



ENERGY STAR Connected Thermostats

CT Metrics Stakeholder Meeting Slides

December 9, 2021 (2-4PM EST)



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

Brad Powell, Carrier

Jason Thomas, Carrier

Diane Jakobs, Rheem

Carson Burrus, Rheem

Chris Puranen, Rheem

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

Mike Clapper, UL

Alex Boesenberg, NEMA

Ethan Goldman

Jon Koler, Apex Analytics

Hassan Shaban, Apex Analytics

Michael Siemann, Resideo

Arnie Meyer, Resideo

Aniruddh Roy, Goodman/Daikin

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

Peter Gifford, Mysa

Aidan Girard, Mysa

Mustafa Elsisy, Mysa

Vrushali Mendon, Resource Refocus

Riana Johnson, Illume Advising

Sylvain Mayer, Sinope Technologies

Caroline Cote, Natural Resources Canada

Elmer Guest, Efficiency Manitoba



Agenda

- Software updates
- Missing data updates and proposed solutions
- Tau friendly regression
- Canadian ENERGY STAR ST specification



Software Updates: V2.0

- Bug fixes and performance improvements
 - Missing data thresholds (will be discussed later)
 - Tau-friendly regression (will be discussed later)
 - Updating climate zone lookups using eeweather rather than CSV lookup
 - Initial tests appear promising (more climate zones returned with eeweather)
 - CSV lookup method uses ZIP codes for lookups, which have some overlap with ZIP codes, but are not 1-to-1.
 - That leads to the next question...

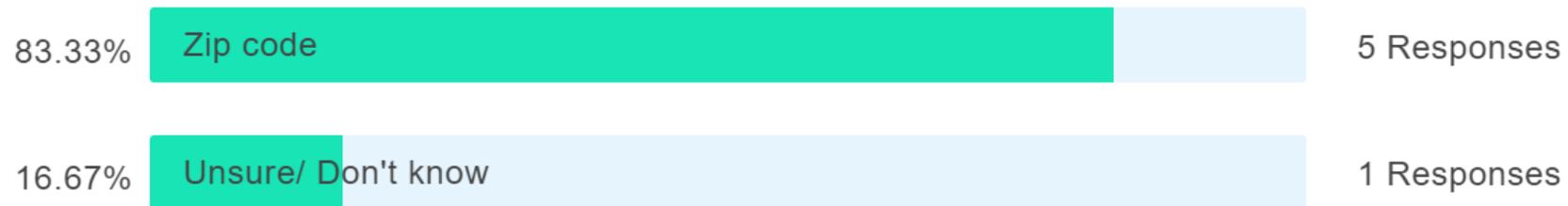


Software Updates: V2.0

- ZCTA vs ZIP Codes
 - 1.7x versions of the software prefer using ZCTA for location lookups with ZIP Code as fallback(weather stations, etc.)
 - Version 2.x uses ZIP Code lookups to get lat/long and uses that for location lookups (weather stations, etc.)
- **Question:** Are vendors providing ZCTA or ZIP Codes to the software? (poll)

1 of 1. What are you using to denote a thermostat's location for the software?

