



# **ENERGY STAR Connected Thermostats**

## **CT Metrics Stakeholder Meeting Slides**

March 7, 2022



## Attendees

Abigail Daken, EPA

Abhishek Jathar, ICF for EPA

Alan Meier, LBNL

Leo Rainer, LBNL

Eric Floehr, Intellovations

Craig Maloney, Intellovations

Michael Blasnik, Google/Nest

Kevin Trinh, Ecobee

Michael Sinclair, Ecobee

Yufeng Deng, Ecobee

Brad Powell, Carrier

Matt Gobson, Carrier

Jason Thomas, Carrier

Theresa Gillette, JCI

Rohit Udavant, JCI

Diane Jakobs, Rheem

Chris Puranen, Rheem

Glen Okita, EcoFactor

Thomas Lorenz, Emerson

James Jackson, Emerson

Daniel Stephan, Emerson

Mike Lubliner, Wash State U

Charles Kim, SCE

Michael Fournier, Hydro Quebec

Robert Weber, BPA

Phillip Kelsven, BPA

Casey Klock, AprilAire

Wade Ferkey, AprilAire

Kristin Heinemeier, Frontier Energy

Ulysses Grundler, Trane

John Hughes, Trane

Mike Caneja, Bosch

Sarathy Palaykar, Bosch

Mike Clapper, UL

Alex Boesenberg, NEMA

Ethan Goldman

Jon Koliner, Apex Analytics

Hassan Shaban, Apex Analytics

Michael Siemann, Resideo

Arnie Meyer, Resideo

Jia Tao, Daikin

Dan Baldewicz, Energy Solutions for CA IOUs

Claire Miziolek, Energy Solutions for CA IOUs

Cassidee Kido, Energy Solutions for CA IOUs

Dave Winningham, Lennox

Dan Poplawski, Braeburn

Natasha Reid, Mysa

Peter Gifford, Mysa

Aidan Girard, Mysa

Riana Johnson, Illume Advising

Sylvain Mayer, Sinope Solutions



## Agenda

- Software updates
- February 2022 data submission analysis
- Missing data updates and proposed solutions
- Tau friendly regression



## Software Updates: V1.7

- Version 1.7.x
  - Realized that we had narrowed the versions of Python to 3.6 (EOL) and 3.7 (old).
  - Updated the 1.7 develop branch to allow for more modern (3.8, 3.9, and 3.10) versions of Python.
  - Will create a new release branch in the coming weeks.



## Software Updates: V2.0

- Version 2.x
  - Modified to use ZIP Codes instead of ZCTA
  - Created a lookup dictionary of ZIP Code to weather station / climate zone
    - ([https://github.com/EPAENERGYSTAR/epathermostat/blob/feature/epathermostat\\_2.0/thermostat/zipcode\\_lookup.py](https://github.com/EPAENERGYSTAR/epathermostat/blob/feature/epathermostat_2.0/thermostat/zipcode_lookup.py))
  - This file will be regenerated when a new release is generated and should allow us to be more consistent with each release on what ZIP Codes map (or don't map) to a weather station / climate zone.
  - Climate zone data is coming from eeweather. (Previously it was coming from a circa 2016 CSV file of indeterminate origin).



## Software Updates: V2.0

- We're looking into better ways to surface why and how many thermostats are "going missing" (not created).
- How we surface that data leads to the following question:
- **Question:** When running the software which of the following do you use?
  - The script under `scripts/multi_thermostat_tutorial.py`
  - A script based on the `scripts/multi_thermostat_tutorial.py`
  - Something else?
- (Why does this matter? Trying to figure out how to get information on why thermostats that don't get processed (from the loading in phase) to be included in the stats file at the end. )



