August 9, 2021

Ms. Abigail Daken
U.S. Environmental Protection Agency
Climate Protection Partnership Division
Office of Air and Radiation
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Topic: ENERGY STAR® Connected Thermostats Version 2.0 Discussion Document

Dear Ms. Daken:

This letter comprises the comments of the Pacific Gas and Electric Company (PG&E), San Diego Gas and Electric (SDG&E), and Southern California Edison (SCE) in response to the United States (U.S.) Environmental Protection Agency (EPA) Discussion Document on Connected Thermostats.

The signatories of this letter, collectively referred to herein as the California Investor-Owned Utilities (CA IOUs), represent some of the largest utility companies in the Western U.S., serving over 32 million customers. As energy companies, we understand the potential of appliance efficiency standards to cut costs and reduce consumption while maintaining or increasing consumer utility of products. We have a responsibility to our customers to advocate for standards that accurately reflect the climate and conditions of our respective service areas, so as to maximize these positive effects.

We appreciate this opportunity to provide comments on the ENERGY STAR® Connected Thermostat Program. Residential space heating and air conditioning account for approximately 43 percent and eight percent of the total energy consumption of U.S. homes, respectively.[1] We view the connected thermostat program as an important avenue for reducing this energy consumption across a wide variety of equipment types on the national level. As such, we provide broad comments on improving the methods to calculate field savings, number, and location of thermostats to sample, extending the scope of the program to new equipment, demand response, and flexible load topic considerations. We strongly urge EPA to consider the following comments.

1. The CA IOUs are supportive of changes to the EPA method to calculate field savings and recommend the following updates.

Below, we discuss several key modifications to the EPA thermostat method to calculate field savings, which we estimate will improve the confidence interval of the savings results from sample to sample and improve comfort temperature estimations. This will lead to more accurate user behavior estimates and improvements in the ability of the metric to account for key savings strategies, such as more efficient set points in a user's home.

[1] EIA RECS 2015, Table CE3.1.

Sample More Thermostats

The field savings metric results are anecdotally estimated to have a variance of one percent in the savings metric derived from sampling variation, as described in recent metrics meetings.[2] We agree that underlying parameters of the model and demographics such as comfort temperature, equipment type, home type, home location, International Energy Conservation Code (IECC) climate zone of the location, and local building practices (i.e., air sealing and insulation), quickly result in the population becoming statistically challenging to draw representative results.

This challenge is amplified once the pre-filtering routines of the software are introduced, where the software requires a linear fit intercept (i.e., Tau parameter)[3] to be positive and within 0-25, and a coefficient of variation root mean square error (CV[RMSE]) above a minimum 0.6 threshold. In a recent metrics meeting presentation, it was shown that five vendors have 50 percent or more heating thermostats discarded, and four vendors have 50 percent or more cooling thermostats filtered. The best performing vendor still discards more than 20 percent in heating and approximately 16 percent in cooling.[4] When these challenges are combined, we see a model that would benefit significantly from a higher thermostat sampling limit.

In evaluating the Energy Information Administration (EIA) Residential Energy Consumption Survey (RECS) 2015 microdata,[5] we note that approximately 5,600 U.S. residences are polled to construct a national representative sample. We acknowledge that the RECS sampling methodology includes complex weighting and representativeness considerations to each selected home; however, each sampled home at minimum represents at least 14,500 U.S. homes. When analyzed by *Building America* climate zone, all zones except Marine have 750 or more samples, with Marine having 420, and the Cold- and Very-Cold climate over 2,000 samples. These sample volumes were achieved after screening criteria to eliminate homes that would not model properly. With a best-case thermostat rejection rate of approximately 20 percent, **we recommend the EPA adjust the maximum sampled thermostats per climate zone to *at least* 1,000 homes per zone with a maximum of 2,500 sample homes per zone** if the data is available.

Regarding this maximum thermostat limit: we note that the initial reasons cited for the development of the maximum thermostat parameter, and the requirement on the lower bound confidence interval of the mean, was to balance the challenge of enough samples to reach a statistical significance of the analysis result and the burden on newer manufacturers in obtaining enough home samples to certify. We recommend that these considerations apply to the initial certification dataset but would not be relevant to the resubmission datasets to continue in the program, since the resubmission dataset allows the submitting of all available thermostats in the zone up to the maximum count. A vendor that had succeeded in certifying initially would not be anticipated to be penalized (i.e., inadvertently) when using more data, as the use of more data was estimated to be an advantage due to smaller calculated confidence intervals.[6] **We recommend this use of an initial certification maximum thermostat count and a separate higher resubmission maximum thermostat count** would improve the ability of program statistics to be robust to sampling error in the population.

---

[2] Metrics Meeting, June, 2021.

[3] Specifically, Tau is the temperature demand (deltaT) associated with no runtime. This is obtained by a linear fit of the non-zero runtimes versus outdoor temperature demand (Toutdoor – Tindoor). This fit also returns the alpha parameter, which is the slope of this fit.

[4] May 2021 Metrics Meeting, p. 11

[5] EIA RECS 2015 Microdata, CA IOU analysis.

[6] The confidence interval size is proportional to 1/sqrt(N) ; where N is the thermostat count.

New Climate Zones

The relative weightings of the climate zones, large variation of climate within a zone, and limited resolution at 250 maximum thermostats per zone have been frequent topics of conversation at metrics discussions. Of note, a stakeholder commented that more than 40 million homes are located in the Cold- and Very-Cold climate zone, such that a random sample of 250 homes is unlikely to be fully representative of the median and more extreme portions of the zone at the same time.[7] We support this reasoning and agree that a more granular climate zone mapping would better allow the metric scores to ensure that consumers in different geographic regions across the U.S. save energy. We suggest that the IECC abbreviated[8] climate zones in EIA RECS provide this practical balance of complexity, climate differences between zones, and interdisciplinary frequency of use in building science analysis. Further, *Building America* climate zones have a direct definition in the IECC methodology, thus they would require minimal analysis to adopt. We show the Building America climate zones overlaid on the IECC continental U.S. map in Figure 1.

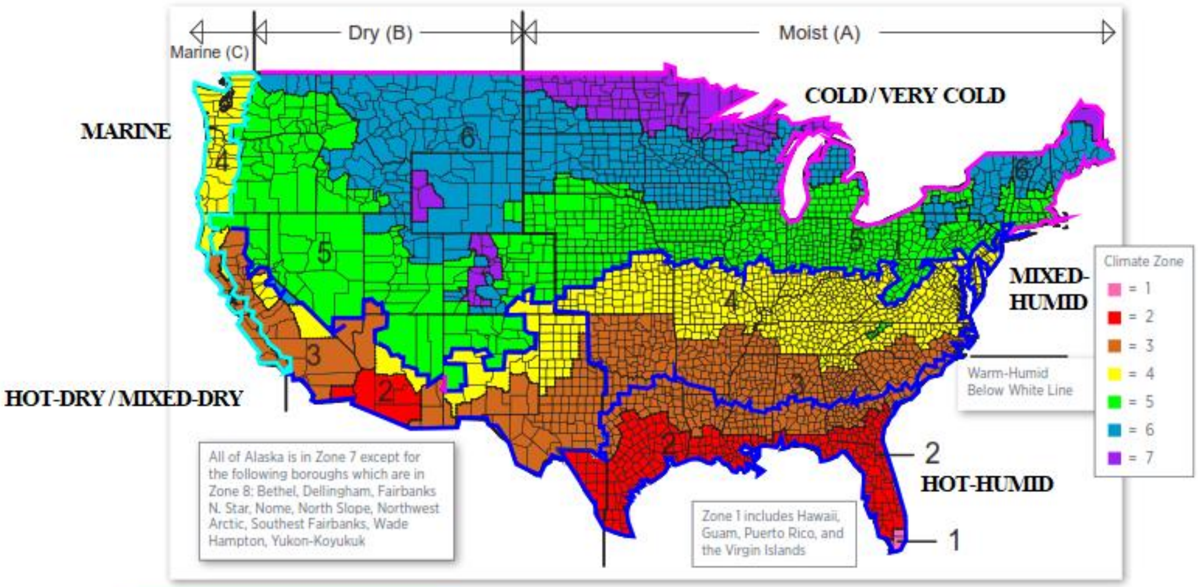


Figure 1: Building America Climate Zones Overlaid on IECC Climate Zones.
Source: US DOE, Guide to Determining Climate Regions by County, 2015; CA IOU annotations of IECC climate zone map on p. 2.

