Energy & Buildings 271 (2022) 112312

Contents lists available at ScienceDirect

Energy & Buildings

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

Understanding patterns of thermostat overrides after demand response events

Valentina Tomat [a], Marika Vellei [b], Alfonso P. Ramallo-González [a,*], Aurora González-Vidal [a], Jérôme Le Dréau [b], Antonio Skarmeta-Gómez [a]

[a] Universidad de Murcia, Murcia, Spain
[b] LaSIE (UMR CNRS 7356), La Rochelle University, La Rochelle, France

ARTICLE INFO

Article history:
Received 22 March 2022
Revised 7 July 2022
Accepted 8 July 2022
Available online 15 July 2022

Keywords:
Thermal comfort
Occupant behaviour
Demand response
Rebound
Space cooling
Clustering

ABSTRACT

Demand Response (DR) strategies represent an innovative option to optimise energy management. In particular, smart thermostats have captured the attention of the scientific community for their effectiveness in achieving energy-saving and peak-shaving by lowering HVAC consumption during critical hours of the year. One way of achieving this aim, is to leave the control of the smart thermostat to a third party for the duration of the DR event in the so-called Direct Load Control (DLC) configuration. Most research focuses on thermostat overrides during DR events; in this work, we use real world data from the Donate Your Data dataset to analyse the interaction of users with the thermostat around the DR event. In particular, this work focuses on users that interact with the thermostat before (anticipative behaviour) or during the DR event (reactive behaviour), leading to a lower efficiency of the load control. Through clustering techniques, different categories of users are identified, and some significant cases are simulated on a building energy simulation tool to quantify the missed power reduction and the impact on energy. The study highlights that the behaviour of some users can reduce or even nullify the efficacy of the DLC strategy. In light of the findings and to prevent this issue, we suggest the need for tailored DR events for different archetypes of users as identified in this work through clustering.

1. Introduction

The current electricity infrastructure is designed based on peak loads, which varies with the climate, the type of usage and the penetration of local production systems. Electricity demand load changes throughout the day and the year according to users' demand, creating a peak in periods of maximum demand. This has consequences at both building and urban scales. The single building can experience power outages and blackouts. Even worse, the single-building peaks sum up with neighbour ones, creating a peak demand at the urban scale, which can lead to a collapse of the grid in the long term [1]. During extreme weather episodes (heat waves, for instance), this issue is even amplified.

Because of the recent increase in power demand (e.g., for the penetration of fast charging electric vehicles [2], but also for the proliferation of electric heat pumps), the maximal load needed now is greater than the one given by the natural renovation of the grid driven by the natural growth of the cities. Loads are becoming 'peakier' and those peaks are more frequent. Even, the situation will become more and more serious in the next decades. Electricity demand is expected to increase by up to 509 % from 2007 to 2050 depending on the country [3]. Maintaining the stability of the supply should be a priority all over the world, but it implies continued investments to achieve the modernisation and the expansion of the actual grid, mainly based on an increase of the capacity, and consequently an increase in the number of hours at which the grid is underused. To make a more efficient system, a reshaping of the peaks over time is needed; and it can be seen as an effective solution to achieve good power management. The ensemble of the mechanisms of shaving peak loads, i.e. reducing the consumption during peak times or shifting it to off-peak times, is called Demand Response (DR). DR programs are considered as part of the Demand-Side Management and are strictly related to customers' engagement and awareness. They are highly encouraged by both energy companies and regulatory authorities to exploit the potential of demand-side flexibility [4].

Direct Load Control (DLC) is a type of DR program that consists in altering directly a customer's energy consumption based on an event issued by a third party [56789]. In particular, direct load con-

* Corresponding author.
E-mail address: alfonsop.ramallo@um.es (A.P. Ramallo-González).

https://doi.org/10.1016/j.enbuild.2022.112312
0378-7788/

trol of the HVAC system is one of the most used incentive-based strategies. Smart thermostats, necessary in order to plan a DLC strategy of the HVAC system, have captured the attention of the scientific community for their effectiveness in achieving energy-saving and peak shaving while maintaining occupants' thermal comfort [10111213]. The users subscribing to the DR program are asked to relinquish control over their smart thermostats, i.e. over the setpoint temperatures that govern their thermal environments, in exchange for bill credits. When a peak demand is expected, the third party takes control over the programmed schedules, increasing or decreasing (depending on the season) the setpoint temperature. This occurrence is referred to as Demand Response Event (DRE). Such management of the setpoint temperature can lead to a substantial reduction in energy consumption, preventing the peak demand from occurring (Fig. 1). Users are an active part of the loop, as they are requested to accept changes of their normal patterns to enhance a better distribution of their consumptions [5].

DREs can have different timescales depending on the degree of urgency the utility company is facing. For what concerns just the HVAC system, the conventional building demand management strategy consists of a low-impact change in the setpoint temperature in order to maintain occupants' comfort. For example, in the hottest days of summer, diminishing the cooling setpoint temperature by a few degrees for the warmest hours of the day, may be enough to avoid the peak load. Nevertheless, sometimes the power imbalance of the grid leads to a more urgent solution. In such cases, an immediate demand reduction is needed, and the HVAC system is shut down for a very short period, of the order of minutes (the so-called ancillary service) [14]. In this article, we will focus on the long DREs, that affect the indoor environment over more than an hour.

Direct control over the thermostats carries the risk of affecting occupants' thermal comfort. It is worth specifying that DREs can be optional or mandatory. In the first case, if an occupant starts to be uncomfortable during a DRE, they can stop it and that is called an *override*. In the second case, customers cannot override the DR event once subscribed to the program. From the literature, it is clear that the majority of DR initiatives consists of interruptible programs [15] since the possibility of stopping an event increases largely the willingness of the users to subscribe to a DR program [1617]. During a DRE, the only way an occupant has to express their thermal discomfort is by adjusting the setpoint temperature, i.e. overriding the event.

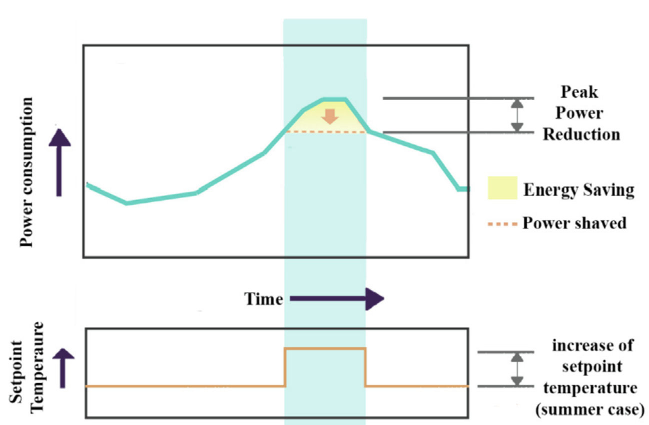

Fig. 1. Sketch of the concept of peak shaving in a Demand Response event. The expected power consumption (turquoise line) is shaved over the duration of the event as a consequence of a change in the setpoint temperature (orange line). For the colour version of the picture the reader is referred to the online version of the paper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Studies in the literature point to different findings of how people react to DREs. In some studies, DREs are found to be almost unperceived by occupants [18]. This is in contrast with other works that analysed the rate of overrides, obtaining quite high results, up to 39 % [19]. Sweetnam et al. [20] analysed this discrepancy through a winter experiment of demand shaping. Although the experiment was characterised by a very modest change in temperature (within ± 2 °C), the feedback received from the occupants revealed a low acceptance of the DRE. Such a conclusion is in line with the theory that the thermal dissatisfaction of individuals is often underestimated [21]. This, together with the fact that an override implies a loss on the effectiveness of the demand response event, highlights the importance of studying overrides during DREs.

Overriding an event has repercussions on the average power reduction, as the total duration will be shorter than expected. We will refer to the difference between the expected power reduction and the effective power reduction as missed power reduction. We have evaluated the magnitude of the missed power shaving depending on both the duration of the override and the change in the setpoint temperature during the override. However, overrides are not the only threat to power consumption. As explained by Conejo et al. [22], if many users participate in the same DR event, the original peak will be shaved but another peak could be created afterwards. This phenomenon is known as the *rebound effect,* and it is due to postponing consumers' consumption to the period after the event [23]. The increase in power consumption due to a rebound can extremely reduce, and even nullify, the power reduction obtained through the DRE [24]. Studies on the rebound effect are often limited to the power increase due to the abrupt increase in HVAC load when brought back to normal operation, right after the DR event [25262728], but the increase in power consumption is also affected by the changes in the setpoint temperature chosen by the users after the event (additional rebounds). To better analyse and quantify the rebound effect, it is important to understand if, after the DR event, the user wants the same temperature that they had before the DRE or a colder/warmer one (summer case/winter case respectively). Ultimately, it needs to be answered if their perception of comfort is different compared to the hours that preceded the DR event.

The rebound effect as a shift in peak-load has been noticed in several experiments, but mostly focusing on the power rebound right at the end of the DR event [29]. The multiple rebounds that can be observed in the following hours (due to additional setpoint changes) has not been commonly studied in the literature. In a recent experiment by Christensen et al. [30], that was conducted in winter, the rebound effect was reported several times in the hours after the DR event. Christensen et al. noticed an increase in the heating power due to the shifting of the demand. The rebound effect phenomenon was observed not just in long-term-DREs, but also in faster events. However, the rebound effect was not fully characterised in terms of occupants' behaviour. In the study proposed by Broka and Baltputnis [31], the rebound effect is only covered from a financial point of view. Moreover, the rebound cases considered, are not real but simulated, analysing several scenarios from the supplier's point of view while the consequences on the energy consumption are not examined.

To the best of the authors' knowledge, a comprehensive study of the multiple rebounds after the users' thermostat overrides and the temperature preferences in the period before and after a DRE has not been carried out yet. This paper aims to cover such a gap in the literature by focusing on how the DREs affect occupants' thermostat usage behaviour before and after the event and on the energy consequences of this behaviour.

Our study is based on the Donate Your Data (DYD) dataset [32]. To help scientists investigate how occupants engage with their thermostats, the thermostat manufacturer Ecobee created the DYD campaign, that collects anonymised thermostat use data directly from the users. Several authors have shown interest in the DYD dataset, approaching it from different points of view and perspectives [33343536]. About users' interaction with the thermostat, Kane and Sharma [37] studied the time that it takes for a user to override a setpoint schedule when uncomfortable. They explained that every manual setpoint change was due to complex interdependencies between the building physics, the users' physiology, the social cognition, the type of thermostat and the actions users are willing to take. The interesting conclusion of their work was that more than half of manual setpoint changes were not a consequence of a programmed setpoint change, i.e. users were overriding decisions they previously made. It has been further observed that indoor temperature control depended on how the users supposed the thermostat worked (their mental model), more than on the actual thermostat's functioning [38394041]. Sarran et al. [29] used the DYD database to study overrides during DR events and found an overall override rate of 12.9 %. They observed a negative impact of households' overrides on the power demand reduction's potential of the DR event. It was further noticed that people that frequently override their scheduled setpoint during normal days are more likely to interact with their thermostat also during a DR event. This conclusion suggests that the overrides cannot be related just to environmental and physical factors: habits play a key role in their prediction. A decision tree analysis confirmed the influence of habitual setpoint change behaviour of the users on the DR overrides.