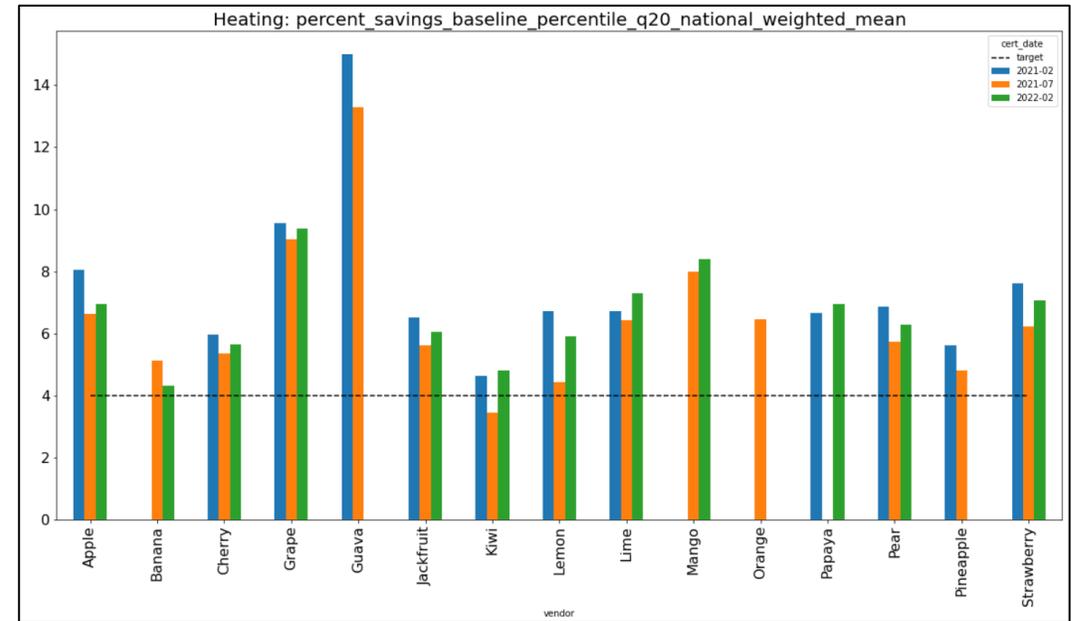
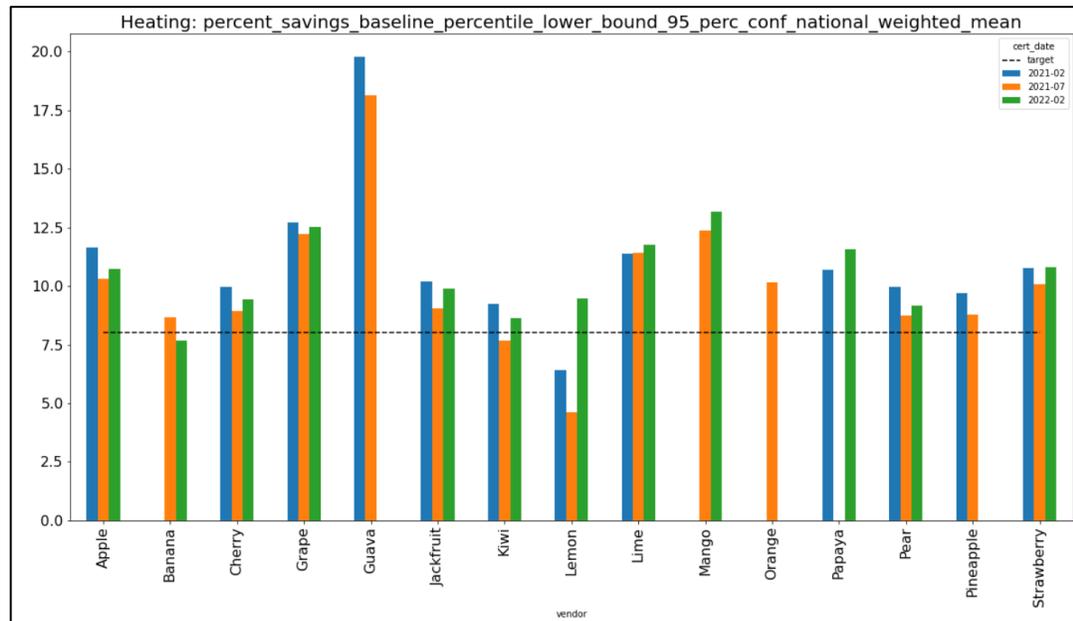
## Discussion: Software Updates: V2.0



## February 2022 data submission analysis

- Resubmission Statistics:
  - 12 datasets received for V 1.7.2/1.7.3 single speed equipment and expecting 1 more soon
  - 1 datasets received using V 2.0 for same sample as V 1.7.2/ 1.7.3
    - Some variances in savings and RHU metrics compared with the v1.7.3 scores particularly for heating. Sample undergoing further investigation from the vendor.
  - 2 vendors were unable to submit data
  - No oversampled sets for RHU2

