ENERGY STAR Connected Thermostats

Stakeholder Working Meeting

March 08, 2019

Attendees

Abigail Daken, EPA
Dan Baldewicz, ICF for EPA
Alan Meier, LBNL
Leo Rainer, LBNL
Michael Blasnik, Google/Nest
Jing Li, Carrier
Tai Tran, Carrier
Brian Rigg, JCI
Kurt Mease, LUX (JCI)
Diane Jakobs, Rheem
Carson Burrus, Rheem
Chris Puranen, Rheem
Glen Okita, EcoFactor
Brent Huchuk, ecobee
John Sartain, Emerson
James Jackson, Emerson
Mike Lubliner, Washington State U
Charles Kim, SCE
Michael Fournier, Hydro Quebec
Ed Pike, Energy Solutions for CA IOUs
Nick Lange, VEIC
Dan Fredman, VEIC
Rober Weber, BPA
Phillip Kelsven, BPA
Casey Klock, AprilAire
Wade Ferkey, AprilAire
Ethan Goldman, OpenEE
Youssef Jaber, IRCO/Trane
Behrooz Karimi, IRCO/Trane
Ulysses Grundler, IRCO/Trane
Mike Caneja, Bosch

Agenda

- Resistance Heating Utilization
  - T Intervals for N <30
- Regional Baselines + Metrics Discussion
  - LBNL: Leo Rainer

RHU Data Recap

- **Previous RHU Datacall:**
  - Statistical significance between datasets
    - (Multiple) Climate Zones
    - (Multiple) Temp Bins
    - (Multiple) Products
  - Oversampled data has the clearest distinctions
  - Low product sample adjustment:
    - Use T Test Confidence Interval for N < 30

- **RHU Open Questions:**
  - Statistically significant differences in products:
    - In Oversampled Data? Standard Data?
  - Differences in certain temp bin groups? Climate Driven?

Charts: **R** –Ove**R**sample. **S** – Standard.
Paired – Only datasets with corresponding Oversample.

