

Long-Term Trends in Internet-Connected Thermostat Settings and Performance

Abigail Daken, Environmental Protection Agency
Leo Rainer and Alan Meier, Lawrence Berkeley National Laboratory
Ethan Goldman, Resilient Edge
Abhishek Jathar, ICF

ABSTRACT

Internet-connected thermostats capture half of annual thermostat sales and as of 2020 were in about 30 million US homes. To be considered by ENERGY STAR®'s certification program, thermostat vendors must periodically submit field performance summaries for a random sample of up to 1250 homes from five climate zones. This procedure has created a unique record of each thermostat's long-term performance in response to changes in weather, customer demographics, building stock, and algorithms, and, recently, the pandemic. Thirteen vendors submitted data up to 11 times across 6 years. These submissions generated relatively stable trends for comfort temperatures and HVAC runtimes over the study period. Heating systems in the very cold climate typically run 700 hours/year, while air conditioners in the hot-humid climate operate about 1400 hours/year. Surprisingly, the pandemic shelter-in-place did not cause any observable changes in the information derived from thermostat data. Major challenges remain in translating reported thermostat data into energy savings, such as lack of information on equipment capacities.

Introduction

The thermostat is the most important device working with occupants to control home energy use. Together, the ~100 million residential thermostats control nearly 10% of national energy use (Peffer et al. 2011). When correctly programmed, the thermostat can reduce a home's heating and cooling energy over 20% relative to a constant-temperature setpoint strategy and 14% for the average home. (Moon and Han 2011; Sanchez et al. 2008). For these reasons, the Environmental Protection Agency (EPA) began certifying and labeling thermostats through its ENERGY STAR™ labeling program. EPA initially focused on programmable thermostats, establishing specifications that required certain features and performance. It also required that the thermostats be shipped with pre-programmed temperature setpoint schedules that were intended to reduce heating and cooling energy.

ENERGY STAR labeling almost certainly accelerated the widespread adoption of higher-quality, programmable thermostats. Over time, however, it became clear that the energy-saving potentials of programmable thermostats were not being achieved (Nevius and Pigg 2000), partly because consumers found them difficult to operate and disabled many of their energy-saving features. In 2009, EPA terminated ENERGY STAR labeling for thermostats until it could

develop a means of ensuring higher utilization of the energy-saving features. Around 2012, several thermostat manufacturers began offering devices with Internet connectivity and other means of communicating with devices beyond the furnace and AC, such as smart phones and home energy management systems. These features opened up many new energy-saving opportunities that skirted some of the earlier problems with unconnected thermostats. In addition, internet connection can provide a window into how users interact with the thermostat to control their home's temperature, thus allowing direct insight into whether features intended to facilitate energy savings are being used.

Internet-connected thermostats (CTs)¹ now control heating and cooling systems in about 30% of U.S. homes and account for about half of current annual sales. All connected thermostats allow users to remotely view and set their home's temperature. Some models monitor occupancy, humidity, multiple inside temperatures, and additional HVAC modes. Popular models also learn patterns of occupancy in a home and automatically establish energy-saving schedules. Sophisticated control algorithms take into account how outdoor temperature, humidity, and insolation are likely to affect a particular home. Researchers have already explored data collected by thermostat vendors to derive insights into such diverse aspects as occupancy prediction (Huchuk, Sanner, and O'Brien 2019), load shapes (Vellei, Martinez, and Le Dréau 2021), creation of digital twins (Hosseinihaghighi et al. 2022), thermal comfort preferences (Wang and Hong 2020), and geo-spatial aspects of power outages (Meier, Ueno, and Pritoni 2019).

Verifying energy savings from these thermostats – or even comparing performance of different models – is difficult because vendors cannot usually gain access to information necessary to perform a conventional evaluation. The vendors rarely have access to the characteristics of the home's heating system, its floor area, or the number of occupants. Most important, they cannot access the home's energy consumption data prior to or after installation of a thermostat. Other parties cannot access the information from the thermostat itself about its use. For these reasons, EPA developed a self-referential methodology to evaluate energy savings. The methodology compares the observed runtime to a counterfactual runtime based on (1) residents keeping the home at their apparent preferred temperature when home and awake (comfort temperature) and (2) the response of indoor temperature to heating and cooling equipment runtime in that home. Both the comfort temperature and the response of the home is based on hourly indoor temperature measurements and runtime collected by the thermostat. This procedure is described in Daken et al. (2016).

