i=j-T+1}^{j-T^s} g_{ij}^{sf} \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \\
 &\quad + P_{\text{ch}}^s N^s \sum_{i=j-T^s+1}^j g_{ij}^{sf} \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \forall s \in S_{T^s \leq T-1}
 \end{aligned} \tag{A10}$$

where $S_{T^s \leq T-1}$ is the set of indexes for the PEV-clusters that satisfy $T^s \leq T - 1$. $\forall s \in S_{T^s \leq T-1}$, we denote:

$$B_1^s = P_{\text{ch}}^s N^s \sum_{i=j-T+1}^{j-T^s} g_{ij}^{sf} \left(1 - \sum_{t=1}^{j-i} h_{i,t}^s \right) \tag{A11}$$

Considering that $i \leq j - T^s$, we obtain $T^s \leq j - i$. Thereby, $\forall s \in S_{T^s \leq T-1}$, B_1^s can be written as:

$$B_1^s = P_{\text{ch}}^s N^s \sum_{i=j-T+1}^{j-T^s} g_{ij}^{sf} \left(1 - \sum_{t=1}^{T^s} h_{i,t}^s - \sum_{t=T^s+1}^{j-i} h_{i,t}^s \right) \tag{A12}$$

Remember that $h_{i,t}^s$ is the probability that a PEV in the PEV-cluster s will be charged for t time-periods ($t = 1, 2, \dots, T^s$), then one can obtain:

$$h_{i,t}^s = 0 \forall t = T^s + 1 \dots j - i \tag{A13}$$

Considering that $\sum_{t=1}^{T^s} h_{i,t}^s = 1$ and (A13), we can derive (A12) as:

$$B_1^s = 0 \forall s \in S_{T^s \leq T-1} \tag{A14}$$

Substituting (6) and (A11) into (A10) and considering (A14) yield:

$$B^s = P_j^s \forall s \in S_{T^s \leq T-1} \tag{A15}$$

Substituting (A8) into (A7) and considering (A9) and (A15) yield

$$A = \sum_{s=1}^S B^s = \sum_{s=1}^S P_j^s \tag{A16}$$

References

- [1] Loisel R, Pasaoglu G, Thiel C (2014) Large-scale deployment of electric vehicles in Germany by 2030: an analysis of grid-to-vehicle and vehicle-to-grid concepts. *Energy Policy* 65(3):432–443
- [2] Donato T, Licci F, D'Elia A et al (2015) Evaluation of emissions of CO₂ and air pollutants from electric vehicles in Italian cities. *Appl Energy* 157:675–687
- [3] Rezaei P, Frolik J, Hines PDH (2014) Packetized plug-In electric vehicle charge management. *IEEE Trans Smart Grid* 5(2):642–650
- [4] Franke T, Krems JF (2013) Understanding charging behavior of electric vehicle users. *Trans Res Part F* 21(4):75–89
- [5] Li CT, Ahn C, Huei P et al (2013) Synergistic control of plug-in vehicle charging and wind power scheduling. *IEEE Trans Power Syst* 28(2):1113–1121
- [6] Alizadeh M, Scaglione A, Applebaum A et al (2015) Reduced-order load models for large populations of flexible appliances. *IEEE Trans Power Syst* 30(4):1758–1774
- [7] Liu P, Yu JL (2016) PEV Charging coordination via group differentiated off-peak tariff schemes. *Chinese Patent* 201610317965.2, 16 May 2016
- [8] Shaaban MF, El-Saadany EF (2013) Probabilistic modelling of PHEV charging load in distribution systems. In: 3rd International conference on electric power and energy conversion systems, Istanbul, Turkey, 2–4 Oct 2013
- [9] Ashtari A, Bibeau E, Shahidinejad S et al (2012) PEV charging profile prediction and analysis based on vehicle usage data. *IEEE Trans Smart Grid* 3(1):341–350
- [10] Harris CB, Webber ME (2014) An empirically-validated methodology to simulate electricity demand for electric vehicle charging. *Appl Energy* 126:172–181
- [11] Rautiainen A, Repo S, Järventausta P et al (2012) Statistical charging load modeling of PHEVs in electricity distribution networks using national travel survey data. *IEEE Trans Smart Grid* 3(4):1650–1659
- [12] Gennaro MD, Paffumi E, Scholz H et al (2014) GIS-driven analysis of e-mobility in urban areas: an evaluation of the impact on the electric energy grid. *Appl Energy* 124(1):94–116
- [13] Lojowska A, Kurowicka D, Papaefthymiou G et al (2012) Stochastic modeling of power demand due to EVs using copula. *IEEE Trans Power Syst* 27(4):1960–1968
- [14] Alizadeh M, Scaglione A, Davies J et al (2014) A scalable stochastic model for the electricity demand of electric and plug-in hybrid vehicles. *IEEE Trans Smart Grid* 5(2):848–860
- [15] Zhang M, Zhuang T (2012) Analysis of the measured charging load of electric vehicle in residential area. *Power Demand Side Manag* 14(3):44–46
- [16] Speidel S, Bräunl T (2014) Driving and charging patterns of electric vehicles for energy usage. *Renew Sustain Energy Rev* 40:97–110
- [17] Azadfar E, Sreeram V, Harries D (2015) The investigation of the major factors influencing plug-in electric vehicle driving

