ENERGY STAR Connected Thermostats

CT Metrics Stakeholder Meeting Slides

April 17, 2020

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
Jing Li, Carrier
Jason Thomas, Carrier
Brian Rigg, JCI
Theresa Gillette, JCI
Rohit Udavant, JCI
Diane Jakobs, Rheem
Carson Burrus, Rheem
Chris Puranen, Rheem
Glen Okita, EcoFactor
Brent Huchuk, ecobee
John Sartain, Emerson
Eric Ko, Emerson
Albert Chung, Emerson
James Jackson, Emerson
Mike Lubliner, Wash State U
Charles Kim, SCE
Michael Fournier, Hydro Quebec
Dan Fredman, VEIC
Robert Weber, BPA
Phillip Kelsven, BPA
Casey Klock, AprilAire
Wade Ferkey, AprilAire
Ulysses Grundler, Trane
Jeff Stewart, Trane
Mike Caneja, Bosch
Sarathy Palaykar, Bosch
Brenda Ryan, UL
Mike Clapper, UL
Alex Boesenberg, NEMA
Ethan Goldman
Jon Koliner, Apex Analytics
Michael Siemann, Resideo
Aniruddh Roy, Goodman/Daikin
Dan Baldewicz, Energy Solutions for CA IOUs

Agenda

- Software update and questions
- RHU data analysis, February resubmission
- Long term trends in resubmission data
- Summary of communicating controllers workshop webinars

Software Updates and Questions

- EPAThermostat Python 2.0 Versions
  - Pandas versions for features, speed improvements and removal of undefined behavior.
- Deprecations
  - Python 2.x : Deprecated (no longer supported by Python developers)
  - Python 3.4 and earlier: No longer supported by Pandas as of version 0.25.0
  - Python 3.5.x: Dropped in Pandas 1.0.x
- Additional support
  - Python 3.8: Added and tested with EPAThermostat 2.0

Discussion: Software Updates

- Please send out email reminding folks that some versions will not be supported. Will od.
- No plans to update Pandas for current software.
- Re Version 2.0: Any more news about what the input file will look like?
  - Yes, on GitHub under https://github.com/EPAENERGYSTAR/epathermostat/tree/feature/epathermostat_2.0
  - Please test
  - Documentation for input file has been updated, still working to make it match reality; feedback welcome
-

RHU2: Status Update

- Previous meetings: shared RHU2 results, discussed need for additional sampling of heat pumps for reasonable results
- Metric proposal: RHU2 in 30F-45F bin, upper 95th confidence limit of the mean ≤ 0.2
- February data resubmission: 10 General datasets and 4 Oversampled datasets
- Questions to address
  - Questions about national vs. regional
  - Questions about 5F or 15F bin(s)

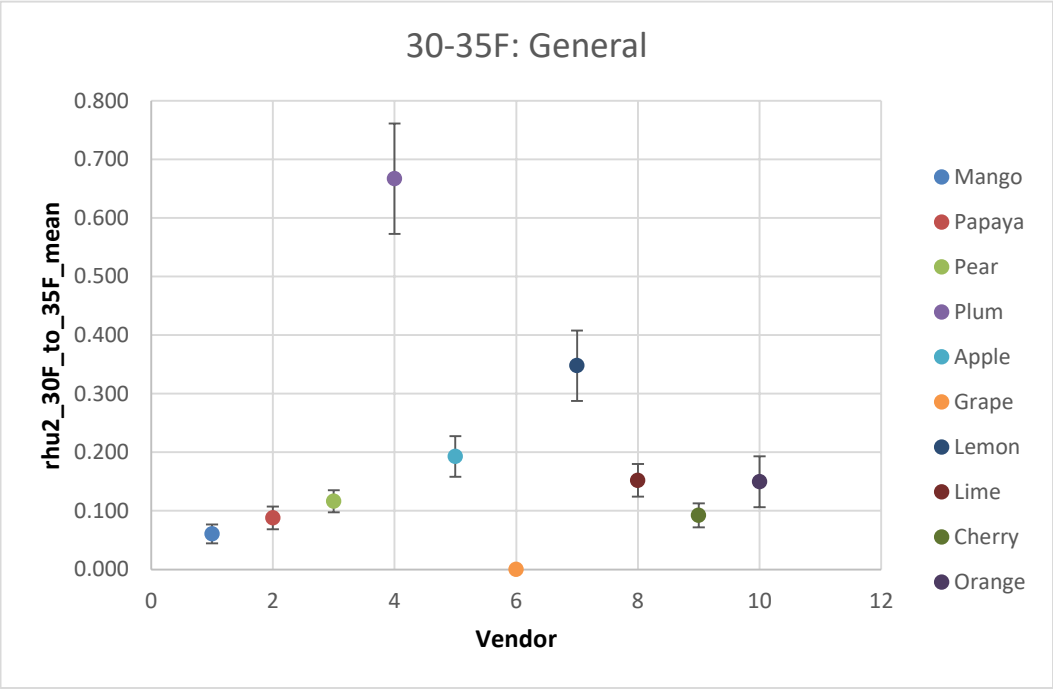
Data from February 2020 Submission

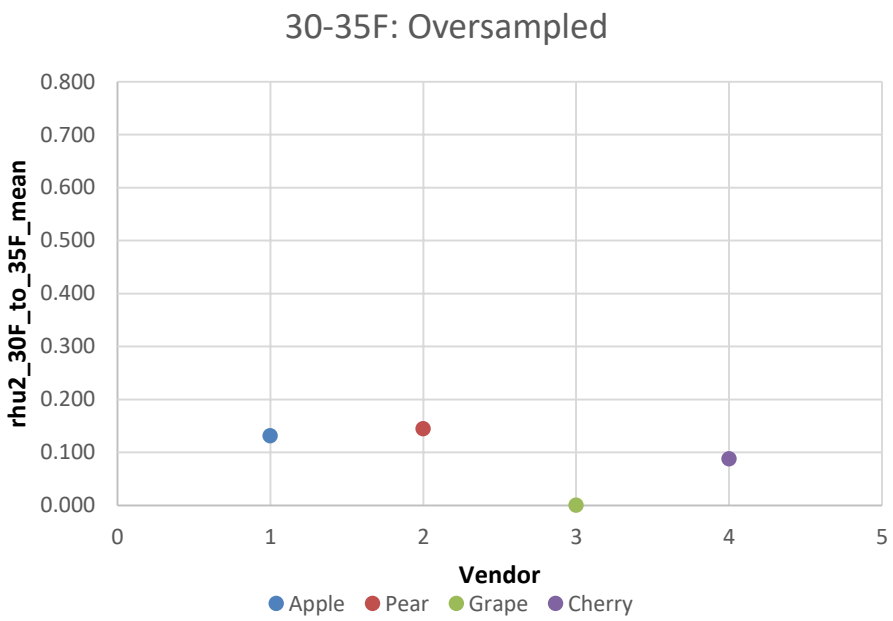
General dataset				
Vendor	30 40F n	30 40F ub	40 45F n	40 45F ub
Mango	69	0.093	65	0.089
Papaya	139	0.1	173	0.111
Pear	69	0.149	68	0.105
Plum	15	0.699	16	0.511
Apple	52	0.208	59	0.147
Grape	7	0	8	0.03
Lemon	21	0.366	28	0.415
Lime	87	0.186	94	0.171
Cherry	57	0.1	57	0.077
Orange	28	0.207	27	0.193

Oversampled dataset				
Vendor	30 40F n	30 40F ub	40 45F n	40 45F ub
Apple	642	0.129	735	0.106
Pear	556	0.149	612	0.107
Grape	130	0.018	180	0.017
Cherry	451	0.09	537	0.088