The vendor first draws a random sample of 250 thermostats in each of the five climate zones (1250 total, minus units that lacked sufficient data for analysis)². The vendors then inputs hourly values for HVAC runtime and inside and temperatures into the EPA-provided software, along with the thermostat zip code. The EPA software obtains historical outdoor temperature

¹ Recently, EPA has proposed switching to the more common (though less informative) name "Smart thermostats". This paper uses simply "thermostats" because no other types are discussed.

² 250 thermostats is intended to be a large enough sample to average over different home types, heating systems, and resident tendencies and circumstances so as to derive capabilities of the product in general. There has been robust stakeholder discussion of whether this sample is large enough, and EPA has proposed in Version 2 to allow larger samples.

from a weather file from a central repository for the location. (In other words, thermostats from all vendors located in the same zip code will use the same weather file.) Using the vendor-provided thermostat information and the hourly outdoor temperature during the year being analyzed, the software generates a set of thermostat metrics—for winter and summer. These metrics are:

- comfort temperature (°F);
- number of core heating days (with >30 minutes heating runtime, and no cooling) (days)
- number of core cooling days (with >30 minutes cooling runtime, and no heating) (days);
- HVAC runtime (hours/season); and
- energy savings metric (% runtime reduction).

Thermostats in mild climates are frequently excluded because relatively few hours of space heating or cooling occur. The Marine climate zone had especially high rejection rates.

The EPA model then aggregates the individual thermostat results into a single file containing summary statistics (including number of thermostats, mean, standard error of the mean, percentiles, and many more) for each metric for each climate zone³ Furthermore, the software calculates a national value for the heating and cooling savings based on the values in the individual climate zones, including the summary statistics.

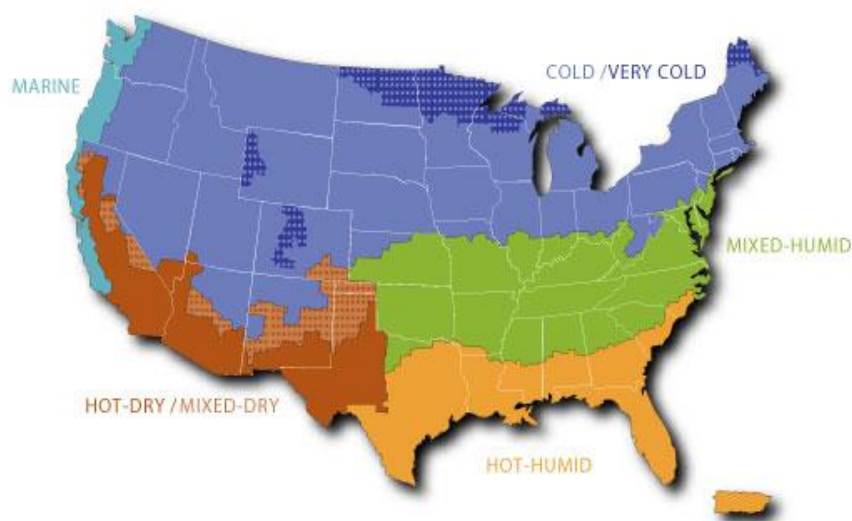


Figure 1: Building America Climate Zones. Source: [\(Baechler et al. 2015\)](#). The dotted areas represent additional climate zones that were merged with zones of the same color.

³ When vendors do not have sufficient thermostats in a climate zone they are permitted to use all available thermostats (that is, less than 250).

Vendors must submit results every six months, derived from a fresh random sample of their installations, in order to maintain their ENERGY STAR certification. Since the certification program began, thirteen vendors submitted data up to eleven times over four years. This procedure has created a unique record of each thermostat's long-term performance in response to changes in weather, customer demographics, building stock, algorithms, and, recently, the pandemic. Together, these submissions capture the heating and cooling habits in American homes in a way that was never before possible.

The ENERGY STAR certification is based on a weighted national average of the performance in the five climate zones. However, vendors must submit data for each of the five climate zones. The following discussion focuses on the results in the individual climate zones because zone-level results reveal more about the thermostats, the homes, and the occupants than national values. Space constraints do not permit displaying and discussing results for all five climate zones.

How Similar are the Homes?

