ENERGY STAR Connected Thermostats

CT Metrics Stakeholder Meeting Slides

November 13, 2020

Attendees

Abigail Daken, EPA
Abhishek Jathar, ICF for EPA
Alan Meier, LBNL
Leo Rainer, LBNL
Nick Turman-Bryant, ICF for EPA
Eric Floehr, Intellovations
Craig Maloney, Intellovations
Michael Blasnik, Google/Nest
Kevin Trinh, Ecobee
Michael Sinclair, Ecobee
Jing Li, Carrier
Jason Thomas, Carrier
Theresa Gillette, JCI
Rohit Udavant, JCI
Diane Jakobs, Rheem
Carson Burrus, Rheem
Chris Puranen, Rheem
Glen Okita, EcoFactor

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
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
Aniruddh Roy, Goodman/Daikin
Jia Tao, Daikin
Dan Baldewicz, Energy Solutions for CA IOUs
Cassidee Kido, Energy Solutions for CA IOUs
Dave Winningham, Lennox
Dan Poplawski, Braeburn
Natasha Reid, Mysa
Peter Gifford, Mysa
Vrushali Mendon, Resource Refocus

Agenda

- Software Updates: V2.0 (5 min)
- NEEA Updates (5 min)
- Variable capacity metrics – continued from previous meeting (60 min)
- Connected Thermostat Use Cases – continued from previous meeting (40 min)

Software Updates: V2.0

- Testing Line Voltage Thermostat Field Data
- Preparing for software alpha release
  - Please test the current development at the following URL: [https://github.com/EPAENERGYSTAR/epathermostat/tree/feature/epathermostat_2.0]
  - Epathermostat/Readthedocs.io has the current input file format, but also has the Version 2.0 input file format and output file format described

    https://epathermostat.readthedocs.io/en/feature-epathermostat_2.0/

- No longer pursuing an anonymized data repository for software testing

NEEA NW Thermostat Study Updates

- Large study in the NW to tie thermostat metrics to savings determined by meter data. End goal: deem savings based on a few metrics
- Now moving forward again!
- Have two anonymous data sets from manufacturers, expect one or two more
  - The data agreements for these may be reused for EPA data
- Expect to have analysis by 2021
- Early results of studying anonymous data, useful for this group
  - Version 2 code appears to work as expected
  - Have tried a couple methodology changes
  - Issues with autocorrelation: whenever the temp is high in the cooling season, underestimating run time, and vice-versa
  - Considering different baselines to correct for this
  - Model fitting set up as an optimization problem with unbounded coefficients – sometimes the coefficients get crazy
  - May address by reorganizing model fitting

NEEA NW Thermostat Study Updates: discussion

- How many units being studied now, and do they control a variety of HVAC equipment?
  - So far 1100 thermostats, with more coming
  - Mix of equipment. It's opt in by customers
- Is A/C a significant feature?
  - Yes, on the East side of the cascades
  - About 40% of installations should have significant cooling
- What are you planning to use to calibrate the model against? Include multiple vendors?
  - Once have non-anonymous data sets, will compare to billing data for multiple vendors
  - Will not be visiting homes or anything else, but have demographic info and house characteristics
  - Note that people with best settings w/smart thermostats may be the ones who had the best settings before; most savings from terrible tstat management to mediocre
  - Expect first scatterplot between metrics and savings expected to be a useless blob, but have hope that adjustments will allow for useful correlation
- Might also get some program targeting information

Variable capacity metrics: Controlling for variability

- The discussion we had about when to concentrate on average capacity factor (see last 2 meetings) is part of a larger discussion