Using the database of the Ecobee DYD program, this paper aims to study the implications of manually modified setpoint temperatures before and after the DRE, in particular focusing on the behaviour of occupants who showed less acceptance of the DRE. Hence, both the users who interrupted the DRE and those who adjusted the setpoint temperature before DRE are included in the analysis. In other words, our work is designed to investigate all the DREs that did not end with a scheduled setpoint. We do this to have a complete vision of the effects of user behaviour on the DR's effectiveness. In the rest of the paper, both types of users will be referred to as *overriders*, since they override the programmed schedule and, in both cases, the thermostat does not return to the scheduled setpoint value. Hence, in our analysis, we propose a classification of overriders through a data-driven clustering analysis, that allows us to learn more about the groups of people that behave similarly, and with that information, improve the design of demand response events. In Section 2, the approach used for the analysis of the database and the methodology used for simulating the different behaviours are explained; in Section 3, results are presented; in Section 4 the main conclusions of the study are delineated.

2. Methodology

2.1. The dataset

The DYD campaign allowed Ecobee's users to share their thermostat data, after a necessary process of anonymisation that removes any personal information. The DYD dataset includes both data and metadata and it is updated quarterly with new participants. Metadata refers to home and householders' characteristics and it includes information such as the home identifier (as an anonymised code), the thermostat model, the country, and the number of occupants. Data refers to the information collected by the Ecobee thermostats (in 5-minutes intervals) and it includes, among others, date, time, schedule selected, event occurrence, indoor temperature, setpoint temperature, relative humidity, outdoor temperature and motion detection.

To understand the analysis presented in this paper, it is useful to first understand how the Ecobee thermostats work, especially during a DRE. When choosing the settings for the thermostat, users have to define a daily profile temperature, by setting different setpoint temperatures at different times of the day through *schedules*. There are some default schedules (*awake*, *away*, *home* and *sleep*) and, in addition, users can create customised schedules. Users are asked to choose a starting hour and a setpoint temperature for each one of the schedules. Any action that modifies the schedule is called *event*. Examples of thermostat-controlled events are *smart recovery*, *smart home*, *smart away* and, of course, *Demand Response events*. Hence, the timesteps in which a DRE occurs are clearly shown in the dataset.

For our analysis, we only used a subset of the DYD dataset made of 13,145 homes (data collected up to 2019). This includes 11,327 (86 %) Canadian homes and 1818 (14 %) US homes. Among the whole dataset, 1398 homes signed up to participate in the demand response events program (1156 from Canada and 242 from the United States). As we wanted to study DREs in the summer period, we decided to first remove winter months from the dataset, as well as homes that did not have an active cooling system. The dataset was filtered according to the following conditions:

- mean monthly outdoor temperature higher than 15°C;
- presence of an active cooling system.

After this filtering, 990 thermostats were left for the study. Among them, 1509 DR events were detected (1.52 events per thermostat on average). Note that in this part of the section we are considering all the DR events of the dataset, including both the ones followed by a scheduled setpoint temperature and the ones with user-adjusted setpoint return.

In our analysis, we chose not to consider occupancy measurement, since some inconsistencies have been observed in the data. In the most common situation this was related to a privacy concern: some occupants did not share information about occupancy. Besides, the literature has shown that a considerable percentage of false detections are registered when it comes to evaluating the occupancy of homes [42]. It is called false positive the case in which the motion sensor detects human presence, but the users are not actually at home, and it is called false negative the case in which at least one user is occupying the house, but the motion sensor is not able to detect this presence. The percentage of false and true detections depends on the detection mechanism [43], being the passive infrared (PIR) motion sensors more reliable in detecting presence than absence. For this reason, the interpretation of the results that consider the 'Occupancy' parameter was excluded.

As anticipated in the Introduction, Sarran et al. [29] focused on the DR events that were interrupted by the users before their natural end. Notwithstanding, their conclusion led us to notice that other users affect the missed power reduction due to the DRE without interrupting it but by changing the setpoint before the event. Indeed, it has been observed in many cases that, prior to the DRE, the setpoint temperature was not the scheduled one, but one chosen by the user through a manual change as in preparing for the event. As the users knew when a DRE is due to start (day-ahead notification), they could adjust the thermostat setpoint temperature before the beginning of the event, in order to mitigate its effect on the indoor temperature. We called this manual change to the programmed schedule to anticipate the DREs a *Hold Action* (HoA). In general, any user's manual adjustment of a scheduled program can be referred to as *Hold* and, depending on how the user has configured the thermostat, the action has a certain duration, that can be fixed (a few hours) or indefinite (a manual interaction is then needed to return to the scheduled program). Consequently, occupant behaviour can affect the DRE's effectivity and the overall missed power reduction in two occurrences: (1) the user chooses to interrupt a DR event, interacting with the thermostat and affecting the duration of the DRE with a User Adjustment (UA), and (2) the user adjusts the setpoint temperature before the DRE with an HoA, affecting automatically the setpoint temperature at the end of the event because the thermostat falls back to that value when the DR event is over. To our knowledge, this is the first work that contemplates this particular, but relevant case. In our analysis of the DYD dataset, we designate as Adjusted Demand Response Events (ADREs) the DR events ending up with a setpoint different from the scheduled one, i.e. cases (1) and (2), and we distinguish them from the Non-Adjusted Demand Response Events (non-ADREs) that return to the programmed schedule (see Fig. 2 for an overview). Finally, we are also interested in analysing and studying the rebound effect in the post-event period, so we defined rebounds in the dataset as the decreases in the setpoint temperature set by the users after the end of the DRE. Thus, we isolated a subgroup of the ADREs that are followed by one or more rebounds (i.e., further adjustments) and we called them ADREs + R (Fig. 2).

2.2. Clustering of post-DRE set-point profiles

Analysing the setpoint temperature series by eye, we noticed that somehow people's behaviour followed some patterns. Hence, we used clustering in order to be able to classify the different profiles according to criteria that cannot be distinguished by eye.

Clustering is one of the most popular data mining techniques. It consists of grouping a set of observations by means of their similarity. As a result, the observations within clusters are more similar to each other than observations between clusters according to the chosen criteria.

The clustering will be used to characterise the setpoint profiles after the DREs. To focus on the problem at hand, we selected from the dataset the subgroup of ADREs + R as defined in the previous section. Thus, we excluded the group of users that, once selected their user-adjusted setpoint temperature for the post-event period, do not interact anymore with the HVAC, i.e. cases without additional rebounds. We applied *agglomerative hierarchical clustering* [44], which starts assigning each value to its cluster and then it proceeds iteratively. Thus, at each stage, it joins the two most similar clusters using a measure of similarity, until there is just a single cluster. The hierarchical strategy was chosen based on the fact that the times series had different lengths, and we could visually inspect how the number of clusters would affect the groupings. The strategy does not require variables to be continuous, contrary to other well-known methods such as k-means. To decide which clusters should be combined, a measure of dissimilarity between sets of observations is required. In most methods of hierarchical clustering, this is achieved by the use of an appropriate metric (a measure of distance between pairs of observations), and a linkage criterion that specifies the dissimilarity of sets as a function of the pairwise distances of observations in the sets. Since our purpose was to classify the overriders' profiles, we decided to use a shape-based metric named k-Shape [45] to group the users by the shape of the curve that the setpoint presents. K-Shape uses a normalised version of the cross-correlation measure, in order to consider the shapes of time series while comparing them. After that, it creates homogeneous well-separated clusters. In order to compute the pairwise distances of observations in the sets, we used the Unweighted Pair Group Method with Arithmetic mean (UPGMA) criterion [46]. Using this, the distance between two clusters is defined as the average distance between each point in one cluster to every other point in the other cluster.

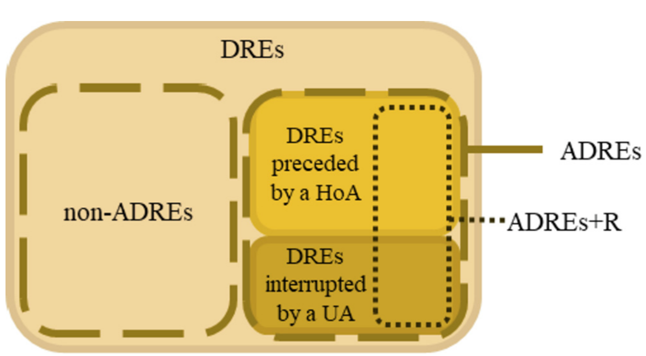

Fig. 2. Sketch of the concept of Adjusted Demand Response Events (ADREs), the demand response events whose effectivity is affected by human behaviour. In the picture, HoA = Hold Action, UA = Users' Adjustment.

2.3. Simulation

To quantify the missed power reduction due to overrides and rebounds, we created a model on EnergyPlus [47] using real building geometries from the metadata available in the dataset. The simulation refers to the cooling energy need during the summer of 2019 in a Canadian home (i.e. the efficiency of the cooling systems was not accounted for). To choose the dwelling used for the model, we isolated homes that were subjected to the same DR event and we chose one case from each identified cluster. Among them, we chose the case that was more complete in terms of metadata available, from which we extracted the geometric parameters (total area of the house = 100 m^2), the number of occupants (3 occupants) and the number of floors of the building (one floor); then, we applied the same scenario to the rest of the simulated cases.

Information about the building fabrics was not available in the dataset. Hence, we acquired the construction elements of the envelope from the default ASHRAE models, available in the OpenStudio's library. In particular, the thermal transmittance of the roof and the external walls correspondent to Climate zone 6 were respectively 0.186 W/m^2K and 0.426 W/m^2K.

For the weather data, the *epw* file of Toronto was first selected [48]. The file was then edited using outdoor temperature values from the dataset [49] to diminish potential disparities due to specific year conditions. According to the DYD handbook, the outdoor temperature of the Ecobee dataset was obtained from the nearest weather station. Also, since we used the same model with different setpoint temperature profiles to show the difference from one another, this uncertainty did not have much effect on the final comparison.

Once the model was created, the real data from the DYD dataset was used to calibrate the model in terms of other factors such as thermal mass or internal heat gains from equipment (gains for the equipment set to 1.5 kW, number of people set to 3 during occupied period). The setpoint temperature used in EnergyPlus is taken from the DYD dataset. Fig. 3 shows an example of the comparison between the real data and the simulation in terms of indoor temperature. Following the ASHRAE Guideline 14 [50], the model was calibrated through the Normalized Mean Bias Error (NMBE) and the Coefficient of Variation of the Root Mean Square Error (CV-RMSE). In this case, the NMBE was equal to 1.12 %, while the CV-RMSE was equal to 1.95 %, so the model can be considered validated with respect to the internal temperature.