- patterns and charging behaviour. *Renew Sustain Energy Rev* 42:1065–1076
- [18] Liu P, Yu JL (2016) Charging behavior characteristic simulation of private electric vehicles for demand response. In: *Proceedings of the 11th international conference on control (Control'2016)*, Belfast, UK, 31 Aug–2 Sept 2016
- [19] Alexander B (2010) Dynamic pricing? Not so fast! A residential consumer perspective. *Electr J* 23(6):39–49
- [20] Mathieu JL, Koch S, Callaway DS (2013) State estimation and control of electric loads to manage real-time energy imbalance. *IEEE Trans Power Syst* 28(1):430–440
- [21] Heinemann GT, Normian DA, Plant EC et al (2007) The relationship between summer weather and summer loads—a regression analysis. *IEEE Trans Power Appl Syst* 85(11):1144–1154
- [22] Qian KJ, Zhou CK, Allan M et al (2011) Modeling of load demand due to EV battery charging in distribution systems. *IEEE Trans Power Syst* 26(2):802–810
- [23] Liu P, Liu RY (2012) Charging load model based on diffusion theory for electric vehicles. *Electr Power Autom Equip* 32(9):30–34
- [24] Shao CC, Wang XF, Wang XL et al (2014) Cooperative dispatch of wind generation and electric vehicles with battery storage capacity constraints. *IEEE Trans Smart Grid* 5(5):2219–2226
- [25] Liu P, Bai XF, Zhang W et al (2016) Diffusion model of PEV charging load and its application on ACE correction. *IEEE Trans Smart Grid* 7(1):501–509
- [26] Huff JO, Hansott S (2010) Repetition and variability in urban travel. *Geogr Anal* 18(2):97–114
- [27] Verma A, Asadi A, Yang K et al (2015) A data-driven approach to identify households with plug-in electrical vehicles (PEVs). *Appl Energy* 160:71–79
- [28] Wen H, Wang SL (2015) Beijing transport annual report. Beijing Transportation Research Center. Available via DIALOG. <http://www.bjtrc.org.cn/JGJS.aspx?id=5.2&Menu=GZCG>. Accessed 12 Mar 2016
- [29] Yang J (2002) Analysis and study of Nanjing Residential load profile. *Power Demand Side Manag* 4(2):24–27
- [30] Jiang Y (2003) Analysis of residents' summer load characteristic in Nanjing. *Relay* 31(4):24–26
- [31] Jiang Y (2003) Characteristic analysis on residential load during winter in Nanjing City. *Northeast Electr Power Technol* 24(4):21–23
- [32] Wen Q, Li J, Zhao J (2005) Air-conditioning load calculation methods and its application. *Power Demand Side Manag* 7(4):16–18
- [33] Lu N, Taylor T, Jiang W et al (2010) Climate change impacts on residential and commercial loads in the western U.S. grid. *IEEE Trans Power Syst* 25(1):480–488
- [34] Xia Q (2014) Number of main durable consumer goods per 100 urban households (1978–2013). In: *Beijing statistical year book, China Statistics Press, Beijing*, p 198
- [35] Lee J (2013) Growth cycle of consumer durables demand and economic growth correlation: 1978–2012. *Ind Econ* 233(7):31–42
- [36] Liang H, Su M, Tian L (2014) Does household electricity use pattern accommodate the requirement of urbanization in China—empirical research based on household appliances. *China J Econ* 1(1):163–168
- [37] Byrd RH, Hribar ME, Nocedal J (1999) An interior point algorithm for large scale nonlinear programming. *SIAM J Optim* 9(4):877–900
- [38] Zhang S (2001) *Studies on interior-point algorithm*. Dissertation, University of Science and Technology of China
- [39] Vehicle Technologies Programs (2013) EV project electric vehicle charging infrastructure summary report. The U.S. Department of Energy. Available via DIALOG. http://energy.gov/sites/prod/files/2014/02/f8/evproj_infrastructure_q22013_0.pdf. Accessed 25 Oct 2016

Peng LIU received the B.S. and M.S. degrees in electrical power engineering and its automation from Harbin Institute of Technology, Harbin, China, in 2010 and 2012, respectively. From 2012 to 2014, he became an engineer in the Jinan Power Supply Company. He is currently pursuing his Ph.D. degree in power engineering and its automation at HIT. His current research interests include the technologies of eco-friendly integrating PEVs and renewable energies into smart grids.

Jilai YU received the B.S. degree in information process and pattern recognition engineering from Harbin Institute of Technology, Harbin, China, in 1988 and the M.S. degree in electric power engineering and its automation from HIT in 1990. He got the Ph.D. degree in electrical power engineering and its automation from North China Power University, Baoding, China, in 1992. Since 1998, he has been a Professor with the Department of Electrical Engineering, HIT. Now, he is the Director of Institute of Electrical Power Science and New Technology. His research interests include power system optimal dispatch and operation, pattern recognition of power system and smart grid, green electricity co-uniting optimization and control.

