1. Updated Terminology

Currently, EPA uses the term “connected thermostats” with partners to identify the products that fall within the scope of this product category. However, the term Smart Thermostat is more broadly used in marketing material for the products and thus is believed to be the consumer-recognized classification. In recognition of this, EPA has used the term in marketing materials from the inception of the program. With Version 2.0, EPA intends to begin referring to the product category as ENERGY STAR Smart Thermostats with all stakeholders. This should also reduce confusion with products commonly referred to as Communicating Thermostats, which use a digital bus to communicate with HVAC equipment. This change will be reflected in related documents, informational material, and throughout the ENERGY STAR website.

1. Will this updated terminology cause any issues or business impact for manufacturers, suppliers, or other relevant parties?

2. Revisions to Specification Metrics for Currently Certified Products

2.1 Improved Submission-to-Submission Stability of Metrics

EPA provides random number seeds to partners before each submission to ensure that the data submitted is auditable and reproducible. However, vendors have noticed that choosing several samples from the same population in the same time period can yield different results and reported the same to EPA. EPA currently requires the maximum sample size that may be submitted per climate zone and is now proposing a minimum sample size per climate zone to reduce the effect of housing, occupant characteristic, and use-case variability on metric stability. EPA seeks partner feedback on the following:

2) What, if any, hindrances are there to moving to allowing a larger sample size per climate zone?

3) Would it be helpful if EPA proposes that partners include a percentage of total population as opposed to a specific sample size?

Yes, Mysa would prefer a percentage of the total.

4) Based on your analyses, are you able to recommend minimum or maximum sample sizes that EPA should include in the test method?

Minimum of 20 thermostats per climate zone.

2.2 Improved Correlation between Metrics and Energy Savings through Metered Data

The current ENERGY STAR heating savings and cooling savings metrics only reflect savings from setback and set-up and assume that no setback behavior would occur without the smart thermostat. Thus, in some ways, it overestimates savings and in other ways, it underestimates savings. Because of this, EPA is not confident that a higher metric score reflects higher savings and advises against using the smart thermostat specification criteria as a basis for deemed savings, making programs harder to set up and administer. (Note that the specification requirements are set high enough that even if the savings are half of that indicated by the

metric, purchase of the product is generally a good decision for consumers, so the current metric suffices for the ENERGY STAR program.) EPA has always intended to work with partners to validate and/or improve the metric such that it is correlated with energy savings for groups of
homes and ranks product performance accurately. However, since the program's inception in 2017, EPA has not been able to make significant progress on this work.

5) Are there data sets correlating meter data and data from smart thermostats for the same group of homes, either with AMI information or large samples, or both? What would it take for EPA (or another party) to be able to analyze this data to examine the correlation and, if needed, improve it? Note that we would ideally need to be able to correlate energy meter data and thermostat data from individual homes or sub-groups of homes, not just to have aggregate energy bill savings and ENERGY STAR smart thermostat metric scores.

6) If such data sets do not exist, what would it take to collect them? We understand there are consumer privacy and proprietary information concerns.

7) A less expensive approach might be for smart thermostat vendors participating in utility incentive programs to calculate metric scores for users in the program, or at least in the same geographic areas, and submit them to evaluators or program managers for comparison with those to the metered savings for each vendors' products in the study. Is this feasible for any utility partners? Would smart thermostats vendors be amenable to participating?

2.3 Weighted Savings

The ENERGY STAR software maps US zip codes with the five Energy Information Administration (EIA) climate zones. Currently the tool uses a weighted average to compute national savings for both heating and cooling, weighting zones by how much heating (or cooling) energy they use according to the EIA. This is appropriate to ensure significant national savings, and also avoids a problem when there are minimal installations with heating in very warm climates or cooling in cold climates. EPA chose this weighting because it's natural for the calculation of program savings we derive from it. However, a more even weighting across the different climate zones would yield a metric more relevant to percent savings for every homeowner in the US.

8) How do stakeholder use and think about the metric? Would one weighting serve your purposes better than another? If a more even weighting serves your purposes better, but not an exactly even weighting, what would be a fair basis for weighting the zones?

