ENERGY STAR Connected Thermostats Data Request

Stakeholder Webinar and Discussion

Abigail Daken, U.S. EPA

December 4, 2015

Agenda

- Welcome & Introduction
  - How data helps EPA
  - Questions we hope to answer
  - Advantages of participation
- Mechanics of the data request
- Specific proposed data elements
- Possible additional data
- Discussion continues

Attendees

Abigail Daken, EPA

Doug Frazee, ICF International, for EPA

Jennifer Kulp, ICF International, for EPA

Alan Meier, Lawrence Berkeley National Laboratories

Ethan Goldman, VEIC

Nick Lange, VEIC

Michael Blasnik, Nest Labs

Dave Cassano, Nest Labs

Raj Shah, Carrier

Phil Ngo, Impact Labs

Brent Huchuck, Ecobee

Michael Siemann, Weatherbug Home

Wendell Miyaji, Comverge

Laurie Sobczak, Comverge

Alex Bosenberg, NEMA

Matt Golden, Open EE

Ed Pike, Energy Solutions, for CA IOUs

Ford Garberson, Ecofactor

Ram Soma, Ecofactor

Chris Smith, IRCO (Trane)

Kurt Mease, Lux Products

ENERGY STAR

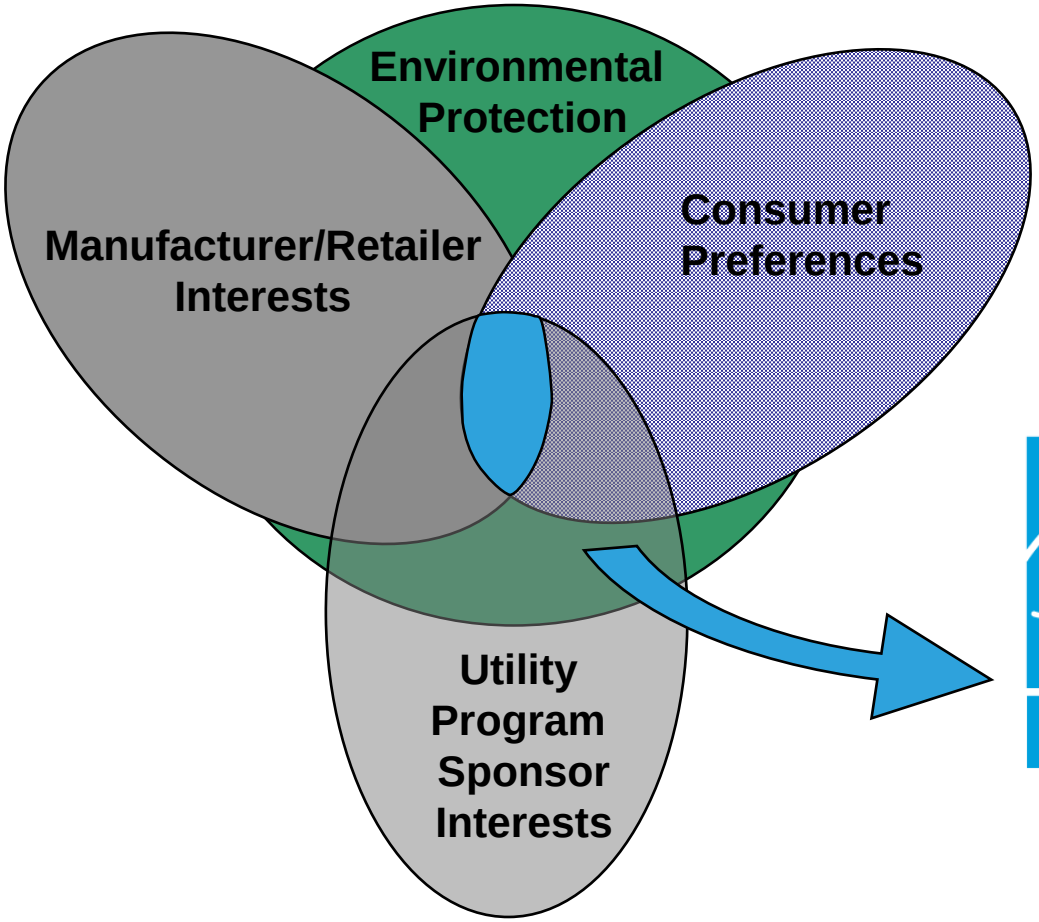
For more than 20 years, EPA's ENERGY STAR program has identified

the most energy efficient products, buildings, plants, and new homes –

all based on the latest government-backed standards and a rigorous third-party certification process.