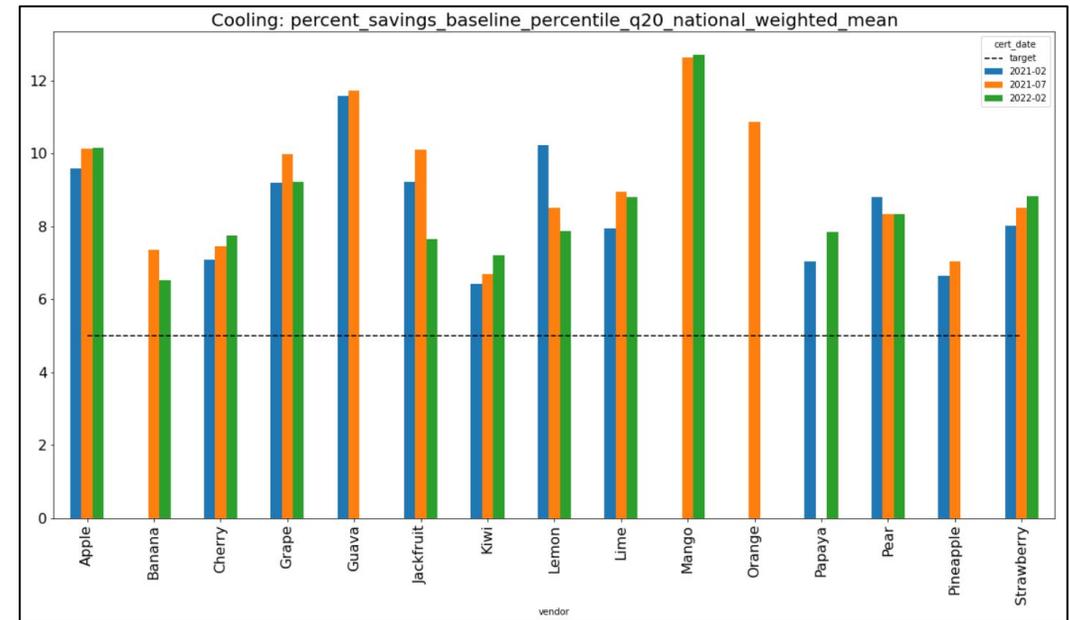
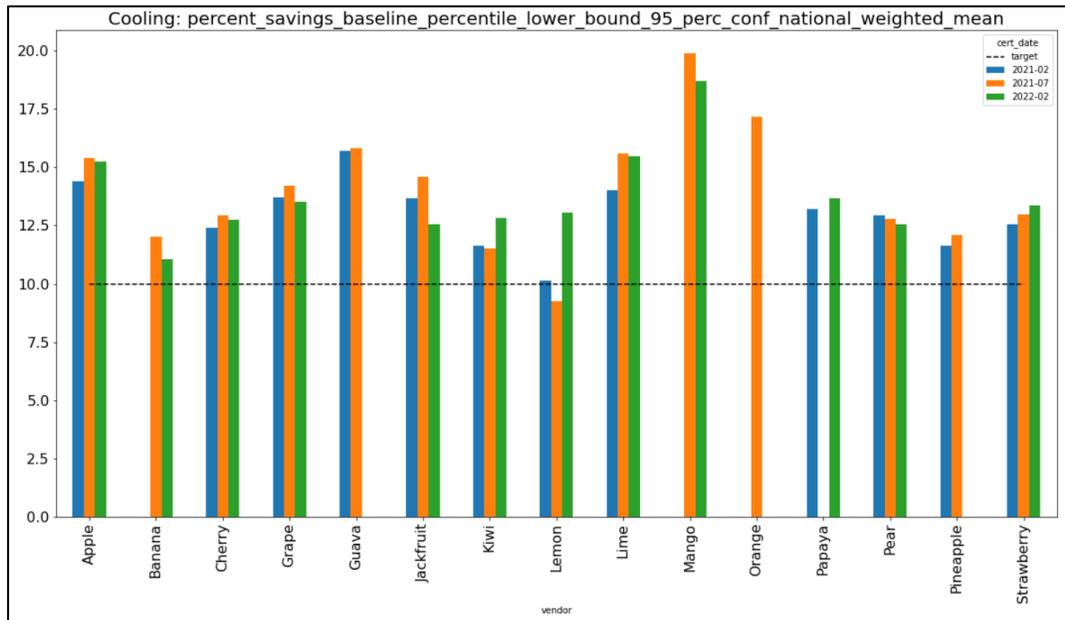
## February Submission: Heating Metric