- Apple Oversample values are significantly lower (60%) than regular

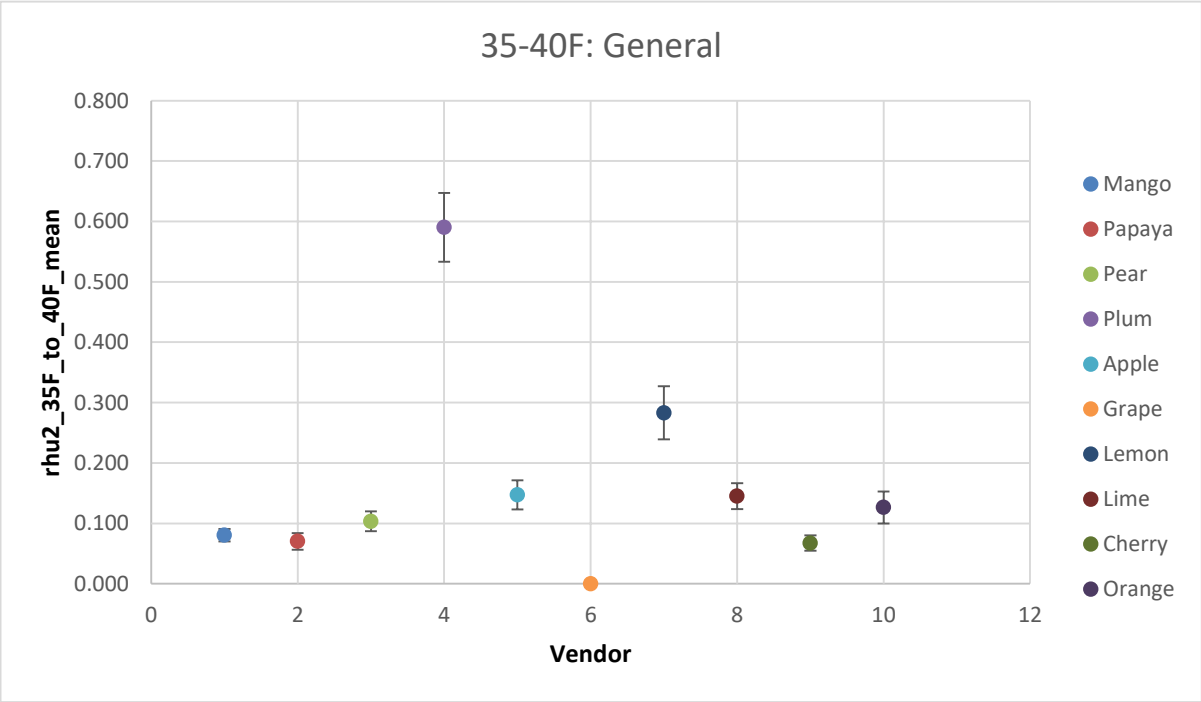
Statistical Difference – All Region

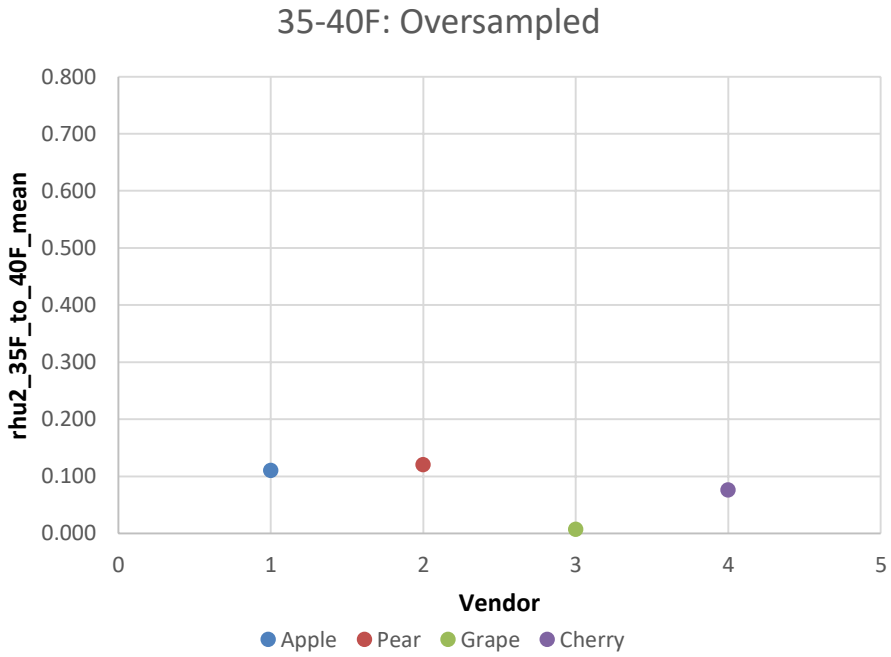




- Mean value for each vendor's sample with whiskers showing standard error of the mean

