Google LLC
1600 Amphitheatre Parkway
Mountain View, CA 94043
650 253-0000 main
Google.com

August 9, 2021

From: Michael Blasnik
mblasnik@google.com

Google Nest Comments on "ENERGY STAR® Smart Thermostats: Version 2.0 Specification and Test Method Discussion Guide June 2021"

Thank you for the opportunity to provide feedback on the development of Version 2.0 of the ENERGY STAR Smart Thermostats Specification and Test Method. Our comments follow the section numbers and titles in the discussion guide.

1. Updated Terminology

We support the shift to using the term "smart thermostats" instead of "connected thermostats" and don't foresee any problems being caused by this change.

2. Revisions to Specification Metrics for Currently Certified Products

2.1. Improved Submission-to-Submission Stability of Metrics

We strongly support larger samples to improve the precision and stability of the metric scores.

q2) We see no problems with expanding the sample size. The only drawback might be the long runtimes that would be required by the current software.

q3) We don't believe that sample sizes should be set as a percent of population.

q4) For sample sizes, we would recommend a minimum of at least 30 thermostats with usable results in each climate region to provide statistically defensible metric scores for new/small vendors. We see no statistical reasons to limit the maximum sample size, but believe 2,000 thermostats per climate region may be a reasonable limit due to software speed/capabilities and also to limit any perceived advantage for larger vendors due to the narrowing of confidence intervals with larger samples.

2.2. Improved Correlation between Metrics and Energy Savings through Metered Data

q5) The correlation between metric scores and observed changes in customer energy usage in cross-sectional data may not be very high due to energy use variability and the key role of pre-installation thermostat setting behaviors on realized energy savings. The largest savings

can be expected from customers with the least efficient prior settings and those customers may not have the best scores using the smart thermostat. Assessing the relationship between metric scores and energy savings will likely require data on indoor temperatures prior to installation.

q6) There are definite privacy concerns in joining utility meter data to thermostat data and any project would need to secure participant agreement to any data sharing. The current NEEA project has shown that it can be done, but with significant effort.

q7) Google Nest has shared aggregate utility and/or state specific thermostat metric scores with utility evaluators in the past. But it's unclear whether utility billing analysis studies based on quasi-experimental designs with evidence of substantial self-selection bias should be referred to as "metered savings".

2.3. Weighted Savings

q8) We agree that the current weighting of climate regions by estimated total heating/cooling loads may be too focused on national average savings rather than the savings an average customer might experience. We would support adjusting the weighting to some balance between population-weighted vs. HVAC load weighted. In addition, it may be worth using smaller climate regions based on IECC zones to help reduce variability related to large weather differences within some current regions like Cold/VeryCold.

3. Other Specification Criteria Changes

3.1. Updating Demand Response Requirements

q9) Given the larger role that demand response contractors / aggregators (aka DERMS) are playing in DR programs, Google Nest believes that Open ADR and SEP compliance at the thermostat device / vendor level is not needed if a suitable API is provided to work with DERMS.

q10) Google Nest is not aware of other standards and believes the focus should be on requiring provision of “an interface specification, application programming interface (API) or similar documentation that enables demand response (DR) functionality”

q11) No, because of the ability to work through DERMS, we don't believe that OpenADR or SEP should be part of the prescriptive device requirements

q12) Google Nest has a program called Rush Hour Rewards that enables our utility and energy partners to scale up demand response programs through our DERMS partners. RHR has a powerful and flexible API that allows DERMS to fine tune DR events while the open ADR interface can be provided to utility partners by the DERMS.

3.2. Product Families

Google Nest supports the idea of requiring separate performance metric calculations by thermostat model when there are significant differences in hardware or software functionality. The current Product Family approach has needed to be broadly interpreted to avoid stifling new product innovation -- e.g., allowing products with and without on-device occupancy detection to be classified as the same product family. But after a model has been on the market for 18-24 months, there should be sufficient data for calculating full year metrics to demonstrate that any models with significant differences can qualify on their own.