As shown in Figure 1, there is alignment at the county level between the climate zone methodologies. However, IECC with more granularity partitions the Cold- and Very Cold region, Mixed Humid region, and Hot Humid regions into two or more segments, reflective of the local climate differences.

We note that the EIA RECS microdata includes both a *Building America* climate zone flag and an IECC climate zone flag, enabling a direct analysis between the home populations of the two climate models. We calculate national housing counts in each methodology, shown in Table 1.

[7] Discussions in Metrics Discussion, June 2021 p 25. Personal communication with M. Blasnik.
[8] This IECC abbreviated model consists of the following zones, **merged zones bold**: **1A-2A**, 2B, 3A, 3C, **3B-4B**, 4A, 4C, 5A, 5B, **6A-B**, **7-8**.

Table 1: Relationship between Building America and IECC Climate Zones.

Building America Climate Zone	US Homes BA CZ (MM Homes) *	IECC Climate Zone(s)	US Homes IECC CZ (MM Homes) *
Subarctic	42.5	Zone 8	**7-8**: 1.0
Very-Cold		Zone 7	
Cold		Zones 5 [5A,5B] and 6 [6A, 6B]	**5A**: 28.1 **5B**: 6.1 **6A-B**: 7.3
Mixed-Dry	12.7	Zone 4B	**3B-4B**: 11.0
Hot-Dry		Zone 2B, 3B	**2B:** 1.7
Hot-Humid	22.8	2A and 3A counties below warm-humid line	**1A-2A**: 18.5 **3A**: 4.4
Mixed-Humid	33.5	4A and 3A counties above warm-humid line	**3A**: 10.2 **4A**: 23.3
Marine	6.7	All counties with a "C" moisture regime	**3C**: 3.5 **4C**: 3.2

Note (*): EIA RECS 2015 microdata merges some smaller IECC zones, similar to the merging of smaller BA climate zones.
Source: US DOE, Guide to Determining Climate Regions by County, 2015.

As anticipated, moving to the IECC methodology improves multiple shortcomings of the *Building America* methodology. With the housing counts by the IECC zone in Table 1, we note that: (1) the Cold- and Very Cold climate are now split into three distinct zones with minimum subpopulations of at least six million homes each, (2) the similarly large Mixed Humid zone is now split into 2 climate zones of at least 10 million homes, and (3) the more temperate region of Hot Humid is now analyzed separately. We anticipate that a model produced from these climate zones will have more representative home performance results of each sub-region and have less dependency on sampling selection.

Climate Zone Weightings

We then consider the potential updates and adjustments that this climate zone methodology would bring to the national climate zone weights, as currently used in the model.

The current national weightings for heating and cooling are calculated by the EIA RECS 2009 space heating consumption tables by climate zone, which result in the weights shown in Table 2.

Table 2. Current EPA national weightings for heating and cooling by climate zone.

	Building America Climate Zones				
	Very Cold / Cold	Mixed Humid	Mixed Dry / Hot Dry	Hot Humid	Marine
Heating Weight	**54.9**	**31.2**	5.4	4.9	3.6
Cooling Weight	9.6	**34.0**	14.4	**42.0**	N/A (0 Weight)

Source: Metrics Meeting April 2019, p 15.
Note: Emphasis added on zones with more than 30 percent weighting of a parameter.

Nationally, both the heating and cooling metric scores are heavily weighted towards two climate zones, which has been noted as a topic of concern in previous EPA discussions.[9] We agree that this significant overweighting of two zones results in a certification criterion that is too dependent on regional performance of the thermostats in more extreme climates, thus does not fully represent U.S. end users and corresponding heating and cooling equipment selection in more mild climates. Since this program is intended to provide a minimum and typical energy savings result for all U.S. consumers, we recommend that a minimum weighting of 10 percent per zone under the *Building America* methodology is applied, and the remainder of the weighting is proportionally redistributed based on heating or cooling number of core days in the region.

For cooling, this recommendation would re-instate the Marine climate zone as a location that does use mechanical air conditioning. The 2017 Northwest Energy Efficiency Alliance (NEEA) Residential Building Stock Assessment (RBSA) II demonstrates that this region has been witnessing an organic uptake in air conditioning, with the RBSA II showing single-family homes at 65 percent equipped with mechanical cooling, up from 42 percent in the RBSA I in 2011.[10] Another data point, the 2019 Census Annual Housing Survey (AHS), notes an air conditioning prevalence increase in metro Seattle from 31 percent to 44 percent in six years.[11] Recent extreme temperatures[12] in U.S. Marine climates have led to an unprecedented recent acceleration in uptake[13] in air conditioning capability in this region, a trend expected to continue in the future. We encourage EPA to eliminate this Marine air conditioning exemption in the connected thermostats program in all future analysis and program metrics.

In the climate zone case where abbreviated IECC zones are provided instead, we recommend the minimum bin weighing can be set as one half of the uniformly distributed weighting (i.e., a 10-zone model would provide a minimum five percent weighting), where the remainder could be distributed proportionally based on regional consumption, equipment runtime, core days in a season, or some other suitable weighting metric. Of this combination, we recommend EPA evaluate a number of core season days as a candidate metric for the new weightings.

New Comfort Temperature Calculation Methodology

We note that the field savings metric uses the 90th and 10th percentile indoor temperatures as the estimate of the end user's comfort temperature for a given home. In practice, the indoor temperature can drift over the setpoint during the heating season without heating equipment runtime via uncontrolled loads (i.e., by the thermostat) such as cooking heat, clothes drying, and solar gain. We note that this temperature above the setpoint is typically produced inadvertently,[14] is not reflective the end user's desired home temperature, and therefore should not be built into the distribution to estimate the comfort temperature.

In a thermostat evaluation study for Commonwealth Edison (ComEd), Guidehouse implemented such a comfort temperature selection algorithm into the EPA thermostat software, with the impact of this adjustment shown in Figure 2.

[9] Metrics Discussion, June 2021, p 25.
[10] RBSA II, Single Family Report, p. 6.
[11] https://www.usnews.com/news/cities/articles/2021-07-09/air-conditioning-trends-in-americas-biggest-cities
[12] https://www.washingtonpost.com/weather/2021/06/25/pacific-northwest-heat-wave-seattle-portland/
[13] https://www.nytimes.com/2021/06/25/us/western-heat-wave.html
[14] https://www.zerocarbonhub.org/sites/default/files/resources/reports/ZCH-OverheatingLeaflet-5-TechnicalSolutions-S_0.pdf, p 4.

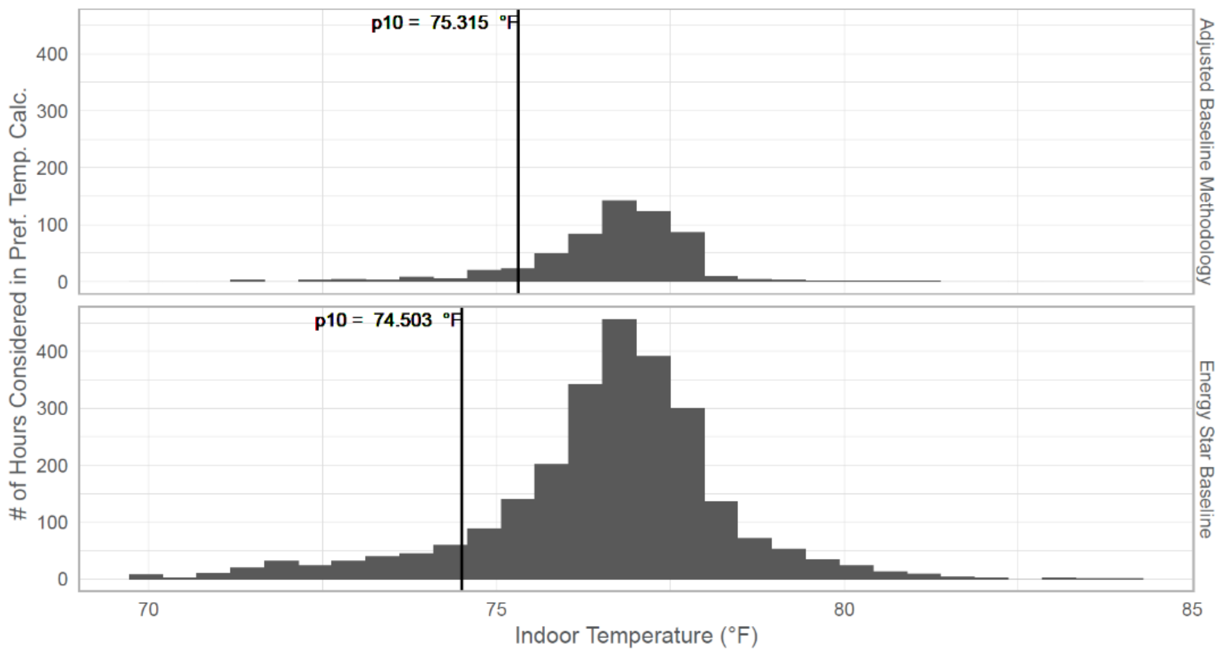


Figure 2: Adjustment in the home comfort temperature when excluding temperature float without runtime above the set point temperature.
Source: Guidehouse, ComEd Advanced Thermostat Evaluation, 2020, p. 62.

