Building and Environment 139 (2018) 199–210

Contents lists available at ScienceDirect

Building and Environment

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

A longitudinal study of thermostat behaviors based on climate, seasonal, and energy price considerations using connected thermostat data

Brent Huchuk[a,*], William O'Brien[b], Scott Sanner[a]

[a] Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada
[b] Department of Civil and Environmental Engineering, Carleton University, Ottawa, Canada

ARTICLE INFO

Keywords:
Connected thermostat
Residential
User preferences
User behavior

ABSTRACT

While previous studies have attempted to understand and predict users' behaviors and preferences for residential thermostats, they have been restricted by a lack of available data. Because of practical constraints, researchers previously relied on short observation periods, small sample groups, and/or participants close in physical location. The advent of the connected thermostat and its inherent centralized data collection now allows for such studies to be performed without the onus of data collection. Specifically, in this article we focus on the 'Donate Your Data' dataset made available by the thermostat manufacturer ecobee Inc. The dataset, consisting of more than 10,000 connected thermostats installed across North America and spanning multiple years, was used to investigate how users' comfort decisions are affected by exterior stimuli such as climate regions, seasonal patterns, and utility rates. Our analysis indicates that seasonality and climate region affected user preferences while utility rates did not contribute to meaningful variation in behavior. Further investigation explored if behavioral user types could be identified based on variation in occupied and unoccupied setpoints, thermostat overrides with holds, or heating and cooling setpoint selection. We did not find distinct user clusters to be identifiable based on any of the metrics; rather, occupant behavior in the population appeared to span more of a continuum across each metric.

1. Introduction

In the United States, 47% of the energy used by the residential sector is consumed by space conditioning [1]. An estimated 85% of American households use a thermostat to control their heating, while 64% use a thermostat for controlling cooling [2]. Given both their prevalence in homes and the magnitude of residential energy, thermostats managed almost 8% of the United States' energy use in 2015 [3]. Hence it is important to study thermostat usage to understand the behaviors which lead to such energy consumption.

Thermostats have evolved in both functionality and user expectation for their role in the home [4]. The most basic in function, a manual thermostat, has a single or sometimes two static setpoints (i.e., one for heating and one for cooling). When the temperature exceeds an upper setpoint or falls below a lower setpoint, the appropriate system is enabled until the space returns to the acceptable temperature range. Active setpoint management is achieved only by an eager and willing user. The first programmable thermostats were released over a century ago, and in general functionality remain unchanged [4]. They allow users to include setbacks (or set-ups) in their schedule based on the time of day. The active setpoint management is performed by the device, once programmed, allowing users to save energy without sacrificing comfort when at home. Theoretically, setbacks on a thermostat were shown to reduce natural gas consumption by up to 25% [5]. However, often schedules are not properly set and energy usage does not achieve theoretical savings [6]. In a survey of thermostat users, Pritoni et al. [7] found that 40% of users were not using the programmable features. While other energy efficient products achieve savings solely thorough installation (e.g., a furnace or air conditioner), programmable thermostats require that users properly configure the device to maximize energy efficiency and savings. The discovery that users were not properly using a programmable thermostat resulted in the Environmental Protection Agency (EPA) suspending the Energy Star program for programmable thermostats [8].

The most recent generation of thermostat devices are referred to as connected thermostats. These devices, which are connected to the Internet, promise to deliver accessible control through web, mobile or even voice platforms, and additional energy savings because of features like occupancy detection or adaptation to occupant schedules. Inherent in all of these capabilities is their ability to measure, transmit, and

* Corresponding author. University of Toronto, Department of Mechanical and Industrial Engineering, 105-5 King's College Road, M5S 3G8, Toronto, Ontario, Canada.
E-mail address: brent.huchuk@mail.utoronto.ca (B. Huchuk).

https://doi.org/10.1016/j.buildenv.2018.05.003
Received 9 February 2018; Received in revised form 3 May 2018; Accepted 3 May 2018
Available online 08 May 2018
0360-1323/

receive data from the time of installation in the customer's home. This gives users, companies, and researchers alike the ability to understand how actual thermostat operation is occurring for the subset of the general population who have installed a connected thermostat. As of 2015, 40% of the 10 million thermostats purchased in the United States each year were of the connected variety [9].

Considering the diverse functionality, historical difficulty in translating theoretical savings to actual savings and broad energy impacts, research into thermostat operation has remained disproportionately limited. Historically, efforts to understand user behavior have relied on small samples (e.g., 10s or 100s) in a localized geographic area [10–13] or on the correct and accurate reporting by the users at one time [1,2,7,14]. Both methods have potential caveats. The small, and often local, samples mean generalizations are difficult to make, while self-reporting can be inherently biased and does not account for adaptations in behavior over time [15]. As found by Pritoni et al. [7], inconsistencies were unavoidable in the data when users were asked to self-report on various questions regarding their thermostat operation even when provided assistance and guidance during the questionnaire. Similarly, Vine and Barnes [14] found that users' reported and actual temperatures could be off by an average of 2 °F (1.1 °C) but that discrepancy could be more than 5 °F (2.8 °C). The results from these self-reported studies are limited in scope because the findings are restricted only to the responses received to questions posed to the participants. They also lack generality because the sample is only for an instant in time for the user.

This paper puts forth an investigation into how a new data source, the interval data (i.e., time series data at five-minute intervals) collected from connected thermostats, could be used as a novel method for understanding occupants' thermostat behavior. In particular, the data source is utilized to solve issues that have limited many investigations in the past: (i) a lack of prolonged study length, (ii) minimal study population size, and (iii) difficulties in data collection. That said, our data is restricted to the subset of the population with connected thermostats who represent a distinct (often early technology adopters) subdemographic of residential thermostat users. Even with potential sample bias, the data source still provides valuable information in understanding how these users are operating their thermostats given the various differences that exist at the individual level. To explore these differences, we ask the following specific questions:

Q1 How does a thermostat user's behavior change as a function of: seasonal variations, climatic adaptations, and utility pricing in the region?
Q2 Are there discernible user types in thermostat operation, specifically as it relates to: engaging in energy savings strategies between occupied and unoccupied periods (i.e., setbacks/set-ups), frequency of holds, and heating and cooling setpoint selection?

The remainder of the article is structured as follows. Section 2 reviews relevant previous studies and findings. Section 3 provides an overview of the ‘Donate Your Data’ dataset which consists of multi-year interval data for over 10,000 thermostats across North America. Section 4 summarizes and discusses results associated with **Q1** and **Q2**. Section 5 presents conclusions of the investigation.

2. Background

As buildings become more energy efficient to operate because of tighter building envelopes and higher efficiency equipment, the effects of user behavior have become more significant [16]. While much has been explored in the domain of commercial buildings and their operation, residential markets have seen less analysis. Occupants of residential buildings often have increased control in how their environments are being managed because they have access to the thermostat. Occupants also have an increased ability to adapt to the environment by taking actions such as changing their clothing level or consuming a hot or cold beverage. In a residential setting, activities are more varied in comparison to the often sedentary activities engaged in at commercial buildings. As a result, residential studies can show diverse behavior and energy results. In Gram-Hanssen's [17] study of identically-constructed homes, the heating energy of the highest energy use home was three and a half times larger than the home with the lowest energy consumption. This variation is the result of the homeowner's behavior and habits; including choice of temperature setpoints. Through simulation it has been shown that individual behaviors and actions exceed the impact to changes in the building envelope such as upgrades to glazing or thermal insulation [18]. The variation, not just in expected user behavior but also in regions (and by extension climates) of study has been shown to significantly change the energy usage of homes [19]. It has been identified that users manage their thermostat settings as a multi-objective optimization problem, with over 25 potential influences being found [20]. Many of the influences deal with elements of comfort (e.g., thermal sensation, temperature control) while others with economics (e.g., heating price, family income).

2.1. Thermal comfort

Thermal comfort is ultimately a state of mind achieved by the subjective evaluation of a number of parameters including: metabolic rate, clothing insulation, air temperature, radiant temperature, air speed, and humidity [21]. The ability for users to adapt to the environment or make sacrifices for energy savings can redefine what is comfortable [22]. Occupants in a residential setting have been observed adapting their thermal behaviors based on local climate, expectation, and available control [23,24]. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) outlines the minimum requirements for acceptable thermal environments. ASHRAE Standard 55 (ASHRAE-55) [21] methods are allowed to be applied on all residential and commercial occupied spaces. The general model links comfort temperature with variables including metabolic rates of users, clothing adaptations, and operative temperatures (which accounts for both radiation and convection in the space). The prescribed method is based on the heat-balance methods determined by Fanger [25]. Fanger's method calculates the Predicted Mean Vote (PMV) which is a prediction of the mean value of thermal sensation as reported by a large group of users. The index ranges from -3 to $+3$ and maps to categories described as ‘cold’, ‘cool’, ‘slightly cool’, ‘neutral’, ‘slightly warm’, ‘warm’, and ‘hot’. ASHRAE-55 suggests keeping the PMV value within the range of $\pm$ 0.5. Unfortunately, thermal sensation and corresponding PMV do not necessarily translate well in practice to residential applications [23]. The PMV method is limited because it only accounts for physical changes by the user as they relate to comfort. For example, a user could experience different sensation based on removing a sweater or engaging in physical activity. Thermal comfort can also be affected by psychological acclimatization to conditions, habituation and/or expectation [24]. These non-physical factors to thermal sensation remain unaccounted for by the model.

