Electrical Power and Energy Systems 133 (2021) 107195

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Flexibility management of electric vehicles based on user profiles: The Arnhem case study

Marc Cañigueral, Joaquim Meléndez

University of Girona, Catalonia, Spain

A R T I C L E I N F O

Keywords:
Electric vehicles
Flexibility
User profile
Smart charging
Clustering
Optimization

A B S T R A C T

The ever-increasing global adoption of electric vehicles has created both challenges and opportunities for electrical grids and power systems as well as the market itself. Smart charging is broadly presented as a relevant opportunity to provide demand-side flexibility, benefiting both the user and the power system through flexibility aggregators. However, coordinating all sessions for the same optimization objective could be inefficient when the flexibility potential mismatches the flexibility demand. Instead, this paper proposes the user profile concept as a tool to group sessions into similar flexibility levels and then schedule the charging sessions of each user profile according to its most convenient optimization objective. Therefore, a clustering methodology based on a bivariate Gaussian Mixture Models is presented and validated with a real-world data set, resulting in seven different user profiles. The simulation of two smart charging scenarios, first coordinating all flexible sessions and second coordinating two selected user profiles, resulted in a more efficient scheduling in the latter case, obtaining similar results with a 35% fewer sessions shifted and the corresponding reduction in exploitation costs.

1. Introduction

The irreversible electrification of the mobility sector will open up an opportunity for a more efficient management of electricity grids thanks to the flexibility electric vehicles can provide. In this upcoming scenario, user participation and engagement are crucial (i.e. the concept of the energy citizen), and the management of EV charging sessions (smart charging) is seen as one of the key enabling technologies since it is technically easy to be incorporated into the energy value chain by intermediate agents such as aggregators or flexibility providers. Smart charging, as an optimization problem, can be addressed from different perspectives (e.g. demand response, energy community management, ancillary services) and objectives (e.g. peak shaving, local use of renewable generation, technical constraint from distribution grid, etc.) that result in different approaches to the problem, but all these use cases require an aggregator, an intermediate stakeholder that creates products capable of engaging the user and schedule flexible demand to satisfy the needs of the grid.

However, two main challenges arise when dealing with day-ahead (or intra-day) scheduling: accurate forecasting of participating charging sessions and reliable performance of smart charging algorithms capable of offering robust (reduced uncertainty) flexibility schedules with low computational costs. To date, the focus of research in this field has been on scheduling all charging sessions for a specific objective. In contrast, the methodology proposed in this paper decomposes the scheduling problem, optimizing each user profile according to its suitable flexibility objective, since the classification of EV sessions among generic user profiles (each with its own characteristic flexibility potential) can be used by aggregators as a tool to deliver smart charging in a more efficient and robust manner.

The first main contribution of this work aims to facilitate the task of aggregators with a clustering method to discover generic user profiles based on two simple attributes of the charging sessions: the start time and the duration of the connection. The objective is to aggregate users among similar daily connection patterns and, therefore, similar flexibility potential. Gaussian Mixture Models clustering is used to identify these rational clusters. The second contribution exploits this knowledge about user behaviour to create a selective smart charging strategy capable of satisfying multiple flexibility objectives through a targeted participation of sessions according to user profile membership. This reduces the uncertainty of the aggregated flexibility potential and the complexity of the optimization problem, reducing at the same time the number of sessions to exploit.

The work is structured as follows. Section 2 gives a short overview of the research into the fields of contribution undertaken in this paper. Section 3 details the methodology and methods used for the clustering

E-mail addresses: marc.canigueral@udg.edu (M. Cañigueral), joaquim.melendez@udg.edu (J. Meléndez).

https://doi.org/10.1016/j.ijepes.2021.107195
Received 4 November 2020; Received in revised form 14 April 2021; Accepted 12 May 2021
Available online 17 June 2021

process, while Section 4 describes the clusters obtained from a real data set and the subsequent characterization of these clusters into user profiles. Finally, Section 5 presents the smart charging algorithm and shows the value of introducing EV user profiles into a smart charging application. In this validation step, actual charging sessions data has been used in order to assess the results separately from quality of forecasting.

2. Related work

This section describes the current state of the art in the areas this paper contributes to the most: 1) clustering EV sessions among generic user profiles and 2) scheduling charging sessions to optimize the aggregated EV load.

2.1. User profile clustering

The main interest of this work is characterising user profiles with potential participation in flexibility programs. Thus, this work uses the terms ‘user profile’ or ‘user behaviour’ to refer to a daily connection pattern, defined by the beginning and end times of the EV session, or in other words, the connection start time and the connection duration. This approach gives a different perspective to the meaning of EV profile or EV user behaviour studied in other works that focus on the electrical demand of electric vehicles, which are based on charging profiles and power curves. There, the EV charging load is modelled using probabilistic density functions (PDFs), considering the initial and final State-of-Charge (SoC) per connection in [1], and the sessions start time, end time and distance driven in [2]. Users’ demographics information such as the gender, the age and the education level is used in [3] to model different probabilistic models for charging load profiles simulation. The user behaviour is also referred in [4] to show the impact in the revenue of a charging infrastructure, defining the user behaviour with the arrival time, the dwell time and the energy demand. Other works have studied user behaviour in terms of charging frequency and the probability to charge at the end of the day. The probability of performing a domestic charge is modelled in [5] based on the final SoC of the daily distance driven. Similarly, the charging decision is modelled in [6,7] through the daily vehicle use, applying K-means clustering with the daily average speed as the feature vector, and using the clusters that the vehicle belongs to for simulating the charging load. The decision of the EV user about whether to participate or not in a demand-response program is also defined as user behaviour in [8,9], where users can choose among three different charging powers according to their time flexibility.

The focus of this work, however, is on the flexibility management of EV sessions; therefore we have explored the potential of *connection profiles* rather than analysing *charging profiles*. The methodology proposed in this work has the aim to discover generic connection profiles (i.e. user profiles) to increase knowledge on the potential flexibility of electric vehicles charging demand. Many studies aim to identify connection profiles according to the daily habits of car owners with different classifications or labelling: charging at workplace, at home or park-to-charge in [10,11]; office chargers, home chargers and visitors/taxis/car-sharing in [12]; full-time/part-time worker, unemployed people or professional driver in [13]; regular and random users in [14]; visitors and local users in [15]; stop&charge, park&charge, work&charge, home&charge and long sessions in [16]; or a deeper analysis in [17] with three types of office hours users, three types of overnight users and three types of non-typical users.

Despite data-driven methods are used in all these works to discriminate between these different EV user profiles, the complexity of the classification method and the variables used vary among them. Kim et al. [14] developed a hazard-based duration model of charging regularity (i.e. inter-charging times). A threshold value of the energy charged to differentiate between locals and visitors was used in [15]. K-means clustering combined with multilayer perceptron to improve classification was used in [18]. In [16], a multinomial logistic regression technique over a single variable (i.e. connection duration) was used, while a multivariate Gaussian mixture model with four variables (i.e. session start time, connection duration, hours between sessions and distance between sessions) was performed in [17]. In [10–12], a two-variable density-based clustering method is addressed with DBSCAN clustering. Session start time and session end time are used for that purpose in [10] [11], and sessions start time and connection duration in [12].

In this paper, the user profiles are not defined *a priori,* Mixture Models (MM) are used to discover them. MM have been chosen as the clustering method for three main reasons. First, [19] exposes that density-based clustering like DBSCAN results in a complex clustering process when different regions of the data space have considerably different densities. In terms of EV charging sessions, this is highly probable since each study case will have some principal user profiles. Second, research carried out into profiling electrical consumption patterns in residential loads shows that MM are better in smoothing out random effects because clustering itself considers the correlation and trends of the variables [20]. In terms of the charging sessions, human behaviour contains an important random component resulting from the different elements that can interfere with our timetables (e.g. traffic lights, traffic, longer-than-scheduled meetings, etc.). And third, another advantage of using MM over K-Means or DBSCAN methods is that the output directly gives both the clusters and the associated models, so modelling each cluster afterwards as done in [11,18] or [1] is not necessary. On the other hand, the decision as to how many variables to include in your data-driven method depends on the availability of the data. The data sets used in [16,17], for example, have an ID variable showing the unique RFID codes of the vehicles. This variable allows a tracking study of the vehicle to be carried out and adds value to the models in terms of travel distance and charging frequency. However, the available data sets may not contain this information, so this work has considered only common variables for any charging infrastructure to undertake a study which is as general as possible. Therefore, the contribution of this paper to the field of user profiles is a methodology reproducible for any charging sessions’ data set, using a robust clustering method scarcely exploited in this field (i.e. Gaussian Mixture Models with Expectation–Maximization) based on two basic variables (i.e. connection start time and connection duration).

2.2. Smart charging algorithm

The classification of sessions into generic user profiles paves the way for the second contribution of this work, a smart charging algorithm based on these profiles. The past decade has seen prolific research into smart charging methods for multiple objectives (e.g. increasing self-consumption of local solar energy, balancing load, reducing energy cost, increasing users profit, etc.), with different configurations (e.g. centralized control, distributed charging, public charging stations, residential buildings, etc.) and diverse optimization methods (linear programming, quadratic programming, meta-heuristics, etc.) [21].

This work presents a smart charging algorithm from the aggregator perspective, with a centralized control in the aggregator figure to schedule and coordinate charging sessions according to a defined objective. A centralized scheduling control is more likely to reach an optimal charging strategy on the system level, since it considers the aggregation of renewable generation and electrical demand in the whole system [21]. Important research has emerged in this field, showing a wide variety of optimization methods and objectives. A two-step Linear Programming (LP) optimization is presented in [22] to reduce the energy demand peak and charging cost shifting sessions from high cost periods to lower cost periods. A LP optimization is also applied in [23] to reduce the power peak demand in a parking lot using valley-filling strategy, and in [24] to reduce the light flicker due to PV fluctuations. A Quadratic Programming (QP) optimization is carried out in [10] for two different scenarios, load balancing and load flattening, to increase the consumption of renewable energy through EV charge. The PV energy self-consumption is also increased with EV optimization in [25], using Particle Swarm Optimization (PSO). The generation costs of supplying the EV load are minimized in [26] with a Mixed-Integer Linear Programming (MILP) optimization. The maximization of the EV aggregator revenue, or minimization of the energy cost, is one of the most used objectives, as seen in [27] [28] using MILP optimization, in [8] with LP, in [29] with non-linear programming or in [30] [31] with QP.

However, most of these works consider all the sessions in the scheduling problem individually, therefore requiring complex mathematical models to solve the optimization problem and define a specific charging power for every EV and time slot. For a large number of sessions in a day-ahead smart charging scenario, obtaining all sessions' schedule from the same optimization problem implies a high computational cost and time. To cope with this complexity of the EV scheduling problem, the Alternating Direction Method of Multipliers (ADMM) is presented in some works as an emerging technique for large-scale optimizations, since it decomposes the original objective function into multiple problems to solve in parallel [31] [24].

On the other hand, the cornerstone of aggregators when participating in demand response programs is the *a priori* quantification of the flexibility capacity. That is, the aggregated power that can be allocated at a specific time to satisfy a certain flexibility demand, directly related with the size and behaviour of the aggregated EV users. The knowledge of generic user profiles among sessions and their characteristic flexibility potential could help the aggregator to define a more feasible objective in the optimization problem suited to each user profile separately. Therefore, in contrast to the complex scheduling optimization models raised in other works, this paper decomposes the smart charging method in the following steps to reduce the flexibility uncertainty and the complexity of the optimization problem: (i) the suitable user profiles to accomplish the aggregator's optimization objectives (e.g. peak shaving, solar use, etc.) are selected according to their flexibility potential; (ii) a quadratic optimization is performed to find the optimal aggregated power demand curve (i.e. setpoint) for each user profile according to their optimization objective; (iii) a postpone algorithm is applied to the charging sessions that have agreed to participate in the demand-response program, until the aggregated power demand matches the setpoint.