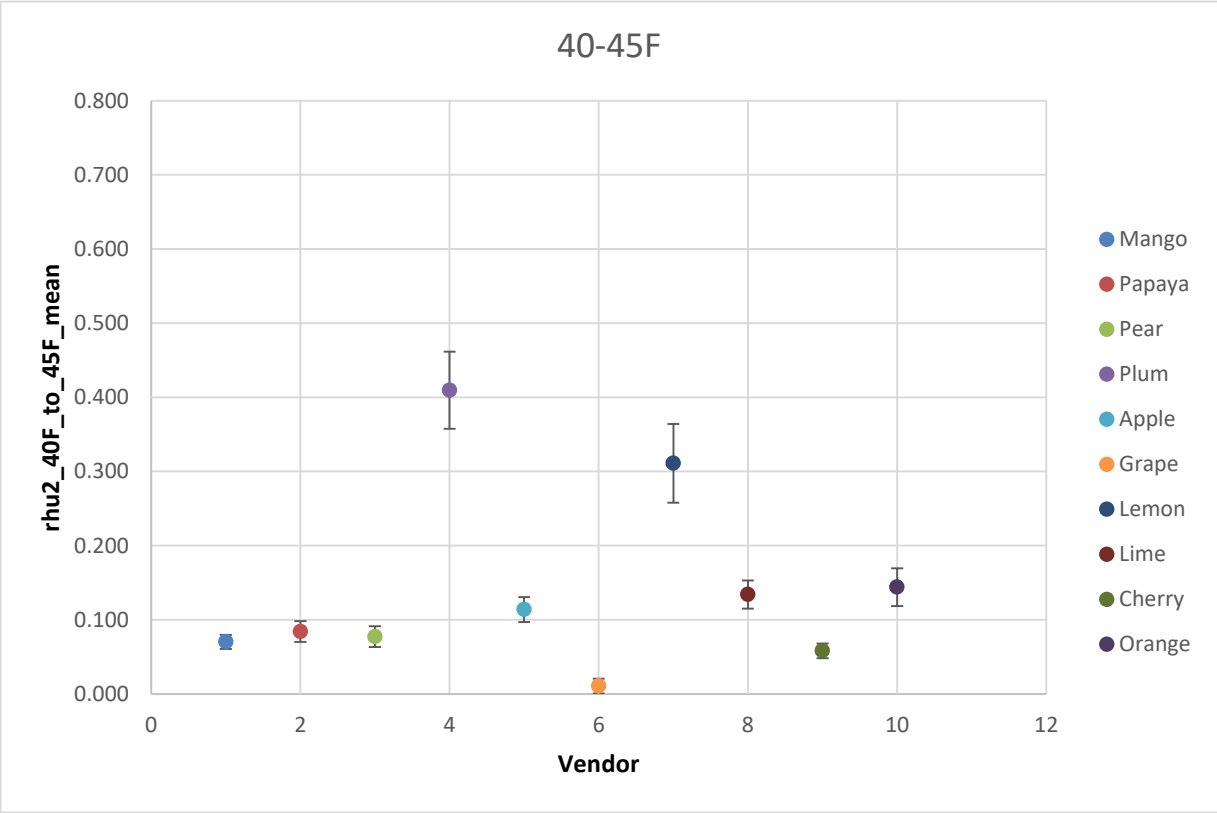
Statistical Difference – All Region

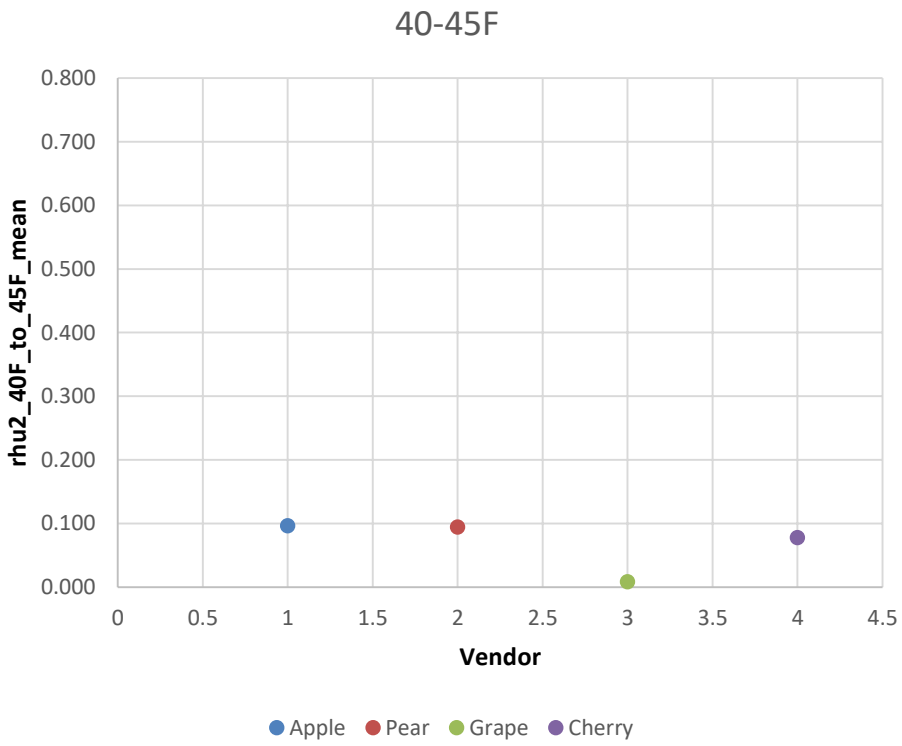




- Mean value for each vendor's sample with whiskers showing standard error of the mean

Statistical Difference – All Region





- Mean value for each vendor's sample with whiskers showing standard error of the mean

RHU2: National vs. Regional

General Dataset												
	All		Mixed Humid		Very Cold		All		Mixed Humid		Very Cold	
Vendor	30 40F n	30 40F ub	30 40F n2	30 40F ub3	30 40F n4	30 40F ub5	40 45F n	40 45F ub	40 45F n6	40 45F ub7	40 45F n8	40 45F ub9
Mango	69	0.093	13	0.089	5	0	65	0.089	13	0.093	4	0.013
Papaya	139	0.1	72	0.092	31	0.114	173	0.111	71	0.067	27	0.130
Pear	69	0.149	43	0.161	3	0.085	68	0.105	41	0.102	2	0.039
Plum	15	0.699	3	0.888	2	0.845	16	0.511	4	0.802	0	0.000
Apple	52	0.208	33	0.221	1	0.033	59	0.147	33	0.155	1	0.023
Grape	7	0	0	0.000	0	0.000	8	0.03	0	0.000	0	0.000
Lemon	21	0.366	10	0.403	1	0.230	28	0.415	12	0.390	1	0.207
Lime	87	0.186	49	0.223	7	0.317	94	0.171	48	0.188	7	0.389
Cherry	57	0.1	33	0.095	2	0.100	57	0.077	32	0.076	1	0.078
Orange	28	0.207	21	0.180	0	0.000	27	0.193	19	0.172	0	0.000

- Impact of more samples in Mixed Humid region: low upper 95th confidence mean values

RHU2: National vs. Regional (New slide added after the webinar)

Oversampled Dataset												
	All		Mixed Humid		Very Cold		All		Mixed Humid		Very Cold	
Vendor	30 40F n	30 40F ub	30 40F n2	30 40F ub3	30 40F n4	30 40F ub5	40 45F n	40 45F ub	40 45F n6	40 45F ub7	40 45F n8	40 45F ub9
Apple	642	0.129	196	0.123	191	0.119	735	0.106	194	0.089	178	0.104
Pear	556	0.149	166	0.149	110	0.151	612	0.107	167	0.095	109	0.115
Grape	130	0.018	0	0.000	1	0.302	180	0.017	0	0.000	1	0.295
Cherry	451	0.09	174	0.081	135	0.086	537	0.088	176	0.065	129	0.073

RHU2: Wider temperature bin analysis – All regions

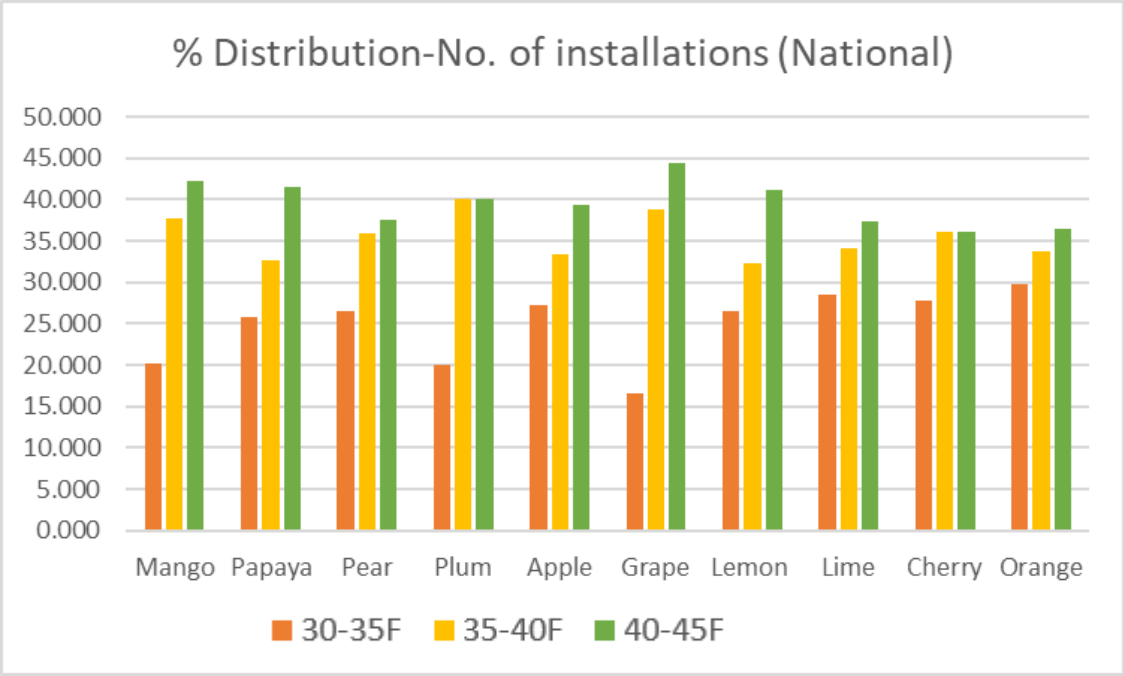
General Dataset								
	All		All		All		All	
Vendor	30 35F n	30 35F ub	35 40F n	35 40F ub	40 45F n	40 45F ub	30 40F n	30 40F ub
Mango	31	0.092	58	0.101	65	0.089	69	0.093
Papaya	107	0.126	136	0.097	173	0.111	139	0.1
Pear	48	0.153	65	0.135	68	0.105	69	0.149
Plum	8	0.852	16	0.702	16	0.511	15	0.699
Apple	41	0.261	50	0.194	59	0.147	52	0.208
Grape	3	0.000	7	0.000	8	0.03	7	0
Lemon	18	0.466	22	0.369	28	0.415	21	0.366
Lime	72	0.206	86	0.187	94	0.171	87	0.186
Cherry	44	0.133	57	0.092	57	0.077	57	0.1
Orange	22	0.235	25	0.178	27	0.193	28	0.207

Oversampled Dataset								
	All		All		All		All	
Vendor	30 35F n	30 35F ub	35 40F n	35 40F ub	40 45F n	40 45F ub	30 40F n	30 40F ub
Apple	547	0.144	633	0.121	735	0.106	642	0.129
Pear	451	0.162	550	0.135	612	0.107	556	0.149
Grape	1	0.000	128	0.019	180	0.017	130	0.018
Cherry	382	0.101	447	0.086	537	0.088	451	0.09