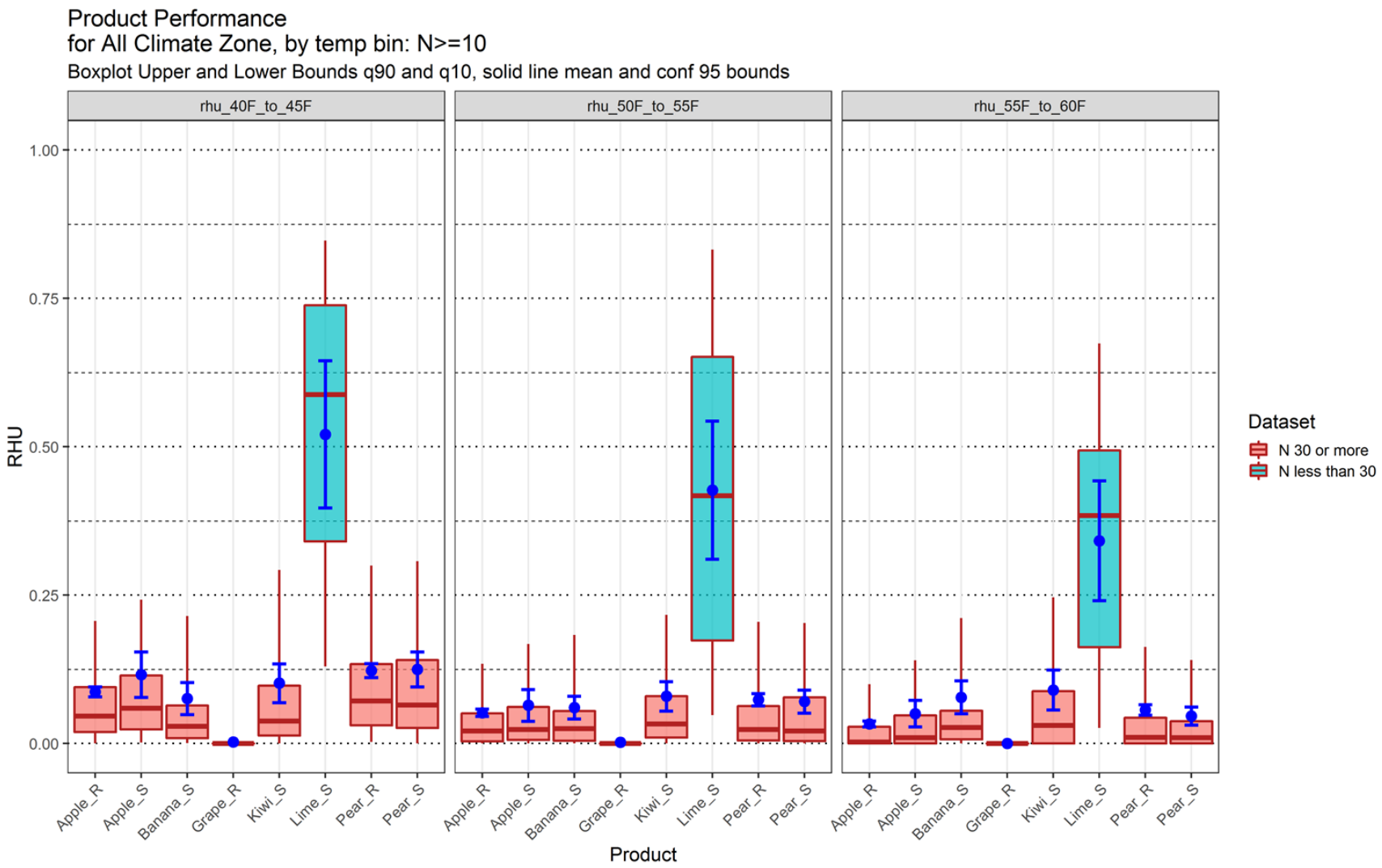
Data Observations

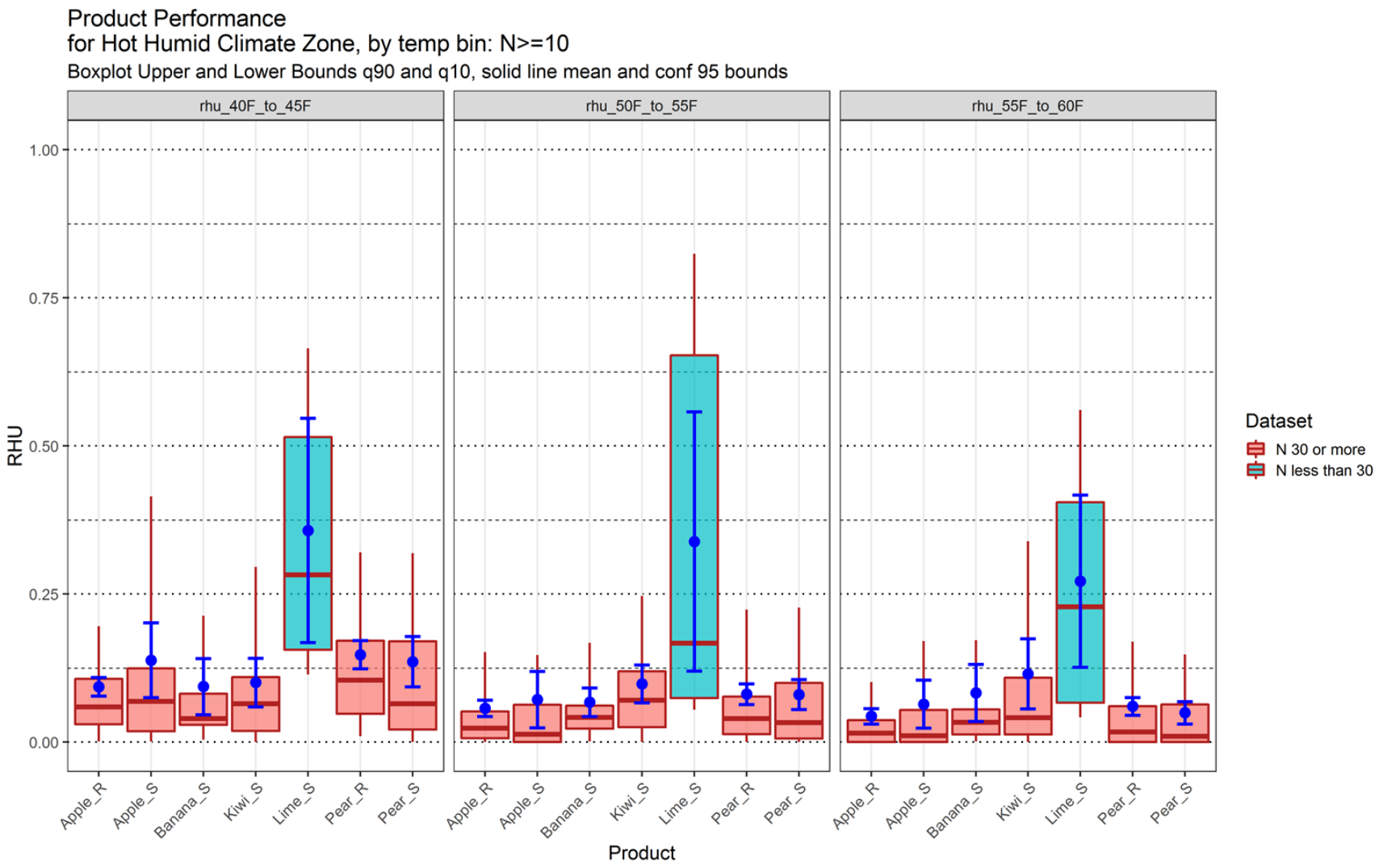
- T Test Confidence Intervals:
  - Wider CI95 than comparable normal (z) CI95 by design
  - N < 30 data requires T Test