Mysa baseboard thermostats have a heavy presence in northern and marine states, and a small fraction of thermostats installed in southern and hot climates. It is more accurate for our customers to have a weighted savings, because energy savings in the hot/humid climate zone is not indicative of what the average US customer would save when they install a Mysa BB thermostat.

3. Other specification criteria changes

3.1 Updating Demand Response Requirements

According to the ENERGY STAR connected thermostat specification, Wi-Fi communication links can use open standard or proprietary protocols. Smart thermostat (CT) products that enable direct, on-premises, open-standards based interconnection are preferred, but alternative approaches, where open-standards connectivity is enabled only with use of off-premises services, are also acceptable. To enable interconnection with the CT product over the communication link, the partners are required to provide an interface specification, application programming interface (API) or similar documentation that enables demand response (DR) functionality. However, as per ENERGY STAR's Grid Communications requirements, OpenADR 2.0 or Smart Energy Profile (SEP 2.0) must be used for communication between service provider's cloud and a Utility's demand response contractor.

9) Is it appropriate for EPA to require OpenADR 2.0 or SEP 2.0 for cases where the CT service provider is acting as a DR aggregator?

10) Are there any other open standard protocols that EPA should consider?

11) Would it be clearer to stakeholders if EPA mentioned OpenADR and SEP 2.0 as a part of prescriptive requirements?

12) What are some of the CT vendor offerings pertaining to demand response that are relevant for utilities?

3.2 Product Families

The definition of a product family in Version 1.0 is deliberately broad, to allow the new program to grow and to gather more information about the variety of products on the market. Most vendors have a single product family certified, with up to a dozen individual product models included. A few vendors have two product families. For the July 2021 resubmission, EPA has requested that vendors provide results from each product in their product family, or in some cases from closely related groups of products. Vendor discussion led EPA to conclude that the most relevant differences between products were likely to be differences in sensors, user interface features, retail vs. pro channel, target market, and default settings (including templates and schedules). EPA will decide whether to propose a more specific product family definition based on the results received from this request. EPA will consider the statistical significance of any differences between models or model groups within a product family in making our decision.

3.3 Reporting Whether Broadband Connectivity is Needed for Savings

As programs become more interested in deploying smart thermostats in rental properties and in neighborhoods poorly served by broadband infrastructure, the ability of products to deliver savings without broadband connectivity becomes highly relevant. Because the current program relies on field data to demonstrate savings, the results from unconnected smart thermostats have no influence on metric results, and EPA would not be able to verify such claims. However, if it were possible to identify features that are necessary for smart thermostats to deliver savings

when disconnected, they could be verified through product and literature examination. For instance, if a product relies on automatically setting back for savings and has both on-board sensors to identify when homes are empty and on-board processing to react to that information, its savings mechanism wouldn't require broadband connectivity.

13) What features or sets of features would be sufficient to give a reasonable expectation of savings without broadband connectivity?

14) Is there some way to identify models that provide savings without broadband connectivity other than by identifying sets of features?

15) If manufacturers claimed this capability, how could it be verified?

4. Expansion and Clarification of Scope

4.1 Smart Line-Voltage Thermostats (LVTs)

EPA recognizes that electric resistance baseboard heaters are far less efficient than many other types of heating methods. As such, the potential savings gained by setting back temperatures in response to demand management are compelling enough to support the inclusion of Smart LVT products within the scope of the ENERGY STAR Smart Thermostat product category. Additionally, EPA has been contacted by three vendors that are interested in certifying LVTs under ENERGY STAR and has been made aware of the general interest of utilities. As the Canadian Standards Association (CSA) has recently completed a baseline standard for LVTs (CSA C828:19), EPA is currently engaged with a vendor to understand whether available data can be used to form sensible conclusions about the potential effectiveness of including LVT under the scope of ENERGY STAR Smart Thermostats Version 2.0.