This result demonstrates a change in the comfort temperature baseline of over 0.8 °F for this home, which can have significant effects on the baseline performance. We note that this effect is not as strong when a home has a bi-modal or more complex distribution of indoor temperatures, which is often generated by users with large setback behavior. This same adjustment in a home with two common indoor temperatures, shifts the comfort temperature estimate under 0.1 °F from the original EPA estimated comfort temperature. See Appendix A-Figure A1 for a plot of this case. We estimate that this metric adjustment will provide more representative comfort temperature for a given home, especially for users that have a strong single setpoint temperature preference.

Regional Temperature Baselines

As part of the development of the field savings metric, we note that a known issue of the current metric is the thermostats that encourage[15] users to operate at more efficient setpoint temperatures either directly or via subtle algorithmic nudges, do not receive the full savings of this methodology reflected in the metric score. The primary reason for this issue is that the more efficient comfort temperatures encouraged by the thermostat are then incorporated into the comfort temperature baseline of this home; thus savings is calculated for the device at this new operating condition.

We recount the earlier development of the thermostat metric, which included the regional baselines and individual home comfort temperature baselines. At the time, the regional baselines were based on self-reported interior temperatures as reported in the EIA RECS 2009 dataset.[16] EPA analysis of this dataset demonstrated challenges in comparisons between the actual sampled home data and RECS values. Plots of this difference are available in Appendix A-Figures A2-3 for heating and cooling, respectively. This

[15] See for example: Google/Nest Seasonal Savings, ecobee eco+: humidity compensating automatic set point adjustment.
[16] https://www.energystar.gov/sites/default/files/DiscussionDoc-ES%20Candidate%20CT%20Field%20Savings%20methods%20implemented%20in%20software_20160915.pdf, p.17, footnote 14.

comparison demonstrated that for this 2017 season, heating comfort temperatures were predominately warmer in all climate zones and cooling comfort temperatures were predominately cooler; consistent with recommendations that self-reported data are more ambitious regarding energy savings behavior than found in practice.

Therefore, we recommend that regional comfort temperatures should instead be developed from actual intra-vendor year-over-year values as collected throughout the duration of the connected thermostat program. One recent metric meeting had presented this average comfort temp by vendor data for three consecutive resubmission datasets, shown in Figure 3.

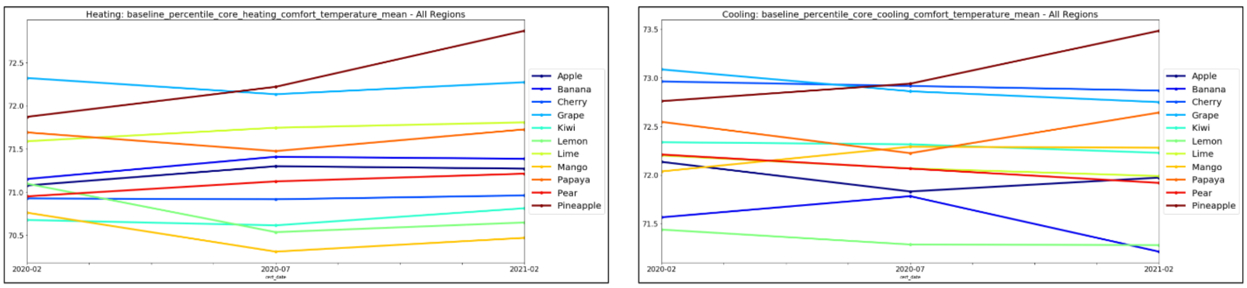


Figure 3: Heating and cooling comfort temperatures by vendor over multiple data resubmissions.
Source: Metrics Meeting, Feb 2021, p. 10.

Of note in Figure 3, there is a clear differentiation between vendors' heating and cooling comfort temperatures on average, with some at more and less efficient temperatures. The relative ranking of vendors is somewhat consistent over time, with most vendors remaining within the same position above or below the intra-vendor median and often in the same quartile of the population. The consistency of vendor rank ordering over time suggests that this effect is not likely to be a sampling artifact. Given these 0.5-1.5 °F differences between vendors, and the known efficient temperature setpoint guidance and algorithms provided by some vendors, we can estimate that these programs are having an effect on the vendor's population.

On this basis, we encourage EPA to develop intra-vendor regional comfort temperature baselines based on the past few years of thermostat data, to build out a savings metric that can account for homes operating at a temperature more efficiently than the typical intra-vendor regional comfort temperatures. In this version of the model, the calculated savings for a home with more efficient operating temperatures than typical of the region can have the baseline runtime increased (e.g., proportionally) to account for this effect.

We note that our recommendations to shift the program from existing *Building America* climate zones to IECC climate zones, and recommendations to collect more thermostats per zone, should further increase the representativeness of the regional indoor temperatures collected from vendors.

Product Family Evolution

We are supportive of the existing capability to bundle a new product, provided it uses existing software savings strategies, with the field savings data of existing certified products to create a joint thermostat dataset. This prevents the significant barrier of entry for new products from requiring numerous national installations when the savings strategies of this vendor have already been successfully demonstrated in a certified product.

However, once certified, these bundled products typically remain paired with the original predecessor product indefinitely. We encourage EPA to develop guidance that requires vendors to develop this new product into an independent, full thermostat count and mutually separate data resubmission file at most after a designated time period, such as 18 months (or, three resubmission periods).

2. The CA IOUs support the inclusion of Line Voltage Connected Thermostats (LVCT) when performance is calculated on a whole home basis.

Approximately eight percent of all U.S. dwellings use electric baseboard heating as the primary heating system, including 18 percent of multifamily dwellings,[17] all of which would have the electric resistance heating coefficient of performance of one. Therefore, we consider the inclusion of this product type in the ENERGY STAR specification a welcome potential addition to the product lineup with broad energy savings implications nationally.

Multiple LVCT per Home

In establishing the methodology for this product, we first note that in a typical home, numerous line voltage thermostats are present.

The Regional Technical Forum's (RTF) measures catalog includes a residential electronic line voltage thermostats measure. Measure documentation[18] includes a field study count of the number of line voltage thermostats per home: the study noted an average of five thermostats per home in a combined population of single-family and multifamily housing. RTF, based on this study, estimates five thermostats for a typical single-family home and three for a typical multifamily dwelling. In practice, this can be even higher, in a 2017 Peak Load Management Alliance (PMLA) presentation[19] on a Hydro Quebec pilot program, an average of 10 line voltage thermostats were reported per home. We consider this supporting the conventional wisdom that homes with baseboard heating have several baseboards amongst the rooms in the home and several independent thermostats that, if changed over to LVCTs, would have significantly more than one LVCT per home.

Sampling the full LVCT population at random[20] from a manufacturer's thermostat population would be expected to only capture a portion of each home, distributed across many homes in the region of sampling. To further develop the need for selecting all the thermostats in a given home, we next evaluate the room-by-room use case variability inherent in this product.

Variability of Comfort Temperature in Zoned Systems

In this same PLMA presentation,[21] project participant feedback indicated that 70 percent of participants would have been comfortable with even more aggressive setbacks and pre-heats in some rooms. Another participant noted that "some of the rooms are unoccupied so pre-heating is not useful [during the DR event time]", and another requested that such a program in the future needs the ability "to set a lower [minimum temp] limit, e.g. do not go below 19 °C in some rooms."