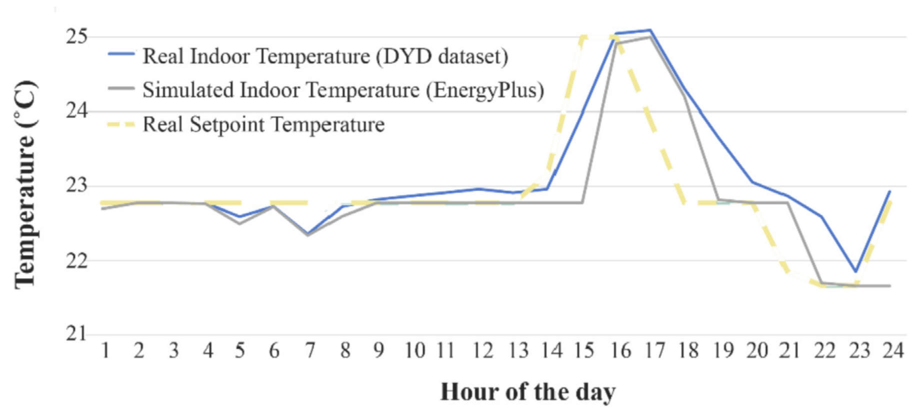

Fig. 3. Comparison between the indoor temperature extracted from the dataset and the indoor temperature obtained through EnergyPlus. The blue line represents the real trend of the indoor temperature, taken from the DYD dataset. The turquoise line represents the indoor temperature obtain as output of the EnergyPlus model. The yellow line shows the setpoint temperature, taken from the real data and used as input in the energy simulation. For the colour version of the picture the reader is referred to the online version of the paper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The model was used to calculate the power consumption during a day in which a DRE occurred and during a day without any event. The missed power reduction obtained was then compared to other cases (one for each cluster) in which the occupants overrode the event and in which rebounds were detected. It is important to have in mind that the aim of the simulation is not to assign a qualitative value to the clusters, but rather to estimate the impact of different users' behaviour on the DRE's effectiveness.

3. Results

3.1. Overview of all DR events

We first concentrated on studying the DREs present in the dataset, in order to have a more complete understanding of the phenomenon before focusing on the overriders' behaviour in the post-event period. Through a visual inspection of the dataset, we detected that, concerning the evolution of temperature throughout the DR event, three main cases could be found, and they have been sketched in Fig. 4. In the first case, the setpoint temperature remained constant throughout the whole duration of the event (Fig. 4.a). In the second case, the setpoint temperature's increased gradually during the event, from a cooler to a warmer temperature (Fig. 4.b). This case suggests an attempt to avoid an abrupt temperature change. In the third case, one can clearly distinguish a first phase in which the setpoint temperature decreases (Fig. 4.c). This evolution of temperature suggests an attempt of cooling down the home to prevent the warm discomfort caused by the DR event. This first phase of preparation is called precooling, while the second phase, the *real* DRE, is often referred to as the setback stage in the literature. Among the 1509 DR events considered, 452 were preceded by a precooling, which means approximately 30 % of the total, while 237 (about 16 %) are characterised by an increasing setpoint temperature. While the remaining majority is the constant case (54 %).

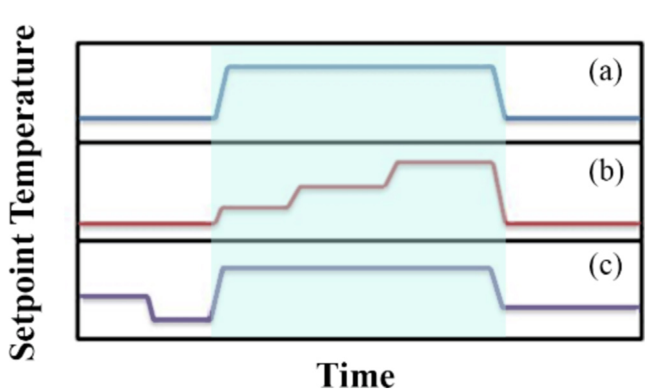

Fig. 4. Sketch of the different evolution of the setpoint temperature during the DR event as designed by the aggregator. a) Constant in time. b) Increasing intensity. c) With Precooling. The light blue timeframe represents the DRE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In the analysis, we focused on all the aspects that can help to describe DR events: when they start, when they end, how long they last, how many DR events can be experienced in the same house, which is the reaction of the users, and so forth. Beginning with the starting time of the day, the great majority of events started between 14:00 and 15:00, with the most recurrent value being 14:30. Also for precooling events, most of the events started between 14:00 and 15:00. Thus, on average the precooling event starts half an hour before the actual DR event. Precooling stages of more than 50 min are also common, and they are likely to be related to the events that started between 15:00 and 16:00. For what it concerns the temperature-increase phase of the DRE, the most recurrent duration is between three hours and three hours and a half, being the mean value of 151 min (Fig. 5.a). As the histogram shows, there are events whose duration exceeds five hours. The two peaks in the histogram (at 120 min and 180 min) represent the most frequent DREs, i.e. events that are sent simultaneously to a greater number of homes. Besides the time, the other characterizing factor for the DR events is the temperature. An increase in the setpoint temperature of 2.2C when the DRE starts is observed as the most recurrent (Fig. 5.b), especially in the events that are constant in time (purple histogram).

Fig. 6 shows the mean values of outdoor temperature, indoor temperature and setpoint temperature during a DRE. The mean outdoor temperature during the DREs is 28C. The indoor temperature's histogram presents a more symmetric distribution, being the mean value equal to 24.4C, which is a reasonable value for a Canadian home in summer. According to the ASHRAE ranges of thermal comfort [51], this is a value that could lead to a slightly cold thermal environement but that can be considered comfortable. When it comes to analysing the setpoint temperature during summer DREs, the graph is less intuitive. Most of the events are characterised by temperatures between 25 and 27C, but higher values are also recurrent. The higher values are likely to be used as set-back temperature, set when active cooling is not necessary although the

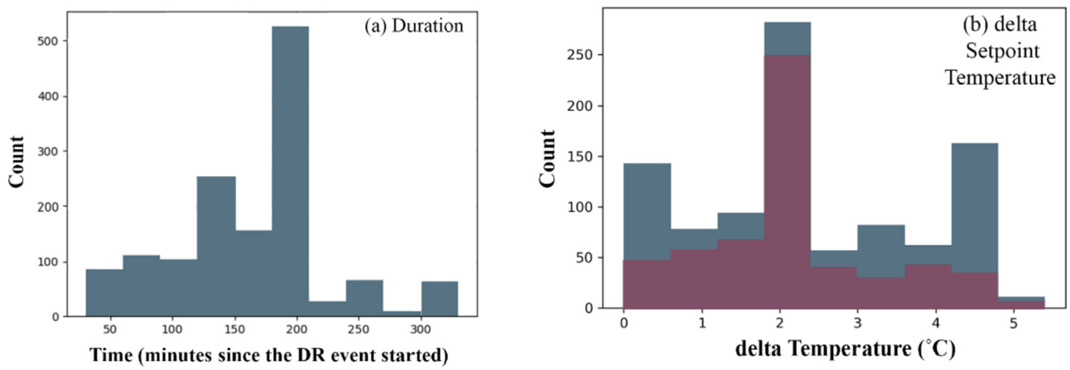

Fig. 5. a) Histogram that represents the duration of the Demand Response events; b) Histogram that represents the increase of the setpoint temperature when the DRE starts. The blue histograms refer to all typologies, the purple histogram refers to the type (a) from Fig. 4 – DREs constant in time. For the colour version of the picture the reader is referred to the online version of the paper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

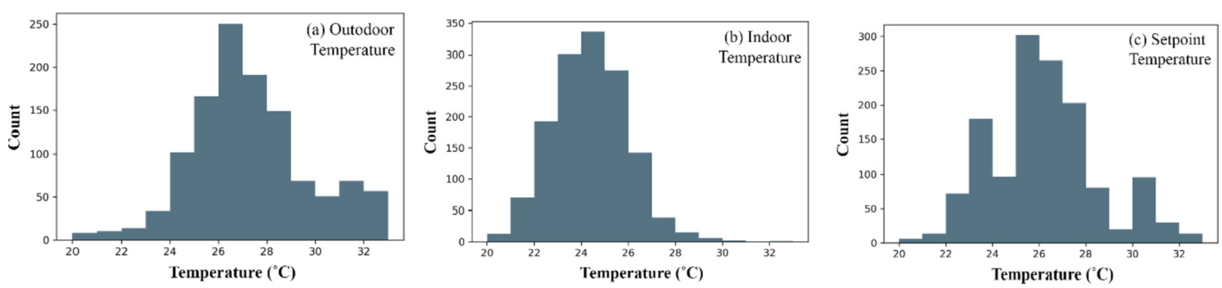

Fig. 6. a) Histogram of the mean outdoor temperatures during the DR events; b) Histogram of the mean indoor temperatures during the DR events; c) Histogram of the mean setpoint temperatures during the DR events.

HVAC is turned on. They are used in particular for the Schedule *Away*, but they have also been detected as a last phase of the DRE in the cases of increasing setpoint (Fig. 4.b). The mean value of the setpoints during the DRE was seen to be 26.3C.

3.2. Adjusted DR events either by HoA or DR interruption

Over the 1509 DREs considered, 38.5 % of them (581) were ADREs. It is common to send the same DREs to several homes at the same time. Hence, to have an idea of the number of ADREs per single event, we isolated the DREs that were concurrently sent to a larger number of dwellings:

- 20–08-2019 (314 ADREs out of 976 DREs, 32 %);
- 16–08-2019 (30 ADREs out of 128 DREs, 23 %);
- 12–07-2018 (21 ADREs out of 55 DREs, 38 %);
- 30–08-2019 (19 ADREs out of 54 DREs, 35 %);
- 21–08-2019 (17 ADREs out of 53 DREs, 32 %).

We will consider the totality of ADRE in the analysis, but studying more in detail this subset of ADREs, we noticed that only an average of 13 % of them are interrupted earlier by a users' adjustment (as seen in previous research by Sarran et al. [29]) while the rest are the ADREs preceded by an HoA. It is also interesting that, apart from the interruptions during the DRE, 16 manual interruptions are registered during the precooling phase (around 3 % of the DR with precooling).

3.2.1. Behaviour during the DR event

Among the ADREs, we analysed the duration of the DR event. Fig. 7 shows the distribution of the cumulative percentage of DR event duration. The abscissa axis represents the total duration of the DR event, where every bin represents the upper limit of the interval (time since the event started). The graph shows a net change of the trend at the bin corresponding to 180 min. This can be explained by considering the biggest peak in the histogram in Fig. 5, where most programs are designed to end after 3 h of activation.

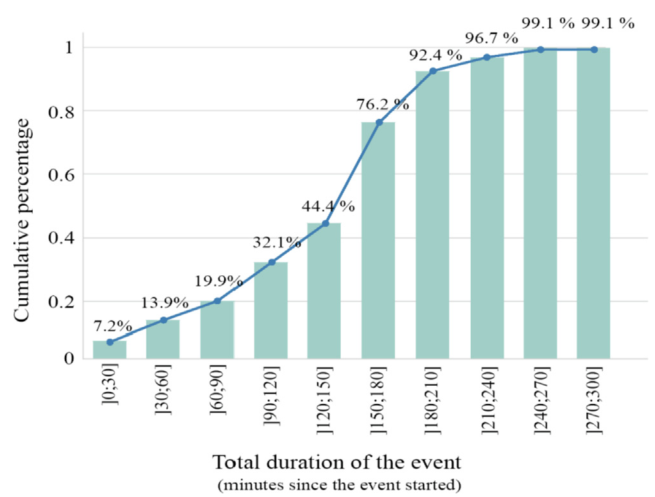

Fig. 7. Distribution of the percentage of adjustments probability at different time bins (among ADREs). Interval notation of the x-axis: reversed bracket indicates that the endpoint number is excluded from the interval, normal bracket indicates that the endpoint number is included in the interval.