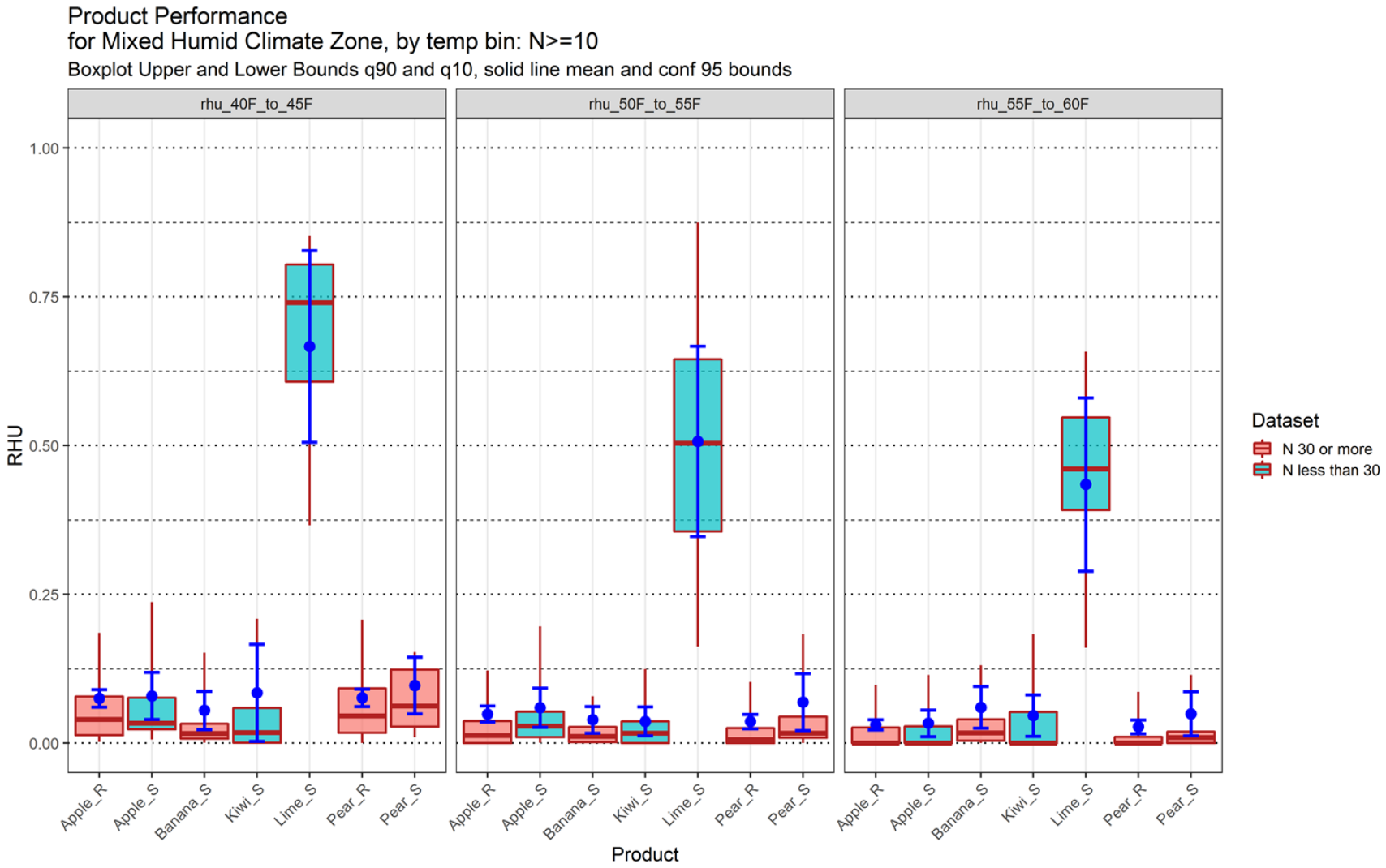
- RHU results
  - Oversampled data has advantage over standard sample on CI95
  - IQR can be helpful at times, some distributions are shifted even in cases of non-sig CI95

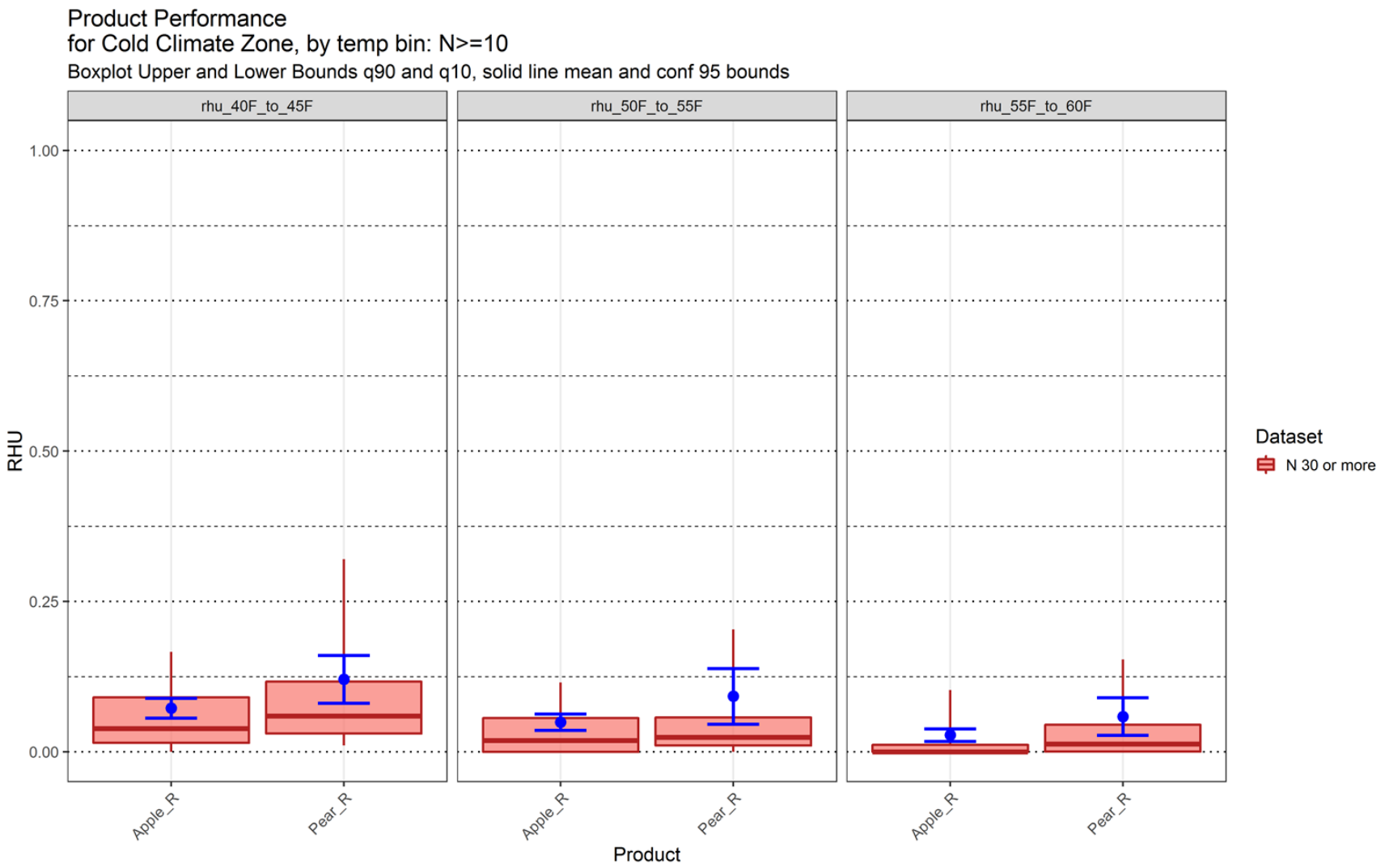
Data Observations: High Temp Bins

- All
  - Min differences: whether q25 (bottom of box) is on zero, or shifted above (~0.05)
  - Large variation on max and q75
  - Sig. means, especially between oversampled data
- Hot Humid
  - Oversample needed for CI95 significance
  - Variations in IQR (box length)
  - Some products can lock out RHU usage in certain bins
- Cold: Oversample only, not enough HP products in this region