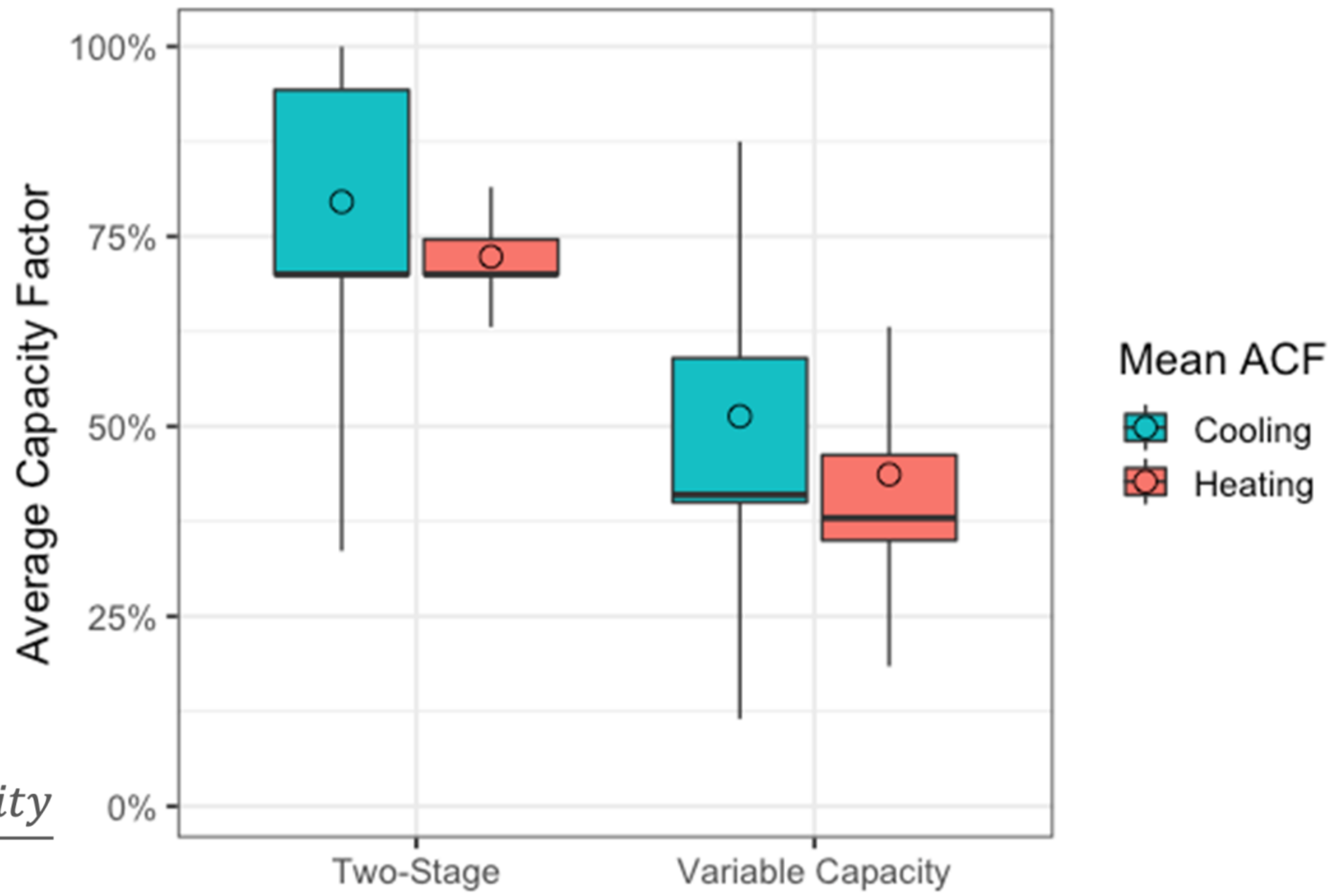
- **How much variability do we need to account for an accurate average capacity factor estimate?**

  - For example, we discussed looking at the outdoor temperature at which the unit starts to run for long periods of time, but that depends on sizing relative to the heating/cooling load

Review of Average Capacity Factor

- Variable capacity systems have lower ACFs than two-stage systems for both heating and cooling

- This comparison is for reference; today we will compare ACFs of variable capacity systems



$$ACF = \frac{ERT}{RT} = \frac{\sum_{time} run\ time * relative\ capacity}{\sum_{time} run\ time}$$

* Data are from an anonymous vendor and represent the average capacity factors across 71 two-stage and 91 variable capacity systems with split AC and gas furnace across 58 cities in 5 states for one year

Variable capacity metrics: Controlling for variability

- For the existing metrics (Heating Savings, Cooling Savings, RHU) we use different methods, grouped pretty much into three categories:
  - Control for household-level factors when necessary
  - Control for general conditions that do not vary significantly across households
  - Average out factors that we can't control for
- In all cases, we need to think about two things:
  - Is it sensible to combine data from multiple houses given the variability?
  - Does controlling for particular variables introduce bias for/against vendors?

Controlling for installation-to-installation variability

- In the heating and cooling savings metrics, there are several ways that variability are controlled for on a house-by-house basis:
  - Consumer preferences are (somewhat) controlled for by using a per-home comfort baseline
  - A heating and cooling runtime model is created for each household that accounts for thermal responsiveness (alpha and tau)
  - Sizing compared to heating/cooling load is controlled for by comparing reduction in runtime as a percent of total runtime for that home (rather than an absolute runtime metric)

Controlling for installation-agnostic variability

- This focuses on external conditions that we expect will affect all homes similarly
- For instance, for the RHU metric, we focus on resistance heat utilization during the outdoor temperature bin from 30F – 45F, because:
  - Outdoor temperature is the primary driver of resistance heat use
  - This is the range where the control has the most opportunity to make a difference