Multiple choice with single answer





Discussion: Software Updates: V2.0

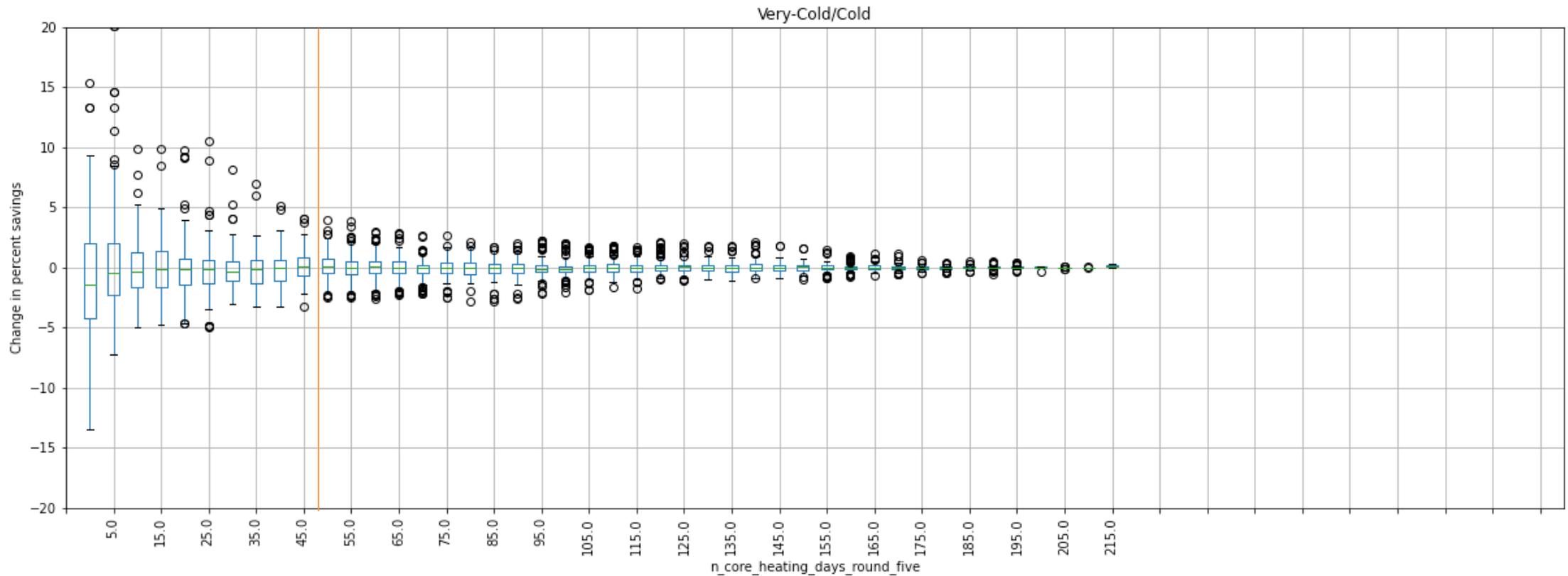
- What does eeweather use for lookup? A: ZCTA, but if we use the “zipcodes” package, we can go directly to lat-long
- Why not just map the zip codes to weather stations and use that file as a resource? Snapshot included in EPA software makes it more reproducible. Another suggestion – if it’s fine that the mapping is slightly stale, then it’s not bad. Could pin the zip code package to a particular version – but creating dependencies may be best to avoid.
- This all started to avoid a memory leak associated with the lookup.
- Also, how does the zipcode library do the lookup? Zips.json file.