Product Performance
for All Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_40F_to_45F
rhu_50F_to_55F
rhu_55F_to_60F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Grape_R
Kiwi_S
Lime_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

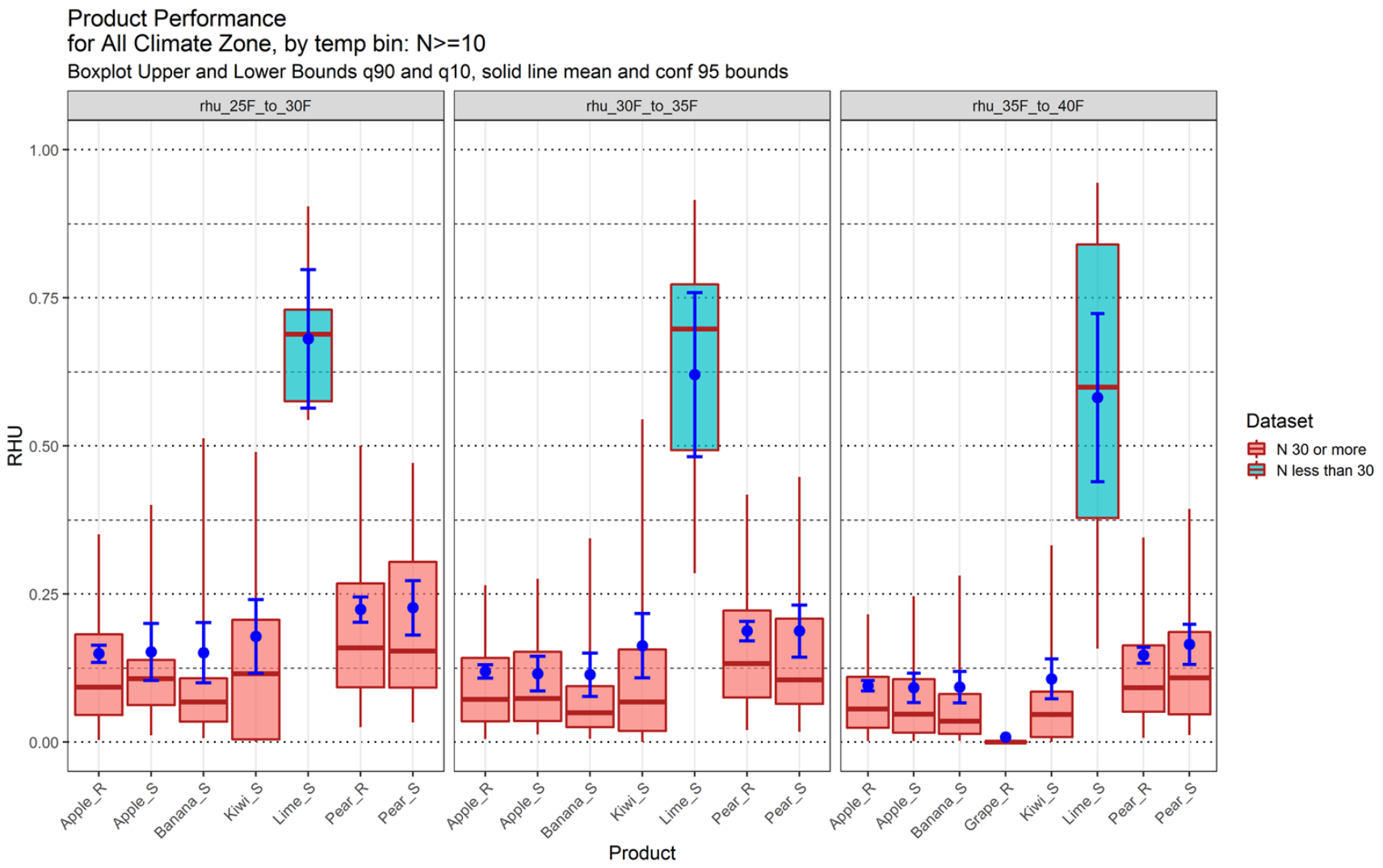

Product Performance
for Hot Humid Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_40F_to_45F
rhu_50F_to_55F
rhu_55F_to_60F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Kiwi_S
Lime_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