The division of the method between a time-series optimization and a scheduling algorithm results in a fast computation, at the same time that separates the objective of the demand-response program and the smart charging deploying strategy (e.g. postpone, power modulation, etc). All references above mentioned optimize the charging power of the vehicles for every time slot. This is an optimistic approach since achieving the optimal charge depends on the charging infrastructure (if the charging point has a power modulation feature) and the vehicle (if the vehicle accepts charging with the desired power). Therefore, this work proposes the *Postpone* method as smart charging strategy since it is more widely applicable.

Finally, the algorithm proposed considers the option for EV users to not participate in the demand-response program. In a real implementation of a demand-response program, not all users are willing to participate even though they could provide flexibility, and this response factor must be contemplated by the aggregator [9]. Therefore, we have introduced a *responsive ratio* parameter in our algorithm to randomly select a percentage of sessions that take part in the smart charging program and simulate a more realistic demand-response scenario.

3. Charging sessions clustering methodology

This section describes the proposed methodology for clustering sessions among representative user profiles. In this work we understand the user profile as a generic daily connection pattern rather than the charging (i.e. demand) profile since we do not focus on the EV demand but rather on the EV flexibility. For example, people who arrive at their workplace every working day around 9:00 and go back home around

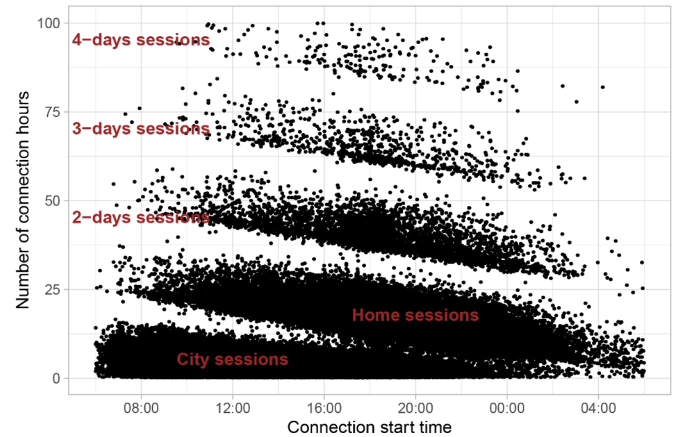

Fig. 1. Charging sessions from the data set of this work.

18:00. Thus, the clustering process has been carried out considering the connection start time and the connection duration as clustering variables. We have assumed that the energy required in one session is not an inherited variable of users' connection pattern and it can vary from day to day and from user to user and consequently this has not been used as a discriminant variable for clustering. Since the methodology proposed considers only these two connection variables, it can be reproduced to any other charging sessions data set, even with anonymous charging sessions without a vehicle or user identifier like the data set used in this case study. The method proposed for discovering and modelling EV user profiles, follows four steps:

1. Division of the data set into sub-sets.
2. Logarithmic transformation of variables.
3. Clustering with Gaussian Mixture Models (GMM) and Expectation–Maximization (EM) algorithm.
4. Characterization of clusters into user profiles.

3.1. Data division

To perform an accurate distribution-based clustering and obtain precise and generic stochastic models for user profiles, the original data set has been divided into several sub-sets according to time cycle and disconnection day.

3.1.1. Time period

The daily habits of citizens change according to the day of the week, the season, holidays, etc., therefore so too does the charging behaviour of the EV users. This is particularly evident for working days (weekdays hereinafter) and weekends, but there are some other cases that could be of interest depending on the community under study. For example, inbound and/or outbound tourism activity during a day or the impact of school holidays. As a result, the distribution of sessions over the day (and the corresponding user profiles) may not always remain the same. Previous studies [10] [11] [1] have analysed the user profiles for weekdays and weekends separately. The method raised in this study does not predefine any time period or duration for the profiles, thus allowing it to discover and model arbitrary user profiles. That said, however, for this study case we have also divided the data set between weekdays and weekends.

3.1.2. Disconnection day

A scatter plot of charging sessions represented according to the connection start time and the connection duration in hours is shown in Fig. 1. Note the visible aggregations of sessions separated by blank ribbons; this is a first sign of different user profiles. Since the addition of the connection duration to the connection start time corresponds to the disconnection time, these blank ribbons represent that vehicles disconnecting at dawn, concretely from 2 to 6 a.m, is not usual. Thus, the different "clouds" in the figure represent the EV sessions that disconnect on the same day of the connection, the day after, two, three, or four days after. As said, this work focuses on daily behaviours, daily connection patterns, so the most relevant groups of sessions are those that disconnect on the same day - labelled *city sessions* - and those that disconnect the following day - labelled *home sessions*.

These different groups of sessions have different density and, since Mixture Models are based on the density distribution of the samples, it is convenient to divide the data into smaller sets in order to obtain better defined distributions shapes to be translated into Gaussian Mixture Models (i.e. clusters).

3.2. Logarithmic transformation

A common practice before applying Mixture Models clustering techniques is to transform the objective variables with the aim of obtaining better distribution shapes to model. Similarly to [17], we apply a logarithmic transformation to our objective variables (i.e. connection start hour and number of connection hours) to reduce sparsity between sessions and increase the density of sessions, so a model-based clustering method such as GMM performs better. Moreover, these variables are defined as only positive, so they present asymmetric distributions. Since Gaussian distribution is by nature symmetric and unbounded, the logarithmic transformation improves normality of time data resulting in better and significant results when applying GMM.

3.3. Clustering

The clustering method used for clustering charging sessions is a bivariate Gaussian Mixture Models (GMM) with Expectation–Maximization (EM) algorithm. On one hand, the strong relationship that Fig. 1 shows between the start time and the duration of EV connections (the later the vehicle connects, the lower the connection duration is), is a sign of a *not null covariance* between these two variables and justifies the use of bivariate Mixture Models, which model the covariance between the two components. Moreover, the Expectation–Maximization (EM) algorithm allows cluster membership to be considered a probability instead of a hard assignment, which makes possible the pertinence to several clusters. This probabilistic classification is convenient due to the random nature of charging sessions and the daily human behaviour. On the other hand, the Gaussian distribution has been selected as the parametric model since the charging sessions are independent and the number of occurrences added at every instant of time is large enough to guarantee a normal distribution, independent of their individual distribution (i.e. the Central Limit Theorem). Thus, considering a user profile as the population and the corresponding data set charging sessions as the sample, and assuming that the sample size is large enough, the density distribution of each cluster can be defined by a bivariate Gaussian distribution.

The use of GMM-EM clustering method requires defining a specific number of clusters. A widely-used strategy to choose the proper number of clusters is to apply the clustering with all desired options of number of clusters and compare their performance using the Bayesan Information Criterion (BIC). The BIC indicator is the value of the maximized log-likelihood with a penalty on the number of parameters in the model. This allows a comparison of models with different parameters or different numbers of clusters. In general the larger the value of the BIC, the stronger the evidence for the model and number of clusters [32]. Once the number of components to explore is defined, then the EM algorithm initializes their parameters, concretely the mixture weight (π), the means vector (μ) and a covariance matrix (Σ) in the case of GMM. After initialization, EM iterates between Expectation–Maximization steps until the log-likelihood function of our model converges with the predefined tolerance. In the following, the main equations of the Expectation–Maximization process are detailed, and the corresponding nomenclature described in Table 1.

Table 1
Nomenclature of Expectation–Maximization algorithm.

Parameter	Description
X	Sample
M	Size of the sample
x	Data point from the sample
i	Index of the data point
K	Number of clusters (Gaussian models)
c	Index of the cluster
π	Weight of the model over the mixture
μ	Means vector of the Gaussian model
Σ	Covariance matrix of the Gaussian model
n	Number of dimensions of the Gaussian model (2 in this case)

The log-likelihood is computed with Eq. (1), refering to each data point as x_i, with i being from 1 to M, and the parameters of each cluster or Gaussian Model, being c being from 1 to K. $N(x_i|\mu_c,\Sigma_c)$ represents the multivariate Gaussian Mixture Model, defined in Eq. (2). The log-likelihood is the logarithmic expression of the weighted description of Gaussian mixture models among all data points. If the Gaussian equation fits the data well, the likelihood increases. The initialization is important in EM iteration, so the log-likelihood is used to select the optimal result of several iterations.

$$log\ p(X|\pi,\mu,\Sigma) = \sum_{i=1}^{M} log(\sum_{c=1}^{K}\pi_c N(x_i|\mu_c,\Sigma_c)) \tag{1}$$

$$N(x_i,\mu_c,\Sigma_c) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_c|^{\frac{1}{2}}} exp(-\frac{1}{2}(x_i-\mu_c)^T\Sigma_c^{-1}(x_i-\mu_c)) \tag{2}$$

3.3.1. Expectation step

In the Expectation step, the probability of each data point being generated by each of the Gaussian models is computed. In contrast to the K-Means' hard assignments, the Expectation assignments are called soft assignments since we are using these probabilities known as *responsibilities*. Each probability or responsibility is calculated with Eq. (3).

$$r_{ic} = \frac{\pi_c N(x_i|\mu_c,\Sigma_c)}{\sum_{k=1}^{K}\pi_k N(x_i|\mu_k,\Sigma_k)} \tag{3}$$

Therefore if x_i is very close to one Gaussian distribution c, it will obtain a high r_ic value for this Gaussian and relatively low values otherwise.

3.3.2. Maximization step

In the Maximization step, the mixture weights (Eq. (5)), the mean (Eq. (6)) and the covariance (Eq. (7)) are updated for each Gaussian mixture model or cluster according to the total responsibility m_c allocated to each cluster (Eq. (4)):

$$m_c = \sum_i r_{ic} \tag{4}$$

$$\pi_c = \frac{m_c}{M} \tag{5}$$

$$\mu_c = \frac{1}{m_c}\sum_i r_{ic}x_i \tag{6}$$

$$\Sigma_c = \frac{1}{m_c}\sum_i r_{ic}(x_i-\mu_c)^T(x_i-\mu_c) \tag{7}$$

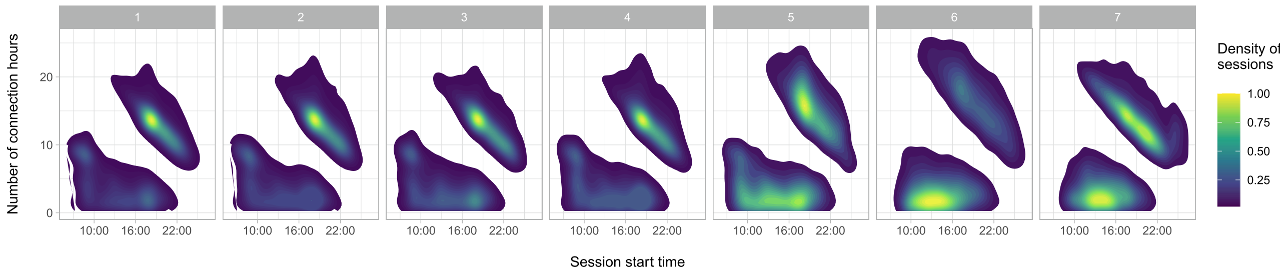

Fig. 2. 2D density plots of sessions by weekday (starting on Monday).