Every single day,
consumers choose
ENERGY STAR
products more than
800,000 times

ENERGY STAR's Focus



- Cost-effective
- No Sacrifice in Performance
- Government backed

Consumer is Key

Specification Development Cycle

Important Process Elements

- Consistency
- Transparency
- Inclusiveness
- Responsiveness
- Clarity

ENERGY STAR CTs – Why a new approach?

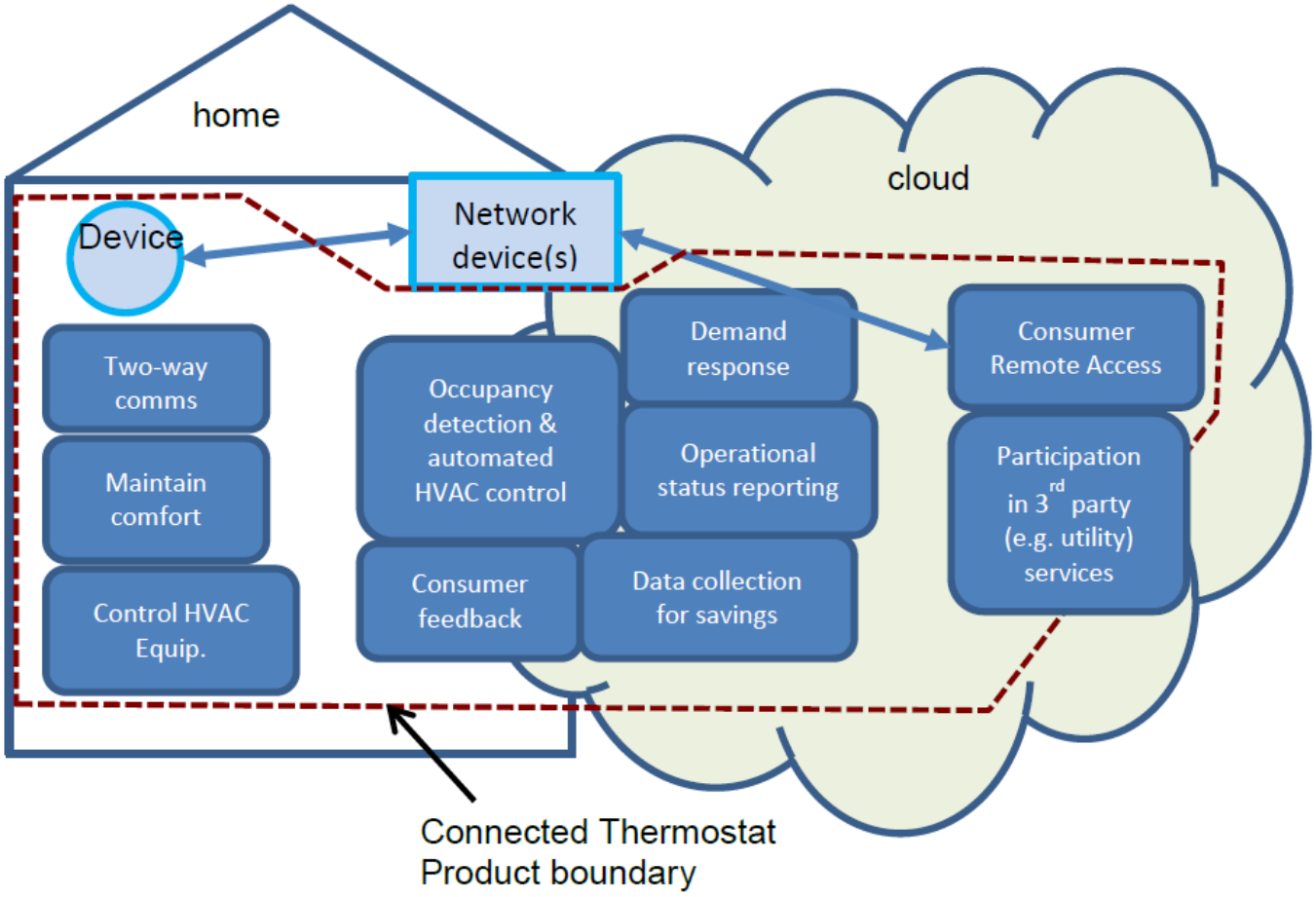
- EPA recognized that CTs were breaking new ground with many entities claiming significant energy savings.
- No standard methodology for calculating savings
- Varied strategies for generating savings
  - Behavioral
  - Occupancy sensing
  - Thermal modeling
  - Automation
  - Integration with other connected devices
  - Weather optimization
- Common denominator was *not ease of use, not consumer engagement, or a default setback schedule* – rather it was the energy savings itself!