Focusing on the indoor temperature reached during the first three hours of the event, Fig. 8 shows a comparison between the ADREs and the non-ADREs, i.e. those DR events that were allowed to run completely and returned to a scheduled setpoint (Fig. 2). The reader is reminded that the amount of data that composed the two

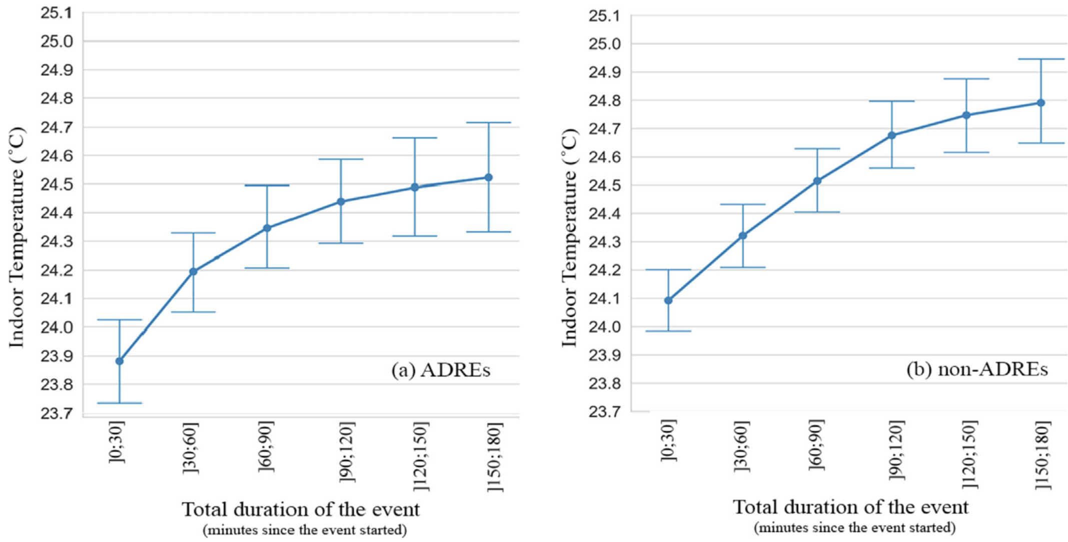

Fig. 8. Mean indoor temperature in the first three hours since the event started a) ADREs b) non-ADREs. The error bar represents the standard error with a 95% confidence level. Interval notation of the x-axis: reversed bracket indicates that the endpoint number is excluded from the interval, normal bracket indicates that the endpoint number is included in the interval.

subsets is different (581 events in a) and 928 in b)). In Fig. 8, the error bar represents the standard error, which measures the statistical accuracy of an estimate (not to be confused with the standard deviation, which measures the dispersion of the data around the mean value). The standard error is strictly related to the sample size: the bigger is the dataset, the smaller is the error. The difference in the indoor temperature when the event started (approximately 0.2 ℃ cooler in the ADREs subset) is mainly due to the adjustments due to a manual change (lowering of the setpoint temperature) prior to the event (HoA). If we focus on the delta temperature, and not on the absolute values, the increase for the two subsets is similar (0.6 ℃ for ADREs, 0.7 ℃ for the non-ADREs). This consideration suggests that the increase in the indoor temperature is not a primary cause of the adjustments, as confirmed in [29].

3.2.2. User behaviour after the DR event

As the other focus of our analysis is the post-event period in the ADREs, it was seen interesting to investigate the setpoint temperature detected right after the event, when the schedule is not followed (for both UA and HoA). In this analysis, we worked with a dataset of 528 samples (slightly reduced compared to the total of 581 events), since anomalies were detected during some DR events, probably due to internet connection failures.

The distribution of the difference in the setpoint temperature can be observed in Fig. 9 (left) for all ADREs. The average temperature difference is 2.5 ℃. The peak in the histogram of 2.2 °C can be explained considering the larger number of DR events that were not adjusted during the DRE, but only before the DRE.

Fig. 9 (right) shows the decrease of the setpoint temperature applied during the user adjustment only (UA). The abscissa axis represents the total duration of the event (time elapsed from the beginning of the DR event). The plot shows only the events overridden in the first 100 min in order to include only cases in which users manually interrupted the DRE before its programmed end. For larger durations of the event, we cannot distinguish without further analysis the ADREs preceded by a HoA from the ADREs interrupted by a UA. The number of samples considered for every box is shown in brackets. The ordinate axis of Fig. 9 (left) represents the downward setpoint variation (i.e. how many degrees the users lowered the temperature when they interrupted the event). It can be seen that the more the duration increases, the more the users want to feel cool rather than neutral. It is important to have in mind that a greater drop in temperature corresponds to a bigger loss in energy efficiency and in the DRE strategy's effectiveness.

Studying in detail the same event (20–08-2019) in two different homes (Fig. 10), it was evident that the override was not the only

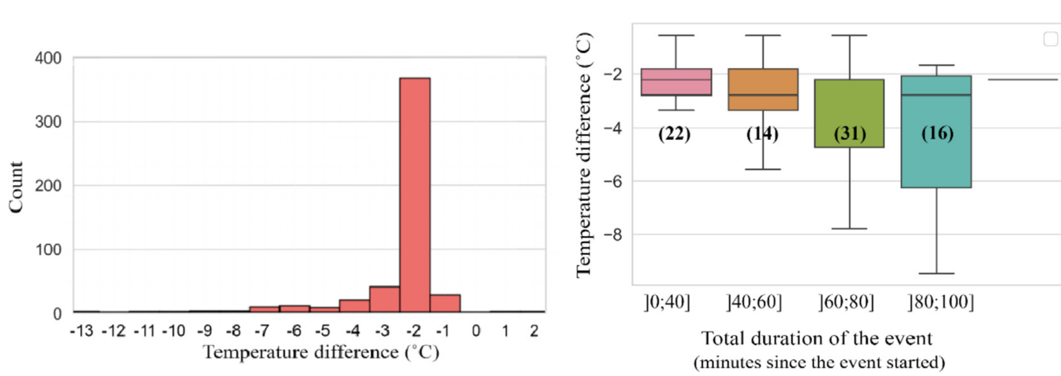

Fig. 9. Decrease of the setpoint temperature for all ADREs (left); Decrease of the setpoint temperature in correspondence of an ADRE (UA), the number of samples beeing in brackets (right). Interval notation of the x-axis: reversed bracket indicates that the endpoint number is excluded from the interval, normal bracket indicates that the endpoint number is included in the interval.

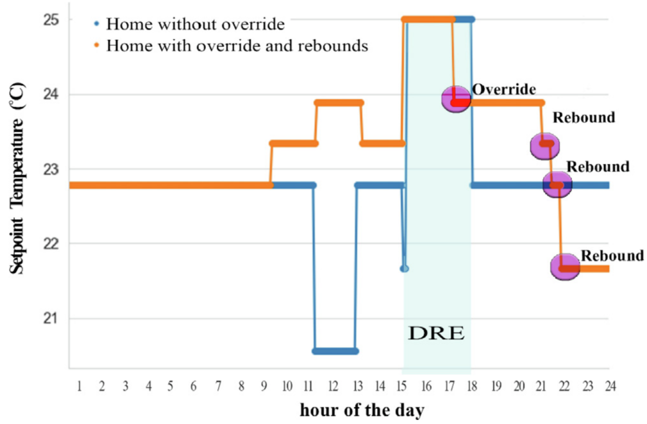

Fig. 10. Comparison between a non-ADRE and an ADRE (the same event in two different homes).

cause of the missing power reduction. Fig. 10 shows the variation of the setpoint temperature during the same day for two different houses, one with the override, and the other without.

The blue line represents a home in which the event started and ended as planned by the third party, while the orange line represents an ADRE (in this case, interrupted by the user). In fact, the increase of temperature up to 25 ℃ represents the non-ADRE, which is supposed to last from 15:00 to 18:00 (there is a little phase of precooling too). The orange line drops to approximately 24 ℃ at 17:00, hence one hour of potential energy saving is lost. Still, occupants in the second home (the orange line) seem to be not comfortable: they keep diminishing the setpoint temperature by ranges until they reach, at the end of the day (i.e. in the coolest part of the day), a setpoint temperature even lower than the one they used to have and accept earlier in the same day. It is a clear case of additional rebounds due to the event. This comparison has been done to show the reader that the reaction of different users to the same DRE can be largely different. In the following section, we will characterise occupants' behaviour after the DR event (through clustering of overriders' profiles) and we will delineate what consequences such actions have on the overall consumption.

3.3. Overriders' profiling (after the DR event)

In this subsection, we will describe the results from our clustering strategy that was applied on the time series of the setpoint

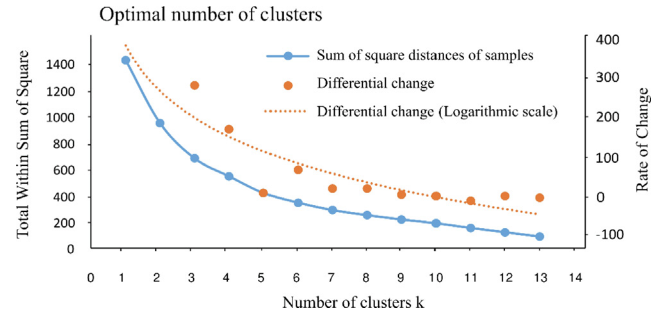

Fig. 12. Elbow method computed to decide the optimal number of clusters to be used in the analysis, before proceeding to the visual inspection of the dendogram. The data points have been shown with a logarithmic trend line to smooth noise. The differential change data refer to the secondary y-axis (on the right).

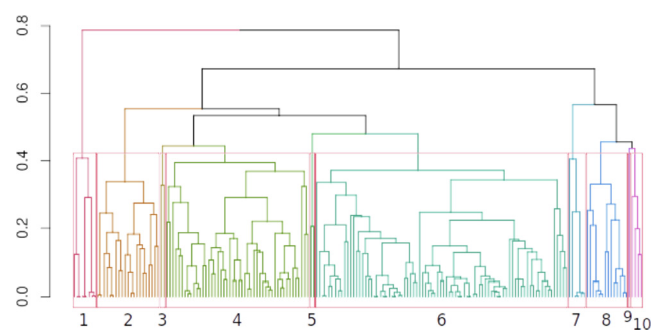

Fig. 13. Dendogram using a division of 10 clusters (only ADREs + R).

temperature after an ADRE (UA and HoA), to study the effect that the DRE has on the user choices over the rest of the day. We will refer to the period between the end of the ADRE and before returning to the set schedule, in which the setpoint temperature is managed by the users, as 'post-ADRE'.

From the 528 detected events, 64 % of them consisted of cases without additional rebounds, which means that the representative shape of more than half of post-ADREs is a straight line. In other words, there is no further adjustment of the setpoint after the DR events. The duration of the post-ADREs without additional rebound ranges between 10 min and 11 h 50 min, being 5 h 40 min the mean (this duration depends on the hold action set by occupants).