Is one brand of thermostat preferred by a certain category of homes or occupants? It would be useful to understand how physically similar the homes are in different vendor submissions. Otherwise, the thermostat data will be revealing differences in the homes or their occupants rather than in the thermostats. Unfortunately, the vendors have little information about the homes using their thermostats beyond what they collect via the thermostats, and what they have is generally self-reported by residents, with the concomitant inaccuracy. It is therefore impossible to perform common comparisons with respect to floor area, fuel type, building age, etc. (Most vendors do not even know if a house has one, two, or more thermostats.) In addition, setback savings would presumably be largest for homes which are often unoccupied, and systems that are oversized may be less able to take advantage of setback savings. These could well correlate differently with different vendors, but we have no direct access to such information and were not able to study public information (e.g. income by zip code) that might be related.

Within a climate zone, it is possible to see how much the average outdoor temperature varies by thermostat vendor. Variations could occur because thermostat ownership is not geographically homogeneous. Some insights into the inter-vendor diversity can be gained by comparing the average outdoor temperatures. The average observed outdoor temperatures are shown in Figure 2 for heating and cooling in three climate zones. Each submission is labelled with its year and A (early in the year, capturing the previous winter) or B (late in the year, capturing that year's summer). Note that the two right-hand charts show winter and summer temperatures in the same climate zone (mixed-humid).⁴

⁴ These climate zones and seasons were selected because they capture the majority of national residential heating and cooling energy as estimated by the EIA.

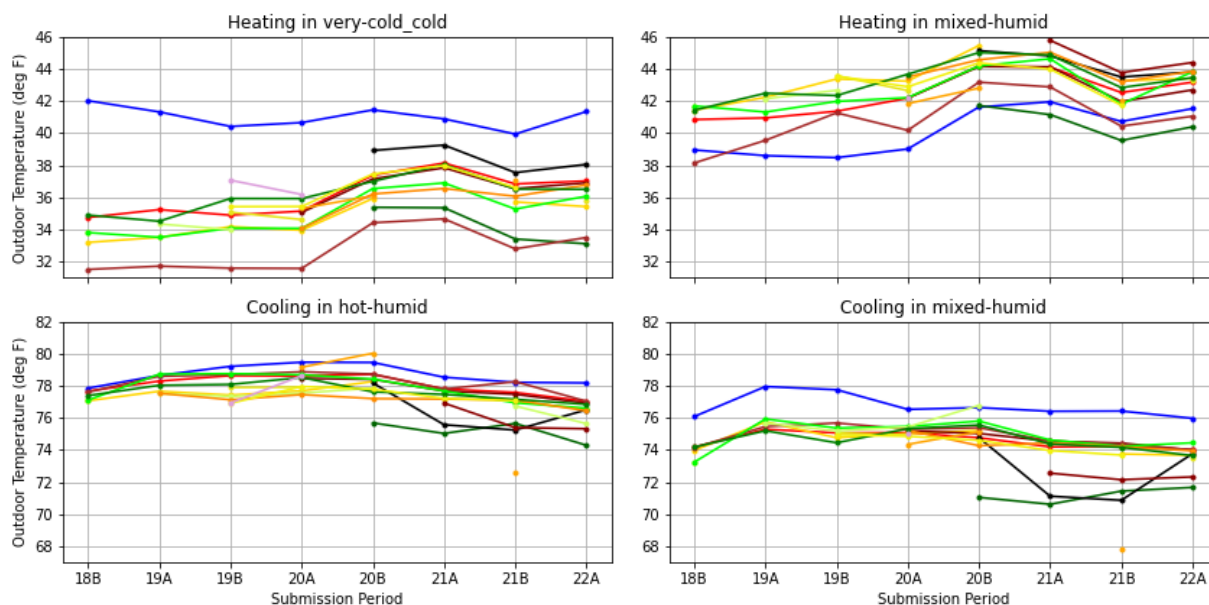


Figure 2. Outdoor temperatures faced by the different thermostats in three climate zones. Each line represents the average outdoor temperature faced by 250 randomly-sampled thermostats of a given thermostat vendor. The average standard error of the mean is typically less than 0.2°F, so the differences between vendors have some significance, but not much, except for the vendor represented by the blue line. This vendor, “Grape”, is atypical in many ways, including the savings strategy they employ.

Several trends are apparent from these data. First, the vendors' submissions respond to seasonal variations in the same direction, that is, a warm winter is experienced by all vendors in the climate zone. These charts are a kind of verification step but they also indicate the extent of non-weather effects. One example of a non-weather effect is when a vendor rapidly increases sales in one part of a climate zone causing its installation-weighted conditions to change.