Other sources of variability are averaged out or excluded

- Additional sources of variability are mitigated by filtering which data are selected for estimating savings and by sampling across sites and regions

- Examples:
  - Missing thermostat or temperature data are excluded
  - More than 5% of days missing HVAC runtime are excluded
  - Variability by region or climate zone are accounted for in sampling
  - We also filter out homes with outsized savings that tend to reduce variation in the data set (e.g., a vacation home that’s unoccupied 85% of the time might be caught by that filter)

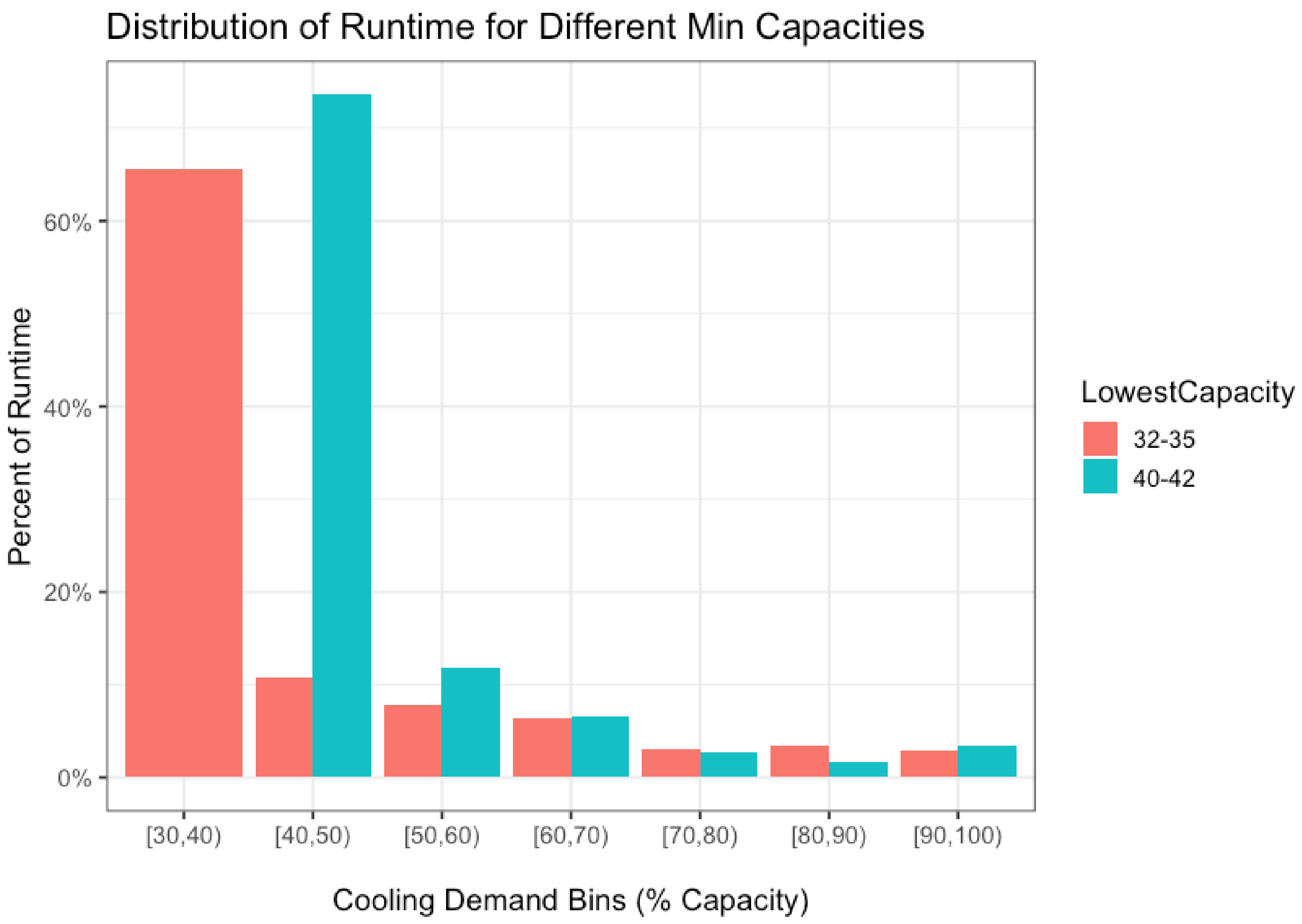
Installation-to-installation variability for variable capacity systems