When reviewing the submitted sample data sets, it may be worth considering differences of practical significance and not solely statistical significance given the typical confidence interval widths of the metric scores.

3.3. Reporting Whether Broadband Connectivity is Needed for Savings

q13) The most important feature that can provide energy savings without a broadband connection is having a good mechanism for creating and maintaining an efficient schedule. Smaller impact features that could provide worthwhile added savings without WiFi include on-device occupancy detection, managing heat pump auxiliary heat (without weather data), and harvesting evaporative cooling through fan over-run under appropriate conditions.

q14) It will be challenging to empirically assess savings for thermostats that aren't online. It may be worthwhile to analyze differences in metric scores between customers who have enabled vs. disabled specific WiFi-based features (e.g., customers who don't use occupancy detection) but any observed differences may be unreliable indicators of feature savings due to self-selection effects.

q15) Claims of off-line thermostat savings functionality would likely need to be based on responses to a set of questions about key features and how the product operates when online vs. offline.

4. Expansion and Clarification of Scope

4.1. Smart Line-Voltage Thermostats

q16) Google Nest does not offer LVTs. But it seems that LVTs could be assessed using the same metric but first aggregating/averaging data across thermostats within each home. Home-level aggregation may also be a desirable improvement for analyzing standard low voltage thermostats. The current approach samples and analyzes thermostats individually, but it may make sense to sample and analyze at the whole home level. Such a change would also reduce the current oversampling of multi-thermostat homes -- if homes are considered the real unit of interest

4.2. Communicating Controllers

q17) Google Nest is potentially interested in continuing to be involved with the development of an evaluation method for communicating controllers for variable speed equipment.

q18) We are not aware of any better venue for developing an evaluation method for these controllers

q19) We agree that it may not be feasible to clearly distinguish the performance of the controller as separate from the HVAC equipment itself for variable capacity systems that rely on proprietary signaling and performance data. It may make sense to simply include the system + controller under the ENERGY STAR CAC/HP category for such systems. But those performance standards should include the impact on efficiency of any limitations in temperature setbacks due to equipment efficiency penalties -- if a variable capacity system can't efficiently recover from a night setback then the overall energy savings is reduced compared to a system that allows for efficient setbacks.

q20) We agree that it would make sense to focus on the contribution of the controller to efficiency. But that goal may not be feasible if the controller employs proprietary signaling based on proprietary system performance data. EPA should promote open communication protocols and publicly available HVAC performance data needed for optimal control so that controllers can compete with each other to prove their efficiency and performance.

q21) Any approach to qualify communicating controllers should consider whether there is any way to define and measure high performance relative to some other controller.

q22) Google Nest believes that some sort of hybrid approach will likely be needed. Field data should be a key component in assessing the real world performance of the controller given the

potential impact of customer interactions as well as variations in part load efficiency due to distribution systems losses. The need for field performance data would require that HVAC equipment manufacturers provide open protocols for collecting the needed data from their systems. The lab testing would be of the HVAC equipment -- part load performance data and optimal setback recovery. The test data should be public so that claims can be verified and other vendors may compete. HVAC manufacturers that don't want to provide this supporting data would not have ENERGY STAR controllers.

q23) CT algorithms and HVAC controls for variable capacity equipment are often at odds. The energy savings provided by CTs are primarily due to employing more efficient temperatures when homes are unoccupied or when occupants are asleep on winter nights. But variable capacity equipment may see significant performance penalties when recovering from more efficient setpoints, negating much or all of the energy savings. HVAC contractors may recommend against setbacks for these systems. If savings from setbacks are reduced or eliminated due to the characteristics of an HVAC system then it may make sense to include this added energy use in assessing overall system performance.

4.3. Mini-Split System Controllers

q24) Google Nest does not offer any mini-split controller. We see the challenges as generally similar to those for variable capacity systems discussed in section 4.2 on Communicating Controllers.