EPA is also currently working to amend the draft Version 2.0 specification language and product requirements to allow for the inclusion of LVT products. This language and the associated requirements include added definitions and device criteria specific to baseboard heaters and will be presented for stakeholder review in the forthcoming Draft 1 specification. EPA cannot yet affirm that there will be no changes required for the current test method or analysis software for Smart LVT products. The need for such changes will depend on the results of the investigation into current data, which is not yet complete. We welcome the involvement of additional vendors who offer LVTs.

16) Do you offer LVTs, and would you be willing to participate in the development of the metric for these products?

Yes, Mysa will volunteer all necessary data and resources to move this initiative forward!

4.2 Communicating Controllers

Currently, the proprietary thermostats designed to control centrally ducted variable speed air conditioners and heat pumps are not able to be certified as ENERGY STAR due to high standby power consumption and limited installations controlling single speed equipment. As to expanding the current specification to include communicating controllers, EPA notes that primary assumptions behind the current metric are not true for variable speed equipment. These include that run time is a good proxy for HVAC energy use, that energy use is approximately

linearly related to indoor-outdoor temperature difference, and that the major opportunity for savings is set back/ set up when no one is home. EPA has explored several ways to evaluate the performance of controls for this equipment, including:

- A field data metric called Average Capacity Factor, which expresses the mean percent of the full capacity a system runs in during a given time period. This is meant to recognize controls that avoid cycling at higher capacity when a load could be met by running constantly at lower capacity. This could also recognize systems that employ intelligent setback and recovery.
- A lab test of recovery from a setback in a way that saves energy, or that shows that the control can hold the system in lower capacity states at an appropriate lower load.
- A hybrid test in which the control is placed inside a small climate-controlled chamber and the response of the controlled equipment is simulated.
- A system specification applying to the combination of controlled HVAC equipment and controller, which would only apply to the highest efficiency equipment. What metrics or ratings would apply to such a specification is unclear at present, and it is likely that some elements of the methods for evaluating controllers mentioned above would need to be incorporated.
- Supplemental characteristics or field data which might include the unit's rated capacity or a means of reporting energy consumption (or a proxy for it).

EPA has not concluded that any of these methods are ready to be tested with vendors. Note that many of these controllers are in installations where they control some equipment that is variable speed and some which is staged. EPA would want the control of the staged equipment evaluated using the standard metric, as for other thermostats.

17) If you supply communicating thermostats and are interested in working with EPA on developing an evaluation method, please let us know.

18) Is there some venue for method development that will be more effective than EPA-convened groups, such as an industry standards development organization, other government agency, or NGO?

19) There is significant overlap between the evaluation of communicating controllers for centrally ducted variable capacity systems and the ENERGY STAR CAC/HP specification. As such, EPA is seeking to include the control algorithms in the evaluation of variable capacity heat pump performance, whether they are in the heat pump control board as in a mini-split, or in the communicating controller as in a centrally ducted system. If you have ideas for recognizing exceptionally efficient HVAC with its controller as a whole system, please let us know.

20) Do stakeholders agree with the rationale behind this approach, which emphasizes contribution of the controller and associated software to efficiency?

21) Are there alternative approaches that stakeholders would like EPA to consider when developing the Draft 1 specification?

22) Do stakeholders favor a field data, lab test, or hybrid approach?

23) Are there any instances where the CT algorithms and the HVAC controls do not complement each other?

4.3 Mini-Split System Controllers

The ENERGY STAR method of demonstrating thermostat savings using field data makes several assumptions. The current Version 1.0 specification is not explicit about these assumptions and that there must be a reasonable expectation that they are met in most homes for products to be certified. EPA intends to propose language rectifying this.

For mini-splits in particular, EPA is not confident in these assumptions, and hence the specification was not intended to cover thermostats used to control them. In 2016 when the specification was finalized, this type of product was just entering the US market and was not addressed in scope. To avoid any confusion, EPA will explicitly mention that mini-split system controllers are beyond the scope of ENERGY STAR specification. EPA would be willing to work with stakeholders to extend recognition to these products but at present has no hypotheses about how the controls would be evaluated for energy savings.

24) Do you offer mini-split controllers and have ideas about appropriate metrics for these products? Would you be willing to invest in making it possible for them to be certified?

Yes, Mysa will volunteer all necessary data and resources to move this initiative forward! What data can we provide?