Interpreting this information in a calculation methodology is challenging, but participant feedback indicated that: (1) zoned systems have an occupied and unoccupied comfort temperature, (2) some rooms

---

[17] EIA RECS 2015, HC 6.1 ; apartments calculated as the sum of items 'Apartment (2-4 unit building)' and 'Apartment (5 or more unit building)'.
[18] https://nwcouncil.box.com/v/ResLineVoltageTstat-v4-1 , tab Summary, cell B32; accessed 07/29/2021.
[19] PLMA Proceedings, Demand Side Management with Line Voltage Communicating Thermostats: a Real Life Experiment, M. Fournier, 2017, p. 6
[20] As suggested in the May 2021 Metrics Meeting.
[21] i.d., [PLMA], p. 14.

are unoccupied enough such that the unoccupied comfort temperature could be the typical operating temperature of that room (e.g., some basements), (3) some rooms are occupied enough such that the occupied comfort temperature of the room predominately drives operating conditions, and (4) night set-back and whole home away setback savings strategies are still valid for LVCTs, as they are for conventional connected thermostats. Further, we could interpret from programmed set-back LVCT set-points which temperatures are approximately the unoccupied comfort temperatures of each location and the home overall.

We note that the performance and user temperature preference of zoned systems is a considerable area of current research, such as the field studies of integrated central and zoned heating and cooling system control. One such Pacific Northwest National Laboratory (PNNL) 2020 study suggests the following advanced control as specified in Table 3.

Table 3: Complex setback schedule in PNNL zoned HVAC experiment

	Central Heating	Zonal Heating	Central Cooling	Zonal Cooling
DHP Conditioning Main Living Area				
Occupied (7am – 9pm)	72°F	85°F	76°F	65°F
Unoccupied (9pm – 7am)	66°F	80°F	81°F	70°F
Central System/Zonal Electric or Window AC Conditioning the Bedrooms				
Occupied (9pm – 7am	66°F	80°F	76°F	65°F
Unoccupied (7am – 9pm	55°F	60°F	90°F	Off

Source: Proceedings of ACEEE Summer Study on Buildings, C. Metzger et. al., 2020, p 1-218, table 2. See T. Ashley et. al, 2020 for the full analysis.

We believe this table estimates a significant opportunity for zonal setbacks in unoccupied rooms. Reports on the use of setbacks in line voltage thermostats is well discussed in the literature; one such example dating back to 1996. In this example, 26 homes in an electronic thermostat pilot (replacing bimetallic) are shown to have a median setback temperature change of approximately 2 °F.[22] See Appendix A-Figure A4 for this data.

For an illustrative example of zoned setpoints in an occupied home, see Figure 4, which is a cooling season home with integrated control between a central and zoned HVAC system.

[22] Lambert, 1996.

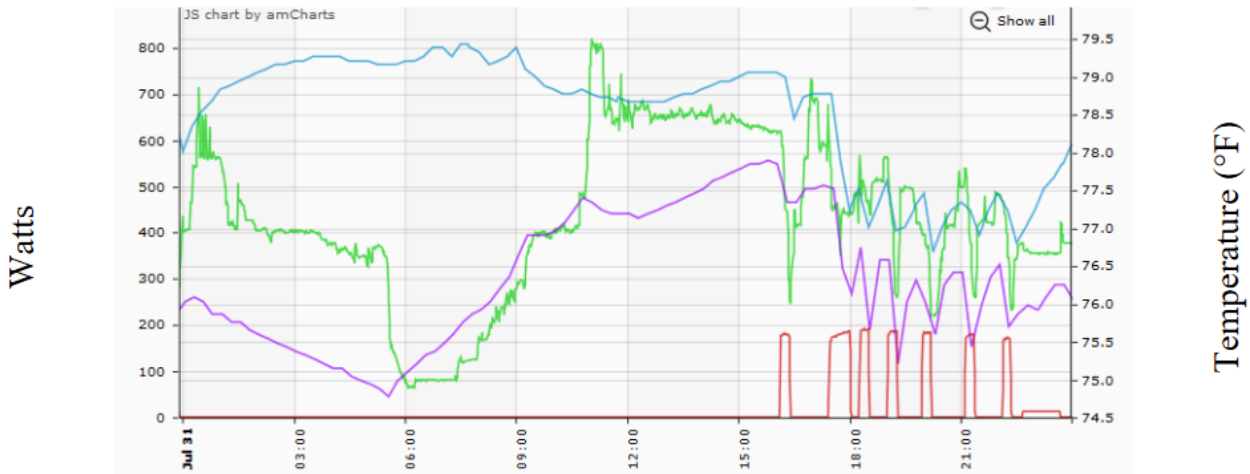


Figure 4: Cooling indoor temperature by hour of a Florida pilot central and zoned HVAC home.
Note: Living room temperature (purple), master bedroom temperature (light blue), central system power divided by 10 (red), zoned heat pump system power (green).
Source: Proceedings of ACEEE Summer Study on Buildings, C. Metzger et. al., 2020, p 1-222, figure 4.

We note that in Figure 4, during the daytime the bedroom is shown to have a notably higher set-up temperature of approximately 79°F, versus a bedtime temperature of 77 °F. Similarly, the living room is constantly maintained below 78 degrees, with an estimated preferred evening temperature of 76.5 °F. This living room temperature fluctuates both above and below this 76.5°F point, indicating the challenge of extracting user preferences and ability to implement optimal control strategies from the actual interval data. The existing 90^{th} and 10^{th} percentile comfort temperature baselining strategies would build a suitable occupied and unoccupied comfort temperature for the bedroom, and a more likely inverted model for the living room since the unoccupied comfort temperature is actually less efficient than the potentially occupied temperature. This unusual fit demonstrates some of the challenges in residential HVAC control commissioning, where some systems may inadvertently be using more energy to satisfy setpoints and indoor temperatures that may not be what the end-user wants from their system. We encourage manufacturers to assist end users in identifying such potential scenarios, especially if the devices are equipped with additional information, such as room occupancy sensing.

Calculations on Zoned Systems

We estimate that, based on these user experience highlights, the most influential variable for a LVCT zoned system is the occupancy at the home and room level, for which most[23] LVCT's are capable of collecting data. We estimate that setback savings when the full home is unoccupied or in night modes could be treated similarly to set back savings of a single zone system. However, rooms with low relative occupancy will tend to experience 90^{th} percentile temperatures (i.e., comfort temperature definition) much cooler than the remainder of the home. Thus this room temperature should receive a low weighting when calculating the overall home comfort temperature. Inversely, rooms with high usage, such as the living room, may maintain a much more constant temperature[24] due to user setpoint requests. We note that this method of developing an occupied vs unoccupied home model is demonstrated in the CalTRACK hourly methodology.[25] Further, this recommendation to collect all zone level data and aggregate into a whole

[23] Mysa-baseboard, uses geolocation for occupancy estimation, Stelpro MAESTRO thermostat and app support geofencing, Sinope product TH1123ZB supports geofencing.
[24] Lambert, 1996, p1:163, Temperature Distributions.
[25] Metrics Meeting Jan 2020, p 18-19.

home model has been performed as far back as estimating the savings impact of electronic line voltage thermostats replacing bi-metallic thermostats in 1996.[26]

Actively occupied zones of the home may vary in temperature demand greatly, whether placed in an interior wall (e.g., inside floorplan bedroom), or a room which has an exterior wall. These effects are further amplified since baseboards are recommended to be installed directly below windows to provide a hot air curtain effect.[27] Due to these differences in placements, zones may experience more or less heating demand relative to each other but will generally be used to provide to occupied zone comfort temperature to end users where they are located at best, and the entire home at worst. Most importantly, some zones can experience considerable 'cross talk', where the direct requested heating load from an adjacent zone with high demand will supply energy (e.g., room door open) to an interior zone, heating it without this controller calling for heat in this zone. Thus, we expect the quality of the temperature demand versus runtime fit to vary greatly in some of these internal spaces, while having a suitable linear fit in others. Illustrating the extent of this challenge, we note that a recent LVCT dataset analyzed with the EPA software, randomly selecting thermostats from the vendor fleet, had 47 percent of this dataset filtered outright due to poor fits. Further, 59 percent of this data was filtered due to an out of range Tau (temperature float) value.[28]

Unfortunately, the heat when directly called for in these spaces is still relevant to the total heating load of this home in aggregate, so a zone with bad model fits is still providing real heat to the dwelling. We estimate that these zones will need their heating runtime accounted for in a joint model, and this whole home should have a workable net runtime versus temperature demand model fit, unless a dual fuel system is present. We anticipate the dual fuel home to have a poor fit between temperature demand and aggregate baseboard runtime, since some other heating source (e.g., furnace) would be providing this heat instead.