ASHRAE-55 has a separate and alternative model to the PMV model for buildings where there is no mechanical cooling, and user comfort is more dependent on the adaptation of user. It is referred to as the *adaptive comfort model*. Fig. 1 graphically shows the prescribed boundaries outlined in the adaptive comfort model. The model linearly relates the acceptable operative temperature inside with the prevailing mean outdoor temperature (which must be between 10 °C and 33.5 °C). The prevailing mean outdoor temperature is the arithmetic mean of daily average temperature in the proceeding seven to 30 days. The prevailing mean can also have an additional weighting applied with a decay curve in which more weight is given to recent days and less to days further away. The standard has a set of boundaries at 80% and 90% acceptability limits (both indicated on Fig. 1) and suggests that users should be kept between the 80% acceptability limits. While the

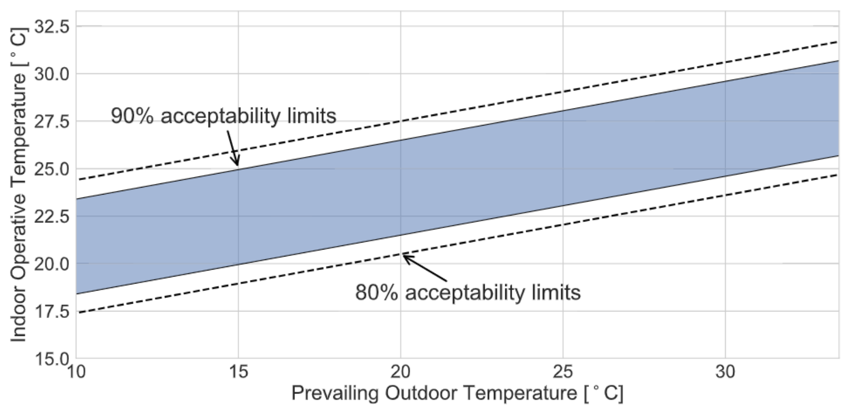

Fig. 1. Acceptability limits for the adaptive comfort model. Figure recreated from ASHRAE-55 [21].

adaptive model addresses psychological effects not covered by the PMV model, it applies only to a subset of residential users (i.e., only those who have no cooling equipment installed) and only during periods of time where those occupants are not using their heating system. Similar adaptive models have been extended to additional seasons of the year in the Netherlands, and varied to include a greater range of building types [26].

In addition to outdoor temperature affecting what is comfortable for the user, observations have been made that user variation between climate regions were larger than what could be explained by differing clothing levels [27]. Past exposure and experience have been found to change individuals perceived levels of thermal comfort [24,28]. Investigations have shown that preferred indoor temperature is dependent on the outdoor temperature [10] and found users would consider themselves warmer in winter than they would be in summer even though the measured room temperatures were lower [29]. In general, the lack of residential data available in previous studies means the understanding of comfort and occupants may need revision when large datasets become available [30].

2.1.1. Cost-based considerations on thermal preferences

How a user operates their home thermally has been investigated as a function of the costs that are being incurred to keep it at that level of comfort or utility. Pimbert and Fishman [10] observed that occupants in the United Kingdom were insensitive to moderate changes in heating costs as it related to their preferred evening living room temperatures. Meanwhile, research into multi-unit residential buildings (MURBs) found those who were aware of the costs they were incurring through direct metering tended to keep their temperatures in more conservative ranges than those who were not [13,31,32]. While costs and their transparency to users were linked to setpoint selection, the fact that MURBs often cater to a specific socio-economic subset of the population cannot be ignored. In a national study of 1000 homes in the U.K., income was identified as a strong indicator of room temperature, with lower income homes found to be 3 °C lower on average in the heating season [33]. Kelly et al. [34] also found income was an indicator of setpoint selection.

2.2. User types

There can be large advantages to researchers when behaviors fall easily into identified user types. Studies show that a user's behavior will remain consistent even when presented with a new control ability. For example, a user who did not manually setback the temperature with a manual thermostat is not likely to then start using a setback when given a programmable thermostat [6,35]. When looking at entire energy usage for homes in the U.K., eight different customer archetypes based on energy usage were identified as: ‘pioneer greens’, ‘follower greens’, ‘concerned greens’, ‘home stayers’, ‘unconscientious wasters’, ‘regular wasters’, ‘daytime wasters’, and ‘disengaged wasters’ [36]. The different user types were based on the combination of the energy efficiency levels of the properties, the ‘greenness’ of the occupants (i.e., their general energy efficiency), and the amount of occupancy during the daytime. Santin [35] looked at developing user types based on housing characteristics and energy usage patterns; identifying five different classes of user. The user types were identified as: ‘spenders’, ‘affluent-cool’, ‘conscious-warm’, ‘comfort’, and ‘convenience-cool’. The models of Zhang et al. [36] and Santin [35] only shared the characteristic of how energy-efficient users appeared. Ren et al. [37] investigated the temperatures and energy usage of affordable housing units and found users could be clustered (by unsupervised methods) into six different categories. The clusters were based on the average and the standard deviation of room temperature for days. The majority of the apartments were found to fall in only three of the six clusters. While many studies have said the behaviors of occupants remain discrete or grouped in nature, O'Brien et al. [38] found that occupants are more accurately modeled as a continuum in preference selection instead of as discrete types.

3. Data

3.1. Data description

The data used for this analysis came from the ‘Donate Your Data’ program administered by ecobee Inc., a connected thermostat manufacturer. The data is ‘donated’ by users who agreed to have their anonymized data released to various research partners. These research partners (e.g., academic labs, non-government organizations, research institutes) have agreed to use this dataset for non-commercial causes and that the results from the studies ultimately enter the public domain. The data consists of five-minute interval data that is measured by the thermostat and any connected remote sensors the users have placed in the homes. An example of the ecobee3 and a remote sensor are seen in Fig. 2. The thermostat is installed on the wall and connects to the home's heating, ventilation, and air conditioning (HVAC) equipment through the same wires a standard thermostat. Exact placement is dependent on where these wires are installed and is unique to each home. The remote sensors are independent, wireless, battery operated devices which measure temperature and occupancy and send information back to the thermostat over a radio channel. The sensors are recommended to be placed within 13.7 m of the thermostat and at a height of 1.5 m. Ideally they should be kept away from direct heating and cooling sources [39]. Table 1 shows the various data values contained in the in the interval data (left column) in addition to a small description (middle column) and associated units where applicable (right column).

The dataset contains data from the thermostat first connecting (or from January 2015 if actual online status occurred before that date) until the release of data in April 2017. Overall, this data contains over

Fig. 2. ecobee3 (left) and a remote sensor (right). The remote sensor is able to be placed in other rooms in the house, while the thermostat is mounted to the user's wall.

three million days of data. The dataset currently possesses 10,250 thermostats and 7,946 thermostat users. This disparity is a result of many users having multiple thermostats in their home, multiple properties, or users who have upgraded their thermostats from one ecobee product to another. For users with multiple thermostats, the most recent thermostat was selected based on the interval data. This reduced the number of available thermostats in the sample to be the same as the number of users. Currently, four generations of ecobee thermostats exist in the data, with two of the four able to connect to up to 32 remote sensors and/or having occupancy detection. In addition to the interval data, the dataset includes voluntary reporting of user-defined metadata including housing characteristics and equipment configuration. While provincial or state information is not available for all thermostats (given that it is user-inputted value), less than 2% of users were rejected based on that omission.

In Fig. 3 we show the thermostat location distributions across North America based on Canadian province and U.S. state. The Canadian territories (i.e., the Northwest Territories, the Yukon, and Nunavut) were removed for space considerations but none had any thermostats. It is seen that the majority of continental North America is covered by the data of at least one thermostat. The largest numbers of thermostats are found in the Canadian province of Ontario, and the states of Texas and California; which corresponds to the most heavily populated province and states in Canada and the U.S.