- Increase in the number of samples with temperature

Percentage wise distribution of samples in All regions

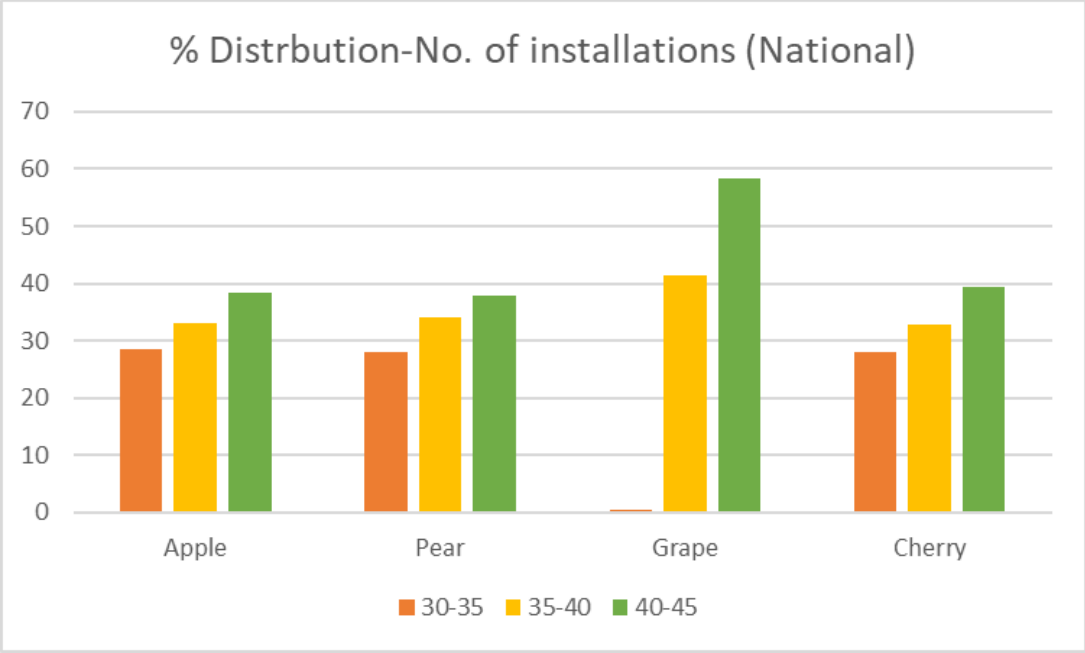
Vendor	30 35F n	%	35 40F n	%	40 45F n	%
Mango	31	20.12987	58	37.66234	65	42.20779
Papaya	107	25.72115	136	32.69231	173	41.58654
Pear	48	26.51934	65	35.9116	68	37.56906
Plum	8	20	16	40	16	40
Apple	41	27.33333	50	33.33333	59	39.33333
Grape	3	16.66667	7	38.88889	8	44.44444
Lemon	18	26.47059	22	32.35294	28	41.17647
Lime	72	28.57143	86	34.12698	94	37.30159
Cherry	44	27.8481	57	36.07595	57	36.07595
Orange	22	29.72973	25	33.78378	27	36.48649



- Most vendors follow similar pattern

Percentage wise distribution of samples in All regions

Oversampled dataset						
Vendor	30 35F n	%	35 40F n	%	40 45F n	%
Apple	547	28.564	633	33.055	735	38.381
Pear	451	27.960	550	34.098	612	37.942
Grape	1	0.324	128	41.424	180	58.252
Cherry	382	27.965	447	32.723	537	39.312



Discussion: RHU2

- Geothermal heat pumps (posited as the result for Grape in the oversampled data set in the 30-35F bin): exclude from sample b/c doesn't show what control can do, but what equipment can do
  - Can vendors distinguish which installations are geothermal? Some cases yes, some no. Can select when install thermostat, but can't tell from wiring. Most GHPs don't have an OB wire, so they wouldn't even show up as a heat pump. Often don't have an aux wire. W1 is always connected, wouldn't have for ASHP unless a dual fuel application. In other words, vendors should know from the wiring that it's either a GHP or CAC and gas furnace.
  - If not, can't distinguish control algorithms either.
- Follow up with email after meeting
- Test for consensus: Yes, GHP should not be in the same RHU data set as ASHP
- May be have their own requirements.

Discussion: RHU2

- Need to look at the oversampled data for consideration of regional vs. national
- Oversampled set seems to support wider temperature bin
- General data set doesn't have obvious problems with 15F-wide bin
- Note sample sizes in general data set supports using an oversampled set to evaluate RHU
- Parking lot: differing performance for GHPs – can we separate that out with the software?

Long Term Trends in CT Submission Data

April 17, 2020

Alan Meier and Leo Rainer, LBNL

Goals

What does the submitted CT data tell us with respect to:

- data quality
- submission procedures
- screening/filtering
- calculations
- new metrics to capture problematic situations
- modifications to data procedures and calculations
- energy savings

CT Metric not discussed yet

  - Let’s first understand the “input data”
  - Focus on the least processed data

Focusing on the least-processed data first

Trends examined

- Comfort Temperatures
- HVAC Runtimes
- Outside temperatures (supplied from elsewhere)

Different levels of aggregation

- National
  - All climate zones treated equally
- Key climate zones
- ~~System type~~

Notes

- 13 vendors
- Up to 7 submissions over 4 years
- N = 250 per climate zone (<u>before</u> filtering)
- National results treat each climate zone equally
  - But 3 climate zones have roughly 86% of heating 76% of cooling energy use.
  - Disproportionate weighting to zones with little energy (like marine)
- Most values based only on ”core” heating and cooling days
- Standard errors for most submissions are very small (barely visible bars)

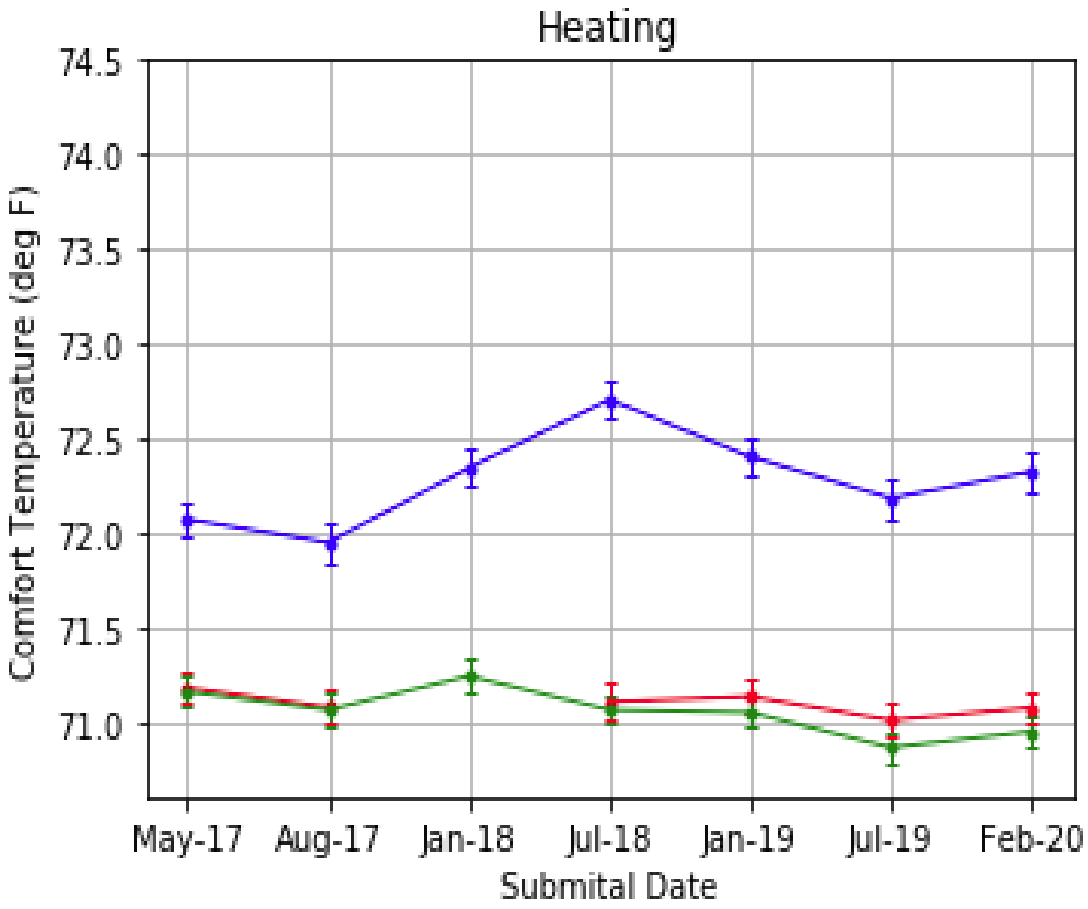
Reminder: Abi, Leo, and Alan do not know which vendor submitted which data