An example of this calculation difference is shown in Figure 5, which demonstrates the calculated run time reduction when using zone by zone comfort temperatures versus a whole home comfort temperature. This whole home comfort temperature was the 90th percentile of all zone data consolidated (unweighted) into a single interval dataset.[29]

[26] Lambert, 1996, p 1:160-161.
[27] https://www.nrel.gov/docs/legosti/fy97/6987.pdf, p2.
[28] May 2021 Metrics Meeting, p. 26. See also p. 34
[29] Personal communication, M. Fournier.

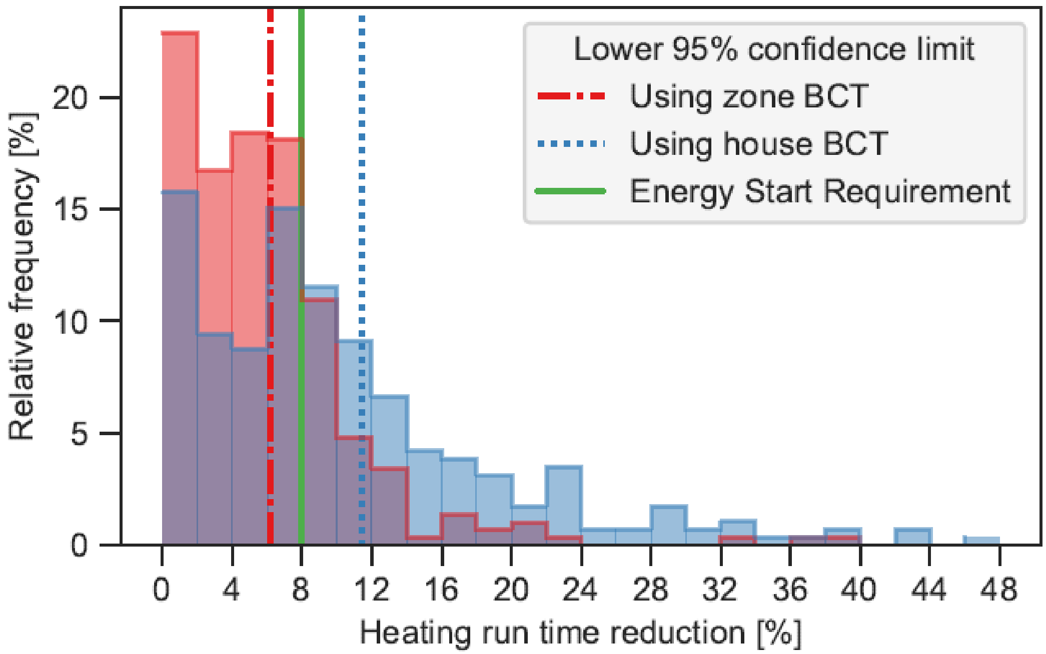


Figure 5: LVCT run time reduction estimates for the set of thermostats with zone by zone comfort temperatures and a whole home comfort temperature methodology.
Source: M. Fournier, Characterizing smart thermostats operation in residential zoned heating systems and their impact, 2018, Figure 9.

As seen in Figure 5, the estimated run time reduction by zone is heavily influenced by the choice of temperature baseline, including the choice of in-zone versus whole home. **We encourage EPA to work with LVCT manufacturers to develop a methodology that considers all thermostats in a sampled home, which:** (1) develops an approximate occupied and an unoccupied comfort temperature, where popular rooms such as the living room are approximately occupied, bedrooms are partially occupied, (2) rank orders and weights the zones by each zone's portion of total energy use, (3) flags and low-weights or removes zones which are apparent outliers on setpoint or extremely low energy use (e.g., rarely occupied basement), (4) evaluates (e.g., CVRMSE) the entire home for inclusion via an aggregate runtime versus exterior demand model, and (5) computes estimated savings on a room by room basis, since some rooms will have notably lower runtime than others.

As described in an earlier section, setback behavior among users of electric baseboard heating is common in the literature. **We strongly encourage EPA to develop savings estimates where the user (1) does not use setback, and (2) uses 2 °F setbacks** in non-common areas (such as bedrooms) where these rooms are identified via lower share of the total home energy and/or evidence of setpoint and indoor temperature setbacks. Conventional users of baseboard heating with modest setbacks may incorrectly assume LVCTs will save them energy if this is not accounted for in a model output.

We estimate that this combined whole home model with and without setback will be suitable for incorporating LVCTs into the connected thermostat program.

3. **The CA IOUs support the development for resistance heating utilization (RHU) metric for heat pump control.**

We are supportive of the development of the RHU2 performance metric and agree with EPA that this metric does not appear to have strong regional trends, as demonstrated in a stakeholder metrics meeting.[30] We further support the use of the 30 to 45 °F bin for this computation, as the originally proposed 5 °F increments of average daily outdoor temperature often eliminated too much data at otherwise informative heat pump installations.

Need for the RHU Metric

One surprising result in thermostat measurement and verification programs was the **negative energy savings of seven to nine percent** experienced by a vendor in the Mixed Humid region when paired with heat pump equipment. This study on smart thermostat performance on heat pump equipment was performed by ICF on behalf of the Southern Maryland Electric Cooperative (SMECO), published in 2018.[31] Further analysis of this data estimated that this negative energy savings was the direct result of high usage of resistance heating during time periods where the other vendor in the study did not apply this backup heating. Later, the severity of this issue was reduced via the impacted vendor pushing an emergency software update to resolve this control algorithm issue.

We consider this scenario to be illustrative of the potential challenges with heat pump control. Had this pilot not been conducted, consumers using this vendor with heat pump equipment would be at significant risk for less than predicted energy savings or even negative savings on their equipment. We agree with EPA's intent to develop this metric to combat this potential issue.

Recommended Changes to the Proposed RHU Metric

We are supportive of the improvements from RHU to RHU2, which includes the ability to drop thermostats in temperature bins which are not commonly experienced for a given thermostat home, by a minimum number of hours the thermostat must experience in that bin.

During the development of both RHU and RHU2 metrics, the thermostat software only provided daily equipment compressor runtime[32] along with hourly auxiliary and emergency heating; thus, the RHU metrics correspondingly were developed around daily average temperatures and daily compressor runtime. We estimate that existing RHU calculation strategies are suitable for catching vendor algorithms that provide extremely warm heat pump compressor lockout temperature defaults.

Now that the alpha (Version 2.0.0 GitHub branch) edition of the *epathermostat* software includes hourly interval data for equipment runtime (i.e., heat pump compressor), we strongly encourage EPA to revise the RHU metric to account for scenarios where the hourly compressor runtime is high in a given hour (e.g., more than 60 percent). This would be likely one of two scenarios: (1) the system is rapidly recovering from a wasteful setback during the heating season, or (2), the system is struggling to maintain the needed indoor temps given the system capacity at that temperature. We recommend that the second scenario could be excluded from the RHU score, since the system capacity issues are not the responsibility of the thermostat vendor. These scenarios can be differentiated by considering the interval data consecutively, checking for deep night setback from which the system ultimately has to recover. The

[30] May 2020 Metrics Meeting, p. 4

[31] BPA writeup, p. 18 on ICF SMECO report; note CA IOUs has a hardcopy of this report but could not find a working online link. Negative savings and confidence interval on p. 26 of SMECO report.

[32] https://epathermostat.readthedocs.io/en/latest/data_files.html, Thermostat Interval Data CSV Format.

product literature of some vendor's products demonstrates an awareness of and account for this exact recovery issue. [33] Such a revised RHU metric would provide even more clarity on the heat pump performance of a vendor's products in the field.

Need for RHU Oversampling

The variation in RHU scores by equipment within a given vendor was found to be highly dependent upon whether the analysis was performed on a standard resubmission dataset, or a special oversampled heat pump data set.[34] In this case, the vendor experienced a 0.08 performance change in the mean 95th confidence interval upper bound of their RHU score (i.e., EPA proposed variable), which was a decrease of over 60 percent and moved the vendor from failing the potential limit to substantially passing it. Note that the proposed requirement score is 0.20 on this ci95ub, so this variation was almost ***50 percent*** of the requirement and **indicates that RHU scores cannot be reliably calculated on a traditional sample**. Analysis of the heat pump count indicates that the number of heat pumps analyzed was increased by more than a factor of ten when a sole heat pump (oversample) was collected.