- We'd like to discuss thinking about this more broadly for variable capacity systems
- The two factors we think might be important to control for, and which we think we might be able to get, are
  - Expected turn down ratio: units with a lower minimum capacity call should be able to save more energy – can't blame that on the controller
    - Should we use manufacture reported minimum capacity call from data?
  - Sizing: it will be harder for an oversized unit to avoid short cycles or run for long periods of time at low loads
    - Estimate from data based on runtime during 5th percentile weather conditions
- We investigated minimum capacity call and are open to suggestions about deriving sizing from data
- Would also be useful to discuss other factors it would be good to control for regarding the metrics we've discussed

Variability in runtime by minimum cooling capacity call

- Clear distinction between systems with different minimum capacity calls when averaged across all systems

- This can help us differentiate between variable capacity systems

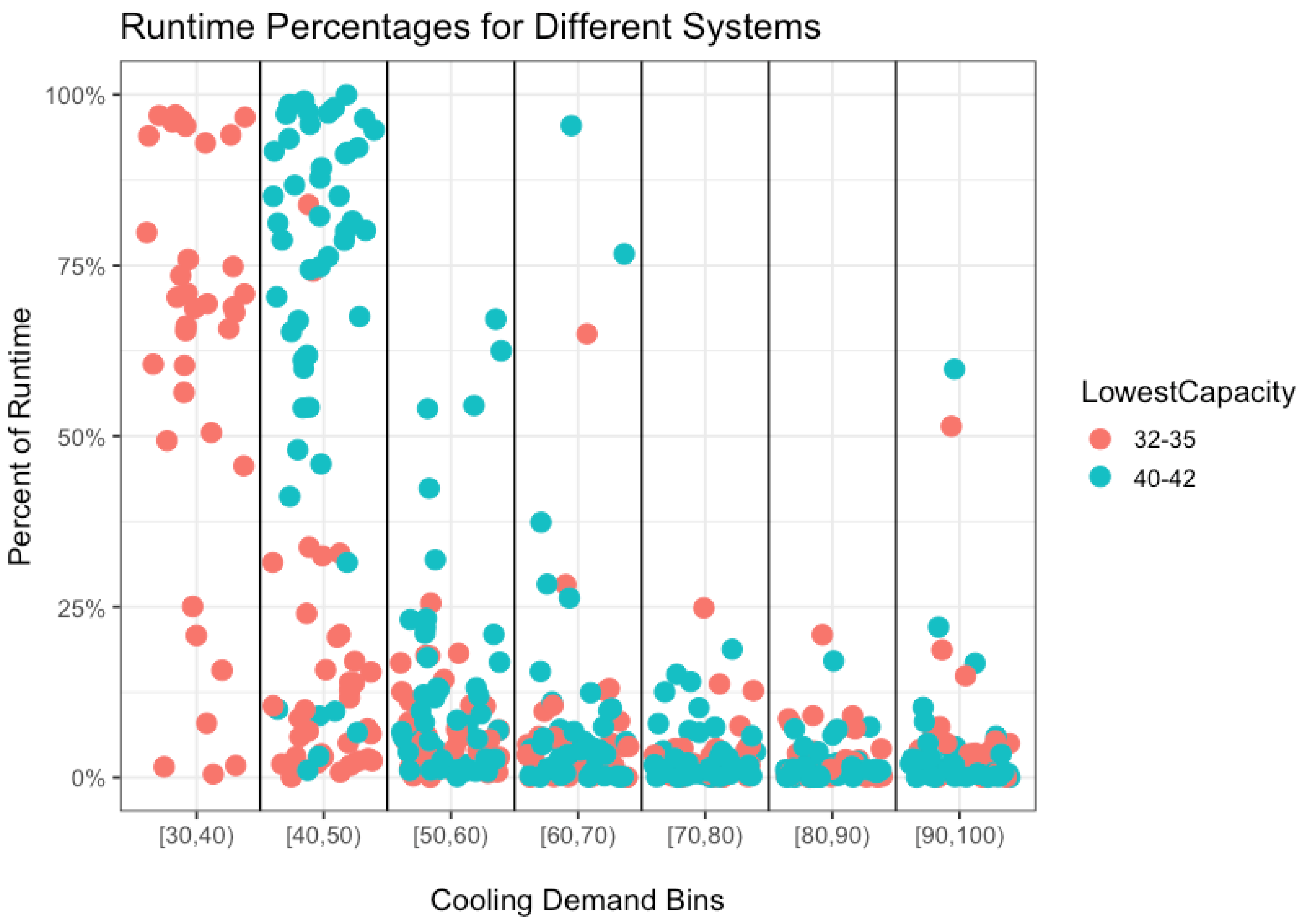


Discussion: Variable capacity metrics

- Clarification: When you say that the units have a minimum capacity call, are you talking about derived from the data or from what the system claims it's capable of?
- It's the minimum capacity call that we saw in the data for that model number
- Rheem: furnaces we always talk about input energy, for HP and AC we always talk about output energy
- Capacity call is relationship of operating set point to maximum set point in Hz, so the output energy depends on other conditions of operation. More closely related to input energy than output energy.
- Agree that thinking about how long units remain at low fire is important
- Note from Rheem: some furnaces seem to stop randomly when the set point hasn't been satisfied
- Does the controller know what the equipment is capable of? Yes, these are controllers with digital serial bus between them and equipment.
- These systems seem oversized to Rheem