- Included only vendor data since Feb 2021 (3 submissions).
- No particular data trends.
- Approximately half vendors witnessed a drop in their metric scores
- One vendor missed the metric mark for lower bound 95% confidence interval



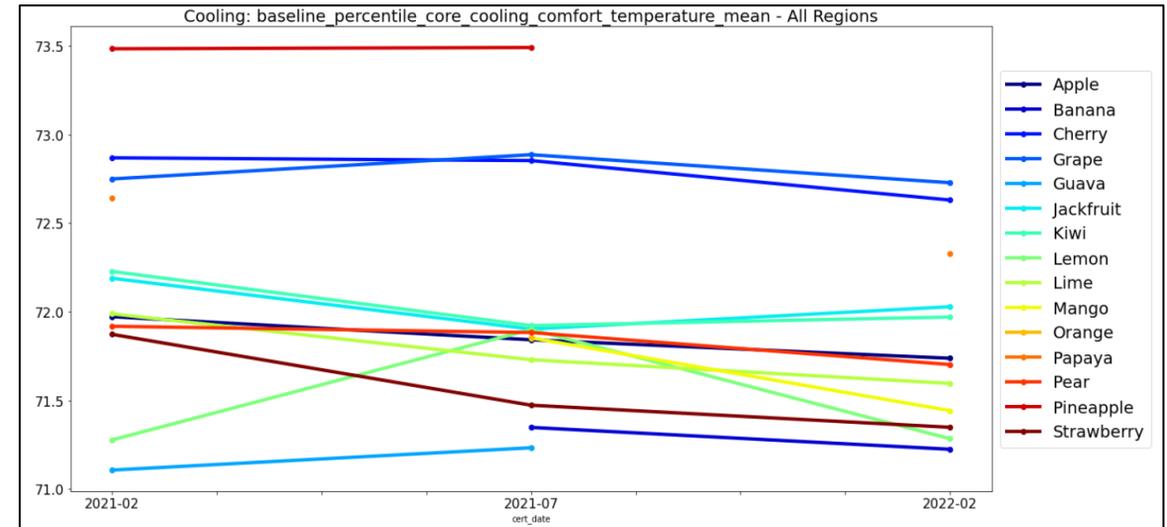
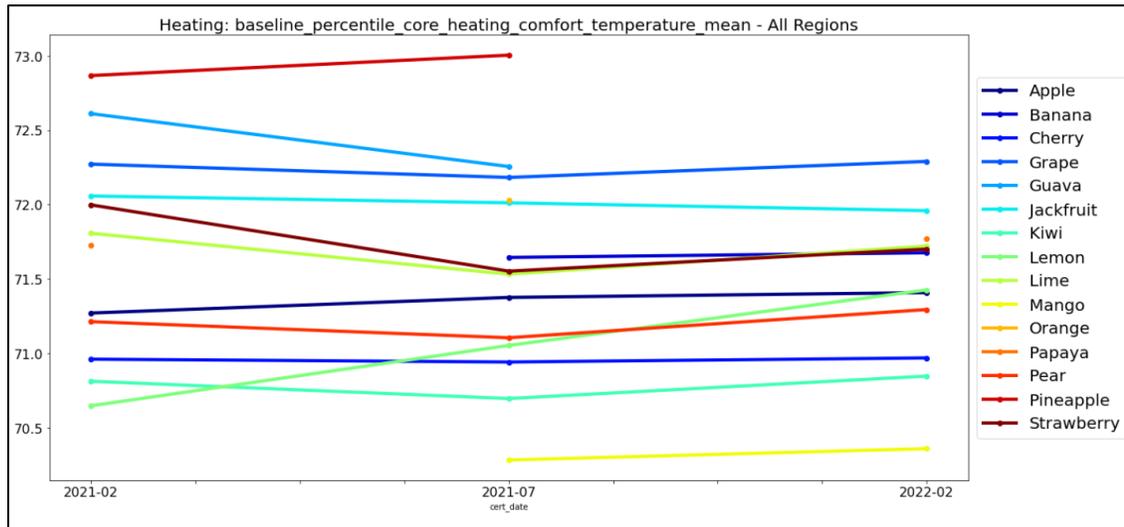
## February Submission: Cooling Metric