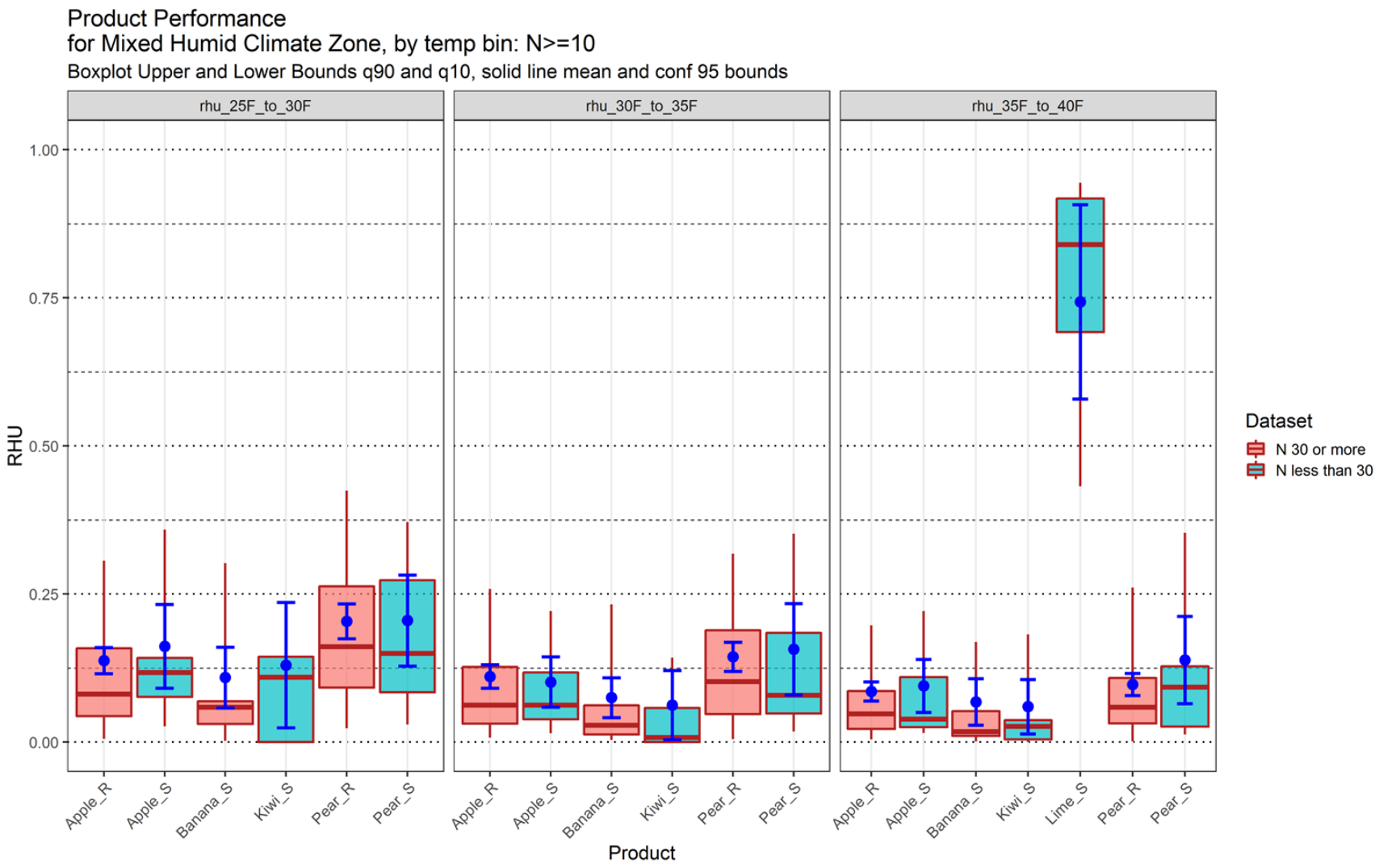

Product Performance
for Mixed Humid Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_40F_to_45F
rhu_50F_to_55F
rhu_55F_to_60F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Kiwi_S
Lime_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