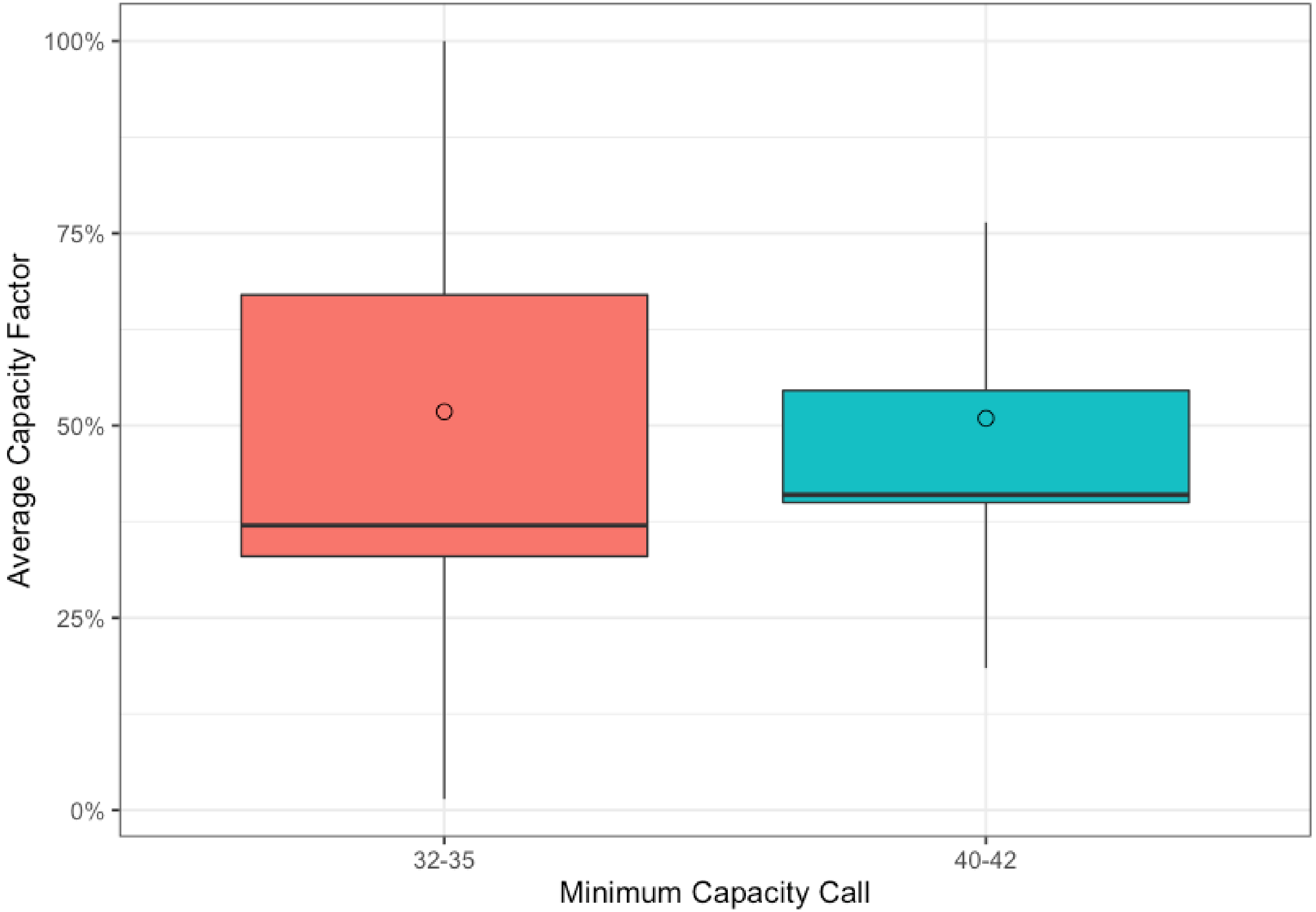
Variability in runtime by minimum cooling capacity call

- More complex when we look at percent of runtime spent by each system in each Cooling Demand Bin
- Still distinct differences between the types of systems, but will it explain enough variability to provide for accurate Average Capacity Factor estimates?