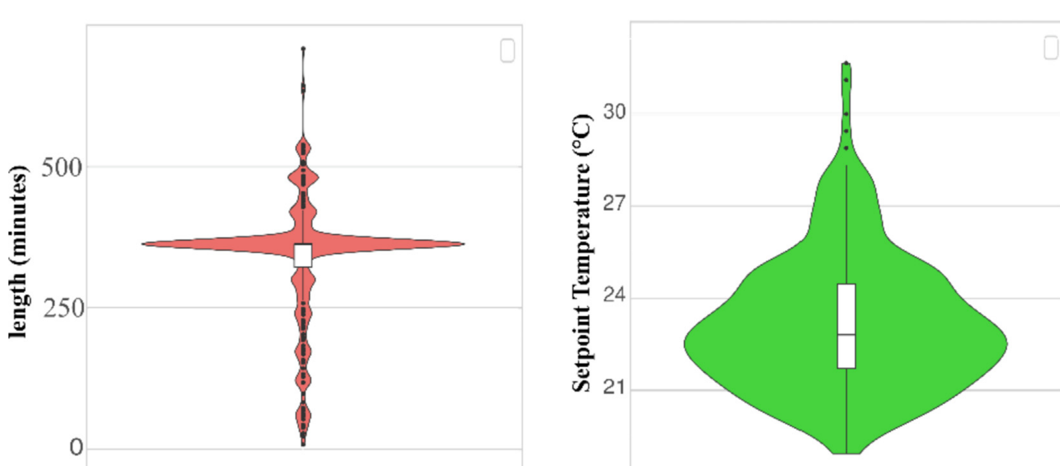

Fig. 11. Length of the series, i.e. duration of the Hold Action after the event (left) and setpoint temperature boxplots(right).

Table 1
Percentage of series that belong to each cluster. In bold those that are greater than 2 %. The sub-sample considered is the so-called ADREs + R, that represent approximately 36 % of the ADREs.

Cluster	1	2	3	4	5	6	7	8	9	10
Observations (%)	1.5	**3.95**	0.38	**9.02**	0.38	**15.79**	1.13	**2.63**	0.19	0.75

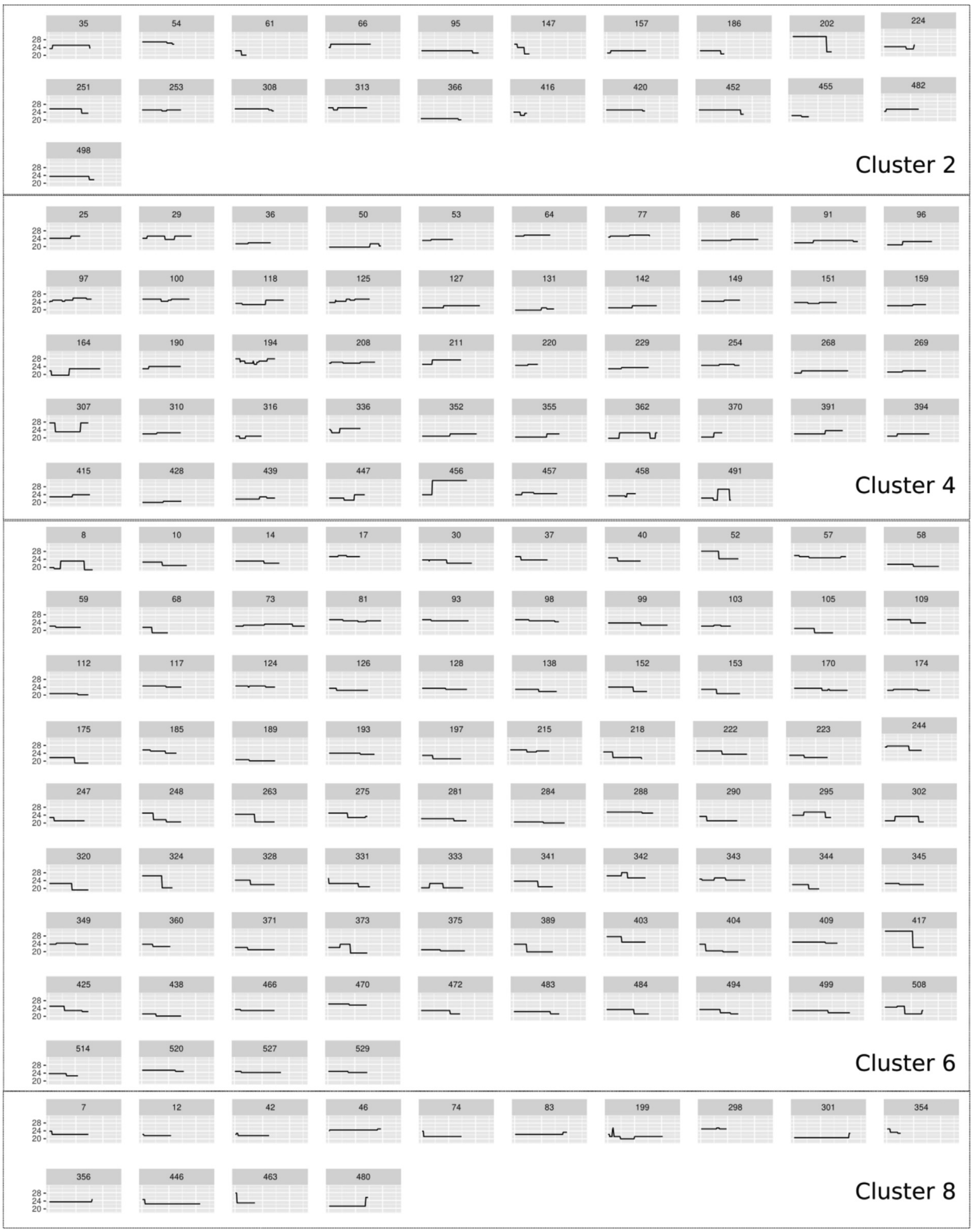

Fig. 14. Time series from clusters 2, 4, 6 and 8, containing more than 2 % of data each. The series represent the setpoint temperature controlled by the users after the DRE, i.e. they started when the DRE finishes and they end when the users return to the programmed schedule. Please notice that the X-axis represents the length of the time series (the vertical grid divides the x-axis into five portions of two hours) and the Y-axis the setpoint temperature in ̊C.

Also, the setpoint ranges from 18.9 up to 31.7 °C and presents an average of 22.8 °C (Fig. 11).

Once we separated the ADREs without any additional rebound, we applied the agglomerative hierarchical clustering with the shape-based metric on the ADREs + R. Hierarchical clustering allows the choice of the number of clusters. The segment of time-series considered as an input for the cluster refers to the period post-DR.

The selection of the number of clusters used for this study was done through the elbow method [5253], which consists of calculating and plotting the sum of squares at each number of clusters, and then looking at a change of slope in the graph to determine the optimal number of clusters. As seen in Fig. 12, the change on slope of the elbow method graph crosses zero when the number of clusters is 10, indicating that 10 is the optimal number of clusters according to this method.

We also made sure that such a number of clusters do not provide groups with too many or too few observations by visual inspection of the dendogram, which is a diagram that shows the hierarchical relationship between items, as it can be seen in Fig. 13.

Also, the percentage of observations of each cluster can be seen in Table 1. Excluding the 64 % of ADREs without rebounds after the event, the ADREs + R considered are 190 (36 %). Among the 10 clusters, we selected only those that contained at least 2 % of the data. In this way, four main clusters were seen to appear (clusters 2, 4, 6 8) for further analysis. 23 time-series are considered outliers based on the fact that they are classified in a cluster with very few other observations (clusters 1,3,5,7,9,10) and are discarded for the following analysis. The cluster selection is independent of the number of observations that each series has, it only depends on the number of series of which a cluster is composed. Every time series can have a different length, and the algorithm obtains the characteristic shape to classify the different groups.

Fig. 14 represents the series belonging to clusters 2,4, 6 and 8, as they contain the bigger percentage of observations. As already said, the time series represent the setpoint temperature chosen by the user, so they end when the setpoint temperature returns to the scheduled value. The time series were recorded every-five minutes and the vertical grid divides the X-axis of Fig. 14 into five portions of two hours.

Looking at the time series, some recurrent characteristics can be found in the shape. We will describe the prevalent shape of each cluster. Cluster 2 seems to be characterised by very mild changes in the setpoint temperature. In some cases, these changes are followed in a short period by others that counterpoise the first ones, returning to the setpoint at the beginning of the time series. Cluster 4 has a prevalence of time series with an increase in the setpoint temperature in a second stage. There is indeed also some decrease, but they are not prevailing. Cluster 6 presents a prevalence of time series in which the setpoint temperature chosen for the override is maintained constant for a long period. Then, approximately at half of the series, a decrease in the setpoint is detected. Cluster 8, the one with the smallest number of samples, presents setpoint temperatures that are constant for almost the entire time series, with a change or right at the beginning or right at the end of the series. We have used clustering to highlight differences that one cannot distinguish by eye, hence it is normal that some of the profiles within the same clusters can look dissimilar if just the shape is considered. To investigate what these distinctions in patterns are due to, we conducted further analysis in the rest of the section..

In Fig. 15 (left), we have depicted violin plots of the length of the series per cluster. The duration of the time-series within each cluster is relatively uniform, meaning that the clustering mainly focuses on other parameters of the time-series and highlighting the robustness of the methodology selected. In the right part of Fig. 15, we have depicted a violin plot with the number of changes per cluster. We observe that clusters 4 and 8 contain some series that present more than three changes.

In order to better understand the differences in the number of changes and behaviour which are somehow related to physical characteristics of the environment at the end of the DRE, in Fig. 17 are shown the violin plots of the indoor and outdoor temperature right before the override, as well as of the delta setpoint temperature in correspondence of the override and the delta setpoint detected during the post-DRE. Fig. 16 clarifies graphically what the deltas are referring to.

To understand the differences of the delta setpoint in correspondence of the user's adjustment, it is worth underlining that this parameter is clearly affected by the prevalence of cases in which the variation is 2.2 °C (i.e. no additional adjustment during the DR event). Another clarification needed is that the violin plots of the post-DREs rebound are representing the algebraic sum of the setpoint change, hence the cases of great drops in temperature that are followed by a return to the previous value, are counterbalanced.

From the violins, we can have a more global vision of the characteristics of the clusters. Cluster 2 and Cluster 6 are alike, they represent the users with the smallest drop in temperature when overriding. The indoor and outdoor temperature values in the two clusters are similar, as well as the number of interactions that the users made. In addition to the initial setpoint drop, occupants

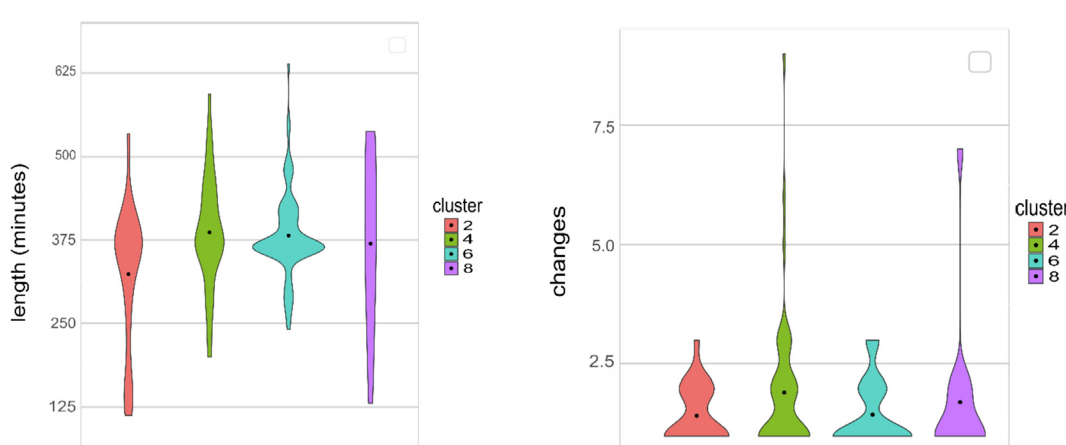

Fig. 15. Violin plots that depict the duration of the series (left) and the number of changes that they present (right).