ENERGY STAR CTs – What new approach

- In the emerging Internet of Things, EPA recognized that CT savings could be modeled using only:
  - publically available weather data, and
  - data reported by the CT itself
- In effect, CT products are able to self-report energy savings

What is a Connected Thermostat Product?

A blend of local hardware and cloud services



ENERGY STAR CTs - EPA Goals

- Significant realized & verified energy savings
- Provide labeling opportunities for both CT manufacturers and service providers
- Provide utilities with a tool to meet Energy Efficiency goals.
- Ensure the methodology limits stakeholder burden while assuring consumers of minimum amount of savings
- Recognize continuous improvement
- Robust participation by:
  - CT manufacturers
  - Service Providers
  - EEPS
  - Utilities
- Prominence of ENERGY STAR CTs in the marketplace

Program Outline

- Recognition for CTs that save energy in the field
- To earn the ENERGY STAR:
  - "CT device" must meet criteria that enables savings
  - Partner must periodically report aggregate consumer savings for each CT product
  - "CT product" includes the CT device and a service component
- Service Provider is the ENERGY STAR partner

Metric for periodic reporting

- Uses only CT data plus outdoor temperature history
- Preserves consumer privacy
- Protects proprietary information
- Practical to calculate
- Method evaluates HVAC system run time reduction relative to baseline run time
  - Step 1 – model the home's relationship between HVAC run time and outside temperature
  - Step 2 – extract heating and cooling comfort baseline temps from the home's CT data
  - Step 3 – calculate the home's baseline run time
  - Step 4 – metric is % run time reduction
  - Step 5 – average over a large number of homes

How data will help EPA make better policy

- Intending to recognize a level of performance that some products on the market are able to achieve
- Get a first sense of how the metric scores reflect savings
- Other information submitted along with metric scores will help EPA understand whether the metric is measuring what we intend to measure (e.g. number of core heating and core cooling days)

Questions we hope to answer

- National data
  - Do all products score similarly on the metric?
  - What is the shape of the distribution of savings? Flat? Bi-modal? Guassian? What kind of tail?
- Regional data
  - Do scores depend more on region than on the product?
  - If so, consider modifying the metric calculation to more fairly reflect product features (e.g. same weighting of regions for all products)

Advantages of participation

- Advanced idea of how your product(s) will score on the metric
  - Find out if metric scores fail to properly reward your product
  - If there are problems it will be much easier to address them now
- Make sure your data systems are working to produce metric scores
  - Work out the kinks in translating your data into an input file for the metric modules
  - Be ready to certify product(s) as soon as the specification is finalized
- A better-informed specification is an advantage to everyone

Mechanics of the data request

- Intended to be as simple as possible – based on module output data file
- Output data file has at least 10 rows; each is a zip code or group of zip codes
  - National summary is the first (and second) row (all zip codes, one row each for heating and cooling.)
  - BA climate zones are the next (up to) 14 rows (heating and cooling in each climate zone) Groupings by weather stations are the next rows
  - Remaining rows are summaries for individual zip codes.
- Each column is a data element, e.g. mean heating savings score, mean days in the core heating season, etc.
- If you prefer, edit the file to include only the rows and columns we ask for before sending to ICF

These are the only rows we need

Output file columns – mean, standard error and decile bins for each

number of days with both heating and cooling	baseline comfort temperature
number of days with insufficient data	actual daily runtime
number of days in the season	actual seasonal run time
number of days between the first and last day of the season	seasonal avoided run time (ΔT method)
ΔT method slope	seasonal avoided run time (daily average CDD or HDD)
ΔT method intercept	
αh (daily avg CDD)	seasonal avoided run time (hourly average CDD or HDD)
αc (daily avg HDD)	seasonal savings (ΔT method)
αh (hourly avg CDD)	seasonal savings (daily average CDD or HDD)
αc (hourly avg HDD)	seasonal savings (hourly average CDD or HDD)
mean square error (daily avg CDD or HDD)	RHU 0F to 5F
mean square error (hourly avg CDD or HDD)	RHU 5F to 10F
mean square error (ΔT method slope)	RHU 10F to 15F
ΔTbase (daily average CDD or HDD)	RHU 15F to 20F
ΔTbase (hourly average CDD or HDD)	RHU2 0F to 25F
baseline daily run time (ΔT method)	RHU 025F to 30F
baseline daily run time (daily average CDD or HDD)	RHU 30F to 35F
baseline daily run time (hourly average CDD or HDD)	RHU 35F to 40F
baseline seasonal run time (ΔT method)	RHU 45F to 50F
baseline seasonal run time (daily average CDD or HDD)	RHU 50F to 55F
baseline seasonal run time (hourly average CDD or HDD)	RHU 55F to 60F

Provide only mean and standard error

Provide mean, standard error and decile bins

Possible additional data

- Also included in the output file
  - Number of thermostat-seasons is, e.g. 10 thermostats each with 3 years of data would provide three seasons each for 30 thermostat-seasons each
  - Number of thermostats that would have been in each average, but were rejected for poor fit, missing data, etc.
- Additional data that could be useful to us:
  - How did you translate your interval data into daily and hourly averages? (Uniform method would be ideal.)