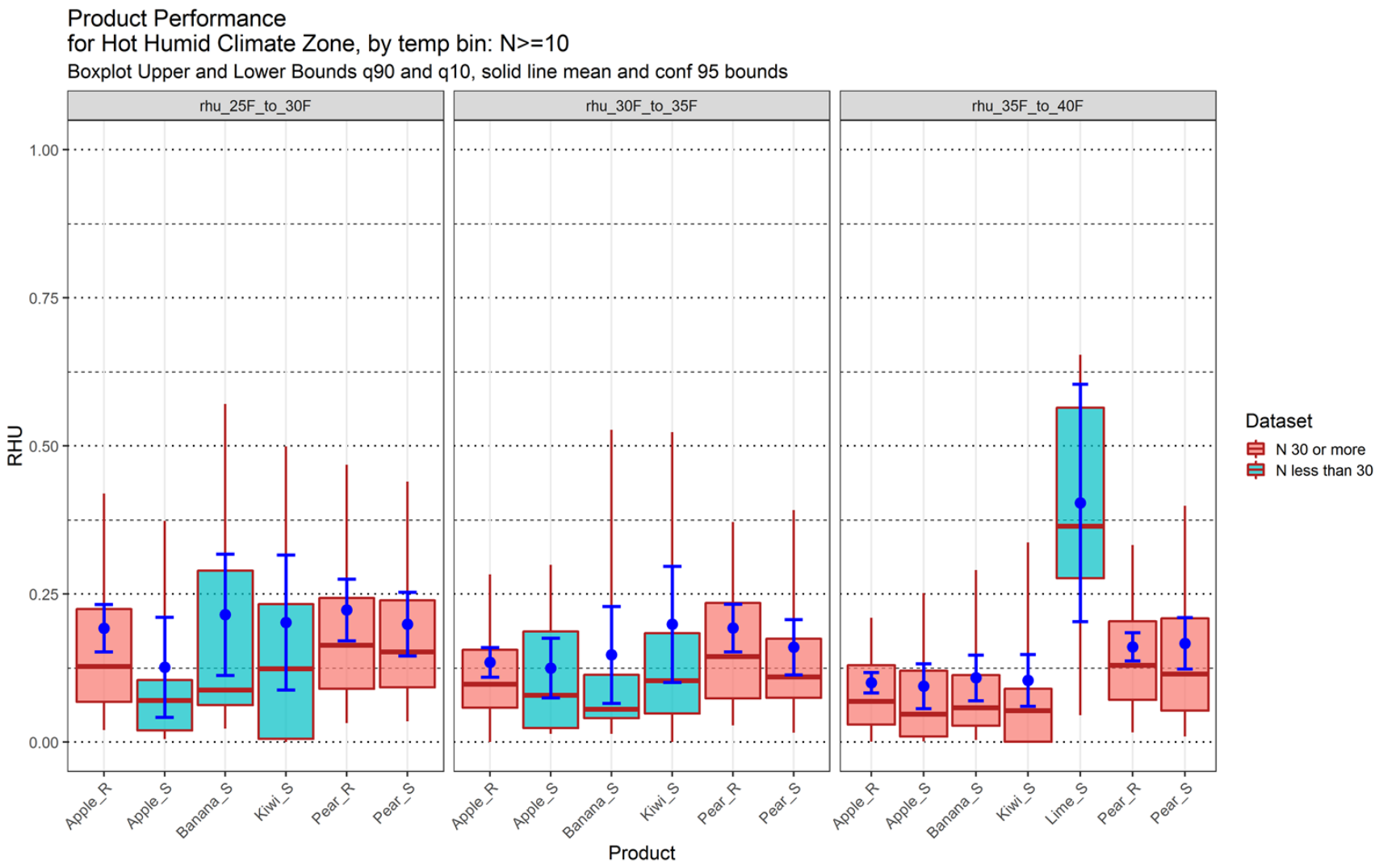

Product Performance
for Cold Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_40F_to_45F
rhu_50F_to_55F
rhu_55F_to_60F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Pear_R
Product
Dataset
N 30 or more

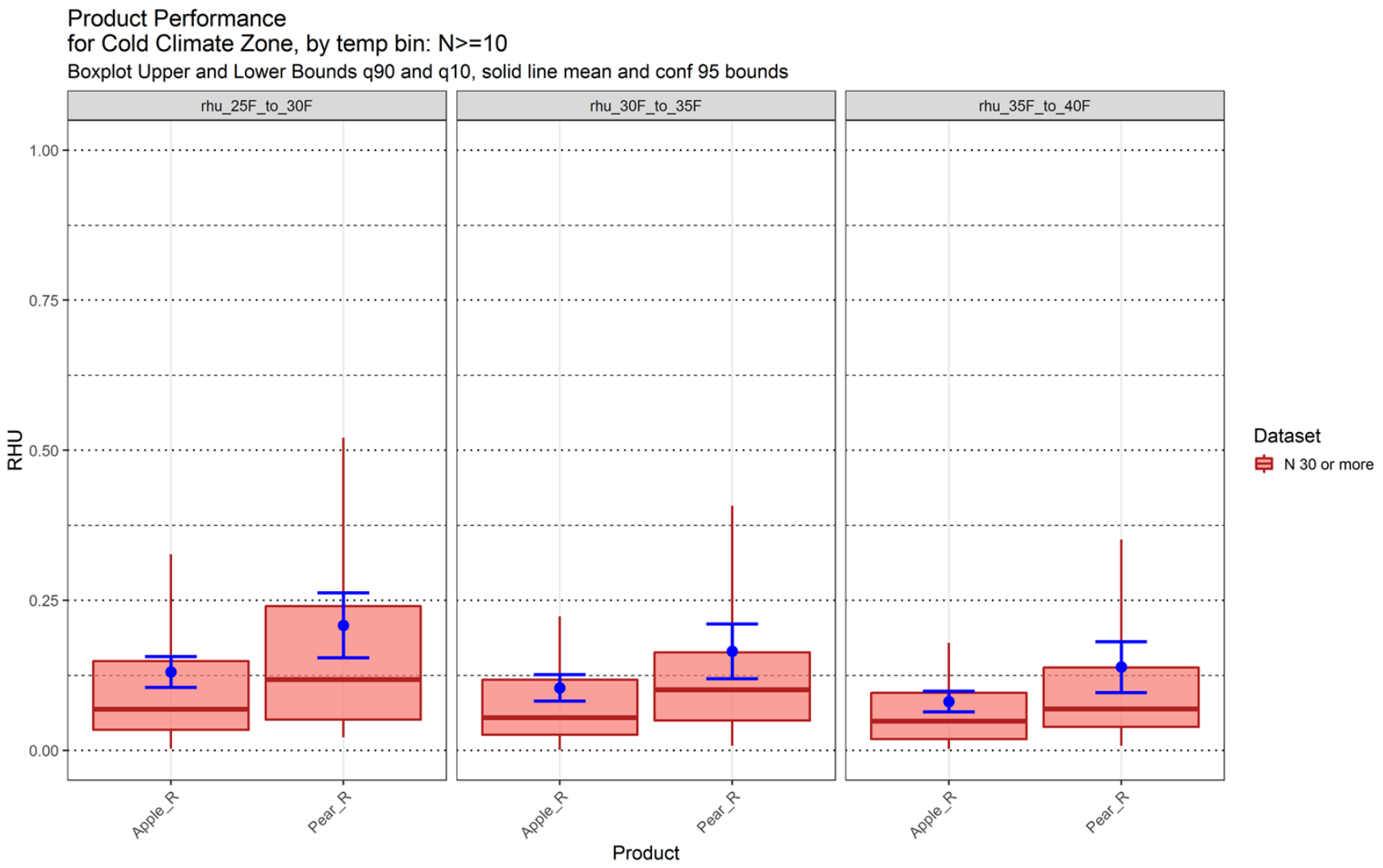
Data Observations: Mid Temp Bins

- All
  - Some cases of standard sample statistical sig. CI95
  - More clear with oversamples
  - Q25 (low RHU quartile) surpasses medians of other distributions
- Mixed Humid
  - Oversample needed for statistical sig. CI95
  - Some non-overlapping IQRs, where CI95 sig. not confirmed.
- Hot Humid
  - Few stat sig. CI95s, even with oversamples.
  - HP's appear to be very competitive at these bins


Product Performance
for All Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_25F_to_30F
rhu_30F_to_35F
rhu_35F_to_40F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Grape_R
Kiwi_S
Lime_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