Average Capacity Factor variability across minimum capacity calls

- More variability for lower capacity systems (also more data)
- Means are almost the same but there is a difference for medians
- Turndown ratio does not explain much variability in ACF across systems



Poll

- What do you think we need to control for in ACF calculation? (single select)
  - Only needs to control for outdoor conditions (as we do for RHU)
  - Find the relevant conditions for estimating ACF (e.g., recovery periods, heating or cooling core days, etc.)
  - Individual household conditions (e.g., indoor temp, humidity, setpoint, etc.)
  - Not sure

Poll

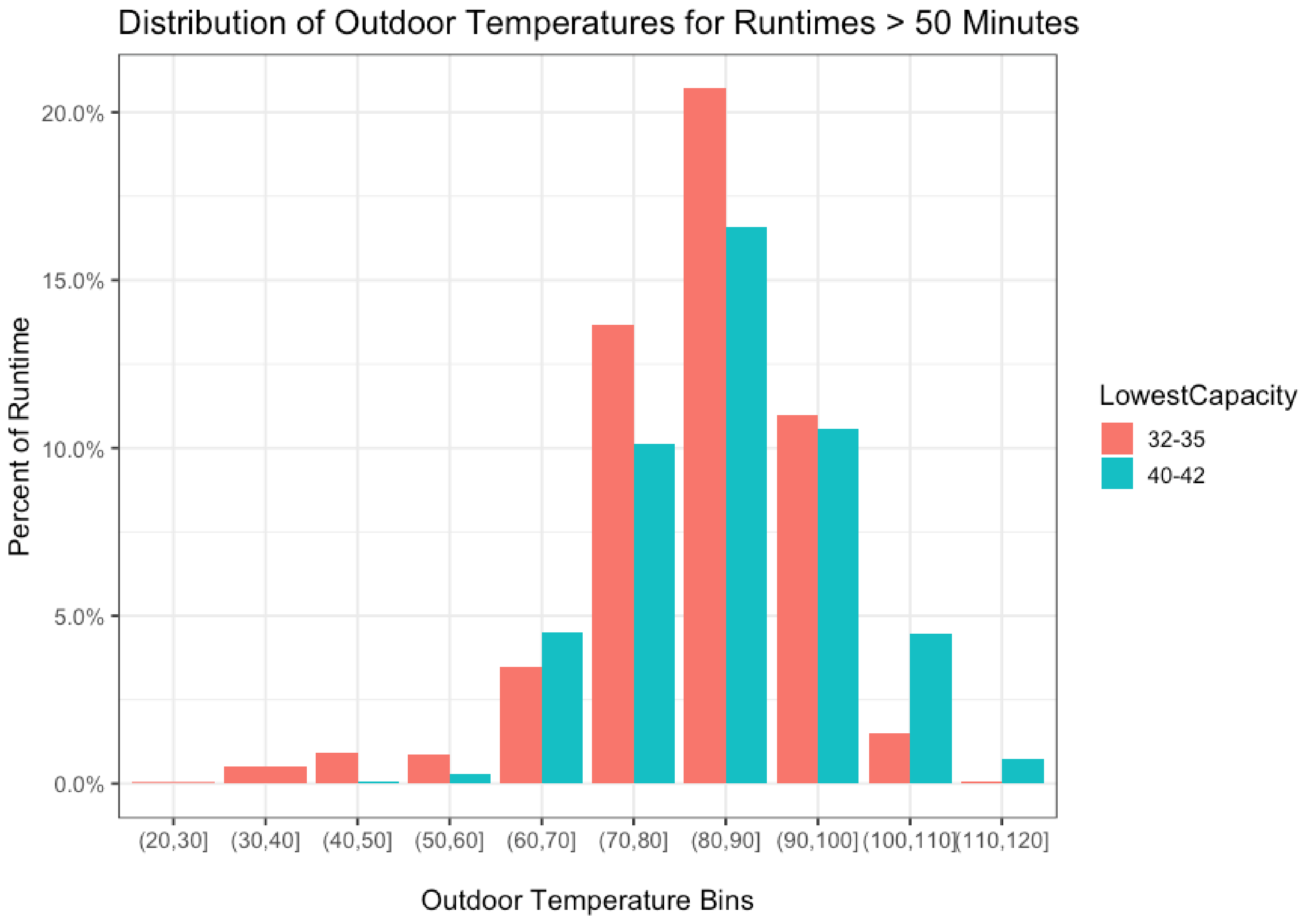
- If we are controlling for individual household conditions, which variables should be controlled for in estimating average capacity factor? (multiple select)
  - System minimum capacity call (turndown ratio)
  - System sizing relative to heating/cooling load
  - Humidity
  - More complicated household model needed
  - None
  - Other