In Fig. 4, the number of days of available data per thermostat as separated by thermostat model is shown. Older generation devices, the Smart and SmartSi, are predominately found as devices with over 800 days of data. This makes sense considering they have been operational the longest and are no longer manufactured devices. The majority of devices in the dataset are the ecobee3. This generation of device has both remote sensing capabilities and occupancy detection. Two relative peaks can be seen in the data near 120 and 480 days. These are from the yearly relative increase in connections from holiday purchases. The dashed line indicates a cut off of one year (365 days). If a thermostat contained less than a year of data it was removed from climate or seasonal analysis. These analyses had a reduced sample of under 2,500 thermostats.

In Fig. 5, we show an example of the temperature readings for a single thermostat. The data shown is for a full year of operation. Fig. 5a (top) shows the outdoor temperature, which is gathered by a local weather station, and is provided by ecobee to the user. Fig. 5b (middle) contains the control temperature, which is the temperature value being used to control the system and compared to the programmed setpoints. Fig. 5c (bottom) shows the temperature readings of all the remote sensors connected to the thermostat and the temperature reading of the thermostat itself. Fig. 5c illustrates the variation in temperature between the thermostat and remote sensors. The remote sensors are designed to measure temperature at various points within a home. These sensor measurements are weighted either by the programmed thermostat schedule or by occupancy in the home. The choice of configuration is made by the user. In the simplest configuration, all available data points are weighted equally from across the home and used as the control temperature (Fig. 5b). It can be seen when comparing the control temperatures (Fig. 5b) and sensor readings (Fig. 5c) how the control temperature can be very similar to one of the sensor readings but generally appears as a blend of multiple sensors. There is also evidence of missing data (e.g., before May 2015) which can happen when power outages occur or if the thermostat loses connection with the user's wireless internet network. The median value of missing interval data per thermostat is 15%.

3.2. Climate region definitions

To be able to better group regions by similar climates, the available dataset was mapped to the Building America Climate Zones [40]. The Building America Climate Zones are based on the average heating and cooling degree days of the region along with the average annual precipitation [40]. The similar zones of Cold and Very-Cold along with the zones Mixed-Dry and Hot-Dry were grouped together further (shown in the left column of Table 2) to reduce the number of unique zones. A similar methodology is applied by the Energy Star metric for connected thermostats [41]. The Building America Climate Zones span over multiple ASHRAE climate zones. For comparison, the associated ASHRAE zones are included in the middle column of Table 2. Of the less than 2,500 thermostats eligible for the analysis, over 350 thermostats (approximately 14%) could not be properly mapped to a climate zone and would include thermostats located in Canada. The number of thermostats mapped to each climate zone is in the right column of Table 2. The Cold/Very Cold zone is the largest region by population of thermostats.

Table 1
Description of the interval data collected.

Data Point	Description	Units
temperature	measurement from remote sensor(s) and thermostat device (± 1.0 °F)	°F
relative humidity	measurement from thermostat device	%
outdoor temperature	measurement taken from local weather station	°F
setpoints	heating and cooling bounds	°F
equipment runtime	measured duration by thermostat device	seconds
motion	state of occupancy detection based on PIR sensor reading	boolean
schedule	user-defined comfort period (e.g., home, away, sleep, etc.)	NA
events	items override the set schedule (e.g., holds, vacations, demand response events, etc.)	NA

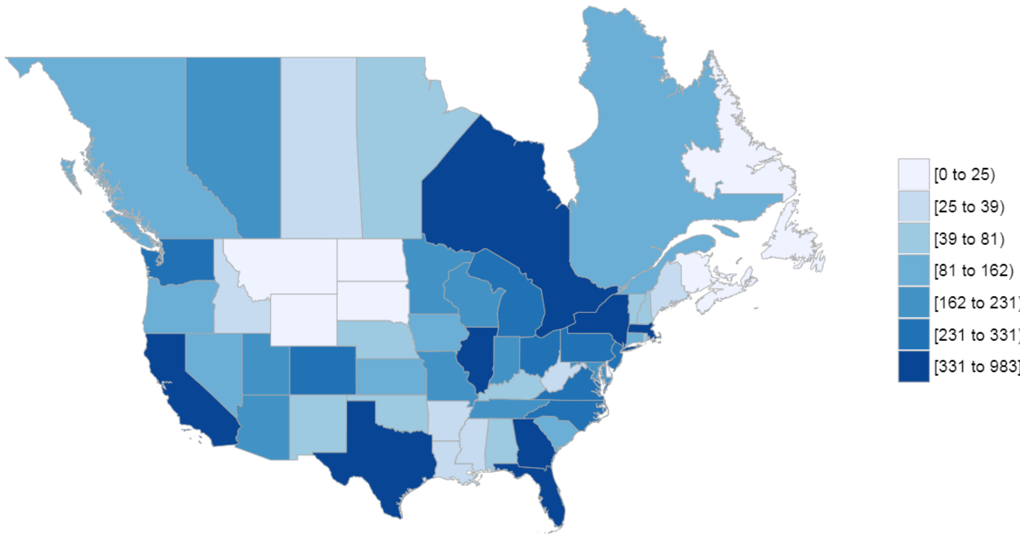

Fig. 3. Number of thermostats in the continental U.S. states and Canadian provinces. The Yukon, Northwest Territories, and Nunavut were removed for space considerations.

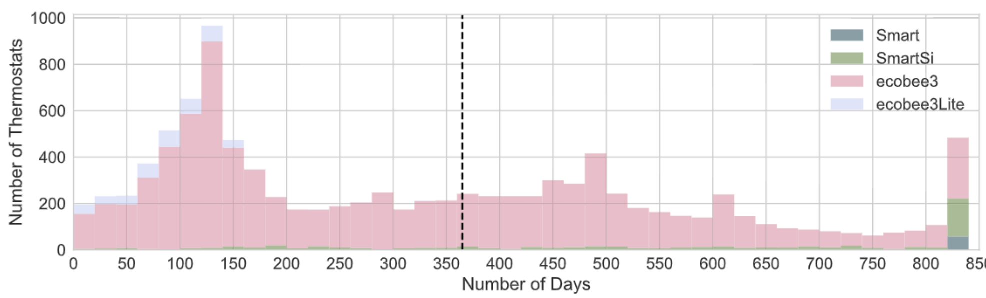

Fig. 4. Number of days of data per thermostat separated by thermostat model. The line indicates 365 days which was used a threshold in climate and seasonal analyses.

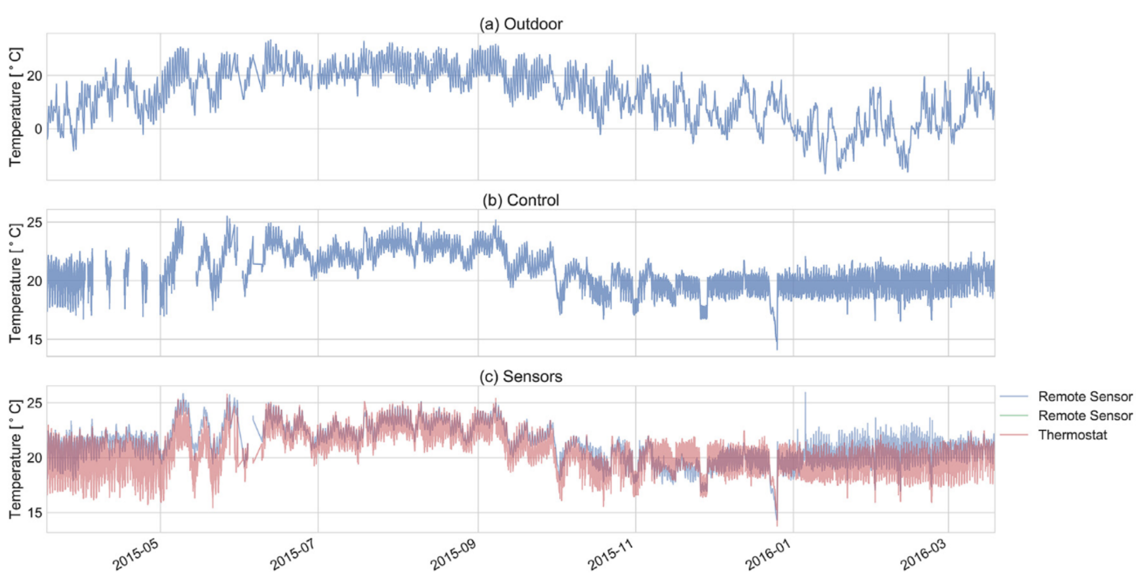

Fig. 5. Sample of the (a) outdoor temperature data, (b) control temperatures, and (c) sensor temperatures for a single thermostat found in the 'Donate Your Data' dataset.