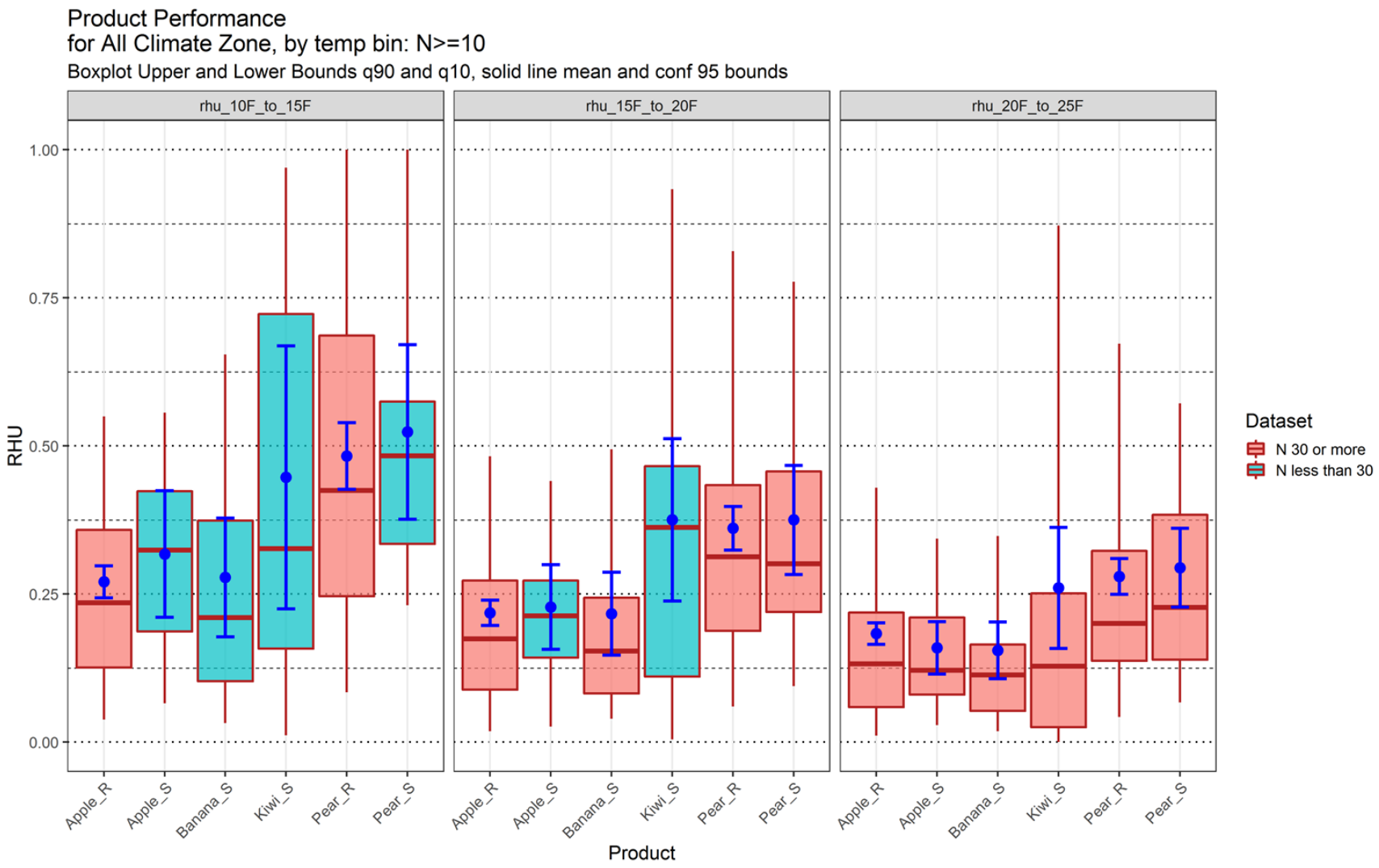

Product Performance
for Mixed Humid Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_25F_to_30F
rhu_30F_to_35F
rhu_35F_to_40F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Kiwi_S
Lime_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