- Included only vendor data since Feb 2021 (3 submissions).
- Cooling metric scores improved for majority of the vendors compared to previous year



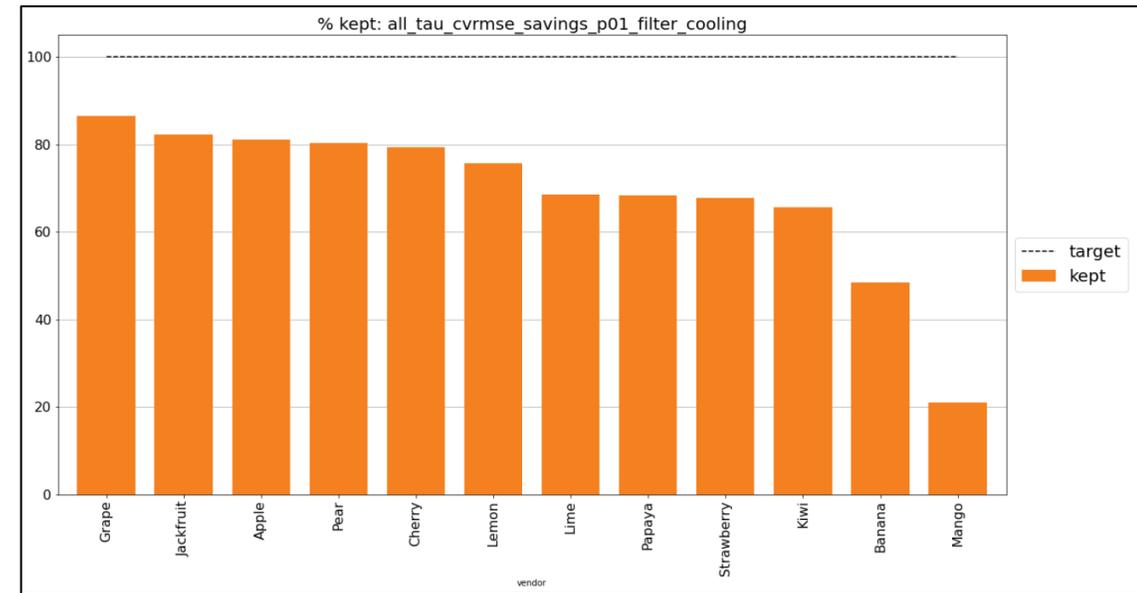
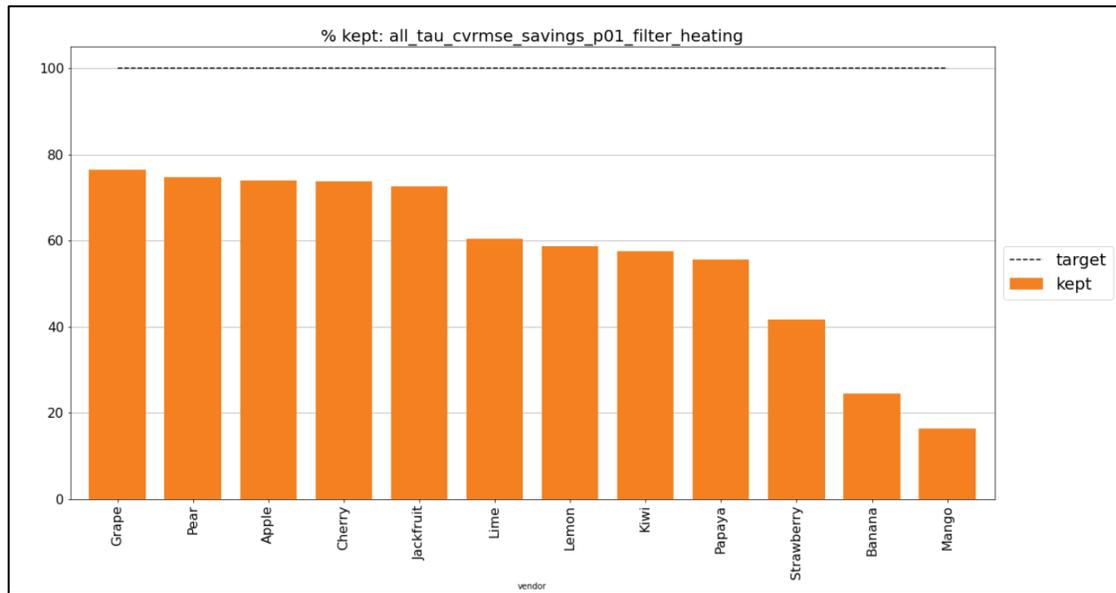
## February Submission: Comfort Temperatures