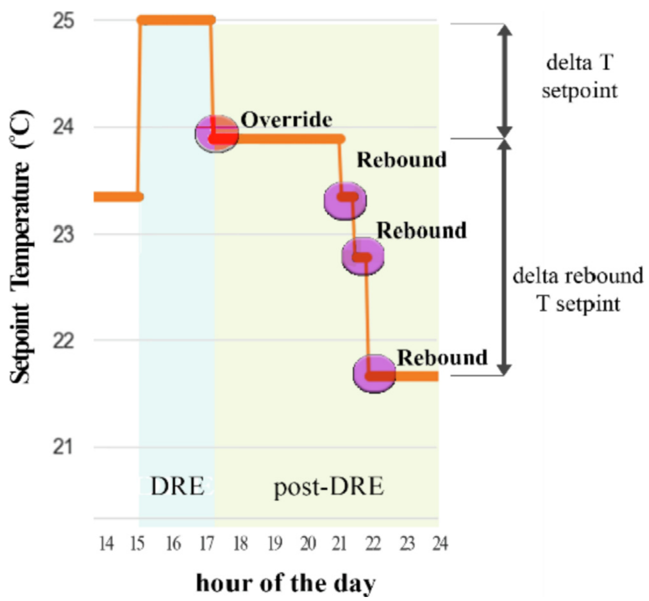

Fig. 16. Time series of an ADRE + R.

in both cases lowered again the temperature of an average of 1 °C for cluster 2 and 1.25 ℃ for Cluster 6, suggesting a long thermal discomfort in these homes. Hence, the difference between these clusters lies in the length and the shape of the time series. Cluster 4 and Cluster 8 are the ones that present the highest outdoor temperature and the highest delta setpoint temperature during the override. Their number of interactions is also alike, while the indoor temperature is quite different, presenting a difference of 1 °C. Hence, it is interesting to notice the differences in the average delta setpoint temperature in the post-ADRE. Users from cluster 4 present an average of 0.63 ℃, contradicting their previous choice of lowering that much the temperature during the override. Differently, users from Cluster 8 end the time series with approximately the same setpoint temperature that they chose to override.

It is also interesting to notice that the users with the greater indoor temperatures experienced at the end of the DR events (i.e. users of Cluster 8) are those with the larger delta T setpoint (difference of about 0.5 °C compared to the others) and with no corrective rebound. While Clusters 2 and 6, which are characterised by a less abrupt change of the delta T setpoint, have a negative mean rebound of around 1 °C. This is in line with the discoveries about thermal alliesthesia: in non-steady conditions, users subjected to thermal displeasure do not search for thermal neutrality, but rather for positive pleasure associated with counteracting thermal sensations [5556]. This result also suggests that rebounds can be more critical than the delta T setpoint in terms of magnitude and that the utility should avoid sending aggressive DREs to those users that could override the thermostat with high rebounds, in particular those of Cluster 6, that can make a net negative effect.

3.4. Rebounds quantification

To quantify these differences in occupants' behaviour in terms of missed power reduction, we considered that the most appropriate strategy was isolating and comparing DREs that had similar characteristics but that represent different users' profiles, i.e. different clusters. In other words, we wanted to find similar situations (in terms of hour of the day, environmental characteristics and intensity of the DRE) in order to observe the different occupants' reactions to a same input. We considered that the sample for the simulation should include one case for each cluster plus a case without override, although the comparison does not pretend to hierarchise the clusters. We managed to find 4 out 5 cases with

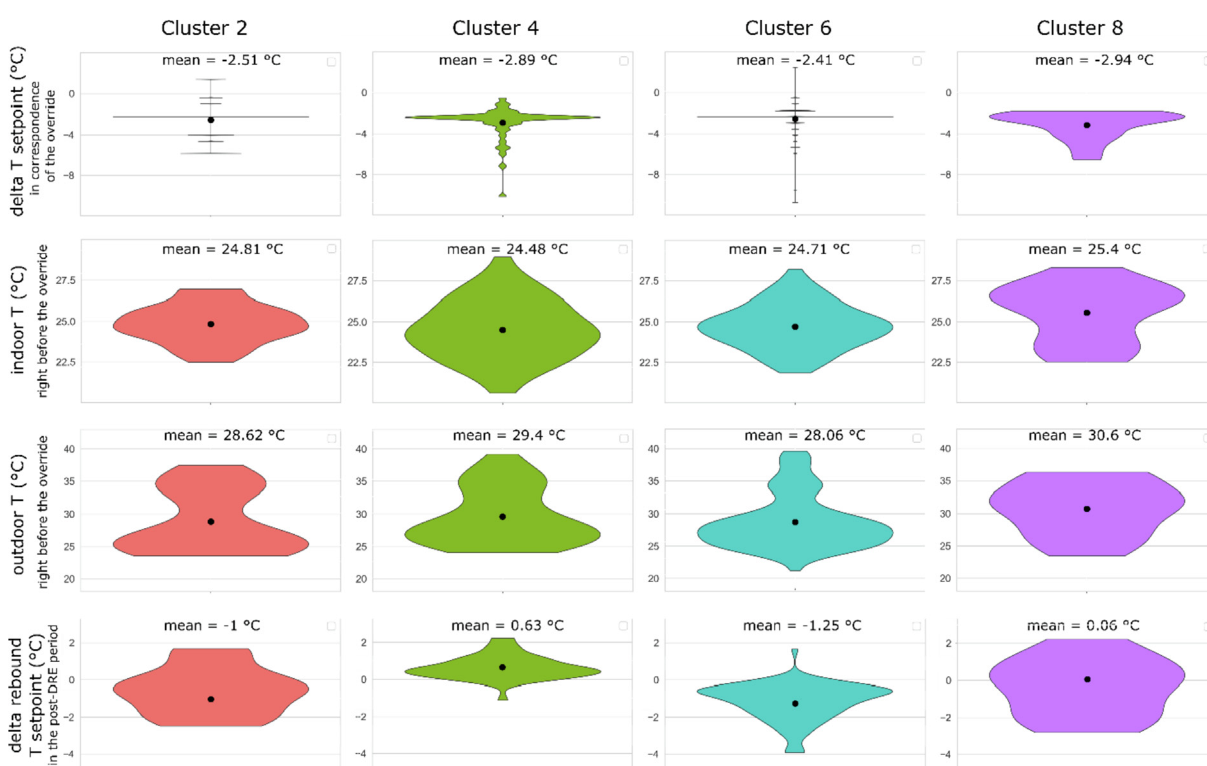

Fig. 17. Violin plots that depict the delta setpoint chosen by the users when overriding, the indoor and outdoor temperature right before the override, and the setpoint rebound after the event, for each main cluster. To a better comprehension of the time frames, please refer to Fig. 16.

the following common characteristics (the case from Cluster 8 presents some minor differences):

- same date: 20–08-2019 (Case Cluster 8: 13–08-2019);
- same province (to maintain similar outdoor conditions): Ontario (Canada);
- same starting hour: 15:00 (Case Cluster 8: 13:00);
- same setpoint temperature during the event: 25 ℃;

In the first one (case without override, Fig. 17.a), the DRE ends at 18:00 as programmed, without any intervention by the users. Apparently, the occupants did not feel discomfort afterwards, since they did not lower the scheduled setpoint temperature. In the late evening, they even increased the temperature over a period. This is the ‘perfect’ case scenario, in which not only it was managed to avoid the consumption during the peak hour, but also to obtain an overall power reduction over the day. In the second case (case from cluster 2, Fig. 18.b) the DRE is preceded by an HoA, so the event has its scheduled duration and the setpoint set afterwards was equal to the one selected before the DR event. Around 21:30 a rebound of more than a degree Celsius is detected, followed by an increase in the setpoint temperature (higher than the one during the DRE itself) around 22:30. In the third case (case from cluster 4, Fig. 18.c), the DRE managed to last just one hour, since the override occurred at 16:00. The temperature chosen was 5 ℃ lower than the temperature during the event, causing a severe energy expenditure. In addition, after some hours with a higher setpoint, the temperature dropped again to 20 ℃. In the fourth case (case from cluster 6, Fig. 18.d) the DRE was interrupted by the users at 17:30, i.e. half an hour before the scheduled end, and the setpoint temperature chosen was equal to the one selected before the DR event. However, a rebound of more than one degree occurred

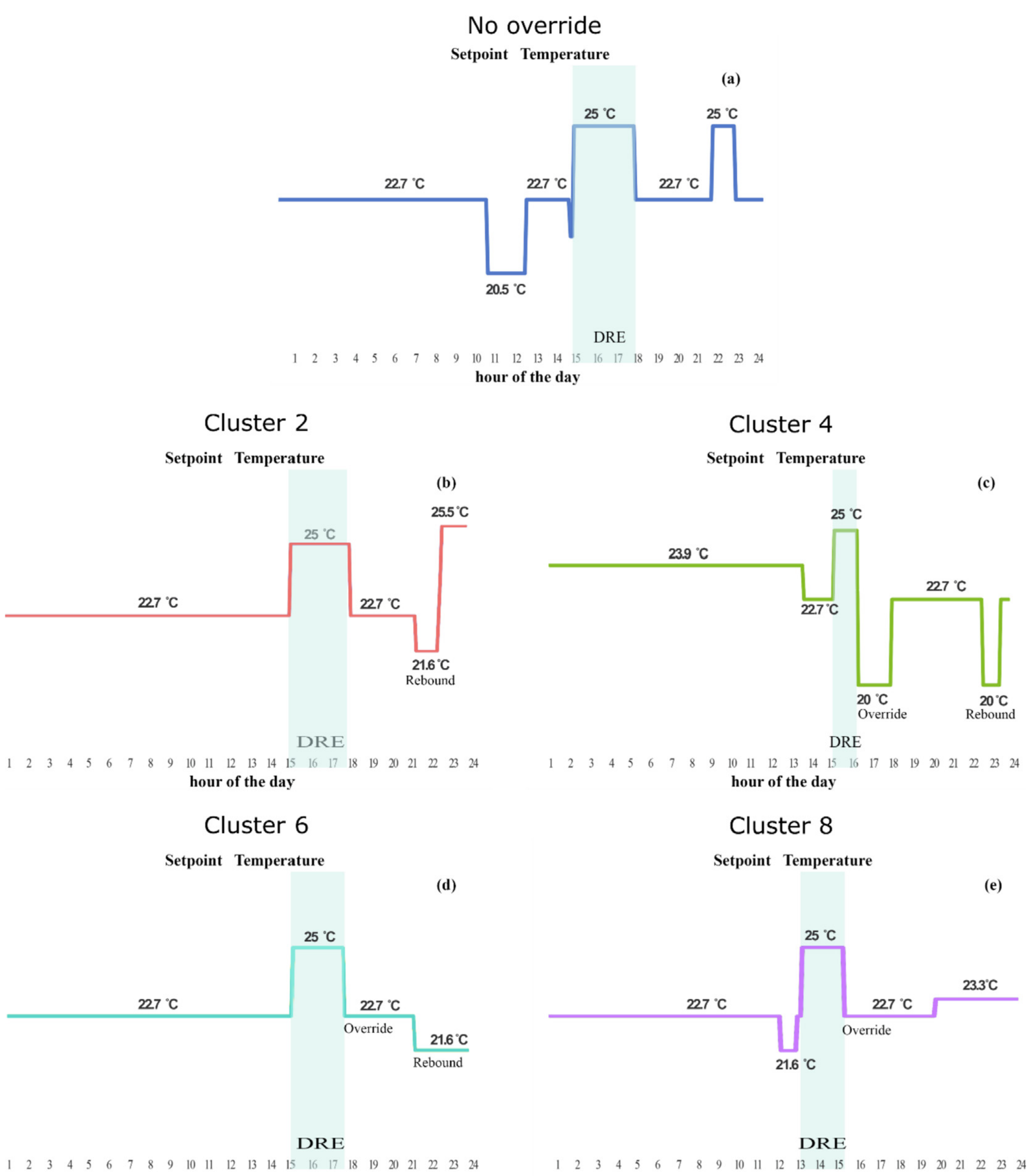

Fig. 18. Setpoint temperature over a DRE day in five different homes: a) Case without override; b) Case from Cluster 2; c) Case from Cluster 4; d) Case from Cluster 6; e) Case from Cluster 8. Note that the blue area symbolises the actual duration of the DRE, that could have been interrupted and therefore finishes before the utility intended (three hours in total). For the colour version of the picture the reader is referred to the online version of the paper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