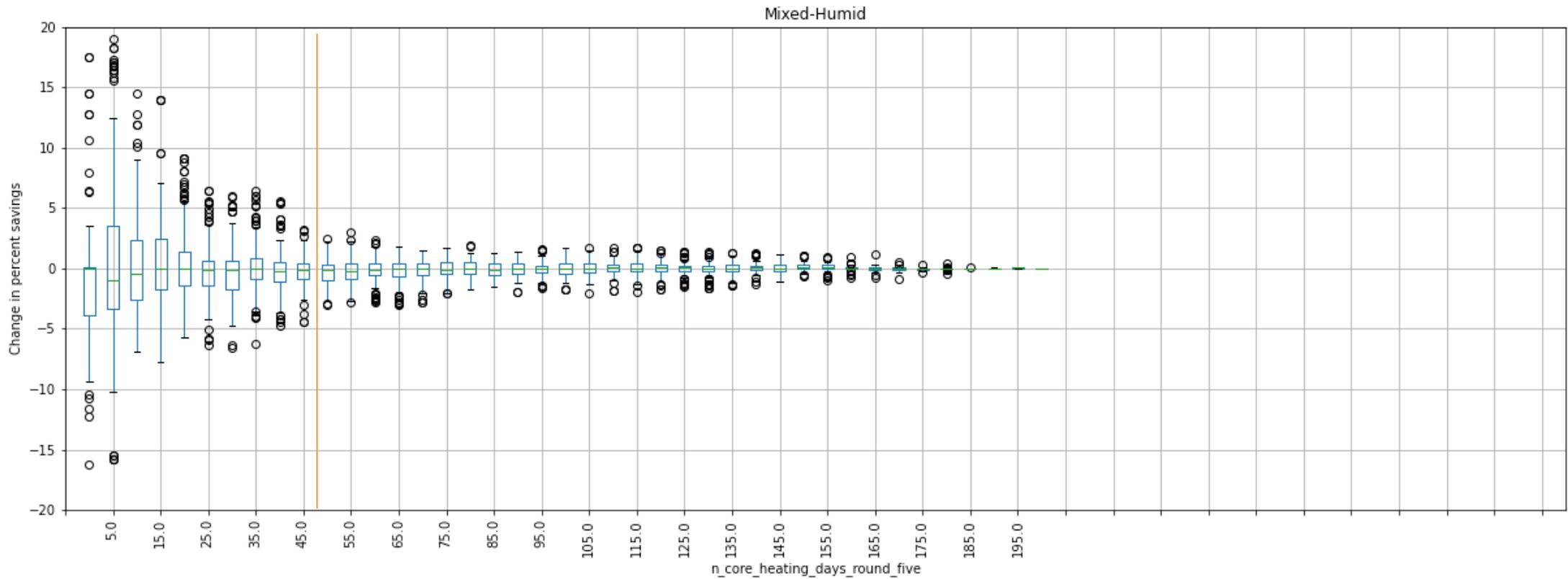
Missing data updates and proposed solutions

- Goal: set a threshold for missing days that preserves accuracy while retaining most thermostats
- Assumptions: missing days are randomly distributed, not correlated with weather
- Proposal: **minimum of 50 core days** in each season
- Tests:
 - Distribution of change in savings metric when core days are randomly deleted as compared to metric score with full data set
 - Distribution of number of core days in full data sets by climate zone and season
- Looking for feedback or additional analysis with your population's data

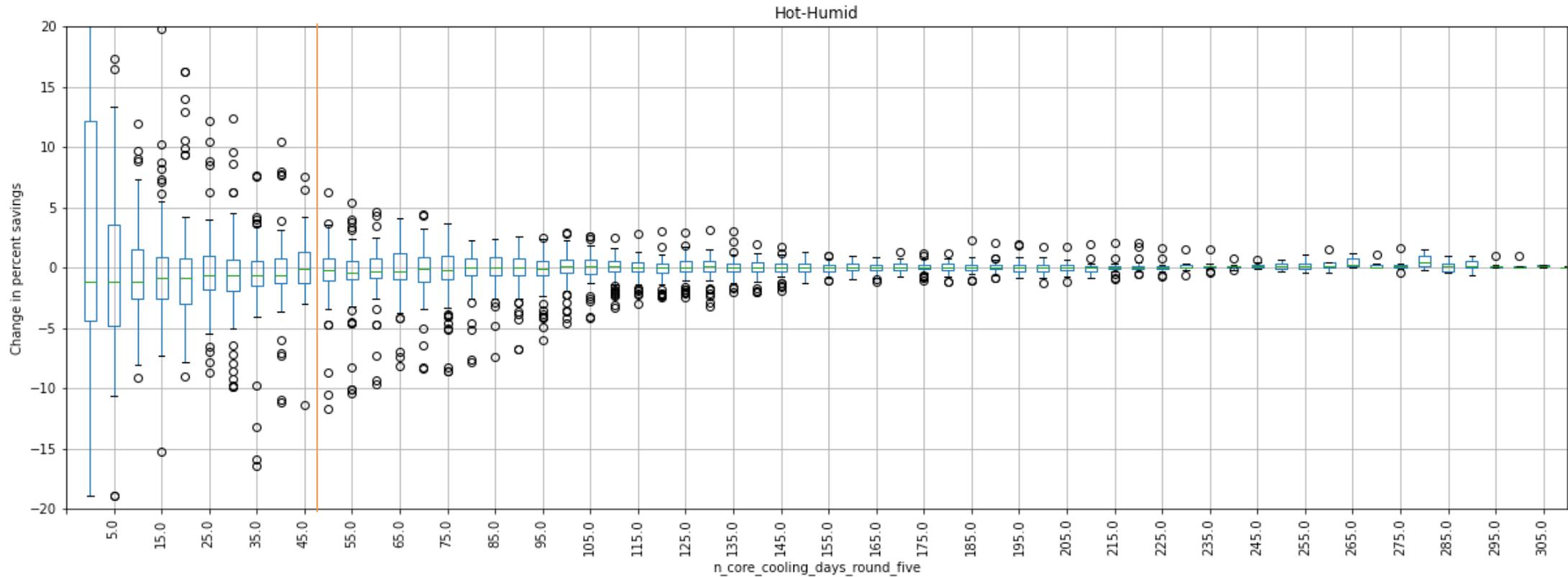
Missing data: Impact on Heating in Very-cold/Cold