Data privacy

- All data will be considered business confidential
  - ICF is able to sign NDA's
- As is typical for EPA data requests, data will be shared with EPA and others only:
  - If at least 3 data sets are received, and
  - Anonymized

Discussion

- Some climate zones may be an issue: just skip those
- Goodness of fit
  - Any goodness of fit measure we include will be largely arbitrary
  - In some climates more homes will drop out
  - Run with several different goodness of fit measures? How many homes drop out for each level

Discussion

- Proposed measures for goodness of fit:
  - CVRMSE: coefficient of variation of root mean square error = RMSE/mean
  - standard error of the slope divided by the slope
  - Could code a couple different measures: CVRMSE, MAPE, and another even simpler method?
  - MAPE = mean absolute percentage error, can have trouble when numbers get small
- Prefer to have a measure that will preferentially eliminate datasets with non-linear or otherwise consistently problematic data, rather than those with a few outliers
  - Do we really want to keep these in the sample, given that the results will still be affected by the outliers?
- Does it work to use an absolute error in run time, for instance?
  - Number is constrained, because HVAC system can only run for 24 hours a day
- Eliminate negative slopes for linear fit method

Discussion

- Ratio estimator for heating, even including outliers, includes from .25% to 10% savings per degree of thermostat setting – well behaved
- Results are MUCH wider for linear fit, including negative slopes, and those up to 61%
- Proposal: keep data request as is, with no goodness of fit requirements; also code up a few measures and provide them for stakeholder to play with
- Simpler approach may be to just have a requirement for a particular certainty and not worry about throwing out outliers unless stakeholders are having trouble meeting the certainty without throwing out thermostats
- Perhaps we should focus on heating where it matters and cooling where it matters
  - Hoping to get a sense of this from the regional data in this data call
  - Are we asking for the right data to address this question?

Discussion

- Requirement of at least 250 thermostats per climate region (from data request) in a sample with geographic distribution of users (from Method for Demonstrating Savings, Version 1 Draft 1) brought up a discussion of weighting.
- Why would we weight by population of users? Significant discussion led to the conclusion that no one supports doing so, with the caveat that we do want an on-ramp for products with smaller user population which may also be more regionally distributed
- Suggestions for weighting instead
  - % of housing stock
  - Average heating and cooling energy use per home (from EIA)
    - May give outsize weight to small populations of homes in extreme climates
  - % of national heating and cooling energy use in the region (EIA)
- Result of discussion: do not need to know number of thermostat-seasons in each climate region, nor proportion of deployment by region. Do not need national sample for data request to reflect distribution of users.

Discussion – do we have all the tools ready?

- Software: yes, excepting goodness of fit, which we are not going to use for the data request
- Zip code mapping: Need to do a little work –
  - anyone other than Nest have zip codes coming out as NA climate zone that you think should have a region? No.
  - Nest and Impact Labs will work together to address those.
- EPA and ICF note that we should have plenty of support available for getting software implemented and running smoothly. At this point, if you have an issue, email Phil Ngo and cc Doug Frazee

Additional question from Alan Meier, re national savings

- Percentage of heat pump vs. fossil fuel heat? Do you know or are you guessing?
  - Nest: as long as the wiring includes an O/B wire, we assume it's a heat pump
  - Ecofactor: same, small but non-zero fraction of heat pumps
  - Ecobee: same
  - Lux: some additional information from regional trends
- For thermostats w/o O/B wire, can we make any assumptions about what the fuel source is?
  - Nest: we ask, but only trust answers of those who enter oil or propane
  - Ecobee: ask, cannot assume. Ask is not during installation process, not likely to be a contractor answering
  - Ecofactor: some by contractor, some by self-report. Very few people say electric as the fuel, so maybe reflects reality

Contact Information

Abigail Daken
EPA ENERGY STAR Program
202-343-9375
daken.abigail@epa.gov

Doug Frazee
ICF International
443-333-9267
dfrazee@icfi.com