4.4 Evaluate 2-Stage Heating/Cooling Installations

The existing ENERGY STAR field savings analysis is limited to installations controlling single speed HVAC equipment types (to the extent it can be determined by the connected thermostat product). EPA understands that up to a third of installations include two-stage heating or cooling and were excluded from the analysis using the V1 software. EPA hypothesizes that two-stage installations can be analyzed using an “equivalent full load run time” where lower stage run time is discounted by the approximate relative capacity of the lower stage before being added to the run time summation. Thus, the equivalent full load run time should be a reasonable proxy for delivered heating or cooling capacity. This method has been included in the Version 2.0 software, and EPA requested results from a sample of two-stage installations from vendors as a part of the February 2021 submission. EPA is still awaiting data from the vendors by product family and expects to receive the same as a part of July 2021 data submission. Accordingly, EPA will analyze the data and revise the process of how a sample is chosen from the population using thermostats in the product family.

25) What is the mix of equipment types and geographic spread/zip codes?

4.5 Filtered data and goodness of fit

The EPA software automatically checks whether installations meet the assumptions of our model, and filters out those that do not, not including them in the calculation of savings metrics. EPA evaluated the 2021 February submission data for filtered thermostats the results of which can be seen in the image below. For most vendors, most installations were usable for calculating metric results, as the filtering is now done. As of now, EPA can tell how many installations are dropped due to poor fit to the linear model but does not have specifics about other reasons installations might be excluded, for instance due to incomplete thermostat data or due to inability to find weather data. Naturally, there will be some installations which will never fit the model, like a single-family home that uses a wood stove to supplement their heat pump

regularly. In addition, there may be vendors with many installations that do not fit the model due to unusual customers or use cases, for instance in hotels or to control refrigeration equipment. (Also see discussion of additional environments below.)

Two possible results of different uses cases include systematic bias in the results (particularly problematic if various vendors have different percentage of use cases) or increased "noise" in the installation data. For the latter, including an analysis of the percentage of data filtered out and the potential for including a minimum percentage unfiltered ensures that the results are meaningful for the product. This would also address vendors with many installations with too much data missing to draw conclusions. To address these concerns, EPA will provide results so that our partners and we can understand how many installations are not included in the summary statistics and why. We are considering the requirement that a certain proportion of the data set shall be included in the statistics. In parallel, we are working to understand just how much missing data our method can tolerate, so that more installations can be included in the analysis.

26) What proportion of your thermostats are filtered out? Do you know why? Can these issues be addressed?

27) Do you agree that the credibility of the ENERGY STAR score depends on the proportion of installations that fit the model it relies on? Why or why not?

Is ENERGY STAR still implementing "core" heating and cooling days and shoulder days that were discussed in the May stakeholder meeting? Mysa feels that the existing model does not accurately account for heating and cooling sources that are seasonal. Thermostats that are offline when the heating / cooling source is turned off should not be disqualified due to connectivity. Is ES still investigating the calculation of tau?

4.6 Clarify Specification Applicability for Additional Environments

EPA has received inquiries from vendors offering smart thermostats for a variety of use cases other than single family homes (e.g., dormitories, rental properties, small commercial spaces). In addition, some current partners have substantial subsets of users in these environments. Lastly, the specification requirement to include an ENERGY STAR certified smart thermostat in ENERGY STAR Smart Home Energy Management Systems (SHEMS) has increased this interest, as some of the first certifications for SHEMS are likely to come from systems marketed specifically for hotels, dorms, senior living facilities, and apartments. EPA has identified several use cases wherein the existing ENERGY STAR metric could give misleading results or penalize a vendor based on comfort temperatures, runtimes, and data irregularities such as contradictory addresses or weather files. EPA is unsure if these use cases biased results but believes that the corresponding installations were likely filtered out in the software or, if present, were a small fraction of the total population. For use cases of models intended for single family homes (e.g., multiple thermostats in a home, small businesses), vendors are inferring use cases or using user submitted metadata, which may not be reflect reality.

28) Would a criterion for the proportion of the data set that is filtered out help to address this problem?

