



# EPA ENERGY STAR Connected Thermostats

Stakeholder working meeting  
Connected Thermostat Field Savings Metric  
7/31/2015

# Agenda

---



- Introduction – anyone new joining the call?
- Administrative updates and issues
- Software module alpha release
  - Metric doc corresponding to module calcs released 7/21
  - Any new results?
- Metric topics related to stakeholder comments
- Starting discussion of DR and of labeling

# Software Module alpha release

- Source code: <https://github.com/impactlab/thermostat>
  - Look for 2<sup>nd</sup> release: 0.1.1
- Documentation:  
<http://thermostat.readthedocs.org/en/latest/index.html>
- Several stakeholders have tried using the modules, and OEE are now tracking down several errors in the input file
- Submissions of errors on GitHub has been very useful – please keep them coming
- Standard iPhython notebook to run modules – any issues?
- Example input files could not be used as inputs to the modules – more to illustrate with usable samples
- No test data means we can't write data tests – would still be very useful

# Software Module process

- Question one: Do the modules faithfully reflect the algorithm we intended? Enter directly into GitHub. Two weeks or so?
  - Bugs
  - Cases where code returns unexpected results
- Question two: Does the algorithm (and the code) measure what it is intended to?
  - Does the choice of algorithm matter for any homes?
  - Homes with different setback behavior – different scores?
  - Homes with similar setback behavior, different comfort preferences – similar scores?
  - Regional bias, Demographic bias?
- Comments on performance can be sent to [ConnectedThermostats@energystar.gov](mailto:ConnectedThermostats@energystar.gov) at any time, or call

# Software Module process, cont.

---

- Eventually we will need to decide if the algorithm (and code) adequately capture savings
- Seeing how you and your competitors land with the current algorithm might help with that
- There is also additional work that might need to be done - stakeholders could volunteer to add code to GitHub
  - Expand code to calculate statistics for large data sets
  - Additional tests of variations on the algorithm

# Software module release discussion

---



- Renewed call for test data
- Ecofactor thinks they can provide sample data – not actual customer, but realistic
- Nest also thinks they can do so by adding random noise to actual data
- The more realistic the data is, the better, and in the right input format – does not have to be from a real home

# Metric doc

---



- Do we really want to do run time modeling?
- Some homes do not have data that fits the model very well – can we average between homes before we do a fit
- EPA and stakeholder to have one on one discussion, report out at next stakeholder meeting.

# Metric topics from stakeholder comments



- Addressing comfort – can we leave that to providers?
- Self baseline vs. regional baseline
  - Regional baselines will capture savings from both setbacks and from more EE comfort settings
  - Self baseline obviates need to for high static temp accuracy
  - How to develop regional baselines? Use CT data?
  - Per ZIP code baselines may be necessary
- Third party verification – related to sampling as well
  - Did not end up talking about on this call
- Certification of newly introduced product
- NEMA DC-3 tests
  - Will address with one on one calls



# Comfort



- Checking – no plan to have a check that the endorsement won't go with algorithms that will make people very uncomfortable? Yes, that is EPA's intention
- How would we include comfort if we needed to?
  - Savings vs average setpoint (as a way to classify users' willingness to be uncomfortable)
  - There are some basic things we know about comfort: ASHRAE psychometric comfort zone; temp swings are a problem
  - Identify frequency of discomfort type events?
  - Some disagreement – CTs whose users implement more energy efficient comfort settings would be penalized
  - Comfort is subjective

# Comfort



- Mistaking whether customer is uncomfortable b/c of stat vs something wrong w/ HVAC equipment? Hard to know whether feedback will get back – might take a while to shake out
- Could we look at RMA rates? Possibly, but not necessarily based on discomfort
- If people are uncomfortable, they'll put it in constant hold mode - do all smart stats have constant hold modes.
  - No, and ENERGY STAR should not mandate existence of HOLD as it leads to wasted energy
- Can we count the # of homes where stats have gone off line? Is it possible that would be biased against some users
- Look at # of user interactions in energy using directions?
- Require ability to turn off energy saving features?

# Comfort discussion continued

- Two concerns with comfort
  - Set of measures to quantify comfort
  - Make sure that choices that sacrifice energy conservation for comfort don't negatively affect ES rating of stat – don't skew results. Disagreement on this topic:
    - Acknowledgement that stat has a role in helping consumers balance comfort and efficiency, though
    - Isn't this the bar we are trying to set and the improvement we are trying to capture
- But comfort of stats *\*is\** visible, unlike energy use

# Baselining discussion – self vs. regional



- Many varying factors:
  - demographics – who buys the thermostat
  - HVAC system type impacts comfort temps (e.g. furnace vs heat pump)
- Some empirical data to see how much these are issues
- Look at demographic data within a single zip code to see if other demographic differences are predictive
- Why are we concerned about noise?
  - Even temp accuracy is just random variation
  - Do demographic effects produce systematic differences?
  - Seniors and low income individuals prefer less energy saving (warmer heating) set points
- Is this noise?
- Didn't we move to a relative metric because there were systematic differences between vendors?