- No significant change in heating comfort temperatures. Marginally up for most cases
- Cooling comfort temperatures showed marginally downward trend compared to last year



## February Submission: Thermostats filtered



- Results for February 2022 submission only.
- 3 vendors have 50% or more thermostats discarded for heating on filtering.
- 2 vendors have 50% or more thermostats discarded for cooling on filtering.



## Discussion: February 2022 data submission

- What should our takeaway be here?
  - A: The metric seems to be stable and consistent.



## Missing data updates and proposed solutions

Current status:

Still looking for a solution to ensure valid models while minimizing number of excluded thermostats.

Trying to set thresholds for:

- Max number of missing total days
- Min number of present core days (in both seasons)

Need to pass at least one of the tests

Looking for more data to empirically determine threshold values.



## Missing data updates and proposed solutions

### Quantitative data request:

Can vendors provide an anonymized copy of the “metrics” file from the most recent certification submission (Feb 2022)? We are particularly interested in the following fields:

- Climate zone
- Number of total days
- Number of days with insufficient data
- Number of heating core days
- Number of cooling core days



## Missing data updates and proposed solutions

### Qualitative data request:

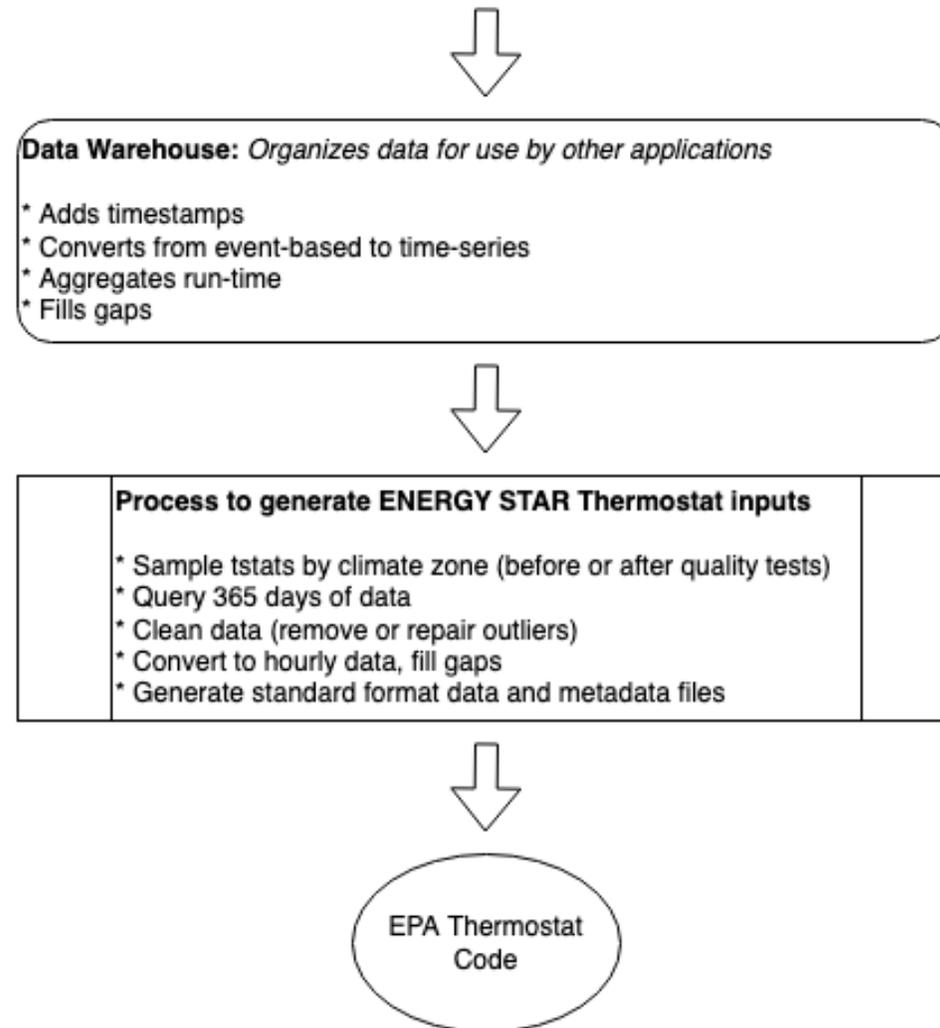
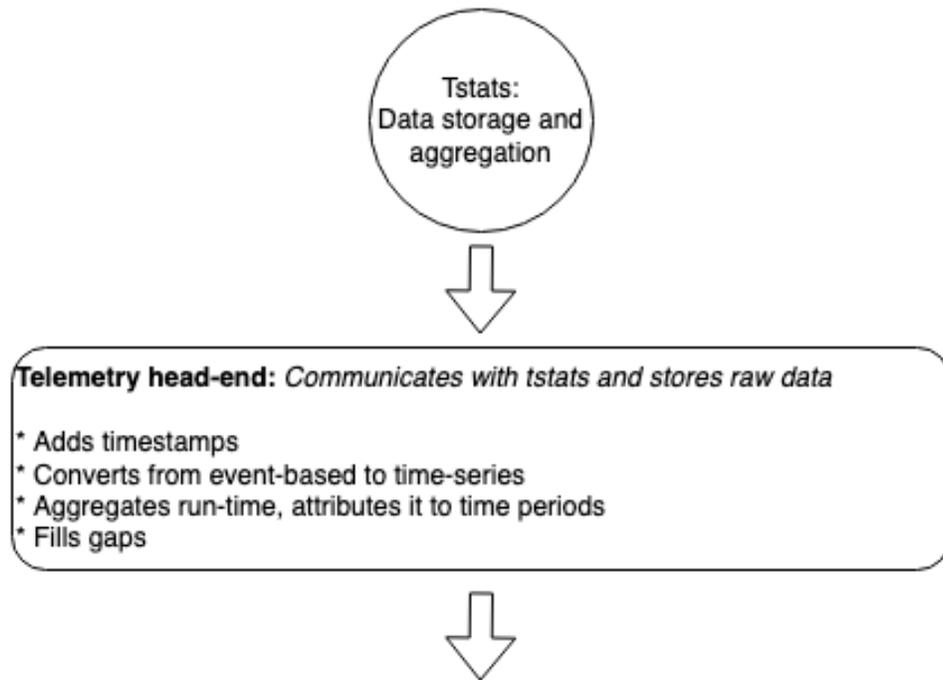
Please describe how missing and outlier data are handled by data-cleaning, interpolation, and data-quality flagging at different stages in the process between your smart thermostat hardware and the EPA Thermostat software.

Please describe how each of your systems/processes:

- Identifies missing or outlier values, particularly indoor temperature and run-time for each heating/cooling control point.
- Flags those missing/outlier values, either using a separate time-series of data quality indicators or by substituting a value of 0, -999, NaN, etc.
- Determines what values should be removed due to data quality concerns.
- Interpolates or otherwise fills gaps caused by missing or poor-quality data.
- Converts event-based data to time-series data, including where data quality tests can trigger different behavior.