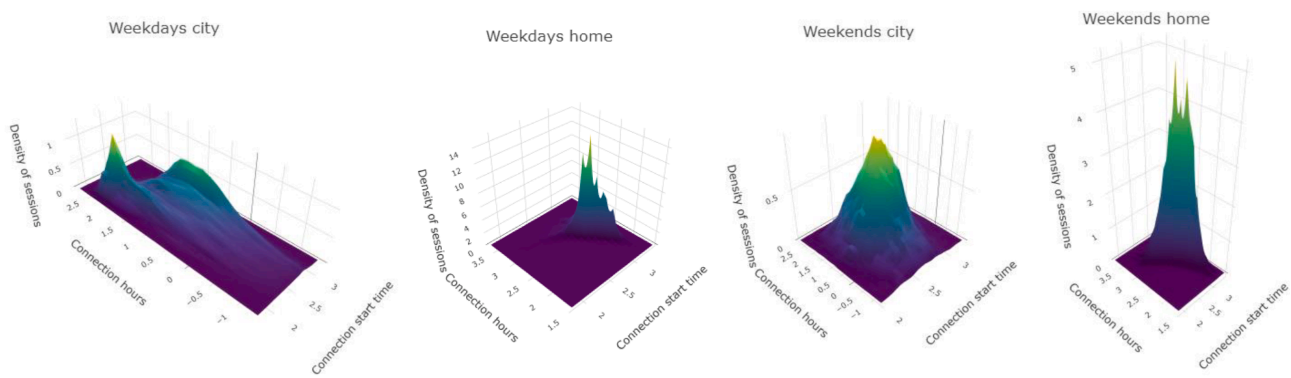

Fig. 3. 3D density distribution plots.

4. Case study: The Arnhem public charging infrastructure

In this section, the methodology presented in Section 3 has been validated with a large data set of EV charging sessions from the middle-sized city of Arnhem, The Netherlands. The data set is composed of 259,419 charging sessions from 2015–08-31 to 2020–06-01 (more than 300 sessions per day in 2020) collected in the Arnhem's public charging infrastructure, which contains 270 different charging poles, each with 2 charging slots with a maximum charging power of 11 or 22 kW (depending on the pole). In fact, analysing just public charging infrastructure data is not a limitation since the 75% of the households in the Netherlands are dependent on public charging facilities and DC charging is not widely used [17].

From the original data set, we discarded 10.3% of sessions considered errors, resulting in a clean data set of 232,583. The sessions with the following characteristics were discarded: energy equal to 0, connection duration less than 15 min, charging duration higher than connection duration, or charging power higher than the maximum power that the public charging infrastructure of the study case can supply (i.e. 22 kW for Arnhem).

Besides, as pointed out in Section 3.1.2, sessions that finished two or more days after the connection, as seen in Fig. 1, have not been considered since they represent only 3% of the clean data set and therefore they are not a generic user profile object of this study. Probably in future research, these long-connection sessions could provide interest on V2G technology due to the potential of being charged and discharged as a battery connected to the public grid. It the end, the final Arnhem's data set consists of 225,040 sessions.

4.1. Preparation of data before clustering

Following the methodology outlined in Section 3, the first step is to explore different density distributions on the data according to time cycles. Even though a large difference in session distributions between years or months is not observed, a relevant difference between weekdays does stand out. In Fig. 2, we can see a similar distribution pattern from Monday to Friday, and different density shapes for Saturdays and Sundays. During weekdays most EVs charge during the evening - probably after working hours - with long connection durations. In contrast, during weekends most sessions have short connections and throughout the day, probably due to brief visits to the city. We have considered two main different time cycles according to the distribution homogeneity in the sessions' distribution, i.e. two different models: *weekdays* and *weekends*.

Besides this, as pointed out in Section 3.1.2, for each time cycle we can distinguish two different groups of sessions, labelled as city and home sessions, according to the disconnection day. In total, four different subsets of sessions will be submitted independently to the GMM clustering process: weekdays city, weekdays home, weekends city and weekends home. Another step before clustering is the logarithmic transformation, explained and justified in Section 3.2. Fig. 3 shows the distribution shapes of the four subsets in the logarithmic scale. This figure justifies the need to divide the data before applying GMM clustering. Each subset contains a clearly different distribution and, moreover, the existence of different peaks of density is an indicator of a mixture of different models (i.e. clusters). Additionally, the big difference between density values in some of these subsets justifies the choice of distribution-based clustering over density-based clustering (e.g. DBSCAN), considering the complexity to find different clusters with a single configuration of parameters when the density differences are so relevant [19].

4.2. Clustering and characterization of user profiles

Arnhem's charging sessions data set has been divided into four subsets (weekdays/weekends city/home sessions), with a transformation of the clustering variables to a logarithmic scale. In this section, a GMM-EM clustering process has been applied to each one of the sub-sets in concordance with the methodology presented in Section 3. The mclust R package [33] has been used for the clustering process. First, the BIC approach has been applied to each sub-set, considering from 1 to 15 clusters (see Figures A.1 - A.4 in Appendix A.1). Considering the number of components from which the BIC indicator stops decreasing, the

Weekdays (74 %)	City sessions (48 %)	Cluster 1 (10 %)	**Worktime (10 %)**
		Cluster 2 (17 %)	**Visit (36 %)**
		Cluster 6 (19 %)	
		Cluster 3 (24 %)	**Shortstay (40 %)**
		Cluster 4 (3 %)	
		Cluster 5 (13 %)	
		Cluster 7 (14 %)	**Dinner (14%)**
	Home sessions (52 %)	Cluster 2 (6 %)	**Commuter (23 %)**
		Cluster 6 (17 %)	
		Cluster 1 (16 %)	**Home (35 %)**
		Cluster 7 (5 %)	
		Cluster 8 (14 %)	
		Cluster 3 (15 %)	**Pillow (41 %)**
		Cluster 4 (9 %)	
		Cluster 5 (17 %)	
Weekends (26 %)	City sessions (53 %)	Cluster 1 (29 %)	**Shortstay (46 %)**
		Cluster 5 (6 %)	
		Cluster 6 (10 %)	
		Cluster 7 (1 %)	
		Cluster 4 (16 %)	**Dinner (16 %)**
		Cluster 2 (8 %)	**Visit (38 %)**
		Cluster 3 (30 %)	
	Home sessions (47 %)	Cluster 1 (14 %)	**Home (59 %)**
		Cluster 4 (18 %)	
		Cluster 6 (7 %)	
		Cluster 7 (19 %)	
		Cluster 2 (11 %)	**Pillow (41 %)**
		Cluster 3 (21 %)	
		Cluster 5 (9 %)	

Fig. 4. Classification of sessions into user profiles with proportions.

Table 2
Average features or user profiles.

Profile	Average start time	Average connection duration (h)	Average charging duration (h)	Average charging power (kW)	Average energy (kWh)
Worktime	09:34	8.66	3.07	3.99	12.93
Visit	13:24	4.26	2.62	4.51	12.23
Shortstay	14:35	1.28	1.16	4.73	5.54
Dinner	19:00	3.74	2.67	4.64	12.81
Commuter	19:14	13.65	3.52	3.73	13.77
Home	18:12	17.89	3.64	3.96	15.32
Pillow	22:13	11.78	3.63	4.18	15.97

following number of components have been selected for each subset: 7 clusters for Weekdays city sessions; 8 clusters for Weekdays home sessions; 7 clusters for Weekends city sessions; and 7 clusters for Weekends home sessions.

In all cases, the convergence in the fitting process came from VVV models (i.e. ellipsoidal distribution, varying volume, varying shape and varying orientation). The corresponding ellipses of every component and every subset are shown in Figures A.5 - A.8. Each ellipse defines a bivariate Gaussian distribution and its centre represents the average start time and average duration of sessions belonging to that group.

Behind the numbers we can interpret a user behaviour in terms of timetable. For a better readability, an exponential transformation has been applied to the centroids of each cluster to translate the logarithmic values into time in hours.

At this point, it is appropriate to add a second-step classification, or profiling step. Each cluster has been labelled with a generic user profile according to their respective interpretations. Thus, each user profile can be assigned multiple Gaussian Mixture Models with the corresponding weights or probabilities. The authors' interpretations of each cluster and the user profiles assigned to them are shown in Tables A.1 - A.4 of Appendix A.2. Our interpretations have not only been based on the values of connection start time and duration of the centroid of each cluster, but as well on the shape of the corresponding ellipses (see Figures A.5 - A.8), which represent the covariance matrix of each cluster. A wider ellipse means a less concrete definition of the user profile. In this way, we have defined very specific user profiles like Worktime (starting around 09:00 for 8-9 h), Dinner (starting around 19:00 for 3-4 h), Commuter (starting after work at 18:00–19:00 for 12–14 h) and Shortstay (duration for less than 1 h), and more general user profiles like Visit (dispersed around the day and varying duration), Home (starting during daytime and connected until the next day) and Pillow (starting during evening-night and connected until the next day). Worktime and Commuter profiles are present only on working days since these are behaviours resulting from work timetables.

Fig. 4 summarizes the clustering process and the user profiling step, showing the different categories found and the corresponding weights of each model. The number of the cluster corresponds to the numbers of Figures A.5 - A.8 of Appendix A.2.

Table 2 shows the average values for the features that define every user profile. Observe that all user profiles, apart from the Shortstay, remain connected longer than charging, and therefore have flexibility hours (i.e. difference between connection and charging times). In fact, for this data set 49.9 % of sessions have more than 5 h of flexibility, and a 56.3% more than 2 h. In the case of the Worktime, Commuter, Home and Pillow profiles, the number of flexibility hours is highly considerable. It can be observed that the charging time is similar for all profiles, with the exception of Shortstay users whose charging time is limited by the connection time. This is a consequence of a similar energy being required for most sessions, concretely between 9 and 16 kWh, independent of their user profile.

Another way to validate the clustering process and the corresponding characterization of each cluster, is to visualize the demand power profile for all EV user profiles. Fig. 5, for instance, shows the demand curves of each user profile for a week in January 2020 The demand curves have been calculated with time intervals of 15 min, using the connection start time, the energy charged and the charging power of each real session. The demand of each profile can be seen to correspond to a specific time-

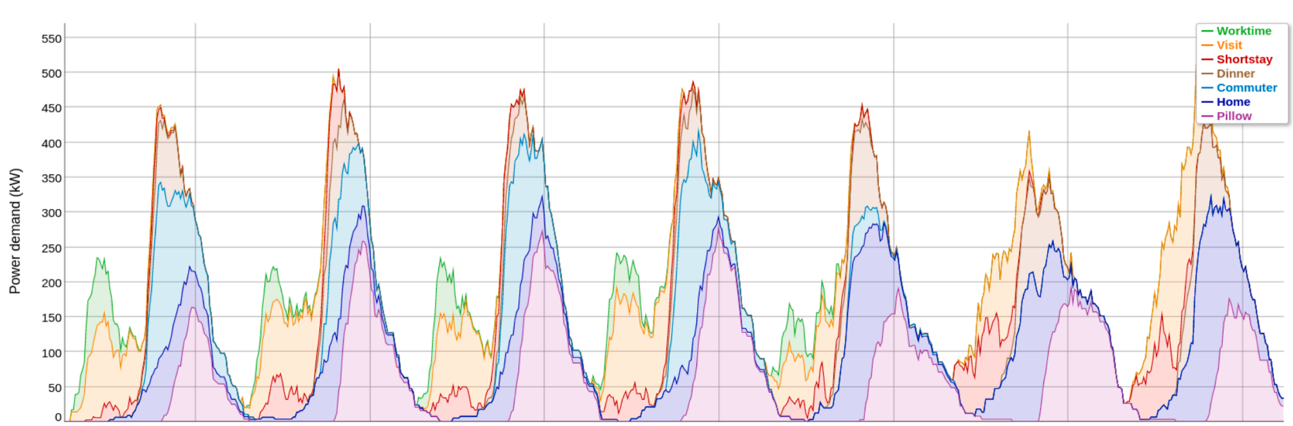

Fig. 5. Arnhem's EV real demand by user profile.