3.3. Season definitions

When considering the effects of seasonality, it is conducive to define core heating and core cooling seasons in comparison to the shoulder seasons (i.e., the transitional periods between the two core seasons). As opposed to defining seasons by fixed date calendar-based changeovers or calculations of heating degree days/cooling degree days, the basis was made using the observed runtime of the HVAC equipment controlled by the thermostat on a given day. On a single day in either of the core seasons, it was assumed a user's equipment would run heating or

Table 2
Number of thermostats identified in each Building America Climate Zone along with the corresponding ASHRAE Climate Zones.

Building America Climate Zone	ASHRAE Climate Zone	Number of Thermostats
Cold/Very-Cold	5/6/7	960
Hot-Humid	2A/3A	276
Marine	3C/4C	179
Mixed-Dry/Hot-Dry	3B/4B	219
Mixed-Humid	3A/4A	471
Subarctic	8	0

cooling equipment but not the other. For example, a day that has heat runtime greater than zero minutes and no cooling runtime would be considered core heating. A shoulder season day was considered any day that did not qualify as either core heating or core cooling. During shoulder seasons, users would also be considered to be less likely to run their equipment and more capable of operating their building as a naturally-conditioned space taking actions such as opening the windows.

3.4. Utility rate information

State-level residential prices for natural gas and electricity were collected from the U.S. Energy Information Administration (EIA) [42,43]. Average yearly data was collected for 2015 and 2016 and averaged over the two years for each state. Thermostats were then matched by their state information to these rates. Similar rate information was not incorporated for Canadian provinces. While cooling is achieved using electricity, heating could be provided by multiple fuel sources. Natural gas and electricity account for 83% of the heating source for the U.S. [2]. with other fuels such as kerosene and distillate fuel oil making up the balance. Based on the metadata, users with a heat pump were assumed to use electricity as their heating source, else natural gas was assumed. The metadata does not provide additional information on equipment configurations such as distinguishing between a furnace and a boiler.

4. Results and analysis

The investigation into the first question (**Q1**), regarding what factors could be contributing to how users are operating their thermostats is explored in Sections 4.1, 4.2, and 4.3. Meanwhile, the investigation into the second question (**Q2**), which focused on potential user types based on behavior is in Section 4.4.

4.1. Seasonal variation in thermostat behavior

As temperatures and seasons change, thermal preferences are expected to change because of factors including changes to clothing level [21]. Hence, the temperatures selected by users on their thermostats are expected to change in order to maintain a similar level of temperature acceptability. For each thermostat which met or exceeded the criteria for minimum number of days online, the daily average control temperature was calculated. Average outdoor temperature was calculated similarly for each day. To determine prevailing outdoor temperature, the previous 30 days' outdoor temperature was averaged with equal weighting. Each day was further identified as being in the cooling, heating, or shoulder season as described in Section 3.3.

In Fig. 6, we show the scatter plot of average indoor control temperature versus prevailing outdoor temperature for all heating and cooling days. The 7 °C temperature band for the 80% comfort acceptability limits from the adaptive comfort model are indicated. The cooling season's prevailing outdoor temperature limit matches well with that of the adaptive comfort model's limits, while the heating

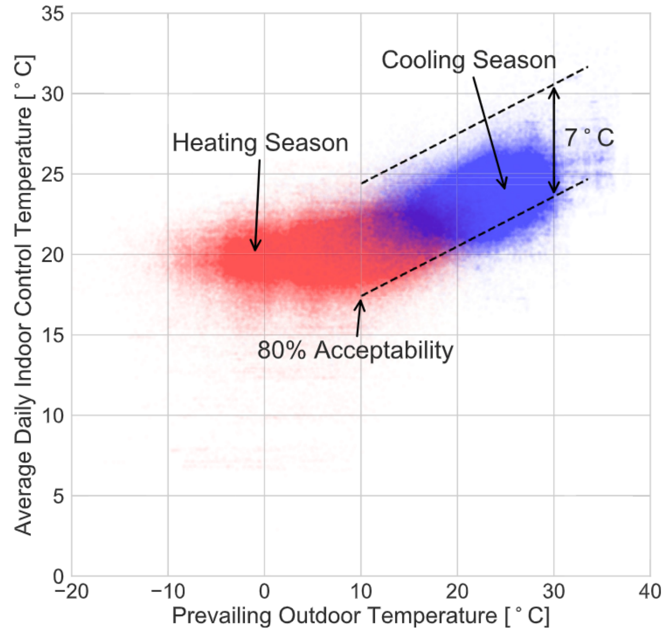

Fig. 6. Average daily indoor control temperatures based on the prevailing outdoor temperature for both heating (red) and cooling (blue) season days. The 80% comfort acceptability limits from ASHRAE-55's [21] adaptive comfort model are included over its prescribed range of temperatures. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

season overlaps only in the region from 10 °C to 20 °C. The heating season's prevailing outdoor temperature is found to extend to a much colder temperature (lower than −10 °C) than the adaptive model was intended. It should not be expected that users accept or continually allow their indoor environment to get colder. A linear regression was performed on the heating and cooling season data independently. The R^2 value for the heating season is 0.089 and 0.159 for the cooling season. The slope for the heating season was found to be 0.075 and the cooling slope was calculated as 0.140. Neither the heating or cooling season relation is found to agree with the slope value of the adaptive model; which is prescribed at a value of 0.31 [21]. This difference in relations between indoor temperature and outdoor temperature alludes to two separate behavior strategies that need to be considered by a model. As an aggregated sample, the switch between the two models would appear at the overlap of the two seasons, (approximately at a prevailing outdoor temperature of 15 °C). This value is similar to the base temperature of various heating and cooling degree day methods [44]. Fig. 6 contains outliers – particularly during the heating seasons – where some homes reached temperatures as low as 12–13 °C. These are likely a result of users who reside at secondary properties or who take extended vacations. In these cases, heating is utilized more as a safety barrier to prevent damage at the property than for occupant comfort.

Fig. 7 shows the similar indoor-outdoor temperature relation but for the days considered in the shoulder season. Once again, the 80% acceptability bounds from ASHRAE-55 [21] are included. The linear regression on the shoulder season data calculated the slope to be 0.222 with an R^2 of 0.267 – the closest to the adaptive models of the heating, cooling and shoulder seasons used for comparison. Similar to the heating season, the prevailing outdoor temperature region extends more than 10 °C lower than the adaptive model.

Fig. 8 presents the heat maps for (a) heating, (b) cooling, and (c) shoulder seasons for average daily indoor temperature and prevailing outdoor temperature along with the 80% acceptability bounds. Similar to Figs. 6 and 7 in content, the heat maps better illustrate the amount of data present in the previous scatter plots. In all three sub-figures, a density of data is seen to be contained by the prescribed acceptability

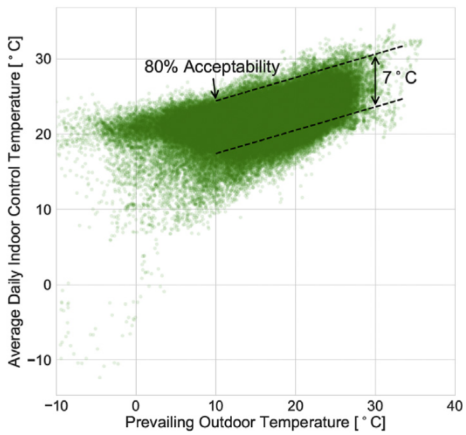

Fig. 7. Average daily indoor control temperatures based on the prevailing outdoor temperature for shoulder season days. The 80% acceptability limits from ASHRAE-55's [21] adaptive comfort model are included over its prescribed range of temperatures.

bounds of the adaptive comfort model but the cooling (Fig. 8b) and shoulder (Fig. 8c) are much better captured.

Based on the analyses, the adaptive model from ASHRAE-55 [21] holds for a large fraction of residential operations in the shoulder season. Of note, the upper bound on prevailing outdoor temperature seen here (approximately 30–35 °C) is very similar to the bounds on the model in the adaptive model (33.5 °C) while the lower bounds are exceeded. Many observations are found at a lower prevailing outdoor temperature or an associated lower indoor temperature than the model would predict. The defined temperature bounds prescribed by the adaptive model do not seem to be violated by the application into the cooling seasons but does not capture well the heating season's behavior. In addition to the temperature regions, the slopes for the seasons are considerably less. The heating slope is particularly shallow, and almost flat. This is similar to the results observed by Pimbert and Fishman [10] who found users reach a limit where they will go no lower in their room temperature as outdoor temperature changes and demand a certain minimum temperature. The low R^2 values of these regressions indicate that the outdoor temperature is not a strong indicator of thermostat setpoint selection for the heating and cooling seasons. However, we must acknowledge that setpoint temperatures do not represent the complete optimal comfort conditions outlined by ASHRAE-55's [21] adaptive comfort model. Those are influenced by variables we are unable to capture such as radiant temperatures or airspeed within the space.