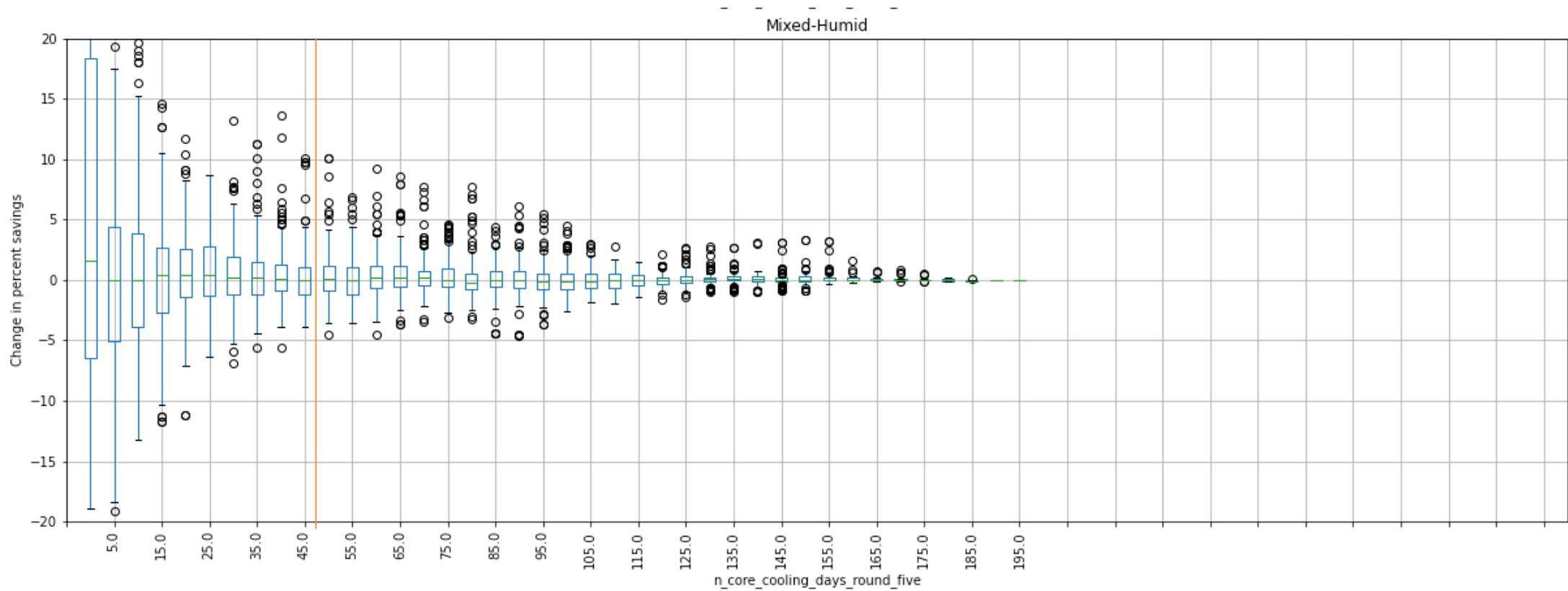
Missing data: Impact on Heating in Mixed-Humid