We acknowledge the additional vendor level of effort to certify and send resubmission data to EPA if a requirement for heat pump oversampling is enacted. Recent trends in the U.S. demonstrate that this knowledge of heat pump performance of a thermostat product will increasingly become more essential in understanding smart thermostat savings capability in the U.S. population.

California has aggressive greenhouse gas (GHG) reductions targets and building electrification through heat pumps is an identified strategy to achieve those goals.[35] A new pilot program, Technology and Equipment for Clean Heating (TECH),[36] recently launched to accelerate the adoption of commercially available, low-emissions space and water heating, primarily focused on air source heat pumps and heat pump water heaters. This is a 4-year project with over $100 million commitment to drive this transformation in California. Heat pumps are growing in popularity across the U.S., outside of the California perspective; the CA IOUs have seen commitments from Massachusetts to install one million residential heat pumps by 2030,[37] and significant financial commitments in New York to grow the deployment of heat pumps, including an April 2021 announcement from New York's Governor Cuomo to support education and awareness efforts to expand the deployment of heat pumps in New York.[38] Maine, one of the coldest states in the nation with less than 600,000 households,[39] is already well on its way to transition towards heat pump technology, with a heat pump in almost 10 percent of households already by 2019, and a 2020 legislative commitment to install another 100,000 heat pumps by 2025.[40] While statewide commitments, pilots, and programs are moving in this direction, the Biden Administration has also declared a strong commitment to growing electric heat pump technology.[41]

While the CA IOUs understand that currently, relatively few smart thermostats are installed on heat pumps (as demonstrated by the current sampling challenges), we are confident that this landscape will be

---

33 https://www.ecobee.com/en-us/citizen/top-10-ways-to-get-the-most-out-of-your-ecobee-thermostat/ , 10. Smart Recovery.
34 April 2020 Metrics Meeting, p. 7
35 https://www.energy.ca.gov/data-reports/reports/building-decarbonization-assessment
36 http://www.aqmd.gov/docs/default-source/clean-air-plans/air-quality-management-plans/2022-air-quality-management-plan/2022-aqmp-residential-and-commercial-buildings-working-group/tech_sc-aqmd_5-6-21.pdf?sfvrsn=10
37 https://www.mass.gov/doc/2030-interim-cecp-march-public-webinar-slides/download slide 8
38 https://www.nyserda.ny.gov/About/Newsroom/2021-Announcements/2021-04-12-NYS-Clean-Heat-Members-Announce-Nearly-10-Million-Consumer-Education-and-Awareness-Campaign-to-Accelerate-Adoption-of-Clean-Heating-and-Cooling-Solutions-that-Combat-Climate-Change
39 https://www.census.gov/quickfacts/ME
40 https://www.efficiencymaine.com/beneficial-electrification-study/
4141 https://www.whitehouse.gov/briefing-room/statements-releases/2021/05/17/fact-sheet-biden-administration-accelerates-efforts-to-create-jobs-making-american-buildings-more-affordable-cleaner-and-resilient/

changing rapidly. As such, we strongly recommend EPA consider analyzing a separate heat pump cohort to ensure utilities, states, and stakeholders are able to understand the interaction between these two important technologies.

4. The CA IOUs support the continued development of modulating and variable capacity system metrics by EPA.

We are strongly supportive of the development of modulating and variable capacity metrics for the connected thermostat program. We acknowledge EPA's comments on the challenge of assembling a full three vendor or more datasets, and the associated challenges of developing a successful metric based on the existing data collected. We note that other recent analysis[42] on variable capacity systems have used equivalent full load hours, as estimated by the 98th percentile of full load power of the system and relative power consumption of the unit at a given capacity. Since the EPA program does not currently collect power data, potentially the reported relative capacity of a given operating point could be looked up in a standardized power to relative full load capacity table EPA develops with manufacturers.

We also encourage EPA to work with manufacturers to pilot a hybrid evaluation procedure, where the communicating controller can be connected to simulated (i.e., test harness) equipment, and theoretical ambient conditions can be applied to the device. This methodology allows EPA and third-party certification bodies to poll the communicating controller at a set of test signals without the manufacturer needing to disclose the exact nature of the equipment control algorithms, a known variable capacity equipment sensitivity.

5. The CA IOUs support the inclusion of equivalent full load runtime (ERT) for staged systems.

We are supportive of the incorporation into the scope and development of the equivalent full load runtime for staged equipment, as discussed in previous metrics discussions. We agree with EPA and industry comments that capacity ratio between two and three stage equipment is often following industry sizing conventions, thus can be analyzed using supplied interval data for each stage and the estimated capacity ratio.

We encourage EPA to conduct a data-call to demonstrate the success of this calculation strategy in providing acceptable CVRMSE fits in aggregate, indicating suitable model performance, for at least three vendors as part of the development process. We anticipate the successful application of this metric on field data.

6. The CA IOUs support key software updates to enable more thermostats and new metric calculations.

Discussion in numerous earlier comment sections recommend the inclusion of significantly more thermostats in each vendor run of the EPA thermostat software than current requirements. Anecdotally, the current software has a multiple-hour runtime on the sample of 1250 thermostats, and this is generally attributed to the long runtime of the *eeweather* sub-module retrieval of National Oceania and Atmospheric Administration (NOAA) weather station data. We encourage EPA to construct a pre-compiled weather station cache file as part of the data resubmission process. We believe EPA should host this file and provide download instructions and a means for the software to import easily, thus removing the primary runtime speed obstacle.

[42] For example, see: C. Metzger, 2020, p. 1-222.

Further, this cache file can be analyzed by EPA and its contractors to ensure all the NOAA station files are at 100 percent data integrity. At least one stakeholder[43] has reported that some weather station files are not downloaded for a file or after download are missing weeks or more of interval data, ensuring any thermostat ran with this file would be rejected immediately. EPA can substitute nearby station data in these circumstances, in a similar manner to how the software is intended to retrieve a nearby weather station if the weather station closest to the Zip Code Tabulation Area (ZCTA) centroid is not available or unusable.

On runtime speed improvements in general, we encourage EPA and its contractors to benchmark the software and identify the next bottleneck after weather data retrieval and resolve that, to ensure a superior vendor program experience with the substantial proposed increase in submitted thermostats.

Lastly, vendors have reported significant challenges in determining the filtering reason(s) for thermostats in their dataset in aggregate. EPA has also indicated the importance of determining the sources of thermostat data file rejection in the population in aggregate, to ensure that vendor submitted data files contain sufficient thermostats and are representative of the vendor overall population. To remedy this situation, we recommend that EPA and its contractors implement an automatic log file, detailed version, and aggregate version, where the detailed version allows a vendor to see individual thermostat rejection criteria by ID number, and the aggregate file would be submitted to EPA as part of the summary statistics data file(s). Under this methodology, both the vendor and EPA can monitor challenges with the dataset population.

7. The CA IOUs recommend further investigation into the challenge of connected thermostats which are in offline-only operation.

While the CA IOUs fully expect that the best energy savings and demand response results will come from a smart thermostat that is connected to a broadband internet network, as smart thermostats grow in popularity, there are many situations that would necessitate a disruption to the full connectivity. Homeowner behavior around smart thermostats at the time of home sale is variable, with a growing trend of homeowners installing smart thermostats shortly before listing their home to help improve the home value.[44] Those homeowners are unlikely to take the thermostat with them, and while the new homeowners will presumably re-connect the installed thermostat to the internet and account eventually, there is an undetermined period of time where a smart thermostat could be installed but not connected to the internet. A similar scenario could occur in rental properties, where tenants would cancel internet service when they leave a rental, and there is an undetermined lag between when new tenants move in, set up internet, and connect the smart thermostat. As smart thermostats grow in general popularity and significant rebate programs are in place, the likelihood that a thermostat may have more than one owner in its useful life increases. Finally, it is inevitable—even for the same thermostat owners in the same home—that the internet will go down intermittently, or routers will periodically get unplugged. Given the potential gaps in connectivity, the CA IOUs recommend EPA consider ways to increase energy savings for non-connected smart thermostats.