Poll

- How promising do you think the Average Capacity Factor is for gauging the performance of variable capacity HVAC systems?
  - Not promising
  - There's promise but it still needs a lot of work
  - It seems promising

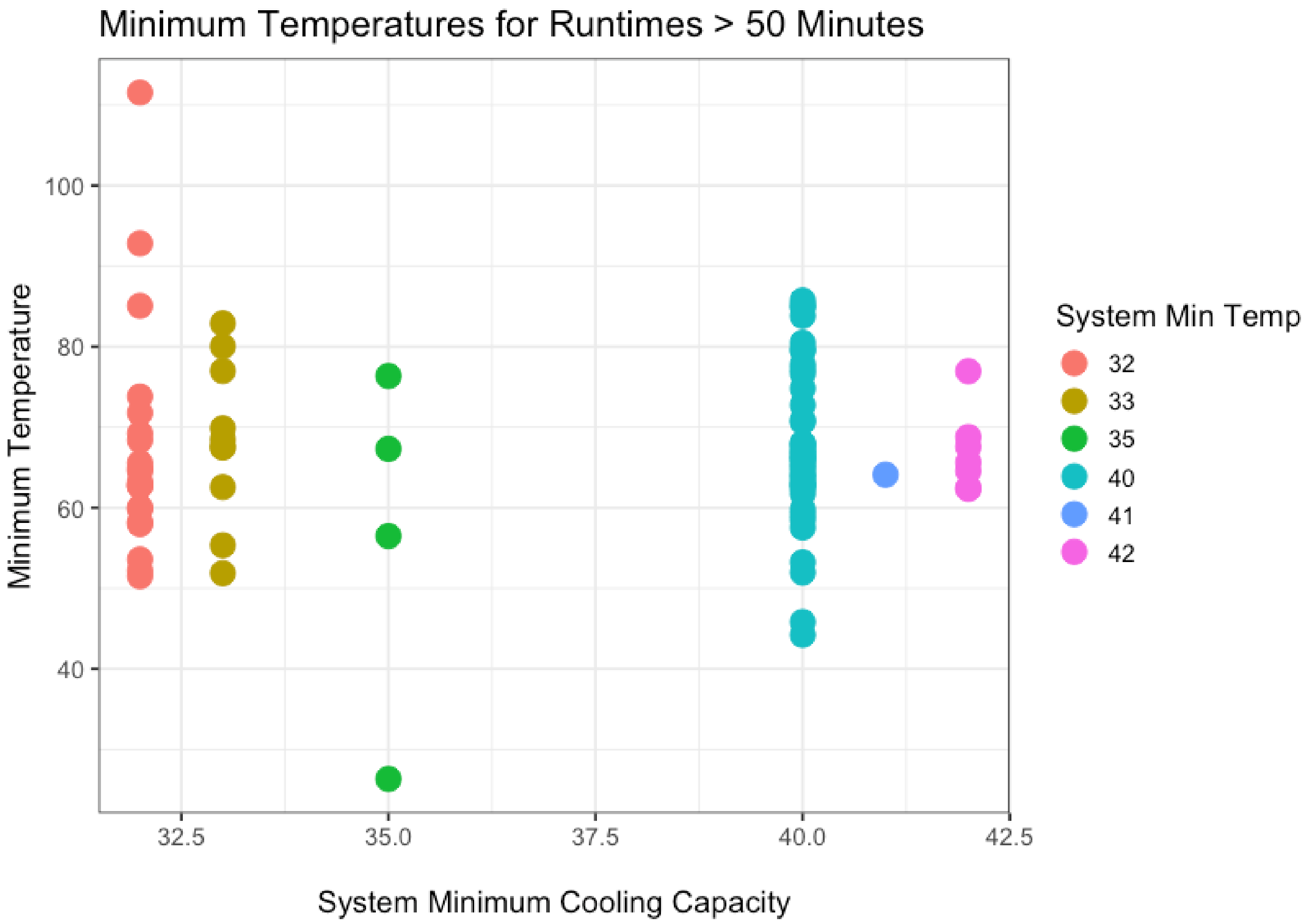
Minimum temperature for uninterrupted cooling

- Normal distribution of percent of runtime associated with outdoor temperature
- Distribution is shifted slightly left for lower capacity systems



Minimum temperature for uninterrupted cooling

- Large amount of variability when we look at individual systems
- No clear minimum temperature at which a system can run for long periods without cycling
- What would we need to control for to get a more accurate estimate of the minimum temperature?



Discussion: Is it possible to estimate system sizing from data?