Missing data: Impact on Cooling in Hot-Humid

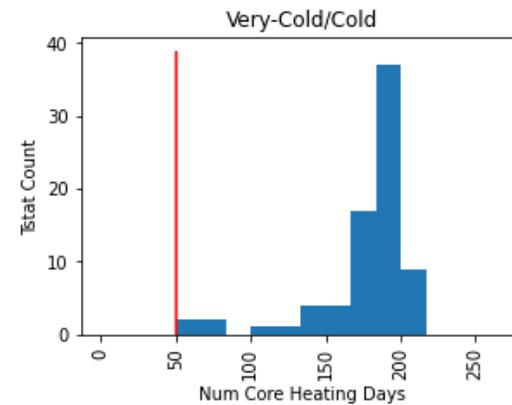
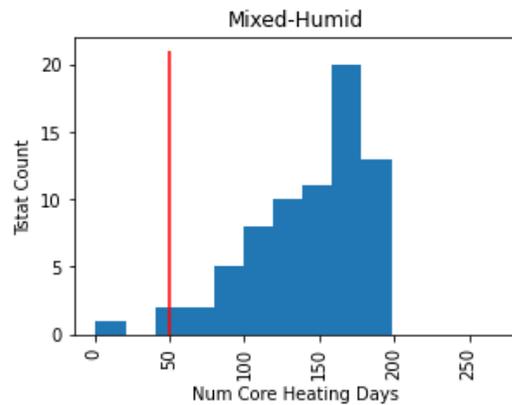
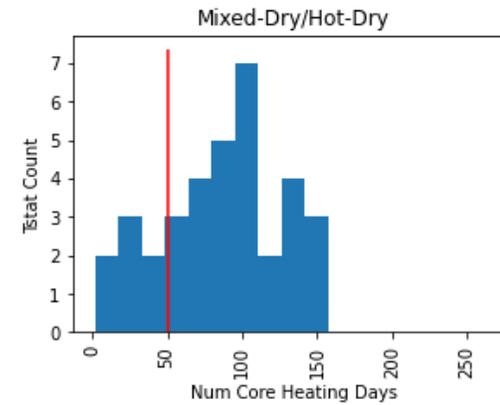
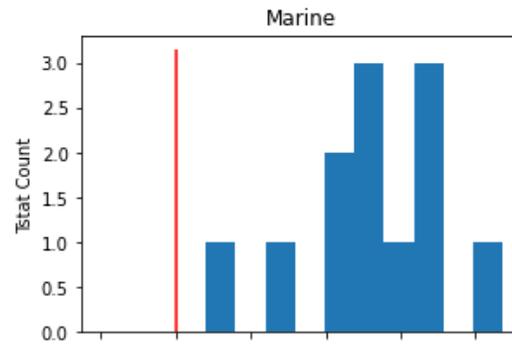
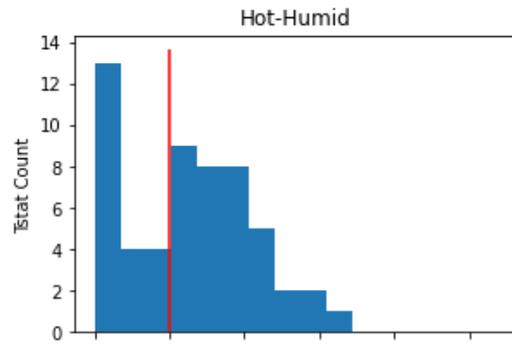