Second, some vendors' thermostats face consistently different outdoor temperatures than others. The “Grape” (or blue) thermostats face different conditions than the others. Grape's thermostats face an average winter temperature 4 - 6 °F higher than the others in the very cold climate range but about 2°F in a mixed-humid climate (Fig 2b). For cooling, the locations with Grape thermostats faced consistently warmer temperatures than other vendors.

We conclude that most thermostats face relatively similar climates and are therefore similarly distributed across the climate zones. However, homes with Grape thermostats are distributed differently – possibly clustered? – in the climate zones and are likely to have different performance metrics than other vendors.

HVAC Runtime

The vendors report equipment runtime for heating and cooling separately, based on what the thermostat is calling for. The submissions are shown for three climate zones in Figure 3 (same zones as for Figure 2).

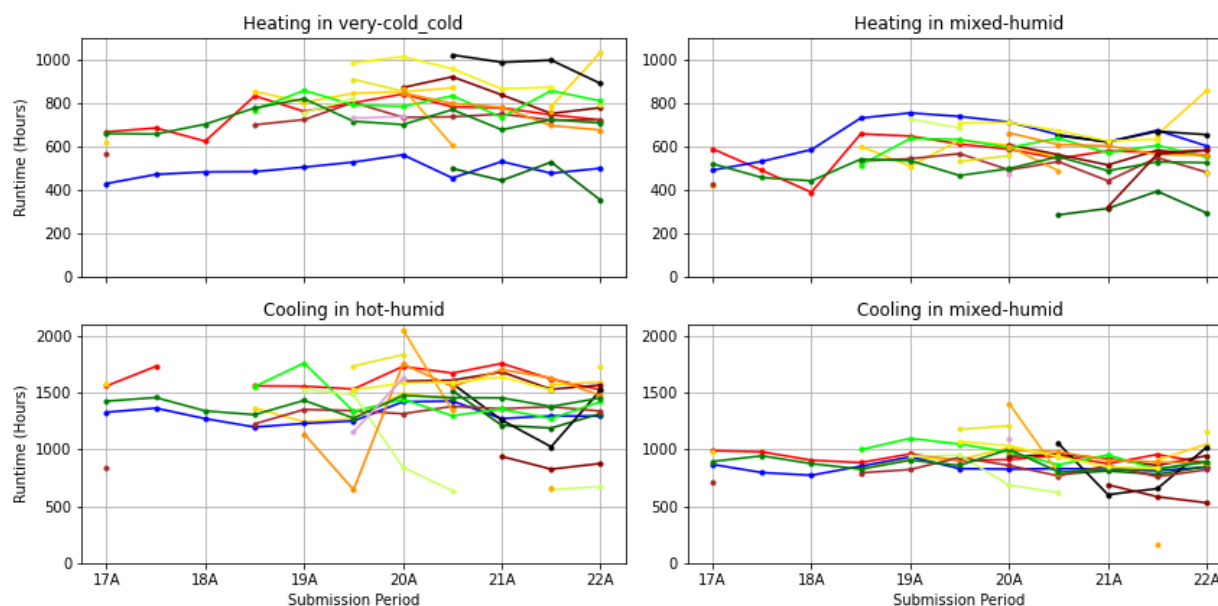


Figure 3. HVAC runtimes in three climate zones.

The HVAC runtimes submitted by the vendors are fairly stable from year to year. In the very cold-cold climate, heating systems typically run 700 hours/year and, in the hot-humid climate AC runtime averages about 1400 hours/year. The standard errors in the reported heating values range from 18 - 27 hours while the standard error in the reported cooling values range from 29 - 56 hours. While the runtimes for most of the vendors are similar,⁵ even small reductions in runtime are important because they could reflect superiority of one vendor's savings strategies over another. A 50-hour reduction corresponds to a 3 - 7% reduction in energy consumption.

Grape thermostats reported much lower runtimes in very-cold climates. This is not surprising given that Grape thermostats are located in significantly warmer parts of the climate zone. The opposite effect occurs with Grape in the mixed-humid zone, where runtimes are longer in response to lower outside temperatures.

The complexity of these off-setting effects – even for a simple runtime metric – explains the need for the model developed by EPA. It also shows the model's limitations. The model can estimate energy savings from a connected thermostat compared to a standard condition, serving to assure potential purchasers that they are likely to save energy with it, but it lacks the precision to identify differences in savings among vendors.