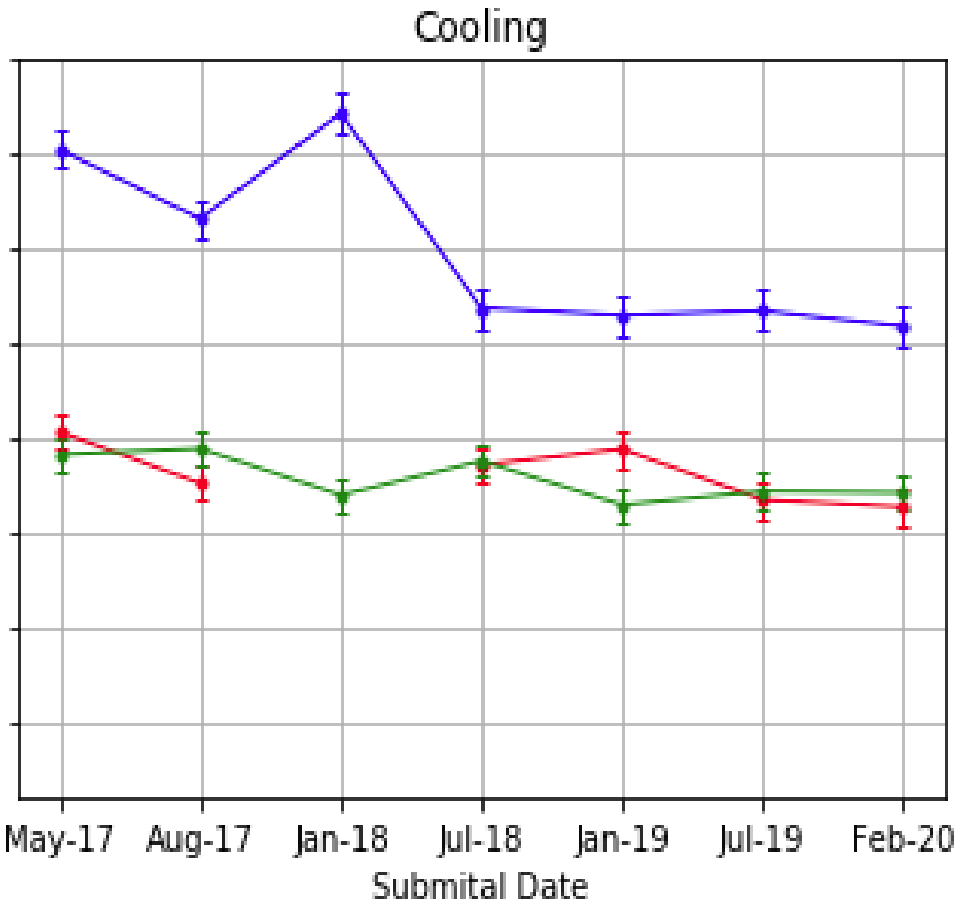
What Should We Be Looking for in the Trends?

Examples:

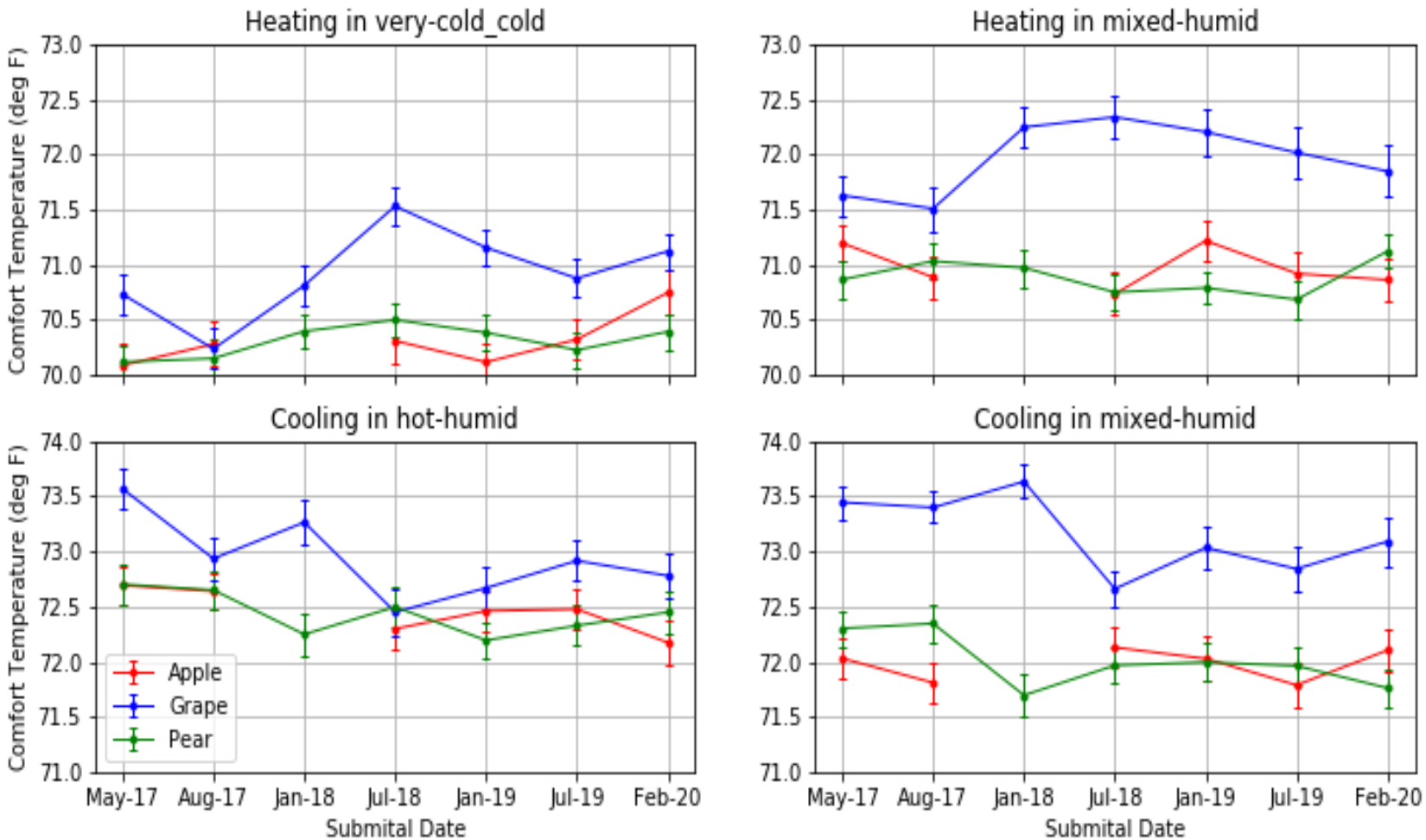
- Are the absolute values of a vendor’s submission reasonable?
  - *Can we believe a heating comfort temperature of 73.5°?*
- Are there large fluctuations in a vendor’s submissions?
  - *Should a reported comfort temperature change 1.5° between two submissions?*
- Do submissions from vendors have similar values?
  - *Are the trend lines on top of each other?*
- Do vendor trends fluctuate in the same direction?
  - *Do all vendors show the impact of a cold winter?*
- Are one vendor’s deviations explainable by deviations in some other characteristic?
  - *Does a vendor with high reported comfort temperature also have long heating runtimes?*