Missing data: Impact on Cooling in Mixed-Humid

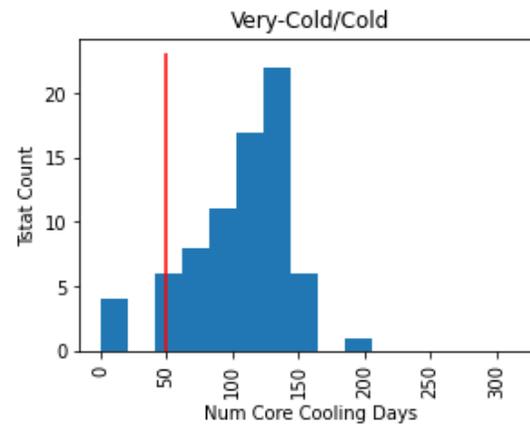
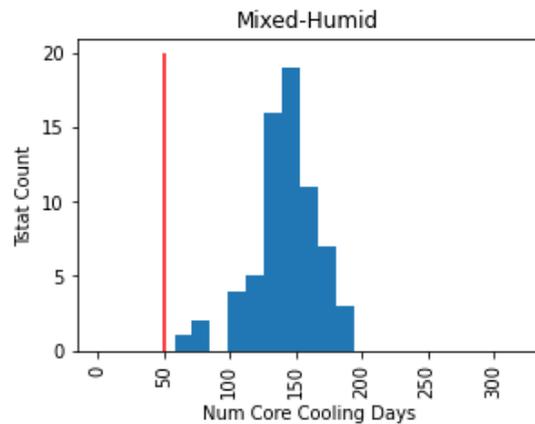
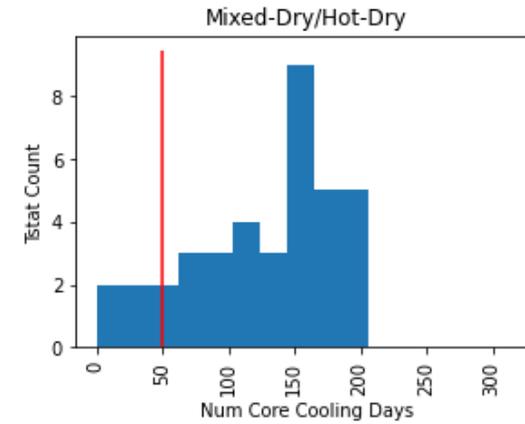
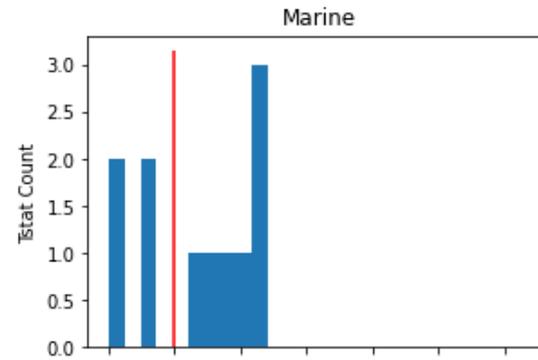
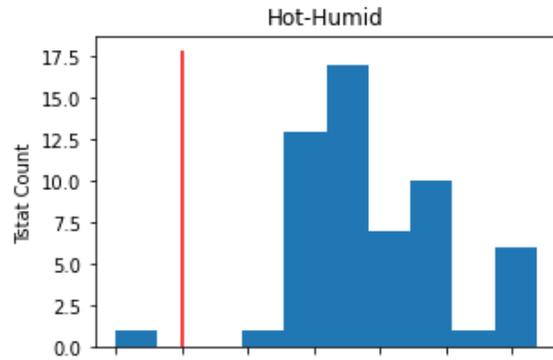