Runtime can be a proxy for heating and cooling energy consumption for single speed units, and can give reliable savings in kWh when the equipment capacities are known. Unfortunately, the vendors do not generally have reliable information about equipment capacities or fuel types, so translating runtimes into annual energy consumptions directly is not possible.

⁵ Some of the extreme variations are caused by small samples.

The only available way to estimate energy savings is using total HVAC energy use released by the Energy Information Administration (EIA).

Comfort Temperatures

The calculation of the ENERGY STAR savings metric begins by estimating each home's preferred comfort temperatures for heating and cooling. The comfort temperature for heating is defined as the 90th percentile of indoor temperature during core heating days and the comfort temperature for cooling as the 10th percentile of indoor temperature during core cooling days. The choice of 10% was arbitrary but provides a transparent anchor for analysis and comparison (Urban and Roth 2014). Comfort temperatures are important for the metric because for the same setback temperatures, a vendor with a more energy-conserving comfort temperature will show less difference between temperatures and therefore less savings.

Comfort temperatures are shown for three climate zones in Figure 4. Winter comfort temperatures mostly range from 70 - 72°F, with colder climate zones at the lower end and warmer zones closer to 72°F. Comfort temperatures during the summer range from 71 - 73°F, with warmer climates having lower temperatures. The temperatures do not significantly change from one year to the next, but for some vendors there may be a slow trend toward comfort temperatures that use more energy. Once again, Grape's temperatures are markedly different in some zones than other vendors, though they become more similar over time. This may be a result of successive adjustments to the setpoint algorithm.

Some of the absolute temperature differences between vendors may be caused by measurement inaccuracies. ENERGY STAR requires the thermostat to have a temperature error below 2°F (EPA 2016), measured in the lab in stable conditions. Given this, our analysis gave no evidence of differences between most vendors.

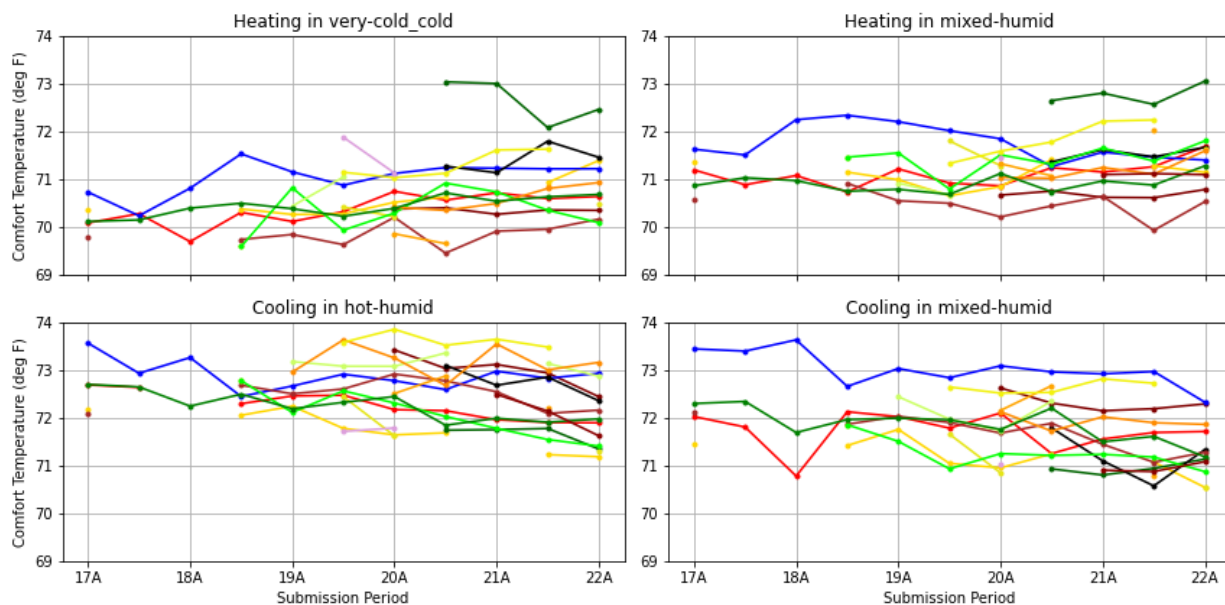


Figure 4. Comfort temperatures derived from thermostat data in three climate zones.