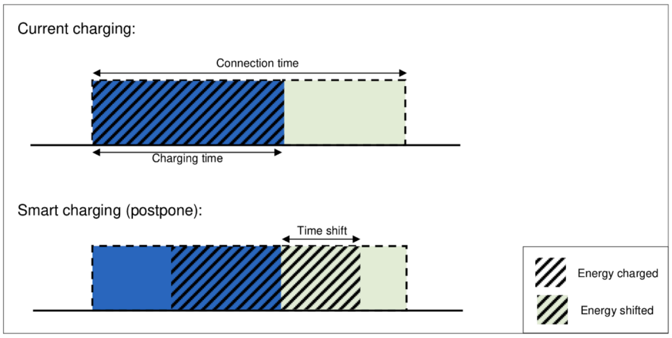

Fig. 6. Smart charging with Postpone method.

Table 3
Nomenclature.

Parameter	Description
T	Number of time intervals within the optimization window
Δt	Time interval, in hours
$TCONS_s$	Connection start time of a session
$TCONE_s$	Connection end time of a session
$TCHS_s$	Charging start time of a session
$TCHE_s$	Charging end time of a session
P_s	Charging power of a session
F_s	Flexible hours (i.e. connection hours – charging hours) of a session
S_{FLEX}	Sessions with flexibility potential
P_{FLEX}	Flexibility potential, in power units
w_1	Weight for grid balance optimization strategy
w_2	Weight for peak shaving optimization strategy
S_t	Solar generation time series
L_t	Static EV load (BAU) time series
V_t	Flexible EV load (BAU) time series
O_t	Optimal flexible EV load time series
E	Total energy demand from flexible EV within the optimization window
δ_{flex}	Percentage of users responsive to the flexibility program
T_{SHIFT}	Timeslots where power demand is higher than the setpoint
P_{SHIFT}	Power to shift from one time slot to the following, considering the setpoint

range according to the sessions' start time. At the same time, morning and evening peaks and a big valley during midday stand out from the total demand curve on weekdays (i.e. 13th – 17th January), while during weekends (i.e. 18th and 19th January) there is a wider and irregular demand profile. This validates employing different models for weekdays and weekends.

5. Flexibility management based on user profiles

In this work, we pursue a smart charging strategy capable of adapting charging sessions to cope with a flexibility demand as a result of participating in a specific program through an aggregator. The purpose of activating such flexibility can accomplish many different goals, from solving technical constraints at DSO level, to peak shaving or simply efficient use of local RES. EV owners are interested in participating in such programs because it implies a monetary benefit, or similar incentive, without affecting their daily habits and thus resulting in a win–win scenario.

However a common drawback is that flexibility potential usually mismatches flexibility demand, so it is extremely important for the aggregator to have information about the typology of EV users and their connection patterns, in order to offer a feasible flexibility demand to the suitable EV users. From here on, instead of rescheduling all sessions according to the same optimization objective, this paper extends the existing EV coordination methodologies by associating each user profile to a particular optimization objective.

Another important point of smart charging is the way that the session is modulated. Traditionally, the charging profile of an EV can be modelled as a power step lasting a certain time and starting as soon as the vehicle is connected. In that sense, a charging session could provide flexibility in terms of time (i.e. the charge is postponed or divided into several shorter sessions), power (i.e. the charging rate is modified) or energy (i.e. the user agrees to finish the session without reaching 100%, or transferring energy to the grid, in the case of V2G). This work only considers flexibility potential in terms of time, i.e. the smart charging postpone method depicted in Fig. 6 showing the exploitation of the difference between connection and charging times to postpone the session.

In this section, first the flexibility potential of every user profile is quantified to offer an overview of the difference between the available flexibility levels (in terms of power and time) among user profiles. Next, a smart charging algorithm is proposed to emulate the individual response and estimate the impact of this flexibility when activated in different scenarios. The nomenclature used in this section is described in Table 3.

5.1. Quantification of flexibility potential

Quantifying the flexibility potential of a power demand curve offers a valuable tool for measuring the impact of shifting a specific amount of power from one time slot to another. Inspired by the definition of demand response potential from Develder et al. [11], we define the potential flexibility of a session lasting a time interval $[t, t + \Delta t]$, within the connection interval $[TCHS_s, TCONE_s]$, as its charging power P_s, if the following statements are true:

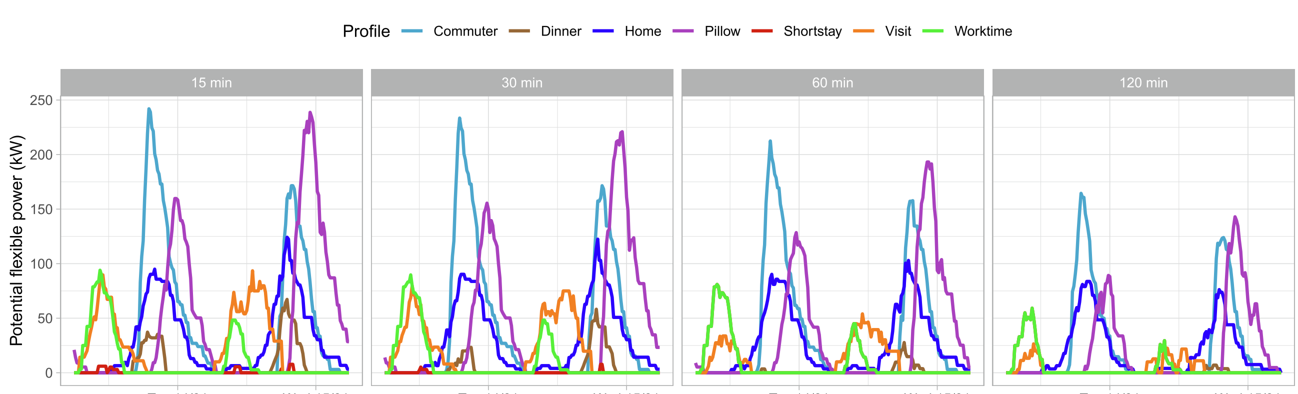

Fig. 7. Flexible power potential by user profile.

1. The vehicle starts charging during this time slot

$$t = TCHS_s \tag{8}$$

2. The vehicle remains charging during the entire time interval

$$[t, t + \Delta t] \subset [TCHS_s, TCHE_s] \tag{9}$$

3. The charging session can be shifted an interval Δ within the connection interval

$$TCONE_s - TCHE_s \geqslant \Delta t \tag{10}$$

When considering all sessions under these statements as $S_{FLEX(t,t+\Delta t)}$, then the aggregated flexible power within the time interval $[t, t+\Delta t]$ is:

$$P_{FLEX(t,t+\Delta t)} = \sum_{s \in S_{FLEX(t,t+\Delta t)}} P_s \tag{11}$$

Eq. (8) has been added to the Develder et al. [11] definition because the only smart charging strategy we consider here is the Postpone method. To postpone the charging start time of a session from time t to time $t + \Delta t$, the session must start charging at time t (i.e. $t = TCHS_s$). Without this constraint, the smart charging method would consider dividing a session into shorter sessions as well, shifting only part of the session instead of the full session.

We have approximated all start times on a 15-min basis since this is a realistic time-base for charging sessions and offers sufficient granularity to participate in different markets and services (e.g. balancing, congestion management). No distribution grid capacity or power system constrains have been considered in this case. Fig. 7 shows the flexibility potential (power vs time) for every existing user profile during two representative days in January 2020, Monday 13th and Tuesday 14th, and considering different time granularity (Δt) of 15, 30, 60 and 120 min. The curves have been obtained by applying the definition of Eq. (11) to the profiles obtained in the analysis from Section 3.

First, it can be observed that the longer the time granularity, the lower the flexibility potential. This is because the probability of having sessions that accomplish the three conditions: Eqs. (8)–(10) decreases with longer time intervals. Note that considering the Postpone strategy is a constraint and for other Smart charging strategies the equations should be modified accordingly. For this study case, there would not be a significant difference if it were to participate in a flexibility market, or demand-response services, with scheduling intervals of 15 or 30 min; but a flexibility management with 120-min time intervals would not be feasible.

Fig. 7 also shows a considerable difference for the potential flexibility between user profiles. As seen in Table 2, most of the user profiles have similar charging duration, while the connection duration varies considerably. This results in higher flexibility potential for user profiles with longer connections. In that sense, Commuter profile has the highest flexibility peak in the evening, followed by Home profile, while Pillow profile has the flexibility peak at night. At the same time, Worktime profile has its flexibility peak at early-morning, while the flexibility potential of Visit profile is more irregular throughout the morning. Finally, Shortstay and Dinner profiles have too short connections in order to deliver a relevant flexibility potential. Note that despite the big difference between the number of sessions of Visit and Worktime profiles (36% for Visit profile and 10% for Worktime profile for Weekdays, see Fig. 4), the level of flexibility is similar. The same happens with Commuter and Home profiles and, therefore, this shows the importance of grouping sessions among user profiles when the objective is to manage their flexibility.

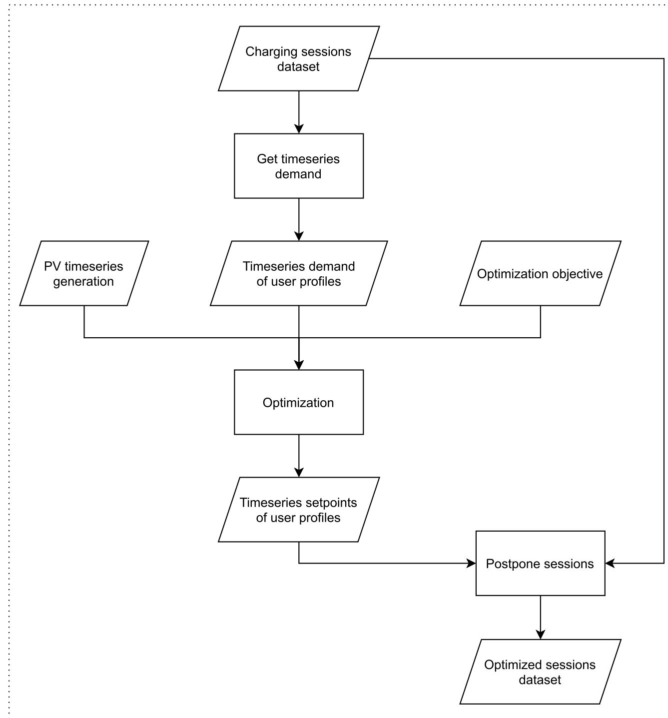

Fig. 8. Smart Charging diagram.

5.2. Smart charging algorithm

The novelty of the smart charging strategy proposed in this section consists in using the previously-identified EV user profiles to address the flexibility management process, specifying a particular objective to a particular user profile. This strategy allows *a priori* estimation of flexibility based on the user profiles and reducing the uncertainty during both the scheduling and activation stages. Thus, the output of the smart charging algorithm modulates the charging start time (postpone scenario) and assumes that the same energy is delivered. The smart charging methodology proposed follows the sequence depicted in Fig. 8 and it is composed by the following three steps:

1. *Get aggregated time series demand*: given a datetime sequence and a charging sessions data set, the demand profile of every user profile is obtained as a time series format.
2. *Obtain the setpoint (Optimization)*: according to user profiles' demand, renewable PV generation and optimization objective (e.g. peak shaving, grid balancing or both), a convex optimization is performed to obtain the best-case optimal demand profile for each user profile (i.e. user profiles' setpoints).
3. *Postpone sessions (rescheduling)*: original sessions of each profile are shifted from time slot to time slot in order to match, if possible, the corresponding setpoint.

5.2.1. Aggregated demand

Considering $S_{CHARGE(t,t+\Delta t)}$, all sessions that remain charging within the time interval $[t, t + \Delta t]$, and therefore satisfy Eqs. (12) and (13), then the aggregated power demand within the time interval $[t, t+\Delta t]$ is calculated with Eq. (14) where P_s is the charging power of a charging session. The aggregated time series demand is calculated then for each time slot considering a time resolution?t and a window of time T.

$$t \geqslant TCHS_s \tag{12}$$

$$[t, t + \Delta t] \leqslant TCHE_s \tag{13}$$

$$D_{(t,t+\Delta t)} = \sum_{s \in S_{CHARGE(t,t+\Delta t)}} P_s \tag{14}$$

5.2.2. Optimization

A quadratic optimization has been developed to obtain the optimal time series demand of each user profile. Though the procedure is general enough to deal with other objectives, the following two objectives are approached in this smart charging simulation:

1. Minimize peaks of demand, shifting demand from peak hours to valley periods.
2. Minimize grid balance, moving demand to hours with local solar generation.