- Proposed method:
  - Look at average runtime for hottest 5th percentile outdoor temperatures for each system
  - Bin into categories of undersized, adequately sized, oversized, or very oversized based on average runtime

Discussion: Variable capacity metrics

- Sizing will affect this, as will the set point, and the characteristic of how much the home self-heats
- There may be differences based on where the thermostat is positioned relative to distribution equipment
- Whatever variation we are seeing is from installation to installation – each teal point is for the same equipment model number and a thermostat from the same vendor

Discussion: Variable capacity metrics (system sizing)

- 5th percentile should be able to detect an adequately sized unit by looking for long run times at 100% capacity
- Do we have multiple years of data? Some years just don't hit an extreme. Better to look for the design temperature for the particular locale and compare that to what is included is in the data trace.
- Design temperatures are based on a looked up in a table based on location, but they may be outdated due to climate change. Note that if system is 100% at design temperature, it will never be able to recover. The actual sizing recommendations take this into account.
- Manual J is known to oversize air conditioners by 30-40%. Most systems run 15-16 hours on design days.
- Rule of thumb: to recover 1F in an hour takes approximately a ton of capacity on an extreme (design-type) day, which speaks to typical thermal capacity of the home
- Note that these thoughts may not apply quite as well the variable capacity systems, because they can be overdriven to provide additional capacity at lower efficiency

Connected Thermostat Use-Cases

ENERGYSTAR's goal is to maintain a level playing field for vendors. To this end, it will consider various strategies to prevent distorted results.

CTs are being used in situations beyond the simple single-family home. These alternative use-cases can influence the vendor's calculation of the metric in two ways:

- Certain use-cases will result in misleading metrics
- A vendor will have a distribution of use-cases that differ from other vendors

Does the metric accurately capture CT performance in common use-cases?

How are CTs being used?

- Building types (single-family, apartments, vacation homes)
- Ownership scenarios (one per home, multiple)
- Other (home and customer different)

Does the existence of these configurations suggest alternative sampling procedures?

Which inputs will generate a misleading metric calculation?

- Unrealistic comfort temperatures?
- Unusual indoor or outdoor temperatures?
- Unusual relationship between temperature and runtimes?
- Other?

#	Use Case	Does Current ES metric make sense? (Y/N/ Maybe)	Notes/ Explanation/ Drawbacks	Fraction of CTs in this category	Type of problem (Sample or metric calculation)

#	Use Case	Does Current ES metric make sense? (T_{90}, runtime, ΔT)	Notes/ Explanation/ Drawbacks	Fraction of CTs in this category	Problem in Sample or Metric?
1	SF detached home (1 tstat)		Our base case, single or dual-speed, unspecified auxiliary heating source,		
2	Vacation home				
3	SF home (>1 tstat)				
4	Multiple thermostats on a single account		(like a motel? Dorm?)		
5	SF home with multiple temperature sensors				
6	Small commercial with own HVAC				
7	Apartment with own HVAC				
8	Duplex home, multiple thermostats, different accounts, same dwelling		Variation on the Apartment idea above		
9	Variable capacity heating or cooling		We're investigating a metric for effectiveness of variable capacity		
10	2-stage system				
11	Dual fuel		They are currently excluded		

Discussion: Connected thermostat use cases (reviewing 1-3)

- Newer homes more likely to be set up for multiple units, especially if they're two stories.
- Larger houses don't work well with a single system
- Vendors who had examined homes with two thermostats found that the two reacted very differently
- Might make sense to just run the metric on a sample that's all 2-thermostat houses and compare to a sample that's all single-thermostat homes
- No guarantee that we can correct for this effect anyway – we definitely won't know all of the homes that have more than one thermostat
- Another interesting question – sample by thermostat or by home?
- Is this something that the NEEA project can/will address? Jon Koliner thinks he has this data in the dataset
- Small effect: looks like multi-thermostat homes have wider variation of scores, slightly better scores for heating and worse for cooling at first look. Is this because multi-thermostat homes will tend to be bigger and newer

Discussion: Connected thermostat use cases (5: SF home w/multiple sensors)

- What fraction of tstats have multiple sensors?
- If sensors are averaged, does it matter?
- Reported temperature is supposed to be the one that's driving the control, whether it's always the same one, an average of all, or determined by some special sauce the vendor uses
- Use of multiple sensors might shift the tau
- If the temperature sensor (or combo) used depends on time, it might just look like noise on tau; less effect on alpha
- If people aren't comfortable wherever they are with the combo of sensors used, they might just jiggle the thermostat setting – also noise
- In general, adding a remote temp sensor makes your home less efficient, but makes the metric score go up (10ths of degrees change in comfort temp relative to average temp)

Discussion: COVID

- ASHRAE recommends more ventilation; will affect energy used for cooling and heating
  - Suggests “higher ventilation” use case

- Ran metric scores for summer 2020 and found relatively small effect of pandemic home use changing