4.4. Evaluate 2-Stage Heating/Cooling Installations

It appears that there may be some typos in this section -- there are two references to product families that should be referencing 2 stage equipment.

Q25) The question "What is the mix of equipment types and geographic spread/zip codes?" is a little unclear. Is EPA requesting the fraction of thermostats in the field that are wired to control 2 stage heating and/or cooling by zip code? What level of geography is actually needed? Also, it's worth noting that many 2 stage systems are wired as 1 stage if installed in a home that had a single stage system. In those cases, staging is handled by the HVAC system, often using a timer for switching to stage 2.

4.5. Filtered Data and Goodness of Fit

q26) Google Nest supports a more detailed exploration of the causes of sample attrition. The software should report causes of attrition for each thermostat. Thermostats that don't

produce any analysis results should be classified as whether it was caused by missing thermostat data, missing weather data, lack of runtime for the season, or something else (e.g., no central AC). We have noticed problems with weather data where NCDC files are either missing or incomplete. It would be very useful if EPA could prepare a single curated weather data table for each resubmission period that avoids the on-the-fly downloads of NCDC data.

In terms of attrition from screening the model outputs, the majority of heating model attrition we see has been due to low values of tau (<=0). These low estimates may be a statistical artifact from the 30 minute runtime requirement leading to truncation bias. Potential solutions include dropping the 30 minute runtime minimum or altering the fitting procedure (e.g., grid search, non-negative regression, Bayesian prior on tau). For cooling, attrition seems more related to noisier data especially when cooling runtime is low . Cooling model attrition has been less than 10% when runtime exceeds 600 hours but is more than 30% when runtime is less than 200 hours.

q27) Google Nest agrees that the credibility of the metric score may suffer if a large fraction of thermostats from a vendor are dropped from the analysis, potentially introducing unknown biases. We support requiring a minimum passing fraction criteria -- preferably some value greater than 50% once methods to reduce attrition are implemented..

4.6. Clarify Specification Applicability for Additional Environments

q28) It's unclear whether having a minimum threshold of thermostats passing the screening would address the issue of the use cases described (hotels, dorms, multifamily). This issue may need further investigation.

q29) The current software summary output file provided for each submission already includes the counts of thermostats filtered out. Is the request to have those summary files shared with EPA including vendor identifiers? Or something else?

q30) Housing type data might be useful to collect and include in the reporting/analysis somehow. But the specific environments listed may not be something vendors are currently tracking and there could be privacy issues. It may make sense to try to find / encourage a study using known instances of these environments to assess the importance of this issue.

5. Additional Metrics for Currently Certified Smart Thermostat Product Types

5.1. RHU Criteria

Q31) Google Nest does not consider the added heat pump sample to be a significant burden. But we think there are some outstanding details that need to be resolved with the new criteria such as treatment of emergency vs auxiliary heat, the specific temperature range included, and incorporating compressor runtime into the analysis as an indicator of problems with the heat pump equipment vs. the thermostat algorithms.

6. Additional Software Changes

6.1. General Improvements

q32) The version 2.0 software seems to include many useful improvements.

q33) Software changes that would be helpful include:

- any changes that improve the execution speed, which will become more important as sample sizes grow
- create a version designed to run in the cloud - which may also help with speed
- provide curated weather data as a downloadable table rather than having the software map and download weather data from NCDC on the fly
- dropping the requirement to use the natsort Python package just for sorting the thermostat identifiers

6.2. Split Output Files

We support dropping the extraneous RHU fields from the output

6.3. Modified Input File Format

We support the changes in input file format -- it makes things easier and faster.

6.4. Additional Outputs

We appreciate the more detailed outputs and logging about why thermostats get rejected from the analysis

7. Version 2.0 Revision Schedule

q34) We don't anticipate any market issues with the proposed timing.