The optimization is constrained to only the Postpone flexibility strategy (see Fig. 6) and and no grid parameters have been used as a constraint in this case. Future work will address other levels of flexibility such as energy or power modulation and the consideration of grid congestion and other constraints linked to geolocation of delivered flexibility, resulting in different aggregations. The objective function of this problem is presented in Eq. (15). See Table 3 for nomenclature definitions. The first term corresponds to the grid balance strategy, while the second term refers to peak shaving.

$$min \sum_{t=1}^{T} w_1 (S_t - L_t - O_t)^2 + w_2 (L_t + O_t)^2 \tag{15}$$

Constrained to Eqs. (16) and (17):

1. Total EV demand must remain the same:

$$\sum_{t=1}^{T} O_t \Delta t = E \tag{16}$$

2. Demand can only be shifted forwards, not backwards (Postpone strategy):

$$\sum_{t=1}^{U} O_t \Delta t \leqslant \sum_{t=1}^{U} V_t \Delta t \quad U = 1, 2, \ldots, T \tag{17}$$

The formulation results in a quadratic problem with linear constraints. Thus, a convex optimization has been applied to the objective function using the CVXOPT Python package [34].

5.2.3. Postpone sessions

In a smart charging application, each charging point would have to decide whether to charge or not when a vehicle starts a new connection. Thus, some in-place computation will be required and in the case of postponing the vehicle's charging, a new schedule proposed by the charging system. Thus, as a more practical approach, rather than optimizing the aggregated user profile demand, the smart charging algorithm presents a new schedule for each charging session. In that sense, Algorithm 1 presented in this section takes an optimal aggregated demand (i.e. setpoint obtained from Eq. (15)) as a reference and postpones each required session to satisfy it, resulting in a modified sessions data set basically with shifted charging start times. Moreover, we have considered a parameter δ_{flex} to represent the percentage of people participating in the flexibility program (i.e. responsive users), since probably not all users will be enthusiasts about this charging system or they simply will be unable to participate on specific days.

Algorithm 1. Postpone charging sessions according to power time series setpoint, for a single user profile

```
Input  : charging sessions schedule S, power setpoint time series O, time interval Δt, percentage of responsive users δ_flex
Output : modified schedule of charging sessions S

while True do
    Get L_t, total power demand timeseries from charging sessions
    Define T_SHIFT, a vector of time slot values where L_t > O_t
    if length(T_SHIFT) = 0 then
        break                                  /* No more flexibility required */
    end
    P_SHIFT = L_t(T_SHIFT) − O_t(T_SHIFT)      /* power to shift in T_SHIFT time slots */
    S_FLEX = sessions with potential to be shifted   /* i.e. TCHS_s = any(T_SHIFT) and F_s ≥ Δt */
    S'_FLEX = randomly select a percentage δ_flex from S_FLEX
    if length(S'_FLEX) = 0 then
        break                                  /* No more flexibility available */
    end
    TI = min(T_SHIFT)                          /* Time slot to shift sessions from */
    S''_FLEX = sessions from S'_FLEX that start in TI
    Sort S''_FLEX from higher to lower F_s
    PI = P_SHIFT(TI)                           /* power to shift in TI time slot */
    s = 1
    while PI > 0 and s <= rows(S''_FLEX) do
        <!– s = S''_FLEX(s)
        –> TCHS_s = TCHS_s + Δt                /* Shift session a time slot */
        F_s = F_s − Δt                         /* Reduce flexible time a time slot */
        <!– S(s_id) = s                        /* Update original schedule with the modified session */
        –> PI = PI − P_s                       /* Update the pending power to shift */
        s = s + 1
    end
end
```

Table 4
Weights of optimization objectives.

Scenario	Profile	w1	w2
Without user profiles	-	0.5	0.5
With user profiles	Worktime	1	0
	Commuter	0	1

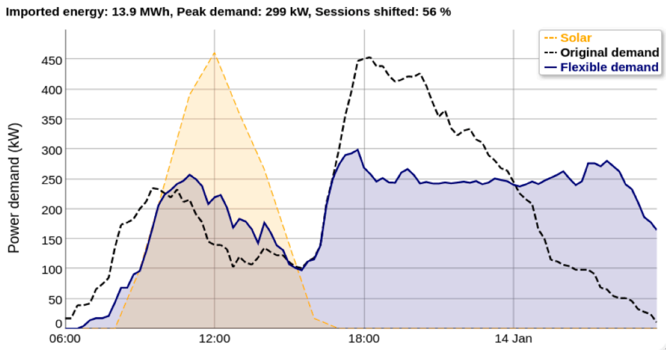

Fig. 9. Scenario 1: optimization without user profiles.

All these points of the algorithm are inside a `while(True)` loop structure that aims to iterate over all potentially flexible sessions (S_{FLEX}) in every time slot where flexibility is required, and leave the loop (i.e. `break` statement) when (1) there is no required flexibility (i.e. the demand setpoint is not surpassed in any time slot), or (2) there is no available flexibility (i.e. sessions have fully exploited their flexibility). Note that this algorithm considers only a Postpone smart charging strategy (see Fig. 6), and further development must take place if other strategies such as power modulation or dividing the session in shorter sessions are to be considered.

5.3. Smart charging simulation

This section proposes a scenario where the municipality of Arnhem aims to supply as much as possible the EV fleet with energy from a local PV field of 500 kWp (i.e. grid balancing minimization), maintaining the aggregated demand curve as flat as possible (i.e. peak shaving minimization). In this scenario, the EV aggregator should optimize the aggregated EV demand by rescheduling charging sessions for a specific time window. For this simulation, real EV sessions are used to decouple results from quality of forecasting. However, in both forecasting and scheduling problems, the knowledge about the existing EV user profiles would reduce the uncertainty since the problem is decomposed and analyzed separately. Thus, the objective of the methodology proposed

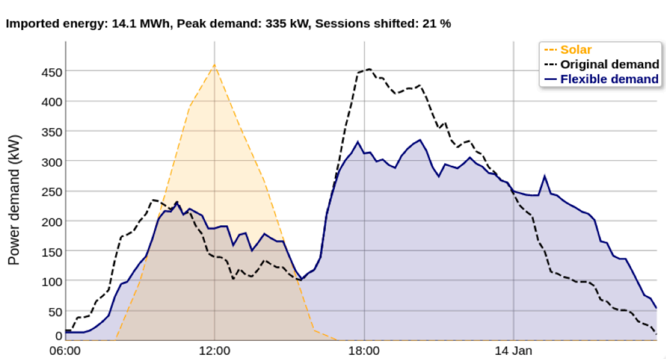

Fig. 10. Scenario 2: optimization with user profiles.

Table 5
Optimization results.

Optimization	Sessions shifted (%)	Reduction of peak demand (%)	Reduction of grid energy (%)
Without user profiles	56	34	6
With user profiles	21	26	5

here is to simplify the decision-making process of the aggregator in the optimization stage, where adjusting the parameters of individual sessions through a single tariff without differentiating the connection profile could be inefficient.

The performance of optimizing a set of EV sessions with and without user profiles is compared by considering two different optimization objectives: evening peak shaving and grid balancing. According to these optimization objectives, and the flexibility potential of each user profile seen in Fig. 7, the profiles used for each optimization objective and their corresponding weights are described in Table 4.

Postponing all potentially flexible sessions without user profiles is a scenario constructed to show the best-case performance. Fig. 9 shows the result of combining the peak shaving and grid balancing optimization objectives (i.e. `w1` $= 0.5$ and `w2` $= 0.5$) and applying the postpone strategy. Postponing 56% of the total number of sessions, we can see a reduction in the peak demand of 155 kW, while the energy imported from the grid has been reduced by 218 kWh. On the other hand, Fig. 10 shows the result of addressing the peak shaving optimization with the Commuter profile (i.e. `w1` $= 0$ and `w2` $= 1$) and the grid balancing optimization with the Worktime profile (i.e. `w1` $= 1$ and `w2` $= 0$). In this case, only 21% of the sessions have been postponed, while the reduction in the peak demand has been 118 kW and the energy

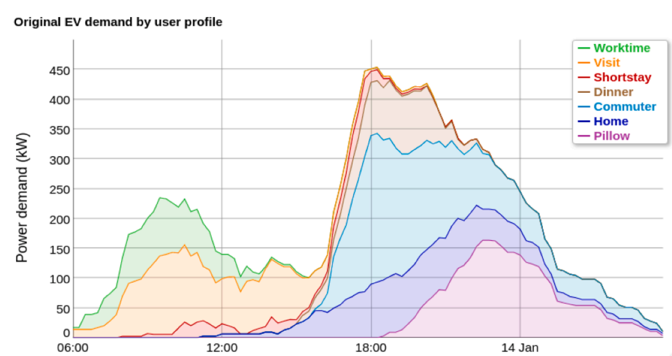

Fig. 11. Original EV power demand by user profile.

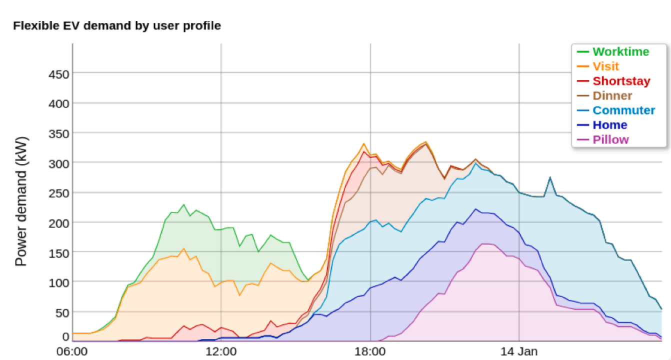

Fig. 12. EV power demand by user profile in Scenario 2.

imported from the grid has been reduced by 182 kWh. To compare both scenarios, Table 5 provides a summary of the results.

Obviously the best-case results are obtained shifting all potentially flexible sessions, independently of user profiles (i.e. business-as-usual case), since more sessions are available to postpone. However, the flexibility management based on user profiles resulted in a relevant improvement on the system efficiency, obtaining similar results than the best-case with practically a third part of the sessions. The optimization with user profiles has obtained 1% more imported energy, 8% higher peak demand and 35% fewer exploited sessions, which implies a relevant lower cost by the aggregator considering a compensation for each postponed session.

For a more in-depth analysis in the optimization with user profiles, Figs. 11 and 12 show the demand curves for each user profile before and after smart charging simulation, respectively. It is visible that early morning Worktime sessions have been shifted in order to charge as much as possible from solar generation, moving the peak to 11:00–12:00. At the same time, the pointed evening peak of Commuter sessions at 18:00–19:00 has evolved to a flatter curve shifting the demand to the night valley.