Missing data: Distribution of Heating Core Days



Missing data: Distribution of Cooling Core Days





Missing data: points for discussion

- Proposal: **minimum of 50 core days** in each season
- This will soften the requirements for data completeness
 - Currently only 5% of days (19 per year) are allowed to be missing
- However, the V.1 software did not enforce this requirement
- Have you looked at missing data metrics for your own data? Could you?
- Do you have any reason to believe missing data is non-randomly distributed?
- Are there any sources of bias or other negative impacts that you foresee here?



Discussion: Missing data updates and proposed solutions

- One vendor is interested in more investigation. Heating-only products tend to go offline in summer, so not random. However, the test is in each season, so this tstat would not be removed from the heating results, only from the cooling.
- This isn't really about missing data – it's about core days. Even among the ones that have decent fits in hot humid don't have 50 heating days (40%). This will focus only on the homes that have the most heat. Introduces bias, though small because of weighting. Different thresholds for different zones? (quoted the percent in each climate zone) Or start with how many days missing in the year first?
- Concern arose about how we could tell what the right threshold is for missing days, and in addition, similar to climate zone weighting, are these homes the most important? Also, a particular number of days missing will have a very different effect depending on which days are missing. We worked on trying to decide if core days were missing, which turned out to be very difficult.



Discussion: Missing data updates and proposed solutions

- Suggestion: capture the number of HDD/CDD included in this trace, compare to total for weather station. Or just look for stats that have at least 300 days or something. (might want to use 325, because by the statistics just quoted, 65 missing days could be a complete heating season).
- A CDD/HDD can be implemented, but it doesn't typically fit into the way the software is organized now.
- Could also be missing $\frac{3}{4}$ of the heating season in a cold climate zone. Works fine in most places, so maybe adjusting the threshold in some climate zones, for some seasons.



Discussion: Missing data updates and proposed solutions

- A couple clarifications of definition of core day, and also how many hours can be missing before a day is missing.
- Morphed into a discussion about how to deal with missing short intervals and how to take an accumulated run time and spread it over intervals when there is no data.
- Some vendors may have intervals with more minutes of run time in an interval than the number of minutes in the interval.
- If vendors are filling in gaps before the data even gets to EPA software, will we even see days as missing that should be, according to our own rules.
- Probably worth talking with each vendor about data filling, particularly on run time data. Will come back next quarter (or sooner) with specific questions.



Tau friendly regression

- Plan: Force search of Tau values from 0-20 rather than full linear regression
- Progress: Created software branch, created new search algorithm
- Next steps: Investigate why some thermostats are producing odd-looking regressions



Discussion: Tau friendly regression

- What's the algorithm for the search? Right now we're just doing each whole number tau, rather than using a binary fit. Once we have things working, we can look for how well-behaved the search space is.
- What do you mean by "bad fits" with the software we're testing? If you plot the data and the regression line together and they visually make NO SENSE, there's a problem.



Canadian ENERGY STAR ST specification

- Goal: Expand ENERGY STAR ST V2.0 program to Canada.
- Barriers: Canadian cold climate requires additional data collection and analysis preventing direct adoption of US tool.
- How to overcome barriers?
 - Validate data points by mapping postal codes to weather data and identifying climate zones.
 - Weighting climate zones by proportion of national heating and cooling energy used in each zone.
 - Determine whether the US metric levels are applicable to Canada for low and line voltage thermostats
 - Update software to incorporate metric system into input and output files for Canadian stakeholders
 - Determine cost savings and GHG emission reductions



Discussion: Canadian ENERGY STAR ST specification

- Sense of timing? Starting in February. Attempting to align with Version 2.0 timeline, though it will depend on the results of the work itself.
- Where are we regarding LVT? They will be in Draft 1 of Version 2.