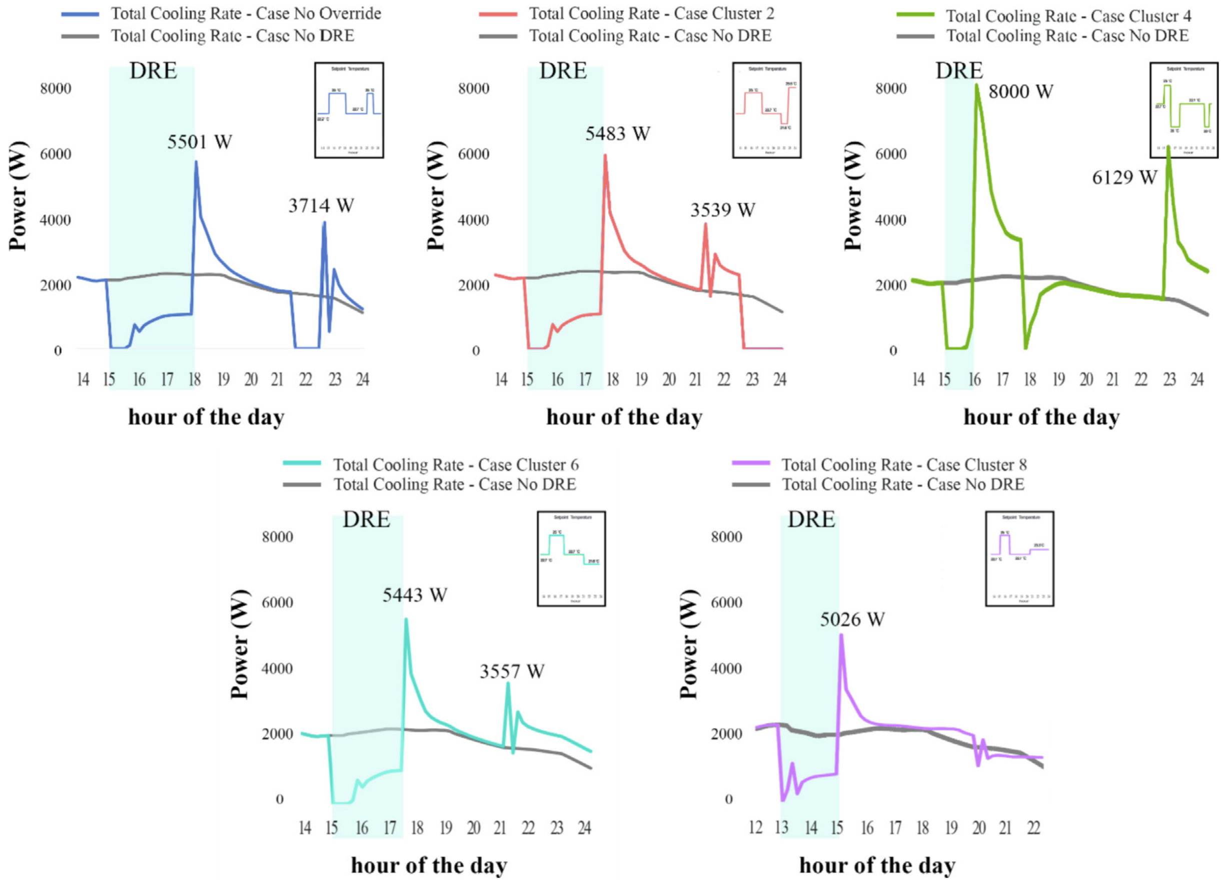

Fig. 19. Total cooling power during and after the DRE in a) Case without override; b) Case from Cluster 2; c) Case from Cluster 4; d) Case from Cluster 6; e) Case from Cluster 8. Note that the blue area symbolises the actual duration of the DRE, that could have been interrupted and therefore finishes before the utility intended. For the colour version of the picture the reader is referred to the online version of the paper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

around 21:00, hence part of the saved energy was lost. In the fifth case (case from cluster 8, Fig. 18.e), the override occurred two hours after the beginning of the DRE. The temperature chosen was maintained for a considerable number of hours and, around 20:00, it was slightly raised (of approximately half-degree Celsius).

It is worth reminding that we are not using the real consumption of the selected homes, because in that case the comparison would be affected by other parameters, e.g. the envelope or to the internal loads, rather than by the users' behaviour. Hence, the schedules are simulated on the same building energy model, in the post-event period (Fig. 19). Every case is compared to the results obtained with a flat schedule (22.7 ℃ over the whole day), that is the hypothetical power consumption if the DRE did not occur.

The average power used in peak hours and non-peak hours is shown in Table 2. In the case without override, it can be clearly appreciated the effectiveness of the Demand Response event on power shifting, since the average power in the peak hours is drastically reduced. The case from cluster 2 presents an override right before the programmed end, a rebound that lasted for a short period (approximately-one hour) and an increase in the setpoint at the end of the day. For this cluster, the overall consumption was reduced and the peak power was shifted, hence the DRE was effective. In Cluster 4 it is particularly evident that the missed power reduction due to the rebound is not negligible. The setpoint temperature chosen after the event (and later on in the evening) is very low (a delta of 5 ℃ compared to during the event). The overall consumption needed to reduce the thermal discomfort of those users, as well as the average power consumption during the peak hours, is enormously higher than the scenario without DRE. The case from Cluster 6 shows that, despite the event being stopped just half an hour before its programmed end, the rebound reduced the power reduction's potential of the DRE. In this case, the power is shifted to a cheaper time frame, but the average power consumption is quite higher than the case without override. The overall energy consumption is slightly higher compared to the case without any event. Finally, the case from Cluster 8 (that is considered until 22:00, in order to have the cumulative sum of the same number of hours for all the cases) is characterised by an increase in temperature, instead of a rebound. Despite that, the overall consumption is quite high, since the DR is overridden after 2 h and the occupants

Table 2
Average cooling power and total cooling energy need of the cases simulated with EnergyPlus from 15:00 till 24:00. Average power is divided into peak hours (15:00–18:00) and non-peak hours (18:24:00).

CASES SIMULATED	AVERAGE POWER (15:00–18:00)	AVERAGE POWER (18:00–24:00)	ENERGY NEED (15:00–24:00)
Case No DRE	2.12 kW	1.73 kW	17.05 kWh
Case No Override	0.68 kW	1.81 kW	13.24 kWh
Case Cluster 2	0.93 kW	1.75 kW	13.63 kWh
Case Cluster 4	2.80 kW	2.07 kW	21.21 kWh
Case Cluster 6	1.09 kW	2.16 kW	16.59 kWh
Case Cluster 8	1.50 kW (13:00–16:00)	1.91 kW (13:00–16:00)	16.29 kWh (13:00–22:00)

maintained for five consecutive hours the setpoint temperature chosen for the override, i.e. they did not return to their scheduled temperature. The power shifting is just partly achieved.

4. Conclusion

In this paper, we analysed patterns of thermostat settings before, during and after demand-response events (DREs), with a particular focus on the behaviour of occupants who showed less acceptance of the DREs. Smart thermostat usage data including around 1500 DREs have been evaluated using data analysis and clustering techniques. From the data, reactive behaviour was observed during the DREs, but to a lower extend (in 13 % of the DREs). Moreover, it was observed that a share of users (one in four approximately) manually set a hold action before a DR event, instead of following the programmed schedule. That can be caused by the day ahead notifications that the occupants receive before a DRE, thus indicating an anticipative behaviour. Depending on when the override occurs and the chosen temperature afterwards, we observed very different patterns. The setpoint decrease after an override was sometimes very high, leading to a large rebound and a low efficiency of the DR program. Finally, clustering was used to analyse the setpoint pattern after the DREs. Additional rebounds could be observed in 36 % of the cases. These patterns can only partially be explained in terms of physical factors (indoor and outdoor air temperatures).

The different patterns highlighted by this analysis can be used to exploit DR strategies' potential since it can lead to the design of tailored DR events. This means, that discriminating patterns of overriders' behaviour can increase the effectiveness of the strategies of direct control. The optimum duration of the DRE to avoid overrides and rebounds should be calculated taking into account the clustering proposed in this work. Also, this can be very useful in predicting how a user would react to an event depending on the cluster to which they belong. Our results suggest that the one-fits-all formula is not appropriate for energy flexibility strategies since it does not take into account the different behaviour of users. As designing a tailored DRE for each user would be impossible and time-consuming for the power company, our paper suggests that a data-driven solution towards personalized DREs can make a great difference in people's acceptance, once the demographics of the participants are taken into account with specific analyses.

The design of DR events tailored with respect to patterns of behaviour can be then enriched by taking into account several other factors, starting from the physical characteristics of the envelope to the connection between the energy flexibility and the thermal mass of the dwelling [57585960]. [61]. Hence, we encourage future researchers to look into the issue of different fabrics, to investigate whether they can induce a difference in the comfort perceived by the users. Grouping the consumers into typologies would open the door to personalised interaction with users and energy contracts: the design would not be limited to sending events with different characteristics depending on the occupants' tolerance, but it could also involve different ways of communication considering the subjective interaction of the users with the thermostat. Beyond the different patterns of the clusters, the main point we want to highlight is that users have diverse behaviours when reacting to a DRE, and that can surely eliminate the effect of the flexibility strategies. Summarising, the work presented here covers an aspect that has been unexplored before and that have large implications with the power shaving and the energy reduction of demand response events. It is firmly believed that the findings of this paper will help on the design of DR programs and it will increase their effectiveness. [54].

Data availability

The data is already published.

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.

Acknowledgement

This project has been funded by Horizon 2020 Project PHOENIX (grant number 893079). AGV would like to thank the Spanish Ministry of Universities by means of the Margarita Salas linked to the European Union through the NextGenerationEU program.