6. Conclusions

The first contribution this study makes is a methodology for characterizing EV charging sessions among generic user profiles, which has been validated with a real data set from the Dutch city of Arnhem. A first analysis of the relationship between the connection start time and connection duration of the sessions showed relevant covariance and multiple density peaks. These characteristics validated the use of bivariate Gaussian Mixture Models as a suitable clustering method. A posterior interpretation of each cluster resulted in seven different user profiles, some of them very specific (Worktime, Diner, Shortstay and Commuter) and other more general (Visit, Home and Pillow). Two main conclusions can be drawn from the flexibility potential quantification: (1) each user profile has its own flexibility potential peak and (2) considering a Postpone smart charging strategy, the time-resolution of the demand response program should not be lower than 30 min. Therefore, the second contribution this paper makes is a Postpone algorithm based on user profiles, with the possibility of configuring the appropriate optimization objective (i.e. grid balancing and peak shaving) to a particular user profile according to its flexibility potential. This approach has resulted more efficient in terms of flexibility management than the best-case scenario where all sessions are considered for the demand-response program. Even though the best-case optimization obtained better grid performance indicators (1% less imported energy and 8% lower demand peak), their differences in comparison with the optimization based on the Worktime and Commuter user profiles are not as relevant as the difference in the flexibility exploitation (35% fewer postponed sessions). Future research should consider designing different policies or market tariffs for each user profile in order to optimize the aggregator profit and the EV user compensation in a win–win scenario. Moreover, other smart charging strategies that differ from postpone should be considered in order to modulate the charging power or energy of sessions, and improve the flexibility management results.

Glossary

Acronym	Description
ADMM	Alternating Direction Method of Multipliers.
BIC	Bayesan Information Criterion.
DBSCAN	Density-based spatial clustering of applications with noise.
DSO	Distribution system operator.
EC	European Commission.
EM	Expectation–Maximization
EU	European Union.
EV	Electric vehicle.
GMM	Gaussian Mixture Models.
ID	Identificator
LP	Linear Programming.
MILP	Mixed-Integer Linear Programming.
MM	Mixture Models.
PSO	Particle Swarm Optimization.
QP	Quadratic Programming.
RES	Renewable energy sources.
RFID	Radio-frequency identification.
V2G	Vehicle-to-grid

CRediT authorship contribution statement

Marc Cañigueral: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft. **Joaquim Meléndez**: Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been carried out within the research group eXiT (http://exit.udg.edu) at the Universitat de Girona - Consolidated Research group (Ref. 2017 SGR 1551) by the Generalitat de Catalunya - and developed under the projects CROWDSAVING (Ref. TIN2016-79726-C2-2-R), co-funded by the Spanish Ministerio de Industria y Competitividad (Agencia Estatal de Investigación), European ERDF funds, and the H2020 project "E-LAND – (Grant agreement 824388). The work has also been supported by the consulting company Resourcefully (The Netherlands) under a specific agreement of scientific cooperation with the University of Girona. The authors would like to thank the city of Arnhem, in particular Peter Swart, for facilitating access to the data used in the validation stage of this research. The author Marc Cañigueral has been awarded a PhD-scholarship (Ref. FPU18/03626) by the Spanish Ministry of Education and Culture through the Training programme for Academic Staff (FPU-programme).

Appendix A. Clustering resources

This appendix was included to show in detail some development steps of the methodology exposed is Section 3.

A.1. BIC analysis

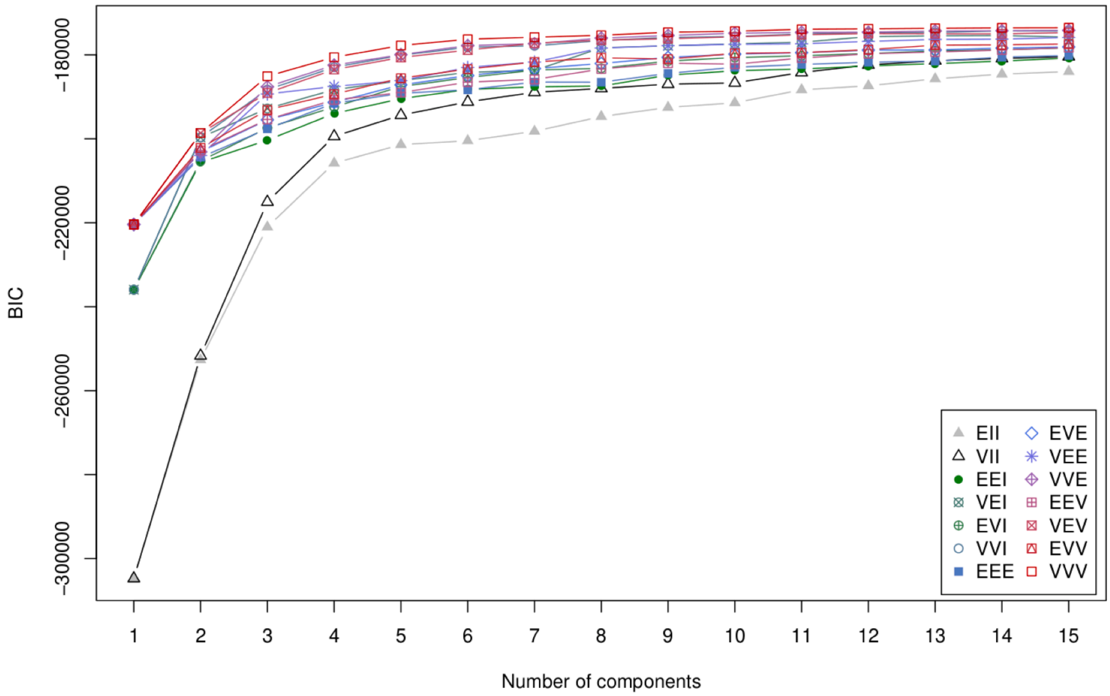

Fig. A.1. BIC analysis for weekdays city sessions.

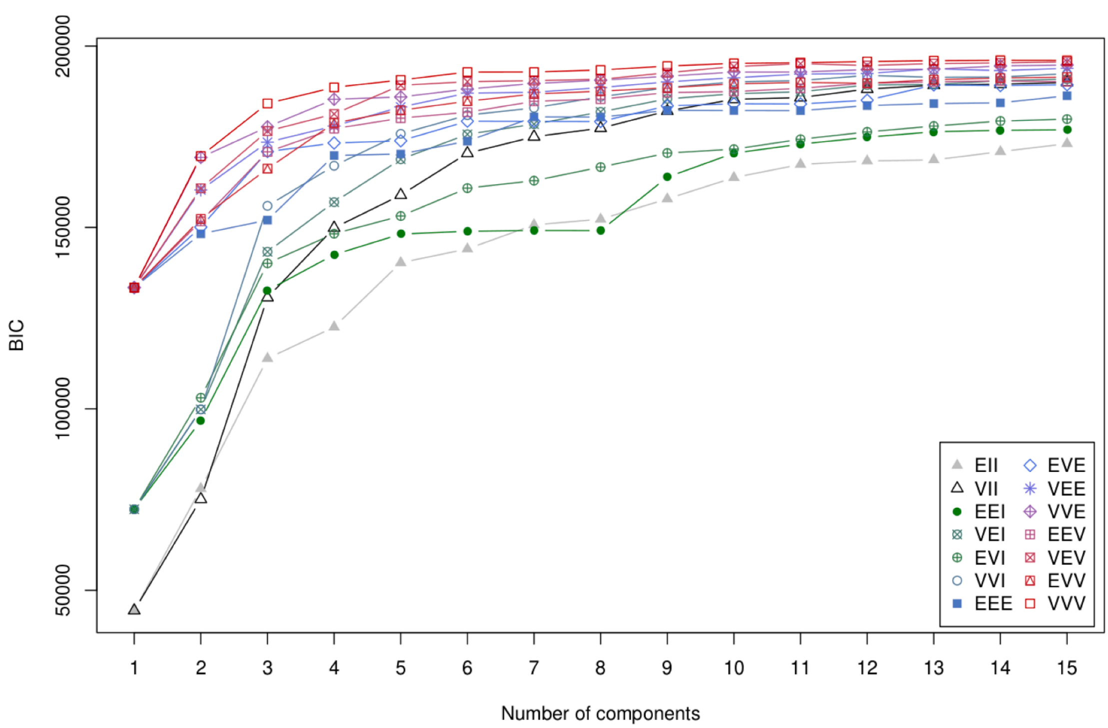

Fig. A.2. BIC analysis for weekdays home sessions.

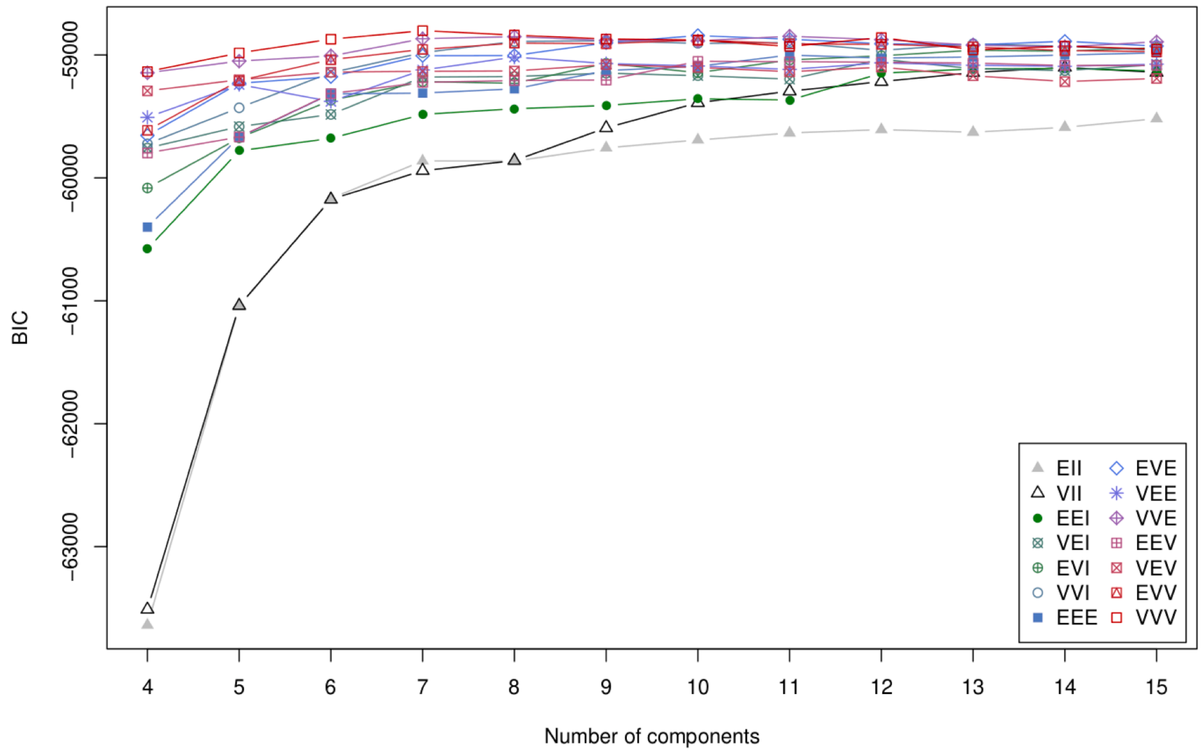

Fig. A.3. BIC analysis for weekends city sessions.