4.2. Regional variation in thermostat preferences

We were able to map a subset of users to the six Building America Climate Zones identified in the left column of Table 2. For each thermostat, each day labeled as heating or cooling season was collected. Days identified as part of the shoulder season were not considered. For each of the heating season days, the median heating setpoint for that day was calculated. For each of the cooling season days, the median cooling setpoint was used. The decision to use the median setpoint was based on the setpoint distribution for an individual thermostat. In Fig. 9, the heating and cooling setpoint distributions for 20 thermostats are shown. Each thermostat is its own split violin with heating setpoints on the left and cooling setpoints on the right. The 20 thermostats were randomly selected from a group of thermostats who were active for at least a year and half. The heating (and cooling) setpoints for each individual thermostat are not normally distributed and instead occur at a small number of distinct values. For example, Thermostat 2 appears to have two main setpoints in heating (left) and three in cooling (right). This setpoint behavior reflects that only a few schedule periods are set (i.e., thermostats may only be set at 'home' and 'away' periods) and/or that they are not actively redefining the setpoints of the scheduled periods over the observation window.

The distributions of both heating and cooling setpoints based on climate zone are shown in Fig. 10. Each climate zone is its own violin plot, with each violin being split. The left distribution of each violin is for heating setpoints and the right the cooling setpoints. For each climate zone and both setpoints, lines denoting the 25th, 50th and 75th percentiles are included. The Mixed-Dry/Hot-Dry has the highest median setpoint for cooling of the climate zones, while Mixed-Humid has the lowest. The Marine climate zone has the lowest median heating setpoint. The Marine region also has the largest comfort band (difference between heating and cooling medians) indicating a greater range of temperature adaptability than other climate zones. The individuals in the Cold/Very-Cold region seem to tolerate a lower heating setpoint than many of the other regions.

To understand the independence of the populations, and if observations could be considered statistically significant, a *t*-test was performed on the combination of heating and cooling setpoint. Table 3 shows the p-values for the testing on heating setpoints. Table 4 shows the p-values for the testing on cooling setpoint distribution. In comparison of the Hot-Humid and Mixed-Dry/Hot-Dry regions (two relatively similar regions) there is a preference for higher cooling setpoints in the less humid (Mixed-Dry/Hot-Dry) of the two regions, presumably because of humidity's effect on comfort. The p-value for heating (Table 3) and cooling (Table 4) between these two regions are both below the 0.05 threshold for significance, indicating the distributions are statistically different. For heating (Table 3), only the *t*-test for

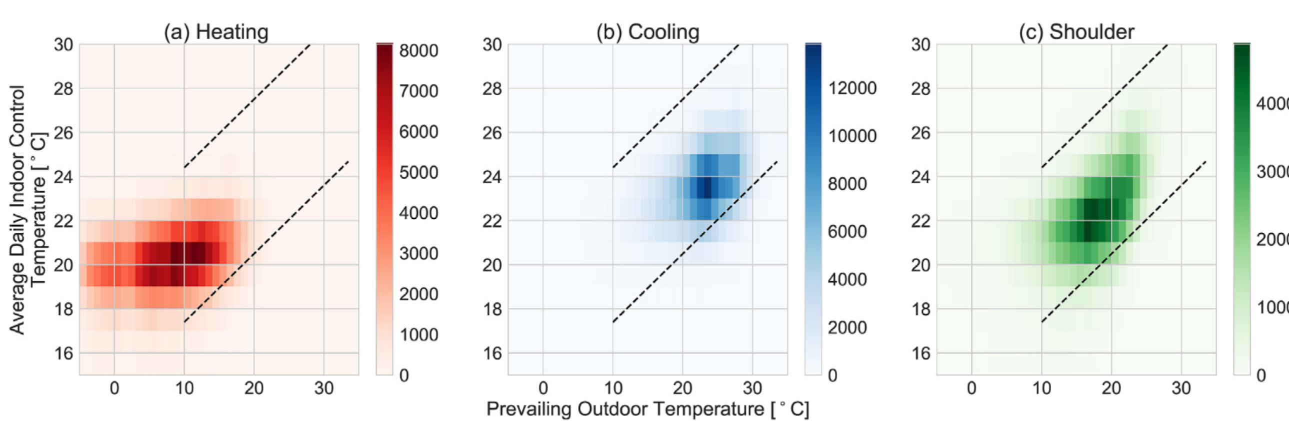

Fig. 8. Heat maps of average daily indoor control temperatures based on the prevailing outdoor temperature for (a) heating, (b) cooling, and (c) shoulder season days. The 80% comfort acceptability lines from ASHRAE-55's [21] adaptive comfort model are included over its prescribed range of temperatures.

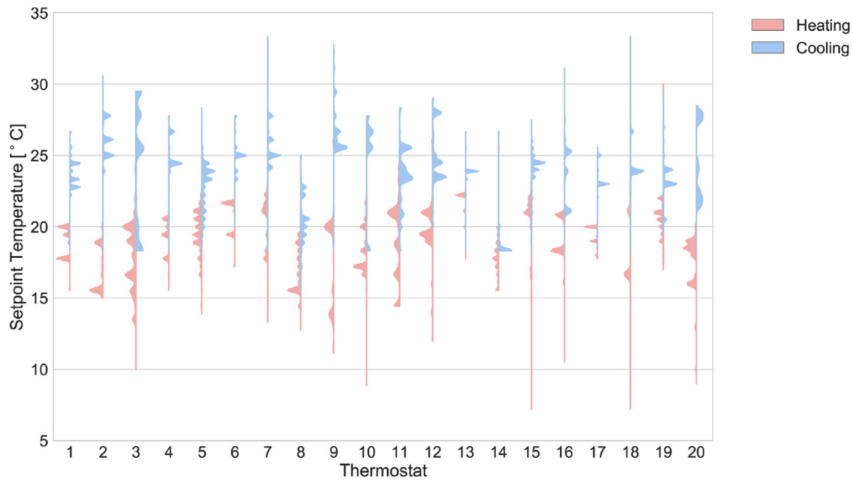

Fig. 9. Distributions in heating and cooling setpoints for 20 randomly selected thermostats.

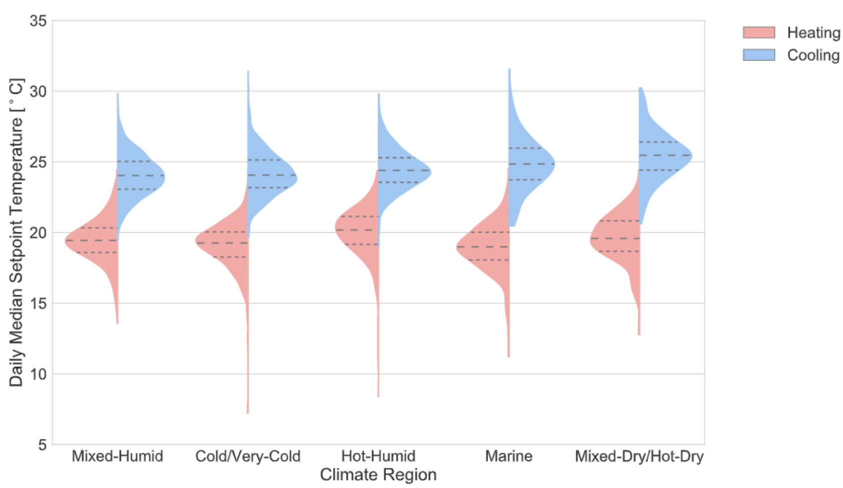

Fig. 10. Setpoint temperature distributions in the various Building America Climate zones. For each violin, the 25th, 50th and 75th percentiles are indicated.

Table 3
P-value results from *t*-tests comparing the heating setpoint distributions. P-values deemed not significant based on a threshold of 0.05 are in bold.

	Mixed-Humid	Cold/Very-Cold	Hot-Humid	Marine	Mixed-Dry/Hot-Dry
Mixed-Humid	1	0.0081	0.0001	0.021	**0.17**
Cold/Very-Cold	0.0081	1	2.3e-09	**0.41**	0.00098
Hot-Humid	0.0001	2.3e-09	1	7.5e-07	0.035
Marine	0.021	**0.41**	7.5e-07	1	0.0024
Mixed-Dry/Hot-Dry	**0.17**	0.00098	0.035	0.0024	1

Table 4
P-value results from *t*-tests comparing the heating setpoint distributions. P-values deemed not significant based on a threshold of 0.05 are in bold.