29) Would stakeholders be comfortable sharing the log files that the software generates that indicate the proportion of the data set that was filtered out?

30) What can we do to confirm that unusual use cases are filtered out or are a small fraction of the population?

5. Additional Metrics for Currently Certified Smart Thermostat Product Types

5.1 Resistance Heat Utilization Criteria

For heat pumps with backup resistance heating, setting back the temperature in winter will result in a higher metric score, but may also result in more energy use due to increased use of backup heat for recovery. Thermostats can use control algorithms that reduce this tendency. To make sure that heat pumps are not achieving high scores at the cost of excess resistance heat use, EPA will propose a performance requirement for resistance heat utilization. EPA and our stakeholders have discussed the specifics of this requirement sufficiently. To ensure statistically significant data, EPA is considering proposing a separate sample of just heat pumps.

31) What would you like EPA to consider that might reduce the added time/effort associated with an additional sample of just heat pumps?

6. Additional Software Changes

6.1 General Improvements

EPA released the alpha version of Version 2.0 software supporting latest version of Python and requested partners to submit data using different versions of the software as a part of biannual data submission process. The Version 2.0 software offers improved Zip Code Tabulation Area (ZCTA) resolution and allows better importing of hourly data. EPA also updated the corresponding documentation to explicitly define the data requirements. Warning messages are tagged with corresponding thermostat ID to simplify tracing down issues, and failed imports are logged in a separate file for easier Troubleshooting.

32) EPA welcomes feedback on these and all other software changes in Version 2.0.

33) Are there additional software changes that stakeholders recommend, to reduce the burden of the program for vendors, increase its usefulness for all stakeholders, etc.?

6.2 Split Output Files

For the ease of certification, EPA has split the output files into a certification file comprising of metric scores for filtered data of national weighted mean and a summary statistics file to evaluate the performance metric and RHU. Only these two files must be submitted by vendors as a part of regular resubmission. The software may also produce other output files such as advanced statistics filtering, a metrics output file (which is passed onto statistics module), and additional error and logging files. These files are not required to be submitted by vendors. As a part of software optimization EPA decided to drop the additional Resistance Heat Utilization

(RHU) results no longer needed such as the old version of RHU along with some of the Duty Cycle information.

6.3 Modified Input File Format

EPA has updated the input file format of both the metadata and interval data files in the Version 2.0 software to account for expanding the scope to include two-stage and modulating equipment. The system types were split into separate heating and cooling types. The metadata file was changed from a single equipment type column to four columns corresponding to heating type, heating stages, cooling type, and cooling stages. Each column uses a set of string enumerations. First stage, second stage, and equivalent runtimes were added. The Version 2.0 interval data file was changed to an hourly time-series format. All columns except the thermostat ID were changed since the hourly columns no longer require the HH postfix. A translation script is provided to ease the transition of older systems to the new format.

6.4 Additional Outputs

As a part of the software update, EPA has also added logging functionality to the importer script. This logging documents thermostats that were present in the metadata file but were not imported. This logging can help aid in troubleshooting why a particular thermostat was rejected. The tool also generates a separate log file with the thermostat ID and the reason why the thermostat was rejected to aid in tracing
any issues with importing the thermostat.

7. Version 2.0 Revision Schedule

Following this discussion guide, EPA expects to release a Draft 1 specification by the end of the year. EPA expects that Draft 2 will come out in early Q2 2022 with the final specification published in fall 2022 and an effective date nine months later.

> 34) Are there any market issues that impact the anticipated timing of this development process that warrant consideration?
>
> Would Energy Star, in light of a lot of stakeholders being currently impacted by this, be pursuing updates to the tau metric calculation, and core heating days eligibility criteria, in the current standard?
>
> If the final specification is published in fall 2022, when is the earliest that can Mysa submit our application for certification?

Please send any written comments and data to connectedthermostats@energystar.gov no later than August 09, 2021. If you have any questions, please feel free to contact Abigail Daken, EPA, at Daken.Abigail@epa.gov and (202) 343-9375 or Abhishek Jathar, ICF, at Abhishek.Jathar@icf.com and
(202) 862–1203.