References

[1] C. Church, W.G. Morsi, M.E. El-Hawary, C.P. Diduch, L.C. Chang, Voltage collapse detection using Ant Colony Optimization for smart grid applications, Electric Power System Research 81 (8) (2011) 1723–1730.
[2] K. Clement-Nyns, E. Haenes, J. Driesen, The impact of Charging plug-in hybrid electric vehicles on a residential distribution grid, IEEE Trans. Power Syst. 25 (1) (2010) 371–380.
[3] Heinen, S., Elzinga, D., Kim, S.K., Ikeda, Y., 2011. Impact of Smart grid technology on peak load to 2050, International Energy Agency, 2011.
[4] C. Eid, E. Koliou, M. Valles, J. Reneses, R. Hakvoort, Time-based pricing and electricity demand response: Existing barriers and next steps, Utilities Policy 40 (2016) (2016) 15–25.
[5] P. Siano, Demand response and smart-grids – A survey, Renewable and Sustainable Energy Review 30 (2014) (2014) 461–478.
[6] P. Palensky, D. Dietrich, Demand side management: Demand response, intelligent energy systems, and smart loads, IEEE Trans. Ind. Inf. 7 (3) (2011) 381–388.
[7] H.A. Aalami, M. Parsa Moghaddam, G.R. Yousefi, Demand Response modeling considering Interruptible/Curtailable loads and capacity market programs, Appl. Energy 87 (2010) (2010) 243–250.
[8] B. Patnam, N. Pindoriya, Demand response in consumer-centric electricity market: Mathematical models and optimization problems, Electr. Power Syst. Res. 106923 (2021) 193.
[9] H. Wang, S. Wang, R. Tang, Development of grid-responsive buildings: Opportunities, challenges, capabilities and applications of HVAC systems in non-residential buildings in providing ancillary services by fast demand responses to smart grids, Appl. Energy 697–712 (2019) 250.
[10] W.T. Sung, S.J. Hsiao, J.A. Shih, Construction of Indoor Thermal Comfort Environmental Monitoring System Based on the IoT Architecture, J. Sens. 2019 (2019) (2019) 2639787.
[11] Feldmeier, M.; Paradiso, J.A. Personalized HVAC control system. In Proceedings of the Internet of Things (IoT), Tokyo, Japan, 29 November–1 December 2010.
[12] D. Li, C.C. Menassa, V.R. Kamat, Personalized human comfort in indoor building environments under diverse conditioning modes, Build. Environ. 2017 (126) (2017) 304–317.
[13] A. Sanguinetti, M. Pritoni, K. Salmon, A. Meier, J. Morejohn, Upscaling participatory thermal sensing: Lessons from an interdisciplinary case study at University of California for improving campus efficiency and comfort, Energy Res. Soc. Sci. 2017 (32) (2017) 44–54.
[14] S. Wang, D. Gao, R. Tang, F. Xiao, Cooling Supply-based HVAC System Control for Fast Demand Response of Buildings to Urgent Requests of Smart Grids, Energy Procedia 2016 (2016) 34–39.
[15] J. Torriti, M.G. Hassan, M. Leach, Demand response experience in Europe: Policies, programmes and implementation, Energy 35 (2010) (2010) 1575–1583.
[16] X. Xu, C. Chen, X. Zhu, Q. Hu, Promoting acceptance of direct load control programs in the United States: Financial incentive versus control option, Energy 147 (2018) 1278–1287.
[17] S. Karjalainen, Should it be automatic or manual—The occupant's perspective on the design of domestic control systems, Energy Build. 65 (2013) (2013) 119–126.
[18] A. Pardasani, Y. Hu, J.A. Veitch, G.R. Newsham, Demand control of baseboard heaters using connected thermostats: lesson learned from a 567-home pilot study, ASHRAE Transaction 126 (2020) part 1.
[19] K.E.M.A. Inc, 2005 Smart Thermostat Program Impact Evaluation, prepared for San Diego gas and electric company, San Diego, California, 2006.
[20] T. Sweetnam, C. Spataru, M. Barrett, E. Carter, Domestic demand-side response on district heating networks, Building Research & Information 47 (4) (2019) 330–343.
[21] V. Tomat, A.P. Ramallo-González, A. Skarmeta Gómez, A Comprehensive Survey about Thermal Comfort under the IoT Paradigm: Is Crowdsensing the New Horizon?, Sensors 2020 (20) (2020) 4647

[22] A.J. Conejo, J.M. Morales, L. Baringo, Real-time demand response model, IEEE Trans. Smart Grid 1 (3) 5607339 (2010) 236–242.
[23] W. Chen, X. Wang, J. Petersen, R. Tyagi, J. Black, Optimal Scheduling of Demand Response Events for Electric Utilities, IEEE Trans. Smart Grid 1949–3053 (2013) (2013).
[24] H. Pombeiro, M. Machado, C. Silva, Dynamic programming and genetic algorithms to control an HVAC system: Maximizing thermal comfort and minimizing cost with PV production and storage, Sustainable Cities and Society 228–238 (2017) 34.
[25] Cui, W., Ding, Y., Hui, H., Lin, Z., Du, P., Song, Y., Shao, C., 2018. Evaluation and sequential dispatch of operating reserve provided by air conditioners considering lead-lag rebound effect, IEEE Transactions on Power System, (2018) 6935-6950, 33 (6).
[26] Sehar, F., Pipattanasomporn, M., Rahman, S, 2016. A peak-load reduction computing tool sensitive to commercial building environmental preferences, Applied Energy (2016), 279-289, 161.
[27] S. Wang, R. Tang, Supply-based feedback control strategy of air-conditioning systems for direct load control of buildings responding to urgent requests of smart grids, Appl. Energy 419–432 (2017) 201.
[28] L. Cheng, Y. Wan, L. Tian, F. Zhang, Evaluating energy supply service reliability for commercial air conditioning loads from the distribution network aspect, Appl. Energy 2019 (2019) 253.
[29] Sarran, L., Gunay, H.B., O'Brien, W., Hviid, C.A., Rode, C, 2021. A data-driven study of thermostat overrides during demand response events, Energy Policy, 153 (2021) 112290
[30] M.H. Christensen, R. Li, P. Pinson, Demand side management of heat in smart homes: Living-Lab experiments, Energy 195 (2020) 116993.
[31] Broka, Z., Baltputnis, 2020. Handling of the rebound effect in independent aggregator framework, 2020 17th International Conference on the European Energy Market (EEM), 2020, pp. 1-5.
[32] ecobee, 2019. eco+ Community Energy Savings [WWW Document] (2019) https://www.ecobee.com/en-us/eco-plus/community-energy-savings/ (accessed 14.04.22).
[33] B. Huchuk, W. O'Brien, S. Sanner, A longitudinal study of thermostat behaviors based on climate, seasonal, and energy price considerations using connected thermostat data, Build. Environ. 139 (2018) (2018) 199–210.
[34] B. Huchuk, S. Sanner, W. O'Brien, Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data, Build. Environ. 160 (2019) 106177.
[35] Meier, A., Ueno, T., Rainer, L., Pritoni, M., Daken, A., Baldewicz, D., 2019. What can connected thermostats tell us about American heating and cooling habits?, ECEE Summer Study Proceedings 4-042-19.
[36] M. Vellei, S. Martinez, J. Le Dréau, Agent-based stochastic model of thermostat adjustments: A demand response application, Energy Build. 238 (2021) 110846.
[37] M. Kane, K. Sharma, 2019, Data-driven Identification of Occupant Thermostat-Behavior Dynamics, ArXiv, 2019.
[38] W. Kempton, Two theories of home heat control, Cognitive Science 10 (1986) (1986) 75–90.
[39] K.M.A. Revell, N.A. Stanton, Case studies of mental models in home heat control: Searching for feedback, valve, timer and switch theories, Appl. Ergon. 45 (2014) (2014) 363–378.
[40] S. D'Oca, C.F. Chen, T. Hong, Z. Belafi, Synthesizing building physics with social psychology: An interdisciplinary framework for context and occupant behavior in office buildings, Energy Res. Social Sci. 34 (2017) (2017) 240–251.
[41] B. Kane, Modeling Human-in-the-Loop Behavior and Interactions with HVAC Systems, 2018 Annual American Control Conference (ACC), June 27–29, 2018, Wisconsin Center, Milwaukee, USA, 2018.
[42] W. Kleiminger, F. Mattern, S. Santini, Predicting household occupancy for smart heating control: A comparative performance analysis of state-of-the-art approaches, Energy Build. 85 (2014) (2014) 493–505.
[43] T. Pedersen, K. Nielsen, S. Petersen, Method for room occupancy detection based on trajectory on indoor climate sensor data, Build. Environ. 115 (2017) (2017) 147–156.
[44] M.L. Zepeda-Mendoza, O. Resendis-Antonio, Hierarchical agglomerative clustering, Encyclopedia of systems biology 43 (1) (2013) 886–887.
[45] J. Paparrizos, L. Gravano, May). k-shape: Efficient and accurate clustering of time series, in: In ProceedIngs of the 2015 ACM SIGMOD International Conference on Management of Data, 2015, pp. 1855–1870.
[46] R.R. Sokal, A statistical method for evaluating systematic relationships, Univ. Kansas, Sci. Bull. 38 (1958) 1409–1438.
[47] EnergyPlus™ v8.9.0. https://energyplus.net/.
[48] https://energyplus.net/weather-location/north_and_central_america_wmo_region_4/CAN/ON/CAN_ON_Toronto.716240_CWEC, accessed on 22/06/2021.
[49] J. Yoon, R. Bladick, A. Novoselac, Demand response for residential buildings based on dynamic price of electricity, Energy Build. 80 (2014) (2014) 531–541.
[50] ASHRAE, Guideline 14-2014, Measurement of Energy, Demand and Water Savings. American Society of Heating, Ventilating, and Air Conditioning Engineers, Atlanta, Georgia, 2014.
[51] ASHRAE, ANSI/ASHRAE Standard 55-2013, Thermal Environmental Conditions for Human Occupancy: Atlanta, Ga, 2020.
[52] R.L. Thorndike, Who belongs in the family, In (1953), Psychometrika.
[53] Marutho, D., Hendra Handaka, S., Wijaya, E., Muljono, 2018. The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News, Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, iSemantic 2018.
[54] 9751, pp. 533-538 González-Vidal, A., Ramallo-González, A.P., Skarmeta, A., 2021. Empirical study of massive set-point behavioral data: Towards a cloud-based artificial intelligence that democratizes thermostats, 2018 IEEE International Conference on Smart Computing (SMARTCOMP), 2018, pp. 211-218.
[55] M. Vellei, R. de Dear, C. Inard, O. Jay, Dynamic thermal perception: A review and agenda for future experimental research, Build. Environ. 205 (2021) (2021) 108269.
[56] T. Parkinson, R. De Dear, Thermal pleasure in built environments: physiology of alliesthesia, Building Research & Information 43 (3) (2014) 288–301.
[57] J. Le Dréau, P. Heiselberg, Energy flexibility of residential buildings using short term heat storage in the thermal mass, Energy 11 (2016) (2016) 991–1002.
[58] G. Masy, E. Georges, C. Verhelst, V. Lemort, P. André, Smart grid energy flexible buildings through the use of heat pumps and building thermal mass as energy storage in the Belgian context, Science and Technology for the Built Environment 21 (6) (2015) 800–811.
[59] J. Kensby, A. Trüschel, J. Dalenbäck, Potential of residential buildings as thermal energy storage in district heating systems—Results from a pilot test, Appl. Energy 137 (2015) 773–781.
[60] S. Wang, X. Xue, C. Yan, Building power demand response methods toward smart grid, HVAC R Res. 20 (2014) 665–687.