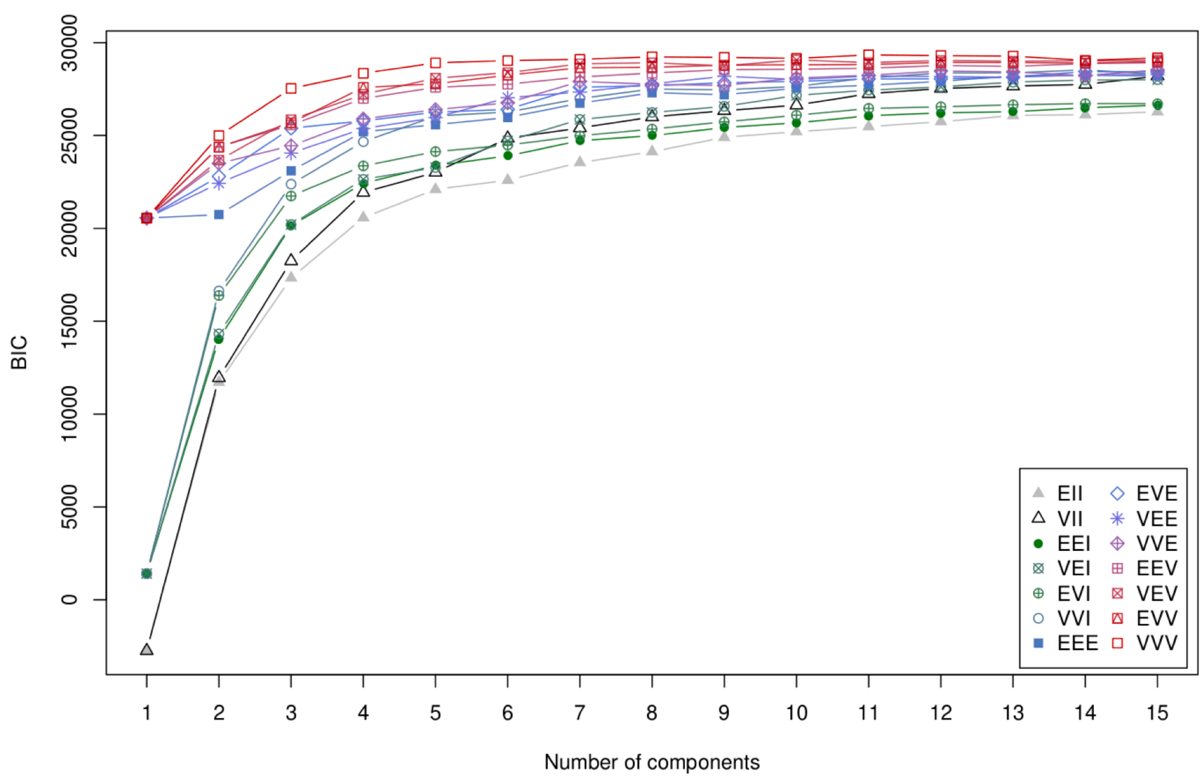

Fig. A.4. BIC analysis for weekends home sessions.

A.2. User profiles from clustering components

A.2.1. Weekdays city sessions

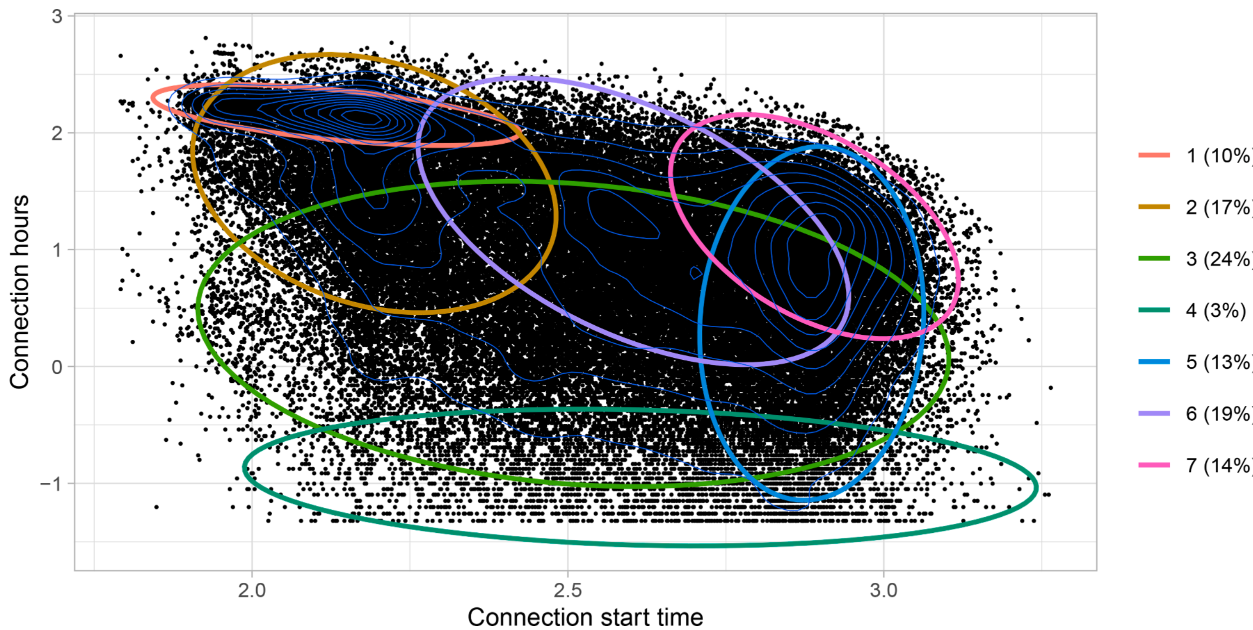

Fig. A.5. GMM clusters of weekdays city sessions.

Table A.1
Weekdays city clusters interpretation.

Cluster	Average start time	Average duration (h)	Interpretation	Profile
1	09:27	8.60	Full-day workers or visitors	Worktime
2	09:58	4.78	Visit the city during the morning	Visit
3	13:17	1.32	Short visits during the day	Shortstay
4	14:40	0.39	Super short connections during the day	Shortstay
5	18:55	1.45	Short visits during the evening	Shortstay
6	14:31	3.46	Visit the city during the afternoon	Visit
7	19:00	3.31	Go out for a dinner	Dinner

A.2.2. Weekdays home sessions

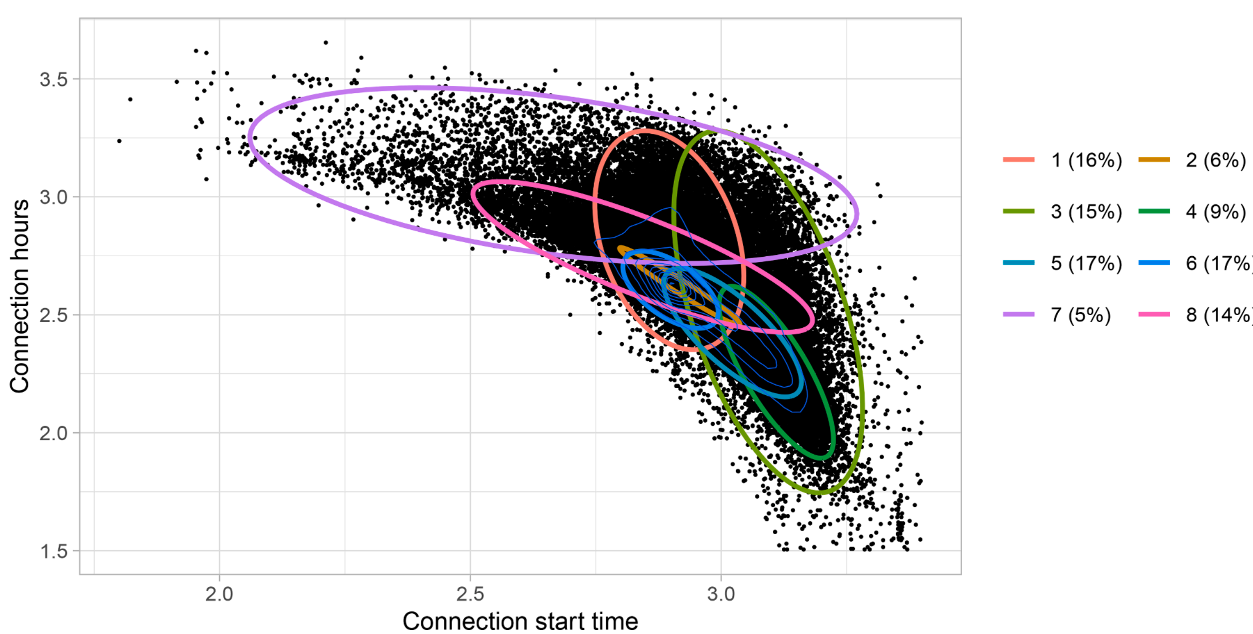

Fig. A.6. GMM clusters of weekdays home sessions.

Table A.2
Weekdays home clusters interpretation.

Cluster	Average start time	Average duration (h)	Interpretation	Profile
1	19:07	16.70	Go home during the afternoon, not necessarily leaving the next morning	Home
2	19:30	13.75	Always go home after work, always leaving the next morning	Commuter
3	23:04	12.33	Go home at night, not necessarily leaving the next morning	Pillow
4	23:25	9.55	Go home at late night, leaving the next morning	Pillow
5	21:33	11.29	Go home at night, leaving the next morning	Pillow
6	19:10	13.55	Always go home after work, always leaving the next morning	Commuter
7	15:23	21.97	Can go home anytime, not necessarily leaving the next morning	Home
8	18:11	15.57	Go home during the afternoon, leaving the next morning	Home

A.2.3. Weekends city sessions

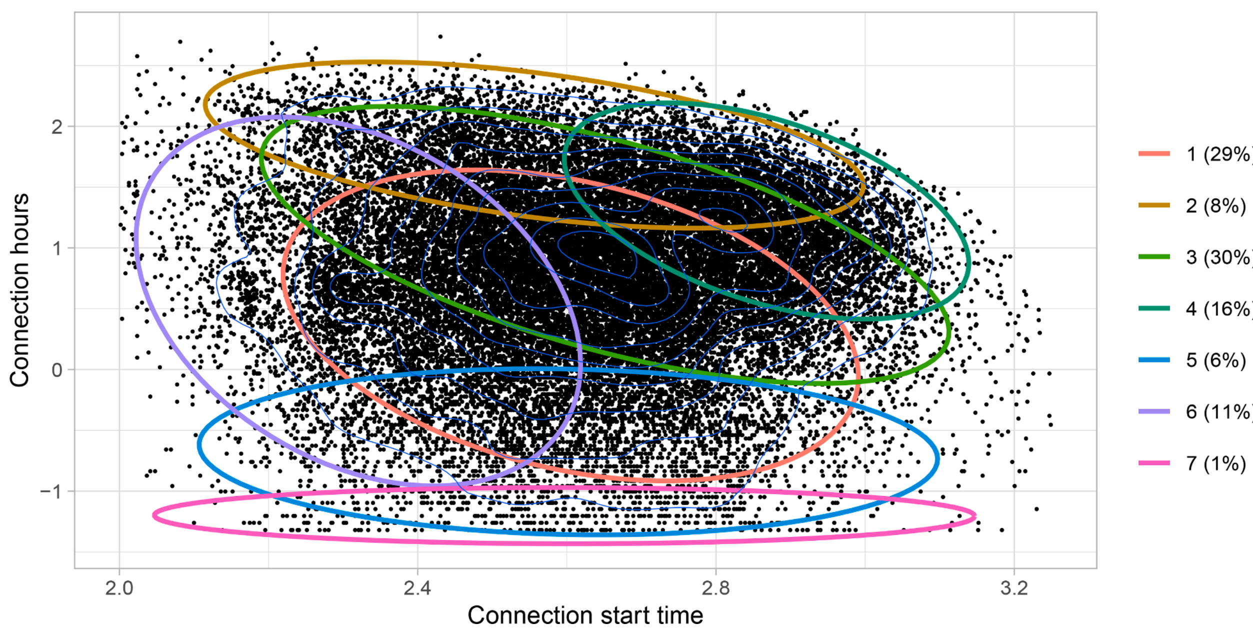

Fig. A.7. GMM clusters of weekends city sessions.

Table A.3
Weekends city clusters interpretation.