The comfort temperature metric was created to translate occupant thermal preferences into a single value that could be derived from available data. The metric appears to be generally successful, that is, consistent among vendors and over time.

Savings Metric

The savings metric is EPA's method of determining compliance for ENERGY STAR certification. The savings metric is the percentage savings from a reference case based on the observed comfort temperature and HVAC runtime. The metric is expressed in terms of percent savings in HVAC runtime caused by the difference between the comfort temperature and the actual indoor temperatures in the home, based on the modeled response of the particular home. This method can overestimate thermostat savings because occupants may have setback in any case, regardless of thermostat choice. But it can also underestimate savings by taking only setback-driven savings into account, not savings due to better HVAC control or more energy-conserving comfort temperatures. Thus, the relationship between the savings metric and actual thermostat savings in the field is unclear. Figure 5 shows trends of metric for all thermostats for all six years.

Over the program's six years, the savings in individual zones have ranged from 3 - 20% for heating and 10 - 20% for cooling. The savings for the vendors have remained generally stable over time and their positions relative to the others have not changed. The savings are gradually declining in many climate zones, possibly due to ownership expanding beyond early adopters. However, EPA does not have sufficient information to draw conclusions as to why the observed changes occurred.

Here, too, Grape's results are exceptional because savings changed dramatically in the early years. Some entries are missing because the vendors lacked sufficient data in that climate zone. The Marine climate zone is especially problematic because it has so few cooling and heating days for evaluation.

To qualify for ENERGY STAR certification, a thermostat must achieve savings listed in Table 1. Each vendor's national savings – the weighted, rolled-up values from all zones – will show less variability than observed in individual climate zones. It's also possible for a vendor's savings to fall below the requirement in one climate zone as long as the weighted average exceeds performance requirement.