Comfort Temperatures – All Climate Zones

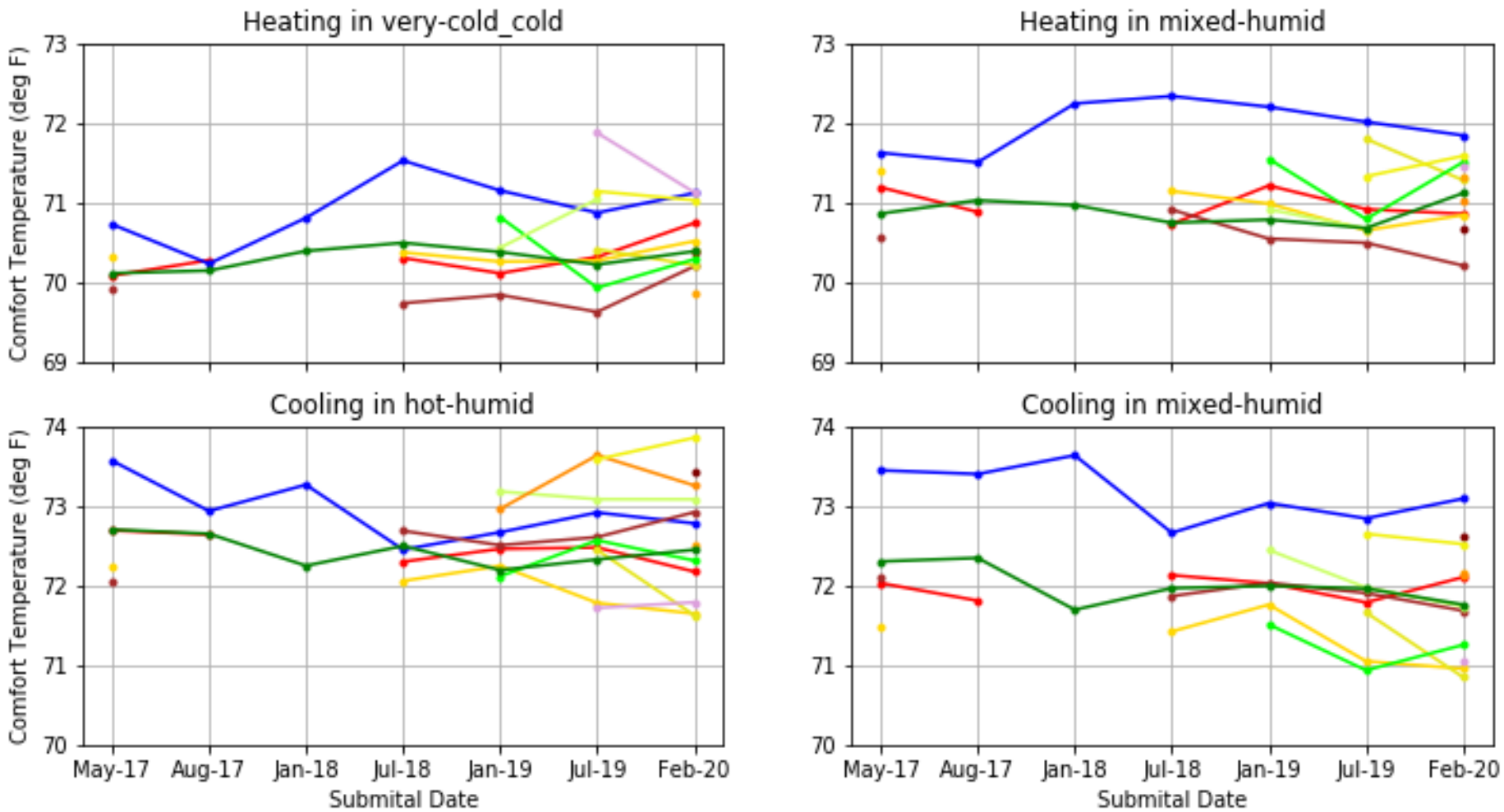




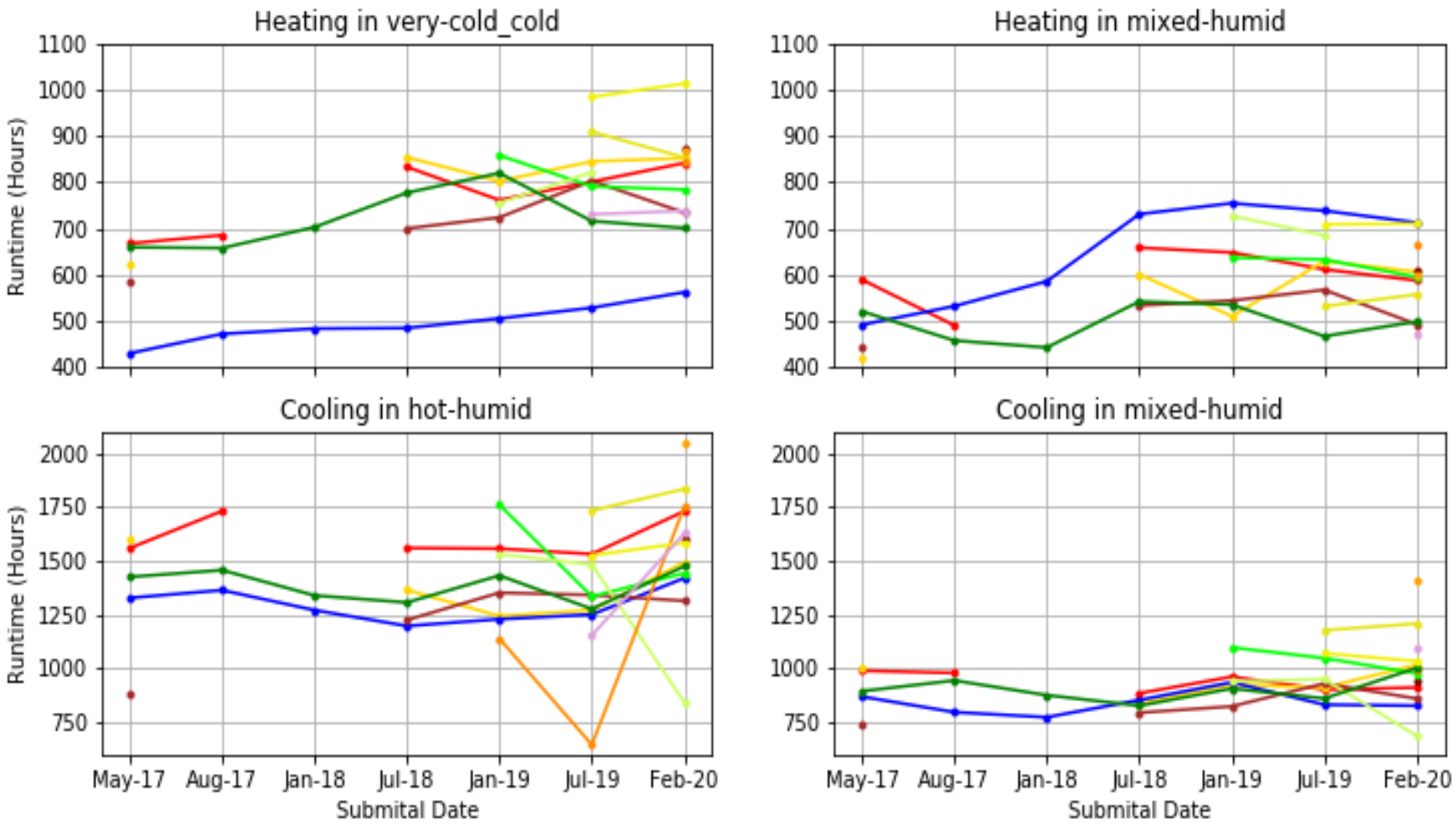
Comfort Temperatures for 3 Climate Zones