One opportunity that the CA IOUs have considered include ensuring any occupancy sensors are using a local network such as Bluetooth or Zigbee to connect back to the central thermostat. While it is our understanding that this is largely an industry practice for functional and cost reasons, there is no requirement that occupancy sensors could operate independently from an internet connection to function as designed. This may be an appropriate way for ENERGY STAR to help codify the best practice to help

[43] Metrics Discussion, May 2021, p. 7

[44] Coldwell Banker, "Smart Homes: An Emerging Real Estate Opportunity" https://blog.coldwellbanker.com/wp-content/uploads/2018/01/CES2018-Smart-Homes-An-Emerging-Real-Estate-Opportunity.pdf

ensure continued energy savings functionality even when not connected to the internet. Another opportunity for continued energy savings in time of un-connectivity could be defaulting to the out-of-the-box schedule, which is likely to be more dynamic than a standard programable thermostat setting. Finally, the CA IOUs encourage EPA to promote manufacturers to make it easy to transfer smart thermostats control from one resident to the next. This can help ensure continued energy savings across new tenants or homeowners.

8. The CA IOUs recommend the following updates to EPA's Demand Response (DR) criteria.

We consider the demand response and flexible load component of connected thermostats to be a significant component of both the product and the specification. We recommend the following updates to further support robust and consistent intra-vendor operation in DR programs.

Requirement of OpenADR Device Certification

We note some discussions around current device 'off the shelf' compatibility and interoperability. Namely, that ADR products that are self-reported as OpenADR "compliant" but are not OpenADR 2.0-certified often require customized solutions on error handling, registration, commissioning, and other topics to integrate with demand response management systems (DRMS).[45] Another vendor described better software integration experiences.[46]

A simple but illustrative example of integration issues is that to transmit and receive DR event signals, OpenADR devices use the *EiEvent* service, which describes how Virtual Top Nodes (VTNs) (for example, a DRMS) communicate DR event information with Virtual End Nodes (VENs) (for example, devices that use OpenADR). The OpenADR EiEvent service can use values such as [0,1,2,3] to denote the relative priority of an event or the level of a simple event signal. Some implementations may store these values as integers, whereas other implementations may use numeric values. A system expecting the wrong datatype may fail to recognize an encoded event. Typically, this sort of challenge is discovered in the field.[47] If two VEN devices in a given program use different strategies, then the DRMS must use special message handling between the systems indefinitely, leading to increased system complexity.

A recent July 2021 article by GridFabric, additional context around cloud VEN interoperability challenges were provided, including: differing ranges of devices controlled, selection of different event types, implementation of event opt-outs, etc.[48] GridFabric notes that one-off development of VEN systems that are not OpenADR 2.0-certified tend to require significant debugging in the field.[49] We also note the requirement of OpenADR 2.0 VEN certification in California's Title 24 Part 6 Building Energy Code.[50]

Some examples of direct CA IOU experience[51] with VTN-VEN integration issues include: several reports of stakeholder challenges obtaining and implementing security certificates, differences in how products implement registration with the DRMS and how they populate VEN ID fields, inconsistency on standard polling intervals, early configuration issues with polling the VTN resolved in the testing phase, handling of events with open end times, handling of different event time zones, handling of near/far event tags, and difference in handling event opt out for OpenADR 2.0A versus 2.0B VENs. As these integration issues

---

[45] Interview with M. Kohanim, Universal Devices, 2020.
[46] Interview with J. Klube, Electriq Power, 2020.
[47] https://energycentral.com/c/iu/how-do-certification-program-right-way, 2014.
[48] https://www.gridfabric.io/blog/cloud-vens-interoperability, 2021
[49] https://www.gridfabric.io/blog/build-vs-buy, Value, 2020.
[50] https://www.openadr.org/assets/symposium/1.Taylor-2019-06-12%20OpenADR.pdf, p. 7-9.
[51] Unpublished interviews of 6 DR technical experts operating in CA IOU DR programs, 2019.

have occurred even with certified VENs, we expect even more issues in cases where the VEN is not certified.

In one particular experience, a vendor used integer values for the VEN ID field, whereas the OpenADR 2.0 protocol allows string values. This is common in practice, thus the vendor had to reconfigure the VEN solution. In another example, a software update to the DRAS inadvertently knocked a vendor fleet offline until the introduced issue was resolved. Occasional show-stopping challenges have occurred in practice, where a non-OpenADR 2.0-certified VEN does not successfully support all DR functionalities that would be supported in a certified VEN.[52]

To benefit intra-vendor, and intra-DRMS system compatibility, we encourage EPA to require OpenADR VEN certification to the OpenADR standard as part of the connected thermostat specification.

Requirement of Major Demand Response Modes

We note that the evolution of thermostat and HVAC equipment demand response generally began with direct load control methodology, where utilities would directly control equipment to implement specific changes during events. One source notes that at least six major utilities piloted air conditioner direct load control strategies as far back as 2010.[53]

More recent developments on HVAC equipment control include the industry standard AHRI 1380, which includes CTA-2045 and OpenADR 2.0 control of equipment via both event-based demand response and price responsive demand response.[54] We strongly support this vision of ensuring the residential HVAC equipment and thermostat devices are capable of both event and price responsive DR, and we encourage EPA to adopt similar requirements for the connected thermostat program, where the device must be capable of: (1) direct load control, (2) DR event response, and (3) price responsive operation.

Similar to other EPA connected devices, we recommend that EPA also develop a method to validate demand response for connected thermostats, where the device's ability to decode and issue a direct load control signal, a DR event request or notification, and a price responsive event request to controlled equipment is verified in a lab test environment. We do not consider this test method to require specific default DR event response strategies, as instead this choice of how to respond to an issued event would remain the challenge of the combination of thermostat vendor, HVAC equipment original equipment manufacturer, utility DRMS administrator, and the local third party aggregator if applicable.

Additional DR Data Collection

We note some topic discussions in the industry on challenges to estimate DR event impact, especially summer peak load curtailment events.[55] We consider data availability one such method to assist in addressing this challenge, which typically benefit from the ability to slice user interval data into pre-event, during event, and post-event time periods if applicable to a given user thermostat.

To facilitate this analysis, we encourage EPA to add an hourly DR event flag to the existing mandatory interval data collection requirements for connected thermostats. This would produce an intra-vendor

---

[52] Personal communication, A. Chiu, PG&E.
[53] M. Siemann, 2013, p 21, cites Newsham and Bowker 2010.
[54] AHRI 1380, see for example, table 2 for CTA 2045 control: DR Events: 'general curtailment', 'critical curtailment', 'off mode'; Price Response: 'Utility Peak Load Price Signal' and table 4 for OpenADR control: DR Events: 'general curtailment', 'critical curtailment', 'off mode'; Price Response: 'Utility Peak Load Price Signal'.
[55] See for example the August 2020 OhmConnect discussions: General article, Recurve Analysis Summary, Recurve Detailed Analysis Report.

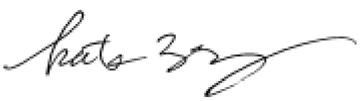
standardized dataset in the region that could be paired with either AMI data if available, or monthly billing data if not, to estimate for example the treatment group, counterfactual group, and the difference in operation to obtain DR event load impacts.

In conclusion, we would like to reiterate our support to EPA's Discussion Document on Connected Thermostats. We thank EPA for the opportunity to be involved in this process.

Sincerely,

Patrick Eilert
Manager, Codes & Standards
Pacific Gas and Electric Company

Karen Klepack
Senior Manager, Building Electrification and
Codes & Standards
Southern California Edison



Kate Zeng
ETP/C&S/ZNE Manager
Customer Programs
San Diego Gas & Electric Company

Appendix A: Supplemental Plots and Data.

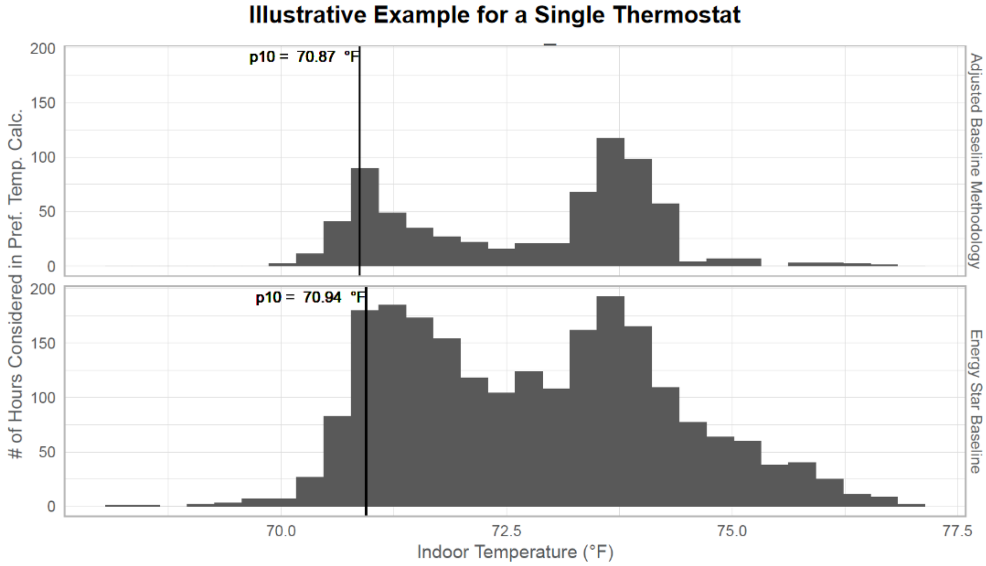


Figure A1: Guidehouse comfort temperature methodology applied to a home with more than one common interior temperature (bi-modal).
Source: Guidehouse, ComEd Advanced Thermostat Evaluation, 2020, p. 63.