Cluster	Average start time	Average duration (h)	Interpretation	Profile
1	14:32	1.44	Short visits during the afternoon	Shortstay
2	13:53	6.34	Visit the city during the day	Visit
3	15:10	2.79	Visit the city during morning or afternoon	Visit
4	18:36	3.68	Go out during afternoon and probably dinner	Dinner
5	14:29	0.51	Super-short connections during the day	Shortstay
6	11:11	1.75	Short visits during the morning	Shortstay
7	14:25	0.30	Super-short connections during the day	Shortstay

A.2.4. Weekends home sessions

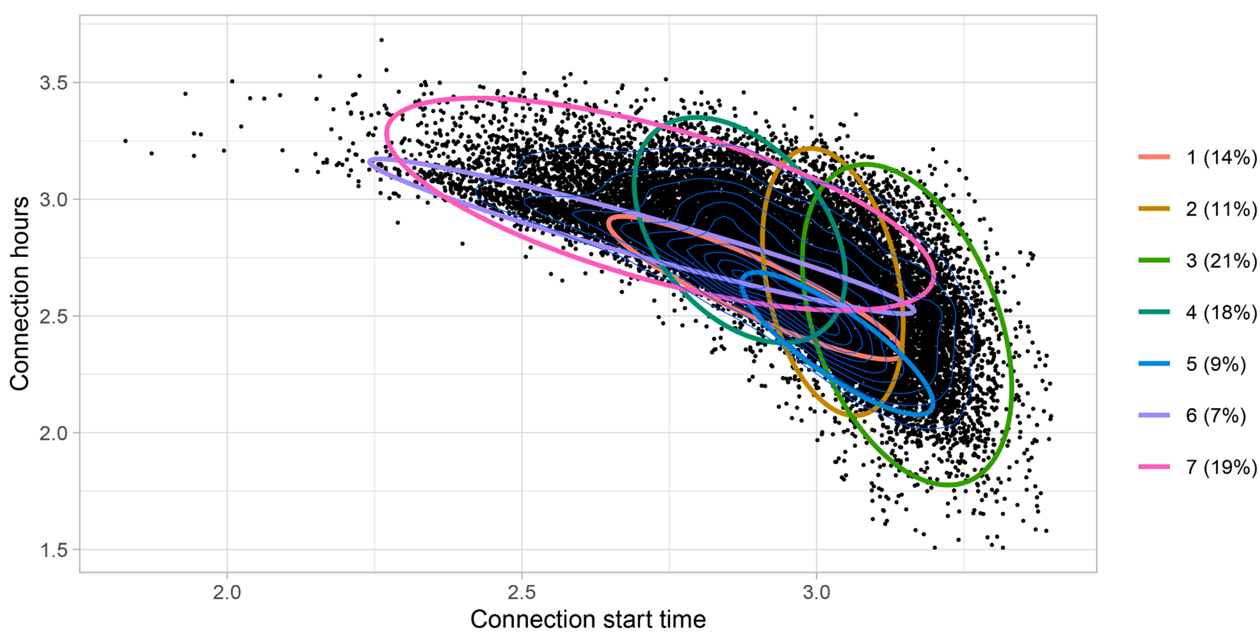

Fig. A.8. GMM clusters of weekends home sessions.

Table A.4
Weekends home clusters interpretation.

Cluster	Average start time	Average duration (h)	Interpretation	Profile
1	19:06	13.74	Go home at evening, normally leaving the next morning	Home
2	21:39	14.09	Go home at night, not necessarily leaving the next morning	Pillow
3	00:25	11.73	Go home at night, not necessarily leaving the next morning	Pillow
4	18:38	17.59	Go home during the evening, not necessarily leaving the next morning	Home
5	21:48	10.83	Go home at night, normally leaving the next morning	Pillow
6	15:56	17.13	Go home during the afternoon, normally leaving the next morning	Home
7	16:24	19.65	Go home during the afternoon, not necessarily leaving the next morning	Home

References

[1] Jairo Quirós-Tortós, Luis F. Ochoa, and Becky Lees. A statistical analysis of EV charging behavior in the UK. In 2015 IEEE PES Innovative Smart Grid Technologies Latin America, ISGT LATAM 2015, pages 445–449. Institute of Electrical and Electronics Engineers Inc., Jan 2016.
[2] Gan Lei, Chen Xingying, Kun Yu, Zheng Jiaxiang, Wei Du. A Probabilistic Evaluation Method of Household EVs Dispatching Potential Considering Users' Multiple Travel Needs. IEEE Trans. Ind. Appl. 2020;56(5):5858–67.
[3] Jing Zhang, Jie Yan, Yongqian Liu, Haoran Zhang, and Guoliang Lv. Daily electric vehicle charging load profiles considering demographics of vehicle users. Applied Energy, 274:115063, sep 2020.
[4] Schmidt Marc, Staudt Philipp, Weinhardt Christof. Evaluating the importance and impact of user behavior on public destination charging of electric vehicles. Appl. Energy Jan 2020;258.
[5] Calearo Lisa, Thingvad Andreas, Suzuki Kenta, Marinelli Mattia. Grid Loading Due to EV Charging Profiles Based on Pseudo-Real Driving Pattern and User Behavior. IEEE Transactions on Transportation Electrification 2019;5(3):683–94.
[6] Constance Crozier, Dimitra Apostolopoulou, and Malcolm McCulloch. Clustering of Usage Profiles for Electric Vehicle Behaviour Analysis. In Proceedings - 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT-Europe 2018. Institute of Electrical and Electronics Engineers Inc., Dec 2018.
[7] Crozier Constance, Morstyn Thomas, McCulloch Malcolm. Capturing diversity in electric vehicle charging behaviour for network capacity estimation. Transportation Research Part D: Transport and Environment Apr 2021;93:102762.
[8] Jean Michel Clairand, Javier Rodriguez Garcia, and Carlos Alvarez Bel. Smart charging for an electric vehicle aggregator considering user tariff preference. In 2017 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2017. Institute of Electrical and Electronics Engineers Inc., Oct 2017.
[9] Clairand Jean Michel, Rodriguez-Garcia Javier, Alvarez-Bel Carlos. Assessment of Technical and Economic Impacts of EV User Behavior on EV Aggregator Smart Charging. Journal of Modern Power Systems and Clean. Energy 2020;8(2):356–66.
[10] Sadeghianpourhamami N, Refa N, Strobbe M, Develder C. Quantitive analysis of electric vehicle flexibility: A data-driven approach. International Journal of Electrical Power and Energy Systems Feb 2018;95:451–62.
[11] Chris Develder, Nasrin Sadeghianpourhamami, Matthias Strobbe, and Nazir Refa. Quantifying flexibility in EV charging as DR potential: Analysis of two real-world data sets. In 2016 IEEE International Conference on Smart Grid Communications, SmartGridComm 2016, pages 600–605. Institute of Electrical and Electronics Engineers Inc., Dec 2016.
[12] Youssef El Bouhassani, Nazir Refa, and Robert Van Den Hoed. Pinpointing the Smart Charging Potential for Electric Vehicles at Public Charging Points. EVS32 Symposium, (May):1–12, 2019.
[13] Tang Difei, Wang Peng. Probabilistic Modeling of Nodal Charging Demand Based on Spatial-Temporal Dynamics of Moving Electric Vehicles. IEEE Transactions on Smart Grid Mar 2016;7(2):627–36.
[14] Kim Seheon, Yang Dujuan, Rasouli Soora, Timmermans Harry. Heterogeneous hazard model of PEV users charging intervals: Analysis of four year charging transactions data. Transp. Res. Part C 2017;82:248–60.
[15] Gerritsma Marte K, AlSkaif Tarek A, Fidder Henk A, van Sark Wilfried GJHM. Flexibility of Electric Vehicle Demand: Analysis of Measured Charging Data and Simulation for the Future. World Electric Vehicle Journal Mar 2019;10(1):14.
[16] Wolbertus Rick, Kroesen Maarten, Van Den Hoed Robert, Chorus Caspar. Fully charged: An empirical study into the factors that in fl uence connection times at EV-charging stations. Energy Policy 2018;123(August):1–7.
[17] Helmus Jurjen R, Lees Michael H, van den Hoed Robert. A data driven typology of electric vehicle user types and charging sessions. Transportation Research Part C: Emerging Technologies 2020;115(April):102637.
[18] Yingqi Xiong, Bin Wang, Chi Cheng Chu, and Rajit Gadh. Electric Vehicle Driver Clustering using Statistical Model and Machine Learning. In IEEE Power and Energy Society General Meeting, volume 2018-Augus. IEEE Computer Society, Dec 2018.
[19] Kriegel Hans-Peter, Kröger Peer, Sander Jörg, Zimek Arthur. Density-based clustering. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery May 2011;1(3):231–40.

[20] Labeeuw Wouter, Deconinck Geert. Residential electrical load model based on mixture model clustering and markov models. IEEE Trans. Industr. Inf. 2013;9(3):1561–9.

[21] Reza Fachrizal, Mahmoud Shepero, Dennis van der Meer, Joakim Munkhammar, and Joakim Widén. Smart charging of electric vehicles considering photovoltaic power production and electricity consumption: A review, May 2020.

[22] Kara Emre C, Macdonald Jason S, Black Douglas, Bérges Mario, Hug Gabriela, Kiliccote Sila. Estimating the benefits of electric vehicle smart charging at non-residential locations: A data-driven approach. Appl. Energy 2015;155(2015):515–25.

[23] Ioakimidis Christos S, Thomas Dimitrios, Rycerski Pawel, Genikomsakis Konstantinos N. Peak shaving and valley filling of power consumption profile in non-residential buildings using an electric vehicle parking lot. Energy Apr 2018;148:148–58.

[24] N.B.G.(Nico) Brinkel, M.K.(Marte) Gerritsma, T.A.(Tarek) AlSkaif, I. (Ioannis) Lampropoulos, A.M.(Arjan) van Voorden, H.A.(Henk) Fidder, and W.G.J.H.M. (Wilfried) van Sark. Impact of rapid PV fluctuations on power quality in the low-voltage grid and mitigation strategies using electric vehicles. International Journal of Electrical Power and Energy Systems, 118(November 2019):105741, 2020.

[25] Liu Nian, Chen Qifang, Liu Jie, Xinyi Lu, Li Peng, Lei Jinyong, Zhang Jianhua. A Heuristic Operation Strategy for Commercial Building Microgrids Containing EVs and PV System. IEEE Trans. Industr. Electron. 2015;62(4):2560–70.

[26] Schuller Alexander, Flath Christoph M, Gottwalt Sebastian. Quantifying load flexibility of electric vehicles for renewable energy integration. Appl. Energy Aug 2015;151:335–44.

[27] Cao Yan, Huang Liang, Li Yiqing, Jermsittiparsert Kittisak, Ahmadi-Nezamabad Hamed, Nojavan Sayyad. Optimal scheduling of electric vehicles aggregator under market price uncertainty using robust optimization technique. International Journal of Electrical Power and Energy Systems 2020;117(November 2018):105628.

[28] Dennis Van Der Meer, Gautham Ram Chandra Mouli, German Morales Espana Mouli, Laura Ramirez Elizondo, and Pavol Bauer. Energy Management System with PV Power Forecast to Optimally Charge EVs at the Workplace. IEEE Transactions on Industrial Informatics, 14(1), 311–320, 2018.

[29] Eldeeb Hassan H, Faddel Samy, Mohammed Osama A. Multi-Objective Optimization Technique for the Operation of Grid tied PV Powered EV Charging Station. Electric Power Systems Research Nov 2018;164:201–11.

[30] Shaolun Xu, Yan Zheng, Feng Donghan, Zhao Xiaobo. Decentralized charging control strategy of the electric vehicle aggregator based on augmented Lagrangian method. International Journal of Electrical Power and Energy Systems 2019;104 (April 2018):673–9.

[31] Zhang Liang, Kekatos Vassilis, Giannakis Georgios B. Scalable Electric Vehicle Charging Protocols. IEEE Trans. Power Syst. Mar 2017;32(2):1451–62.

[32] Fraley Chris, Raftery Adrian E. Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association 2002;97(458):611–31.

[33] Fraley Chris, Raftery Adrian E. Enhanced Model-Based Clustering, Density Estimation, and Discriminant Analysis Software: MCLUST. J. Classif. Jan 2003;20(2):263–86.

[34] Martin S Andersen, Joachim Dahl, and Lieven Vandenberghe. CVXOPT: A Python package for convex optimization. Technical report, 2020.