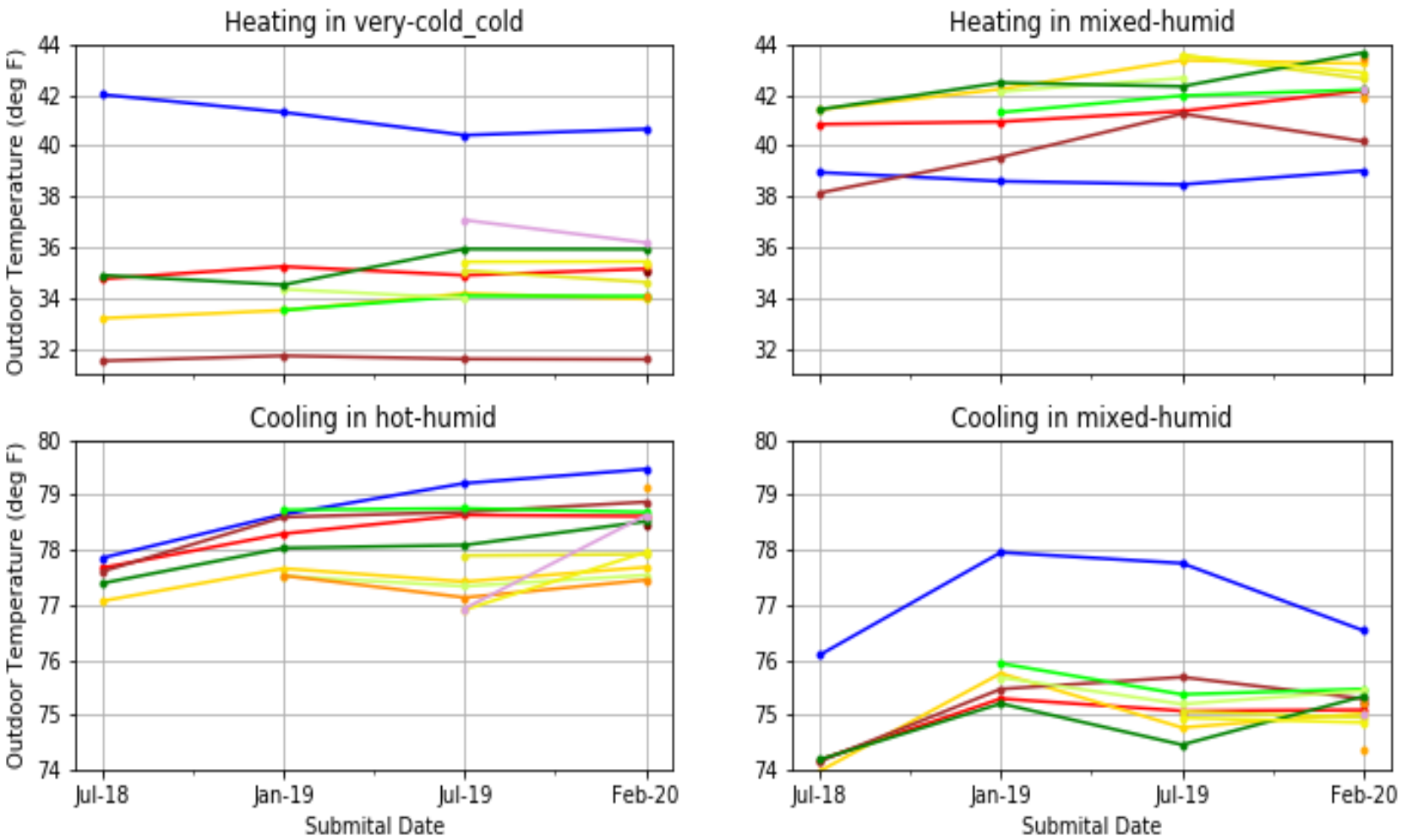
Comfort Temperatures for All Vendors



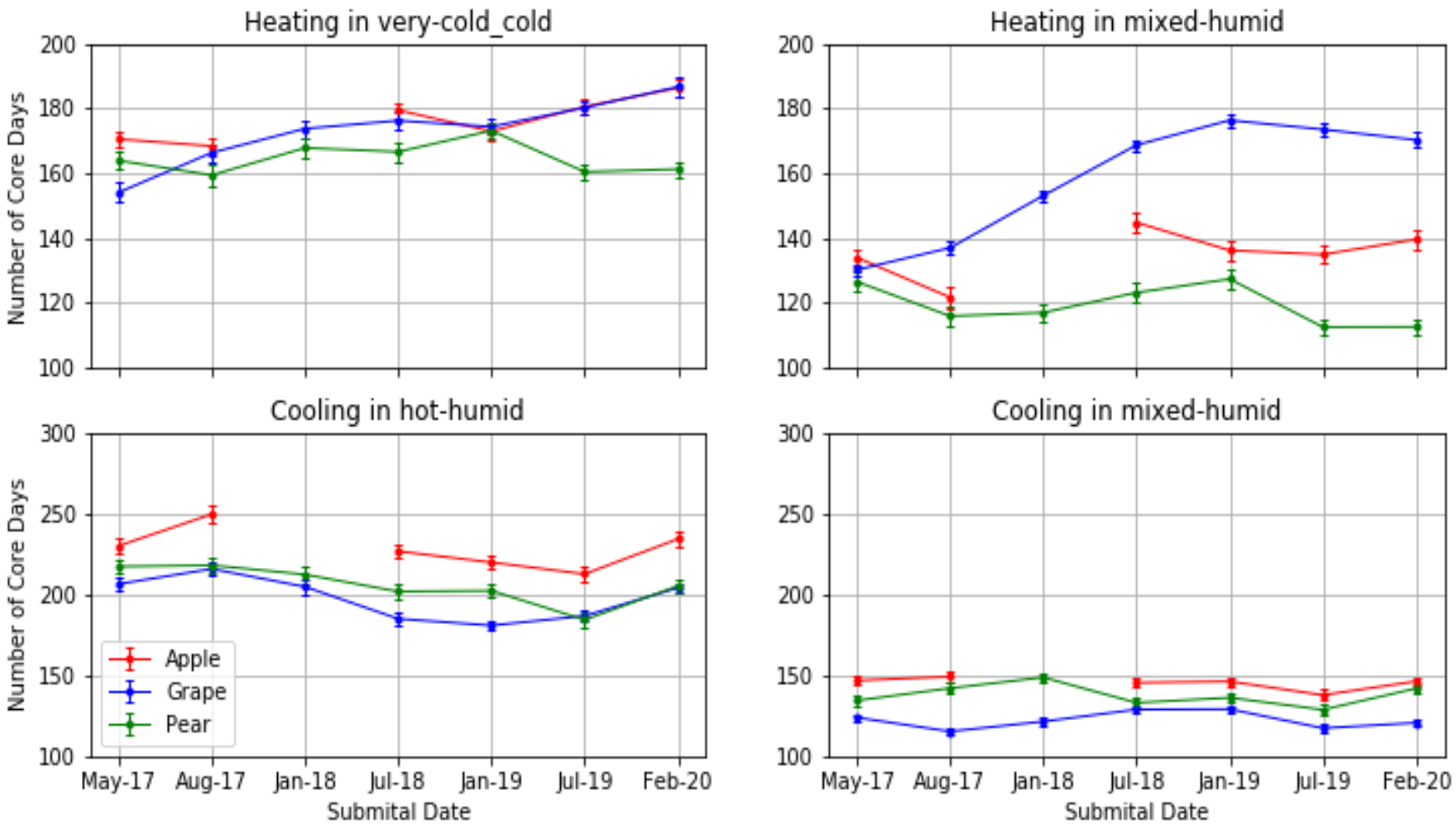
HVAC Runtimes for All Vendors



Average Outside Temperatures During Core Days



Core Heating and Cooling Days – 3 Climate Zones



Observations

- Some data problems – address them later
- National data are misleading – regional analyses are important
- Comfort temperatures are reasonable, best agreement for extreme zones
    - Should we expect greater differences among vendors?
- HVAC runtimes track well, especially for extreme zones, but significant differences in number of hours
    - Do differences in runtimes correspond to energy savings from CTs?
- Significant differences in outside temperatures experienced by each vendor’s customers (in some climate zones)
- Is grape’s data incorrect or is its customer base very different?
    - How can we decide?
- New filters would minimize data problems

Discussion: Long term trends

- Resampling the same population at the same time results in data that are at least as different as each 6 month sample. For instance, run with both old and new random number seeds a few weeks apart.
- Use a larger sample? 250 per CZ not enough
- May be affected also by unrealiability of access to weather data. Can EPA collect weather data in one place and/or make it more easily available?
- Any control on which firmware versions are in each data set? NO.
- Note that its possible that changes over times reflect different algorithms
  - More likely to be because the cohort of customers for smart thermostats in general are moving beyond early adopters. If we see comfort temperatures get less efficient, it might be that
  - Have observed this with several products – more permanent holds happening for smart thermostats
  - Can the temperature sensor accuracy drift over time? Static temperature accuracy test for certification only requires ± 2F.