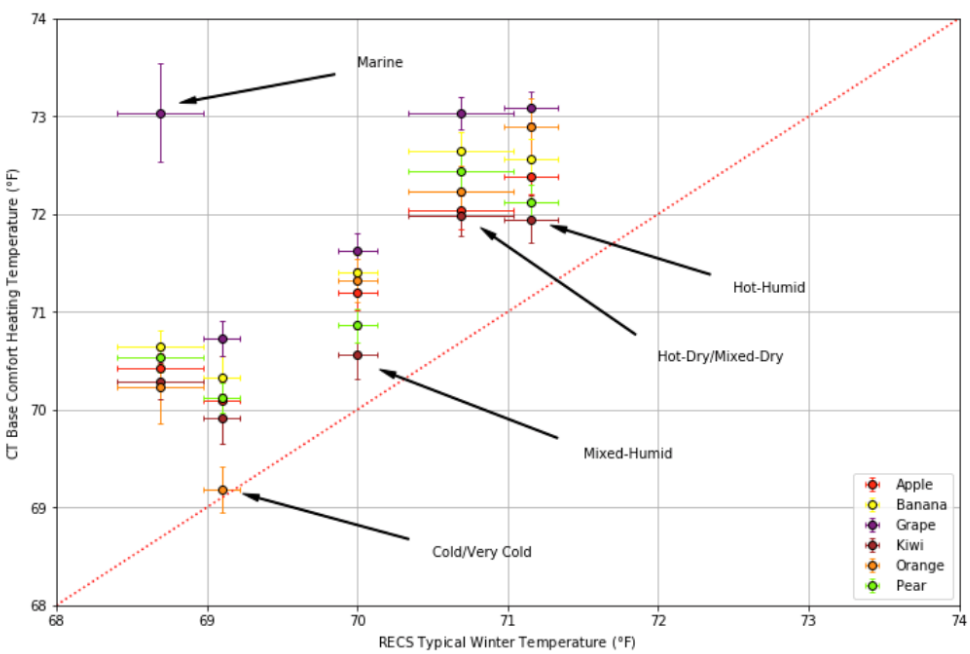


Figure A2: RECS 2015 heating temperatures versus thermostat vendor population comfort heating temperature (mean) in the 2017 season.
Source: EPA Metrics Meeting November 2018, p. 33

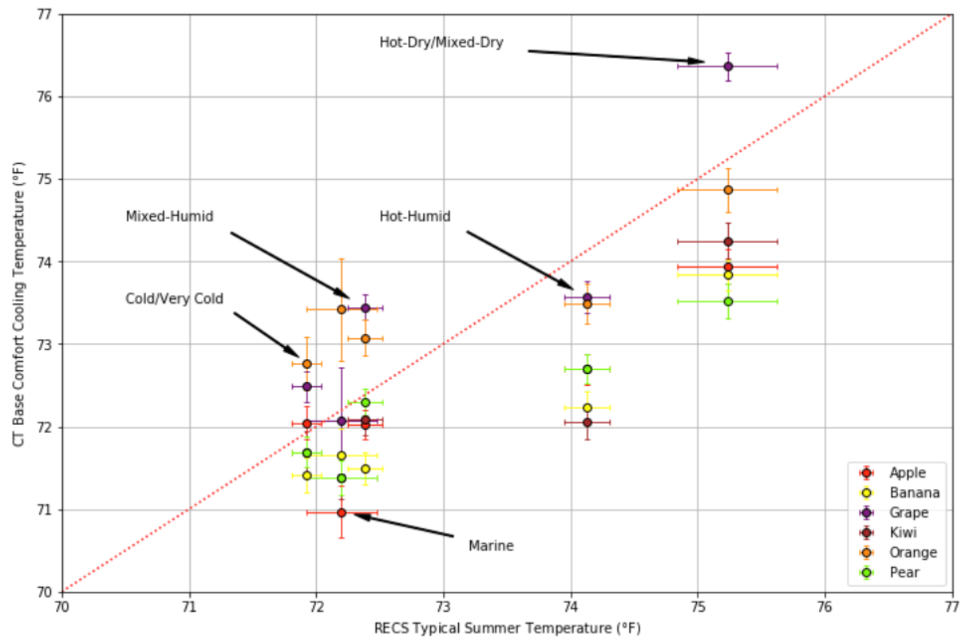


Figure A3: RECS 2015 cooling temperatures versus thermostat vendor population comfort cooling temperature (mean) in the 2017 season.
Source: EPA Metrics Meeting November 2018, p. 34

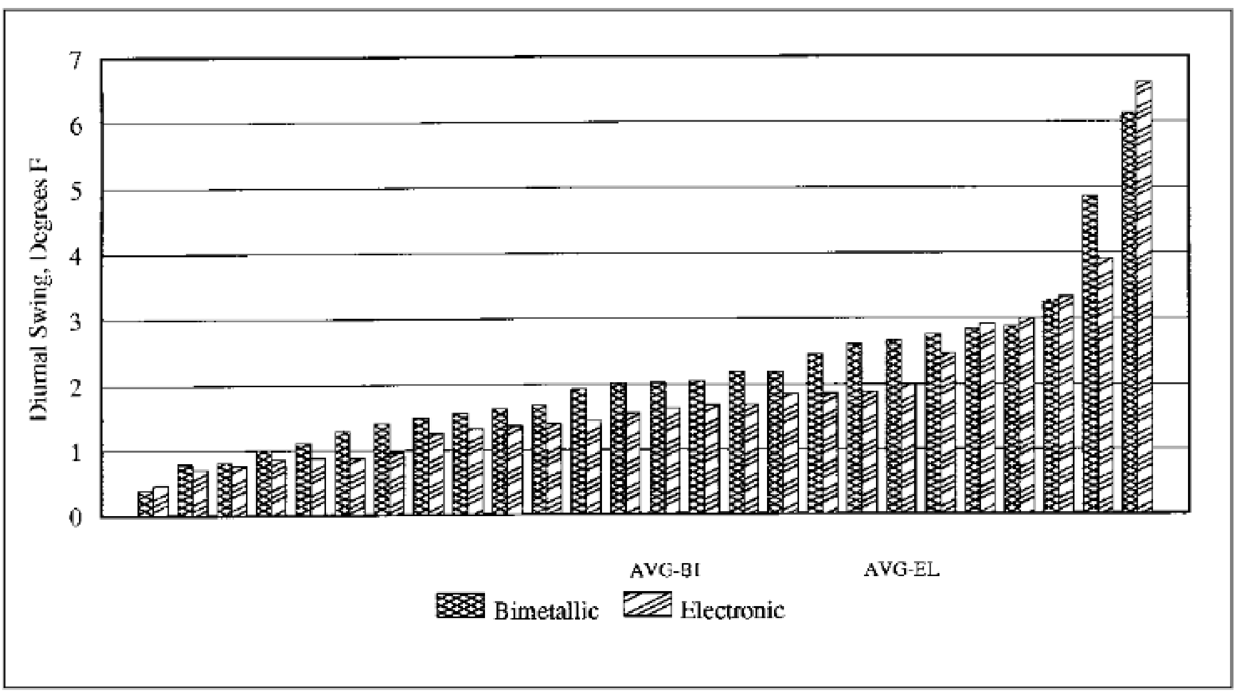


Figure A4: Daily setback temperatures of 26 households with line voltage thermostats in bimetallic to electronic replacement pilot study.
Source: Lambert, 1996, p. 1-163, figure 3.