	Mixed-Humid	Cold/Very-Cold	Hot-Humid	Marine	Mixed-Dry/Hot-Dry
Mixed-Humid	1	**0.32**	0.00058	4.7e-05	6.8e-20
Cold/Very-Cold	**0.32**	1	0.0031	0.00019	2.2e-19
Hot-Humid	0.00058	0.0031	1	**0.061**	6e-10
Marine	4.7e-05	0.00019	**0.061**	1	0.0042
Mixed-Dry/Hot-Dry	6.8e-20	2.2e-19	6e-10	0.0042	1

Mixed-Humid with Mixed-Dry/Hot-Dry, and Cold/Very-Cold with Marine did not meet the p-value criteria of 0.05. Meanwhile, for cooling (Table 4) only the results for Cold/Very-Cold with Mixed-Humid, and Marine with Hot-Humid did not exceed the threshold.

Previous research has indicated that people may adapt thermally to long-term exposure to different climates [24,28]. The distributions shown in Fig. 10 show this to be the case for connected thermostat users. For example, in extreme climate zones, such as Cold/Very-Cold, setpoint preferences shows individuals who have become acclimatized to, or prepared for, colder temperatures. As such, setpoint preferences

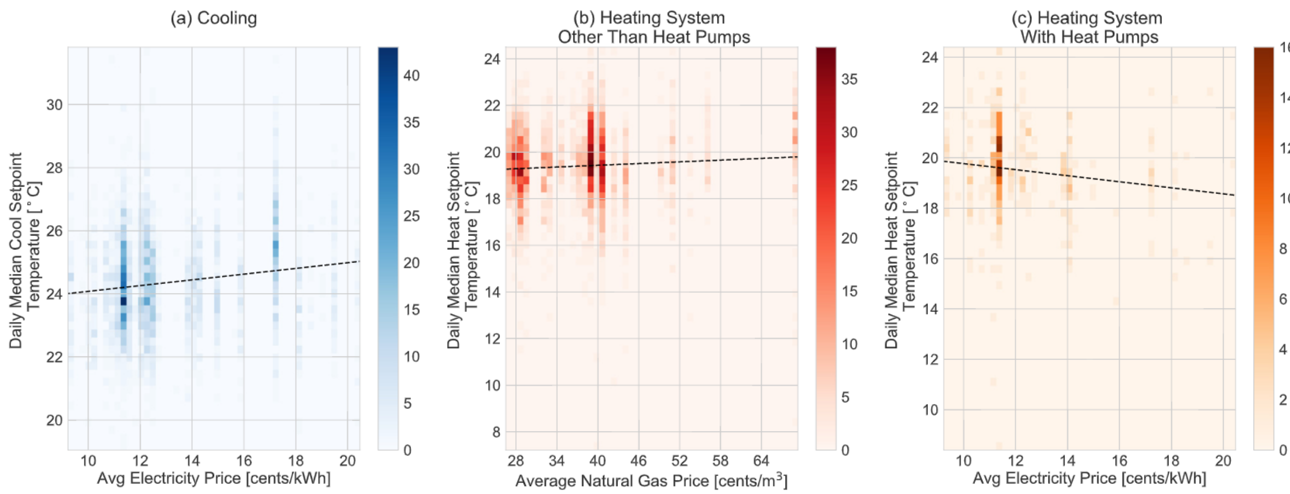

Fig. 11. Heat maps for indoor setpoint temperature for (a) the cooling season and electricity prices, (b) heating season and gas prices for homes with a heating system other than heat pumps, and (c) heating season and electricity prices for homes with heat pumps. Regression lines for all three relations are included.

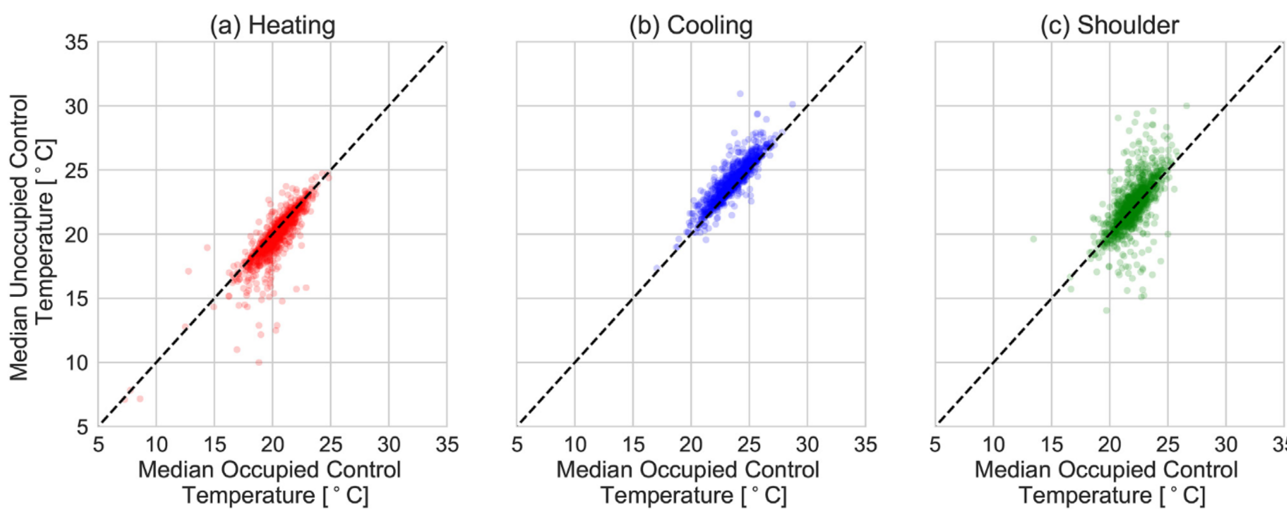

Fig. 12. Median control temperatures for each thermostat in occupied and unoccupied periods in each of the (a) heating, (b) cooling, and (c) shoulder seasons. The identity line is included on each subfigure.

for connected thermostat users should be considered as a function of their climactic region.

4.3. Cost considerations effecting thermal preferences

While energy reducing behaviors generally coincide with a reduction in spending on space conditioning by the household, we investigated a more direct relation between the costs associated with heating and cooling. For cooling it was assumed all systems were electric. For heating, if the thermostat was listed as having a heat pump by the provided metadata it was assumed electricity was used for heating; otherwise it was assumed that natural gas was the main heating source. Electricity and gas prices were taken from the U.S. Energy Information Administration [42,43]. The cooling setpoint temperatures were taken from cooling season days and heating setpoint temperatures from heating season days. Fig. 11a shows the median cooling setpoint temperature as a function of the state's average electricity price. A linear regression was conducted and the line plotted. No statistically significant relation was found. Fig. 11b shows the median heating setpoints and average natural gas prices for homes without heat pumps. The regression results here show an increasing slope which would be counter to a cost aversion strategy as heating setpoints are increasing with price, however the R^2 value at 0.003 is too low to consider this a meaningful result. Finally, in Fig. 11c we show the median heating setpoint and the average electricity prices for homes with heat pumps. The regression trend is of a negative slope, indicating higher costs and lower setpoints are related. This regression had an R^2 value of 0.02, and similar to Fig. 11b, is too low to be considered significant.

There is in general a lack of variation in users' preferred temperatures as a function the utility prices; a finding counter to other previous studies. Unlike previous studies, we are unable to say for certain if individuals were being made aware of their prices in an approach that would have actively encouraged conservation. A potentially large effect could be the self-selection bias that exists in participant groups in these past studies. For example, previous studies have relied on MURBs which often are occupied by a less affluent socio-economic demographic. Meanwhile, connected thermostats, given their relatively high price cater often to a more affluent demographic and who give more priority to comfort than cost savings. However, as the penetration of connected thermostat devices increases in market, particularly though revisions to building codes and through utility programs and partnerships, the implicit solution bias in the sample should improve. Even with the potential bias in the data, the data remains an invaluable resource and should yield further insights with additional investigation.

4.4. Identification of user types based on how users operate their thermostat

How a user operates a thermostat can be categorized by a number of different tendencies in their actions. Given actions in how users treated occupied and unoccupied periods, allowed schedules to run, and generally selected setpoint temperatures, distinct user types such as energy efficient/aware occupants were attempted to be identified. Initially it was explored how users operate a thermostat differently between their 'occupied' and 'unoccupied' periods. When analyzing the data from a single thermostat, if an occupancy state was detected by any available motion sensor, the period for an hour before and hour after was considered occupied. All periods from midnight until 6am were considered occupied. It was assumed that people were home but asleep and not trigging occupancy sensors. The median control temperature for both occupied and unoccupied times was calculated for all the heating, cooling, and shoulder seasons for each user. Fig. 12 shows the median temperature for the occupied and unoccupied periods for (a) heating, (b) cooling, and (c) shoulder seasons. A data point on this line would indicate no change between the occupied and unoccupied periods in how the thermostats were being set to keep the temperature. In Fig. 12a the data is found beneath the unity line meaning users do allow a lower temperature when not home. Fig. 12b has the data above the unity line which is the result of temperatures increase when the home is not occupied. Both of these are consistent with setbacks (and set-ups) being applied for the unoccupied periods. Fig. 12c shows that during the shoulder season the unoccupied and occupied temperatures remained similar based on their centering on the unity line. Since shoulder seasons may have both or neither of cooling and heating requirements the temperature can drift both above and below the setpoint for a user over the sample history.