# Baselining discussion – self vs. regional



- Stats can do a lot of things to help people be more comfortable at higher temps in summer (e.g. humidity control) and lower in winter and it would be good to reward
- Other factors (could we do multivariable regression to develop ZIP code-specific baselines)
  - Fuel type
  - System type (can we segment on this)
  - Income
  - Age of occupants
  - Income
  - Housing type
  - Area of home
  - Rate structure
- Preliminary analysis shows significant effects

# Baselining discussion

- Is bias from ignoring more efficient set points at all times larger than bias from these other putative factors?
  - Maybe not, because tstat market is driven by utility programs
  - In addition, likely product differences are likely to produce correlation between stat models and demographic factors
  - No real data in hand to determine this
- ZIP-code specific constant baselines are also potentially more useful to energy efficiency program sponsors – but not perhaps if we use CT data only
- Regional baselines with CT data may also be tautological in the sense that if certain products are very prevalent in some zip codes, their capabilities will pull the average for that zip code.

# Baselining discussion

---

- If we have to do a pass/fail, will we still be able to set a requirement on the metric?
  - Yes, we should still be able to set a level
  - Have been successful with this before

# Certification of newly introduced products discussion



- How large a group are beta testers usually?
  - Can vary dramatically depending on resources
  - Hundreds maybe, but can imagine less
- Also, who does the testing – is it vendor's own staff
  - Request metrics run from past beta tests from the manufacturer, so we can see how accurate past beta testing is?
  - But different products may have different beta test groups?
- What is “new product”? If it's just hardware running the same algorithm, perhaps it can be qualified using old data? Perhaps three broad categories:
  - Entirely new CT and service
  - New CT works with existing service
  - New (or substantially changes service) with existing CT



# Certification of newly introduced products discussion



- What about major software revision? Is beta testing easier and more likely to be random sample?
  - Not likely to be a random sample. Production beta, maybe, if done.
  - Would testing even be necessary in light of semi-annual reporting?
- Can also do A/B testing for software updates
- What if we just update one part, like phone app?
- Narrative about software revision – can it be detailed enough to determine whether metric results should carry over, but not reveal proprietary info?
- How do we define new product

# Certification of new products discussion



- 3 cases
  - New product from existing energy star partner
  - Updating product that is already certified
    - Grandfathering – if data validates after 6 months, keep label
  - New product from vendor that is not current E\* partner (this is the hardest case)
- There will be some changes that do not require re-certification, particularly because of periodic resubmission of savings

# Beginning Stakeholder Discussions on Labeling and Demand Response Requirements



- Labeling:
  - Talk individually to stakeholders and develop updated proposal
  - Then invite stakeholders to a call specific to labeling requirements
  - Continue with as many subsequent calls as necessary
- Demand Response:
  - EPA will come up with proposal
  - Invite stakeholders to DR specific call series
  - Continue as necessary; if stakeholder consensus proposal is presented to EPA, will consider strongly

# Running parking lot

- Zoned systems? Usually not integrated. Multiple systems in one home? Ask for statistics about how common this is.
- Definition of a “product” – e.g. enrollment in peak control service makes it a different product
- Verification and gaming the system?
- Does the customer base bias the metric results, aside from the qualities of the products?
- Add on today’s parking lot items...

# Contact Information



Web site for these notes and all public discussion/comments:

[http://www.energystar.gov/products/spec/connected\\_thermostats\\_specification\\_v1\\_0\\_pd](http://www.energystar.gov/products/spec/connected_thermostats_specification_v1_0_pd)

Abigail Daken  
EPA ENERGY STAR Program  
202-343-9375  
[daken.abigail@epa.gov](mailto:daken.abigail@epa.gov)

Doug Frazee  
ICF International  
443-333-9267  
[dfrazee@icfi.com](mailto:dfrazee@icfi.com)

# Metric topics from stakeholder comments



- Set threshold for minimum # of homes meeting requirement, so that a few homes with very high savings do not skew results.
- Interested in regional breakdown of scores
- Droop corrected in software, not necessary to measure; at minimum waive for low voltage thermostats
- Static temp accuracy tolerance too tight; display only shows 0.5F increments [can use temp reported on WiFi?]
- Regional comfort baselines
- Third party verification of no gaming
- Does not address comfort
- Certification of newly introduced product – data from beta release useful?