Discussion: Long term trends

- Have we looked at long term or vendor-to-vendor change for average temperature? Would show same offset if its hardware based.
- Would be interesting to look at differences in tau and alpha across vendors (will do, once we get past input data)
- Average temperature, difference between comfort and average
- Worthwhile to provide more tools for sanity checking? Yes, provide to EPA, maybe for vendors as well.
- Is it known how many thermostats are kicked out b/c we couldn't get weather data? Not included in output file – would it be useful?
- Multiple thermostats in the same home are treated separately, even if they are from the same vendor (though fairly common to have an old non-smart thermostat present as well).
- Would it be helpful to get more "raw" data (run times, temperatures) before filtering? Might fit in well with providing data quality outputs to vendor running it.

Report out: Workshop Webinars on Communicating Controllers

- April 3 12pm – 2pm and April 6 1pm – 3pm
  - 59 attendees 4/3, 42 attendees 4/6
  - >50% HVAC and controls product providers, 10-ish% program sponsors, 5-ish% certification body, 30-ish% other (smearing 4/3 and 4/6 numbers together)
- Communicating controllers provisional definition: two-way digital comms with controlled equipment. Usually proprietary, but need not be.
- High level discussion topics:
  - Context: what is keeping communicating controllers out of the program now? What data are available from controllers and systems to build metrics from?
  - Where does the controller end and controlled equipment start? Can we devise a program that recognizes the whole system?
  - What savings mechanisms are we most interested in? Is field data useful to check they are present?
- Links to slides and mural board (discussion tool) from metrics website

What keeps communicating controllers from participating now?

- Can't use installations controlling variable capacity equipment for field data evaluation
- Unclear what value certification would bring
- Standby loss criteria – complicated by confusion about what is considered part of the controller and what isn't

Data available from controllers and systems

- Same things available from other connected thermostats, plus…
- Compressor speed or gas firing rate; can derive capacity fraction and (sort of) energy; CAC/HP energy estimate probably not equivalent between vendors
- Usually fan cfm; can derive airflow
- Fault codes (unique to manufacturer); aux heat and compressor lockout temps; system defrost time; often local outdoor temperature

Sometimes hard to distinguish controller from controlled HVAC

- Table below for centrally ducted; ductless may not HAVE a separate controller
- Control functions divided differently between controller and HVAC control boards for different models and brands
- Most savings mechanisms depend on both control capability and HVAC features
- Given this, labeling system (control + HVAC together) reflects reality better

	Fixed capacity	**Two-stage**	**Variable capacity/modulating**
Relay-based thermostat	typical	typical	Most manufacturers recommend against for compressor based HVAC
Communicating controller	Usually can; adaptor needed in HVAC closet?	Usually can; adaptor needed in HVAC closet?	Most manufacturers recommend highly for compressor based HVAC

Potential for ENERGY STAR System (Control + HVAC) program?

- Criteria not so hard to imagine; as for CTs, mix of hardware criteria, system capability criteria, and field data analysis
- However, universe of possible combinations is enormous
- Discussed how many of those combinations are really on the table; agreement that it's a much smaller number
- More thought needed about how much smaller, and whether it would be manageable
- More thought also needed on relationship to proposed CAC/HP V6.0 specification, which includes installation capabilities and optional connected capabilities that may rely on a controller

Savings mechanisms and metrics

- Extensive discussion of what determines whether a system is actually saving energy in the field
- Polling on most promising savings mechanisms:
  - **Avoid cycling between high and low capacity (or on/off), closely related to keeping system in lower capacity states**
  - **Set back and set up when possible**
  - Energy saving lockout temperatures
  - Fan control (avoid running constantly, especially on high speed)
  - Humidity control
- Extensive discussion of favoring lower capacity states, and whether it really saves as much as we think given distribution losses. EPA and most stakeholders convinced its an important energy (and performance) metric, though may not save as much as lab tests predict.

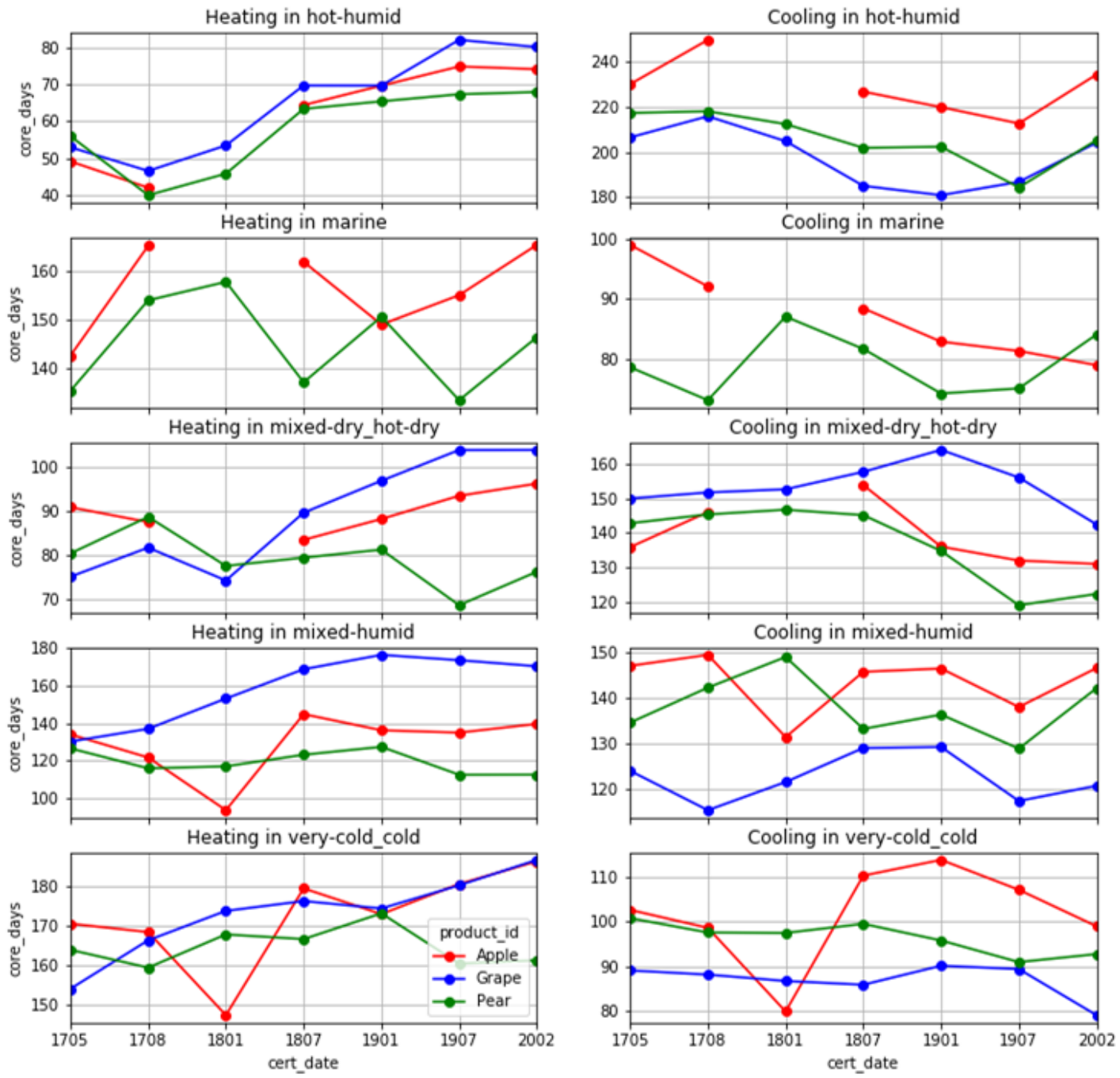
Next Steps for Communicating controllers

- AHRI working to process questions raised during workshop, will come back (with proposal) May 12
- EPA team will think more about possible metrics for top savings mechanisms
- EPA to consider theoretical system specification

Discussion: Communicating Controllers

- Ran of time for discussion

Core Days by Vendor – All Climate Zones



Metrics

Fraction of thermostats eliminated by goodness of fit filter?

Fraction of submissions = heat pump?

Ratio of aux runtime/emergency runtime?

The minimum N for a climate zone?

Observations

Lime had same N for heat runtime and emergency runtime (all coded as HPs)