Fig. 12 indicates setbacks and set-up strategies are generally being applied by the population of the thermostats but does not indicate how users are achieving these results. For example, some users rely on keeping to a schedule while some users rely on prolonged holds that they adjust. For each thermostat, the fraction of time that they had a hold applied and overriding their schedule was calculated. A hold is a manual and deliberate override of a scheduled program on the thermostat. Depending on how the user has configured the thermostat the duration is variable with some people electing for the hold to be in place for a few hours while some users elect to have it in place indefinitely. Once the hold has been lifted, the thermostat resumes its normal scheduled routine. In Fig. 13 we show a histogram of these fractions. Only approximately 100 thermostats were seen to have a fraction above 0.9. In fact, 75% of users spend less than one third of the time in a hold. The most common fractions (on the left of the histogram) appear at less than 0.1. Approximately 10% of users rely heavily on setting a hold to control their thermostat and spend more than 50% of their time in a hold.

Fig. 14 presents the median cooling setpoint and the median heating setpoints for the thermostats. An identity line is included to show the relation of no change between the two setpoints. Users on this line would be setting a single setpoint and leaving it there indefinitely. It is seen that the grouping of users is found above this line, indicating that users are setting different temperatures for both heating and cooling. No distinct clusters are observed which could be used to define user types. For example, users who were all found in the top left corner, would have a very wide comfort band and would be behaving in an energy saving way. While there are some users who are in that region, they are not distinctly separated from the general population.

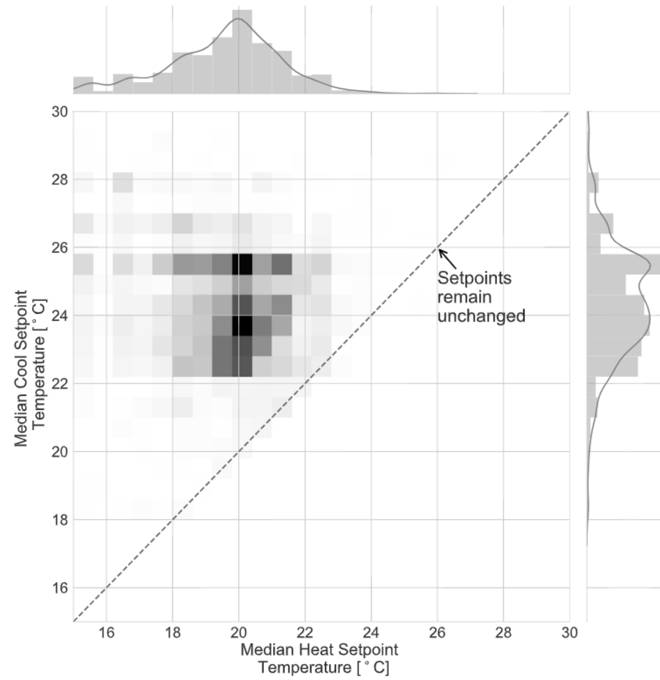

Fig. 14. Heat map for median heating and cooling setpoint for each thermostat. The identity line is included.

Observing the three different behavior analyses (Figs. 12–14), the connected thermostat population appears to have fairly uniform behavior. The data appears in a single tightly grouped region of the various feature-spaces defined. Applying a label, such as 'energy-savers', to a subset of the population as was done by others [35,36] simply by identifying their location on the feature-space would only classify a small fraction of the overall population. Similar to the observation made by O'Brien et al. [38], the users appear more as a continuum across the variables of occupied versus unoccupied setpoints (Fig. 12), holds (Fig. 13), and heating and cooling setpoints (Fig. 14). With the continued growth of this data set and continued penetration of this and other smart devices into the home, the ability to identify these behavior-based user types could still exist with broader investigations and is left as future work. Specifically, the use of other variables or inclusion of more latent characteristic features such as house size or age of occupant are considered as next steps.

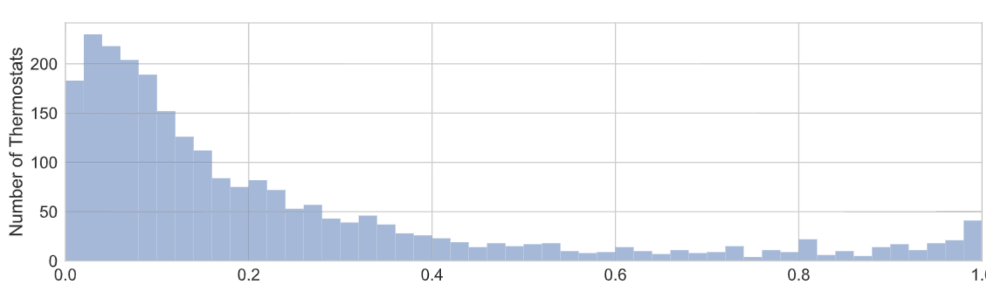

Fig. 13. Fraction of time spent in holds for each thermostat.

5. Conclusions

In this article, we analyzed ecobee's 'Donate Your Data' dataset. This open dataset contains more than three million days, for more than 10,000 connected thermostats. The usage of this data source overcomes many challenges that researchers have previously faced when conducting residential behavior analyses which often were related to the data sampling and collection. Two specific questions were posed. Question 1 (**Q1**) asked if thermostat behaviors changed based on seasons, climates, and utility pricing. We found the thermal preferences of user's displayed seasonal variations, which generally agreed with existing models of users' thermal preferences. Indoor temperature in heating season, was found to not correlate significantly with prevailing outdoor condition. Evidence of users having regional variation indicating varied expectations and levels of adaptation was observed based on climate zone to a statistical significance in most cases. For temperature preferences as a function of pricing signals, no relations of statistical significance were observed for users in their cooling and heating energy prices. Question 2 (**Q2**) asked if user types could be identified using the interval data based on differences in occupied and unoccupied setpoints, how holds were being set, and generally the heating and cooling setpoint selection. It was determined that users' temperature control behaviors generally reflect proper use of a thermostat based on scheduling and override actions using holds. Their behavior showed a consistent usage of setback (or set-ups) during unoccupied periods; which is ultimately part of the value proposition of connected thermostats. The majority of users do not rely on holds to operate their connected thermostat but a small fraction of users do still appear to adjust their thermostat manually. Finally, we showed that the general heating and cooling setbacks used do not vary drastically between users, nor do certain visible clusters of users appear. While the dataset provided a new and unique opportunity for investigation, limitations in our analyses remain. The sample is biased towards homeowners, early adopters, and a potentially wealthier subset of the population. The thermostat setpoint settings and other interactions are an indicator for optimal comfort, but they do not incorporate factors like cost. Moreover, thermal comfort is a function of air temperature, mean radiant temperature, relative humidity, and airspeed. In this paper, we only focused on air temperature (including aggregated air temperature from multiple sensors in the home); the ecobee thermostats nor their remote sensors measure mean radiant temperature or airspeed. Lastly, the thermostat setpoints are partly a function of occupant behavior, but also a function of the control logic. Thus, results must be interpreted in the context that they cannot be directly compared to studies involving simple thermostats without programmable or advanced logic features.

Acknowledgements

This work is funded by a research grant provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) (CRDPJ 508857-17) and ecobee Inc.

The authors wish to make it known that The Corresponding Author has previously been a full-time employee and remains an employee in a diminished role for the organization (ecobee Inc.) while pursuing his studies. The organization provided financial support to the research and do administer the open dataset used in the investigation.

References

[1] Energy Information Administration, Residential Energy Consumption Survey (RECS), (2009) https://www.eia.gov/consumption/residential/data/2009/.

[2] Energy Information Administration, Residential Energy Consumption Survey (RECS), (2015) https://www.eia.gov/consumption/residential/data/2015/.

[3] Energy Information Administration, Annual Energy Review, (2011) https://www.eia.gov/totalenergy/data/annual/.

[4] T. Peffer, M. Pritoni, A. Meier, C. Aragon, D. Perry, How people use thermostats in homes: a review, Build. Environ. 46 (2011) 2529–2541, http://dx.doi.org/10.1016/j.buildenv.2011.06.002.

[5] L.W. Nelson, J.W. MacArthur, Energy savings through thermostat setback, Am. Soc. Heating, Refrig. Air-Conditioning Eng 84 (1978) 319–334.