# Missing data updates and proposed solutions





## Discussion: Missing data updates and proposed solutions

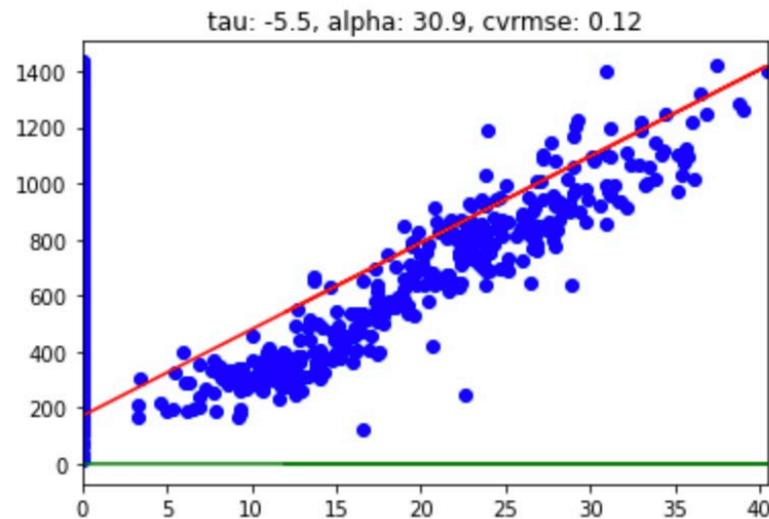
- For the quantitative request, can we just send those columns you've specified? Because those are easier to get approved to share, and would speed up getting data to EPA's team.
  - Yes, probably – we'll go back and make sure we've named everything we need and use the actual column names to eliminate ambiguity.
- From Ethan: what do you think, will this be straightforward or complex?
- Better to get in steps or should we wait until we have everything?
  - Definitely send the quantitative first if you can, which we expect to be more straightforward. For the qualitative, if you find a step where a whole lot happens, please send that information ahead.
  - Could we aim for a month? Two vendors feel they can do this within a month., others are silent.

## Tau friendly regression

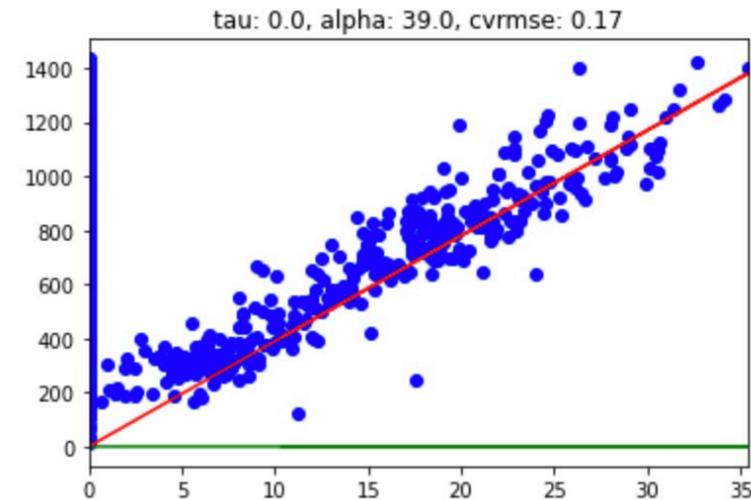
Status: Bugs banished, testing results for improvement to Tau metric (goal is fewer thermostats disqualified for negative Tau without impacting model fit).

Request: More anonymized sample data for testing.

OLD tau search



NEW tau search





## Discussion: Tau friendly regression

- Did EPA consider “releasing” a version of the software that would do this and offering for vendors to run it themselves?
  - We have a branch that does this, which we’d be happy to point you to. However, we aren’t sure that we know what we need to look at in the results. In addition, we expect to be doing a significant amount of re-writing and re-running in this process.
  - We can package up the comparison of the regression models and if vendors are able to contribute the significant work involved in iteration, we might be able to make that work.
- Can keep the anonymized datasets outside FOIA jurisdiction.



## Discussion: The Floor Is Open

- Mike Lubliner: in Washington State U supporting the residential state energy code. Code has a certain number of prescriptive points to meet code based on specific prescriptive measures. Now looking at the possibility of finding new energy credits (about 1200 kWh/year) – what about providing credit (at least half a credit) for an E\* connected thermostat? Need at least 600 kWh/year. Gas furnaces, single speed heat pumps w/ER and gas backup, maybe VRF heat pumps and cCHP with ER backup. WSU will be vetting a proposal with NEEA and other stakeholders – feedback welcome. Probably in effect 2023. Proposal deadline 4/15.
  - May be opportunities to break out by climate zone, but not right now. Average over the whole state (weighted by stock and new builds) for climate-dependent.
  - Can credits vary by the type of equipment in the home? Goes by primary heating system, probably only applicable to centrally ducted (or hydronic) systems.
- Also, likely to be moving to working at ORNL, and will want to look at this for HUD code manufactured housing.
- Part of the problem is that the current base case assumes a manual setback tstat.



## Discussion: The Floor Is Open

- Next steps on Version 2?
  - Draft 1 Q2 22
  - Soonest we'd anticipate specification being finished is the end of 2022, which means it would be effective late in 2023. An effective date near the beginning of 2024 is more likely.