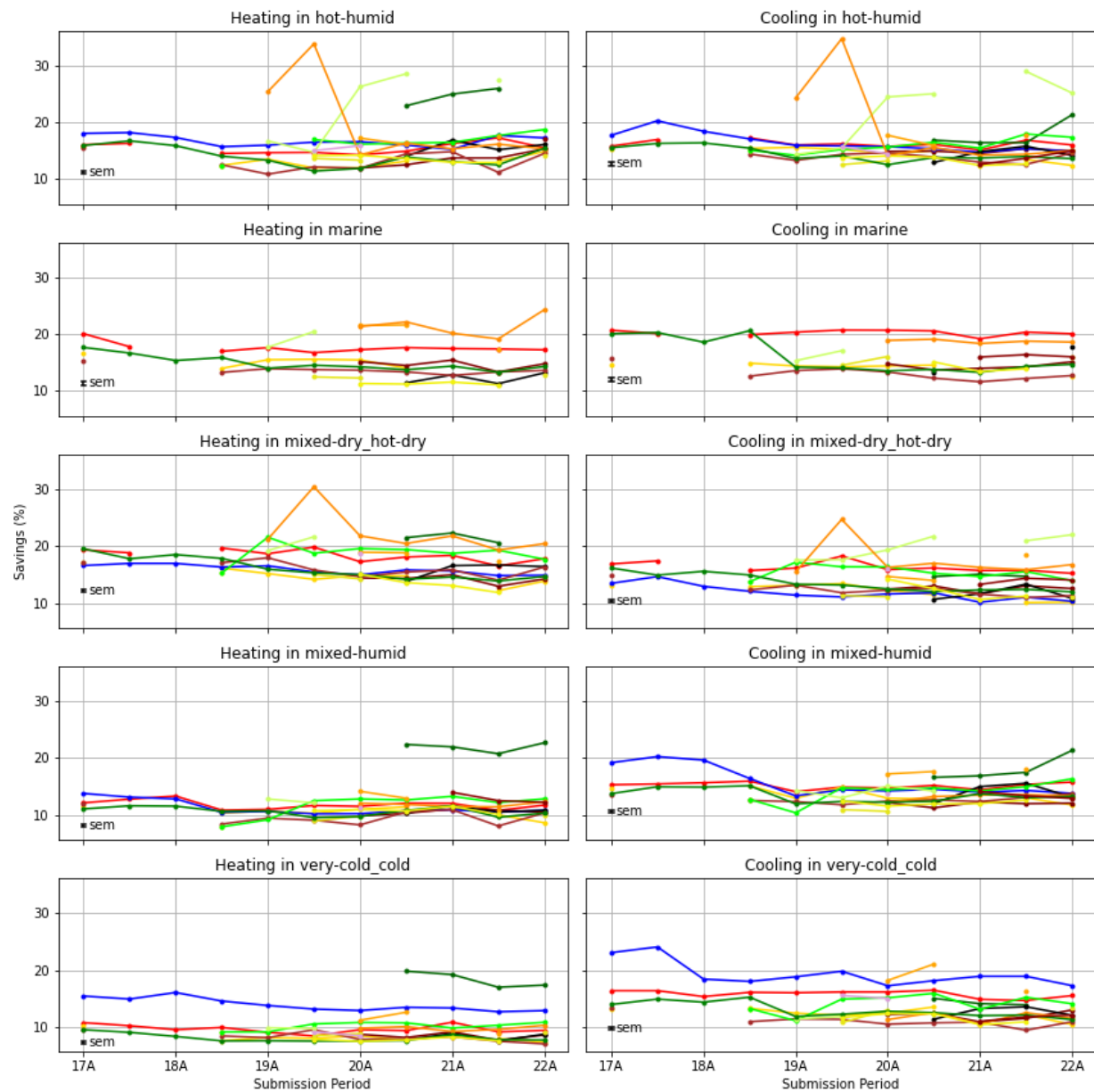


Figure 5. Trends in savings metrics over time by vendor. All points shown represent more than 30 thermostats, and majority are over 100 thermostats.

Table 1. ENERGY STAR Thermostat Energy Savings Criteria (adapted from EPA (2016))

Metric	Statistical measure	Savings performance requirement
Heating: Annual % runtime reduction	Lower 95% confidence limit of weighted national average	$\geq 8\%$
	Weighted national average of 20 th percentiles	$\geq 4\%$
Cooling: Annual % runtime reduction	Lower 95% confidence limit of weighted national average	$\geq 10\%$
	Weighted national average of 20 th percentiles	$\geq 5\%$

The Pandemic's Impact

The pandemic caused a tremendous upheaval in residential occupancy patterns as many office workers and students began working and studying at home. Nevertheless, increases in energy use were less than might be expected and energy use may even have declined (Energy Information Administration 2021). Other trends, such as milder weather, appear to have masked some of the initial impacts (Burleyson et al. 2021). We investigated changes in the thermostat metrics as a result of the pandemic.

The pandemic shelter-in-place began in March 2020 so its impacts would appear in submission periods 2020B and 2021A. An examination of HVAC runtimes revealed no dramatic trends; the runtimes were generally flat, though some increases and decreases occurred. There may be a small effect of less energy-conserving comfort temperatures, though it is difficult to tell amid a general years-long trend. Some increases in runtime may be masked by changes in average temperatures during these periods (as shown in Figure 2). Comfort temperatures were similarly unaffected by the increased occupancy during the shelter-in-place. Finally, the savings metric may have very slightly increased for space heating but was mostly unchanged for cooling.

We conclude that any changes in occupancy and behavior caused by the national shelter-in-place were not strongly reflected in the thermostat metrics. This finding is surprising because we expected, at a minimum, to observe longer HVAC runtimes but our results are consistent with evaluations of energy consumption.

Conclusions

We reviewed six years of data submitted to EPA by thirteen vendors of internet-connected thermostats. These submissions represent the behavior of about thirty million homes. Vendors submitted these data as participants in ENERGY STAR partnership, in order to gain

confidence in the ENERGY STAR savings metric. However, the information also gives us insights into the way homes are heated and cooled.

We first sought indications that the vendors' samples of homes face relatively similar climates, which they did. It would also be valuable to know whether one vendor's customers tend to have very large homes or specific types of heating systems, but vendors generally do not have details about the homes or their occupants.

The operating hours of heating and cooling equipment are good proxies for heating and cooling energy use, so we examined trends in these submissions. The HVAC runtimes submitted by the vendors are surprisingly stable from year to year. We also found that heating systems in the very cold climate zone typically run 700 hours/year and, in the hot-humid climate AC runtimes average about 1400 hours/year. This kind of information could not be obtained elsewhere and can assist in interpreting energy use and forecasting. Sadly, equipment runtime can't be converted directly into actual energy use without complementary data on equipment capacities.

The comfort temperature metric was created to translate occupant thermal preferences into a single value. The metric was generally successful, that is, consistent among vendors and over time.