[6] M.J. Nevius, S. Pigg, Programmable thermostats that go berserk? Taking a social perspective on space heating in Wisconsin, Proc. ACEEE Summer Study Energy Effic. Build, vol. 8, 2000, pp. 233–244 https://www.eceee.org/.

[7] M. Pritoni, A.K. Meier, C. Aragon, D. Perry, T. Peffer, Energy efficiency and the misuse of programmable thermostats: the effectiveness of crowdsourcing for understanding household behavior, Energy Res. Soc. Sci. 8 (2015) 190–197, http://dx.doi.org/10.1016/j.erss.2015.06.002.

[8] U.S. Environmental Protection Agency, Programmable Thermostats Suspension Memo, (2009) https://www.energystar.gov/.

[9] Over 40% of Thermostats Sold in 2015 Will Be Smart Thermostats, (2015) http://www.parksassociates.com/blog/article/pr0715-smart-thermostats.

[10] S.L. Pimbert, D.S. Fishman, Some recent research into home heating, J. Consum. Stud. Home Econ. 5 (1981) 1–12, http://dx.doi.org/10.1111/j.1470-6431.1981.tb00388.x.

[11] M. Shipworth, S.K. Firth, M.I. Gentry, A.J. Wright, D.T. Shipworth, K.J. Lomas, Central heating thermostat settings and timing: building demographics, Build. Res. Inf. 38 (2010) 50–69, http://dx.doi.org/10.1080/09613210903263007.

[12] D.J. Nevrala, S.L. Pimbert, Living room and bedroom temperatures in homes: recorded values, occupant preference and energy conservation potential, Build. Serv. Eng. Technol. 2 (1981) 109–118, http://dx.doi.org/10.1177/014362448100200301.

[13] H. Burak Gunay, W. O'Brien, I. Beausoleil-Morrison, A. Perna, On the behavioral effects of residential electricity submetering in a heating season, Build. Environ. 81 (2014) 396–403, http://dx.doi.org/10.1016/j.buildenv.2014.07.020.

[14] E. Vine, B.K. Barnes, Monitored indoor temperatures and reported thermostat settings: how different are they? Energy 14 (1989) 299–308, http://dx.doi.org/10.1016/0360-5442(89)90101-1.

[15] W. O'Brien, S. Gilani, H. Burak Gunay, In situ approaches to studying occupants, Explor. Occupant Behav. Build, Springer International Publishing, 2018, pp. 129–167, , http://dx.doi.org/10.1007/978-3-319-61464-9_6.

[16] P. Hoes, J.L.M. Hensen, M.G.L.C. Loomans, B. de Vries, D. Bourgeois, User behavior in whole building simulation, Energy Build. 41 (2009) 295–302, http://dx.doi.org/10.1016/j.enbuild.2008.09.008.

[17] K. Gram-Hanssen, Residential heat comfort practices: understanding users, Build. Res. Inf. 38 (2010) 175–186, http://dx.doi.org/10.1080/09613210903541527.

[18] F. Haldi, D. Robinson, The impact of occupants' behaviour on building energy demand, J. Build. Perform. Simul 4 (2011) 323–338, http://dx.doi.org/10.1080/19401493.2011.558213.

[19] S. D'Oca, V. Fabi, S.P. Corgnati, R.K. Andersen, Effect of thermostat and window opening occupant behavior models on energy use in homes, Build. Simul 7 (2014) 683–694, http://dx.doi.org/10.1007/s12273-014-0191-6.

[20] S. Wei, R. Jones, P. De Wilde, Driving factors for occupant-controlled space heating in residential buildings, Energy Build. 70 (2014) 36–44, http://dx.doi.org/10.1016/j.enbuild.2013.11.001.

[21] ASHRAE, ANSI/ASHRAE Standard 55-2017: Thermal Environmental Conditions for Human Occupancy, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc., 2017.

[22] R.L. Hwang, M.J. Cheng, T.P. Lin, M.C. Ho, Thermal perceptions, general adaptation methods and occupant's idea about the trade-off between thermal comfort and energy saving in hot-humid regions, Build. Environ. 44 (2009) 1128–1134, http://dx.doi.org/10.1016/j.buildenv.2008.08.001.

[23] R. Becker, M. Paciuk, Thermal comfort in residential buildings - failure to predict by Standard model, Build. Environ. 44 (2009) 948–960, http://dx.doi.org/10.1016/j.buildenv.2008.06.011.

[24] G.S. Brager, R.J. de Dear, Thermal adaptation in the built environment: a literature review, Energy Build. 27 (1998) 83–96, http://dx.doi.org/10.1016/S0378-7788(97)00053-4.

[25] P. Fanger, Analysis and Applications in Environmental Engineering, (1970).

[26] A.C. Van Der Linden, A.C. Boerstra, A.K. Raue, S.R. Kurvers, R.J. De Dear, Adaptive temperature limits: a new guideline in The Netherlands: a new approach for the assessment of building performance with respect to thermal indoor climate, Energy Build. 38 (2006) 8–17, http://dx.doi.org/10.1016/j.enbuild.2005.02.008.

[27] A.K. Mishra, M. Ramgopal, Field studies on human thermal comfort - an overview, Build. Environ. 64 (2013) 94–106, http://dx.doi.org/10.1016/j.buildenv.2013.02.015.

[28] J. de Dear Richard, S.B. Gail, J. Reardon, F. Nicol, E. Al, Developing an adaptive model of thermal comfort and preference/Discussion, ASHRAE Trans. 104 (1998) 145.

[29] N.A. Oseland, A comparison of the predicted and reported thermal sensation vote in homes during winter and summer, Energy Build. 21 (1994) 45–54, http://dx.doi.org/10.1016/0378-7788(94)90015-9.

[30] L. Peeters, R. de Dear, J. Hensen, W. D'haeseleer, Thermal comfort in residential buildings: comfort values and scales for building energy simulation, Appl. Energy 86 (2009) 772–780, http://dx.doi.org/10.1016/j.apenergy.2008.07.011.

[31] S. Andersen, R.K. Andersen, B.W. Olesen, Influence of heat cost allocation on occupants' control of indoor environment in 56 apartments: studied with measurements, interviews and questionnaires, Build. Environ. 101 (2016) 1–8, http://dx.doi.org/10.1016/j.buildenv.2016.02.024.

[32] A. Levinson, S. Niemann, Energy use by apartment tenants when landlords pay for utilities, Resour. Energy Econ. 26 (2004) 51–75, http://dx.doi.org/10.1016/S0928-7655(03)00047-2.

[33] D.R.G. Hunt, M.I. Gidman, A national field survey of house temperatures, Build. Environ. 17 (1982) 107–124, http://dx.doi.org/10.1016/0360-1323(82)90048-8.

[34] S. Kelly, M. Shipworth, D. Shipworth, M. Gentry, A. Wright, M. Pollitt, D. Crawford-Brown, K. Lomas, Predicting the diversity of internal temperatures from the English residential sector using panel methods, Appl. Energy 102 (2013) 601–621, http://dx.doi.org/10.1016/j.apenergy.2012.08.015.

[35] O.G. Santin, Behavioural patterns and user profiles related to energy consumption

for heating, Energy Build. 43 (2011) 2662–2672, http://dx.doi.org/10.1016/j.enbuild.2011.06.024.
[36] T. Zhang, P.O. Siebers, U. Aickelin, A three-dimensional model of residential energy consumer archetypes for local energy policy design in the UK, Energy Pol. 47 (2012) 102–110, http://dx.doi.org/10.1016/j.enpol.2012.04.027.
[37] X. Ren, D. Yan, T. Hong, Data mining of space heating system performance in affordable housing, Build. Environ. 89 (2015) 1–13, http://dx.doi.org/10.1016/j.buildenv.2015.02.009.
[38] W. O'Brien, H.B. Gunay, F. Tahmasebi, A. Mahdavi, A preliminary study of representing the inter-occupant diversity in occupant modelling, J. Build. Perform. Simul 10 (2017) 509–526, http://dx.doi.org/10.1080/19401493.2016.1261943.
[39] Ecobee Inc, Room Sensor Installtion Guide, ecobee Inc, 2015, www.ecobee.com.
[40] Volume 7.3 Guide to Determining Climate Regions by County, (2015).
[41] U.S. Environmental Protection Agency, Method to Demonstrate Field Savings Version 1.0 (Rev. Dec-2016), (2016) https://www.energystar.gov.
[42] U.S. Energy Information Administration (EIA), Natural Gas Residential Price, (2017) https://www.eia.gov.
[43] U.S. Energy Information Administration (EIA), Average Retail Price of Electricity, (2017) https://www.eia.gov.
[44] M. Isaac, D.P. van Vuuren, Modeling global residential sector energy demand for heating and air conditioning in the context of climate change, Energy Pol. 37 (2009) 507–521, http://dx.doi.org/10.1016/j.enpol.2008.09.051.