The ENERGY STAR certification relies on vendor-submitted data which has been processed in a model developed by EPA. The metric criteria are not applicable to savings estimates from individual homes and have limited utility in comparing different thermostats; however, they are still a useful means of ensuring sufficient savings to justify consumer confidence. The savings in individual climate zones ranged from 3 - 20% for heating and 10 - 20% for cooling.

Finally, we examined the impacts of the pandemic and shelter-in-place on the vendor-submitted data. To our surprise, we observed no significant changes in the metrics. Other, off-setting trends may have masked the full impacts.

The value of the thermostat data could be greatly increased by research and future work in three areas. First, utility and HVAC measurements should be collected for homes with thermostat data. This information would better correlate temperature settings and energy consumption. Second, public information correlating with factors that could be relevant to thermostat savings, such as income and age of residents, and properties of housing stock by zip code, could be studied and correlated with thermostat metrics. Finally, collecting more technical and demographic data directly about the thermostat users would help us explain unusual usage patterns.

Acknowledgement

This work by Meier and Rainer was supported by the U.S. Environmental Protection Agency and the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technology, U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Work by Goldman and Jathar was supported by the U.S. Environmental Protection Agency under Contract. No. GS-00F-010CA.

References

- Burleyson, Casey D., Aowabin Rahman, Jennie S. Rice, Amanda D. Smith, and Nathalie Voisin. 2021. "Multiscale Effects Masked the Impact of the COVID-19 Pandemic on Electricity Demand in the United States." *Applied Energy* 304 (December): 117711. <https://doi.org/10.1016/j.apenergy.2021.117711>.
- Daken, Abigail, Alan Meier, and Douglas Frazee. 2016. "Do Internet-Connected Thermostats Save Energy?" In *ACEEE 2016 Summer Study on Energy Efficiency in Buildings*. Pacific Grove, Calif.: American Council for An Energy Efficient Economy (Washington, D.C.).
- Energy Information Administration, Mickey. 2021. "Despite More People Staying at Home, U.S. Residential Energy Use Fell 4% in 2020." *Today in Energy*. May 17, 2021. <https://www.eia.gov/todayinenergy/detail.php?id=47976>.
- EPA. 2016. "ENERGY STAR® Program Requirements For Connected Thermostat Products." Environmental Protection Agency. <https://www.energystar.gov/products/spec>.
- Hosseinihaghighi, Seyedehrabeeh, Karthik Panchabikesan, Sanam Dabirian, Jessica Webster, Mohamed Ouf, and Ursula Eicker. 2022. "Discovering, Processing and Consolidating Housing Stock and Smart Thermostat Data in Support of Energy End-Use Mapping and Housing Retrofit Program Planning." *Sustainable Cities and Society* 78 (March): 103640. <https://doi.org/10.1016/j.scs.2021.103640>.
- Huchuk, Brent, Scott Sanner, and William O'Brien. 2019. "Comparison of Machine Learning Models for Occupancy Prediction in Residential Buildings Using Connected Thermostat Data." *Building and Environment* 160 (August): 106177. <https://doi.org/10.1016/j.buildenv.2019.106177>.
- Meier, Alan, Tsuyoshi Ueno, and Marco Pritoni. 2019. "Using Data from Connected Thermostats to Track Large Power Outages in the United States." *Applied Energy* 256 (December): 113940. <https://doi.org/10.1016/j.apenergy.2019.113940>.
- Peffer, Therese, Marco Pritoni, Alan Meier, Cecilia Aragon, and Daniel Perry. 2011. "How People Use Thermostats in Homes: A Review." *Building and Environment* 46 (12): 2529–41. <https://doi.org/10.1016/j.buildenv.2011.06.002>.
- Urban, Bryan, and Kurth W. Roth. 2014. "A Data-Driven Framework For Comparing Residential Thermostat Energy Performance." Cambridge, MA: Fraunhofer Center for Sustainable Energy Systems.
- Vellei, Marika, Simon Martinez, and Jérôme Le Dréau. 2021. "Agent-Based Stochastic Model of Thermostat Adjustments: A Demand Response Application." *Energy and Buildings* 238 (February). <https://doi.org/10.1016/j.enbuild.2021.110846>.

Wang, Zhe, and Tianzhen Hong. 2020. "Learning Occupants' Indoor Comfort Temperature through a Bayesian Inference Approach for Office Buildings in United States." *Renewable and Sustainable Energy Reviews* 119 (March): 109593.
<https://doi.org/10.1016/j.rser.2019.109593>.