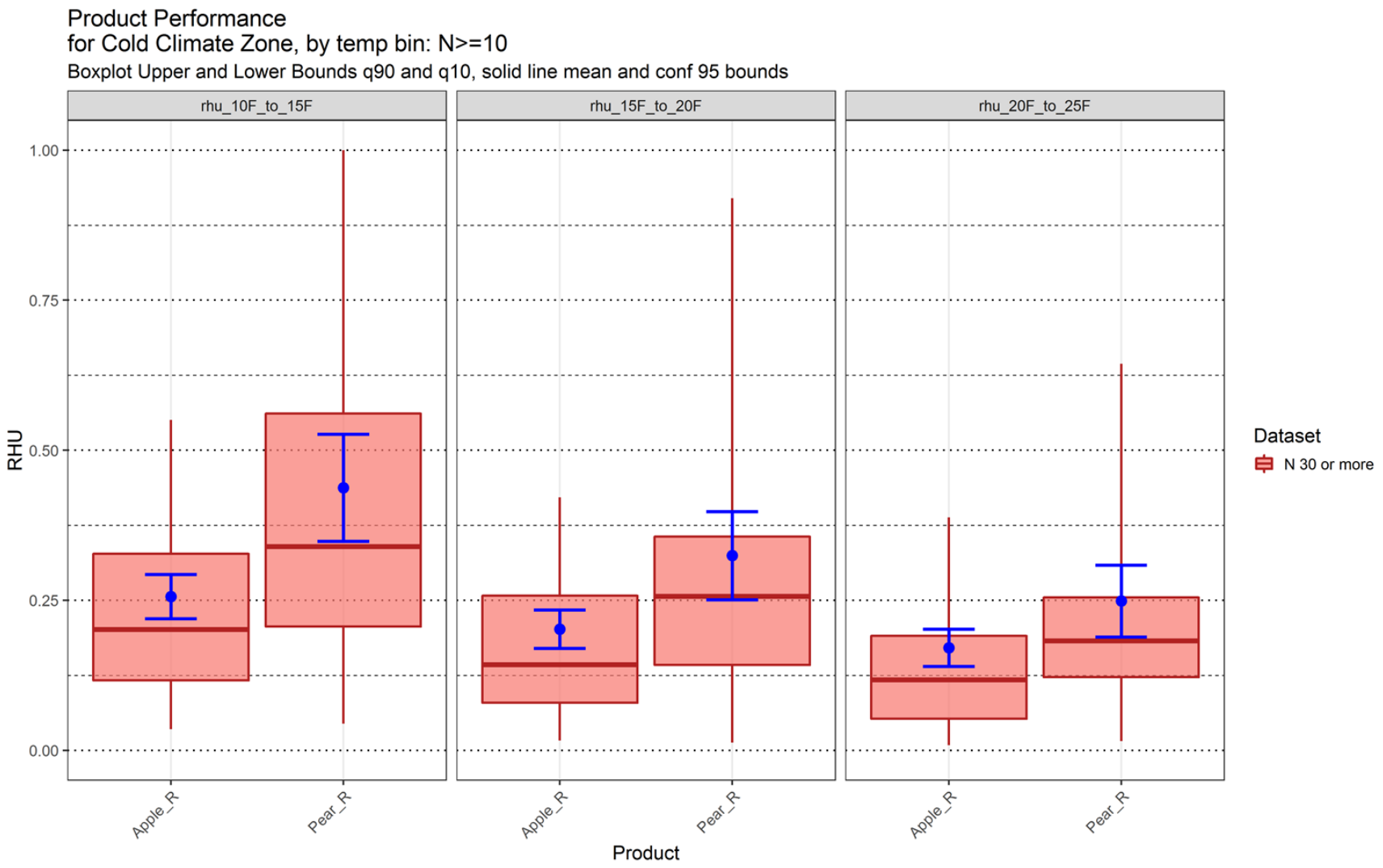

Product Performance
for Hot Humid Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_25F_to_30F
rhu_30F_to_35F
rhu_35F_to_40F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Kiwi_S
Lime_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

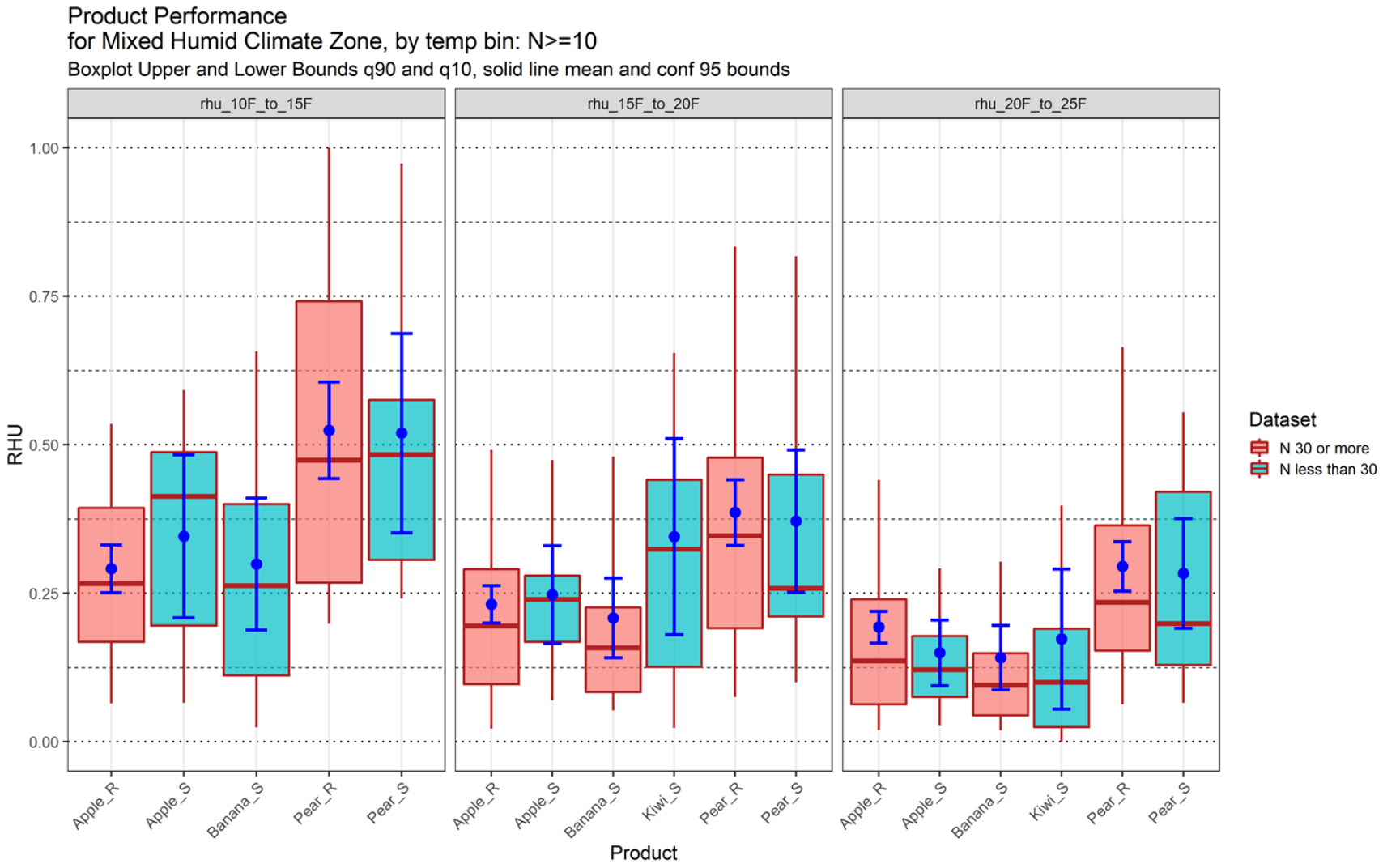

Product Performance
for Cold Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_25F_to_30F
rhu_30F_to_35F
rhu_35F_to_40F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Pear_R
Product
Dataset
N 30 or more

Data Observations: Low Temp Bins

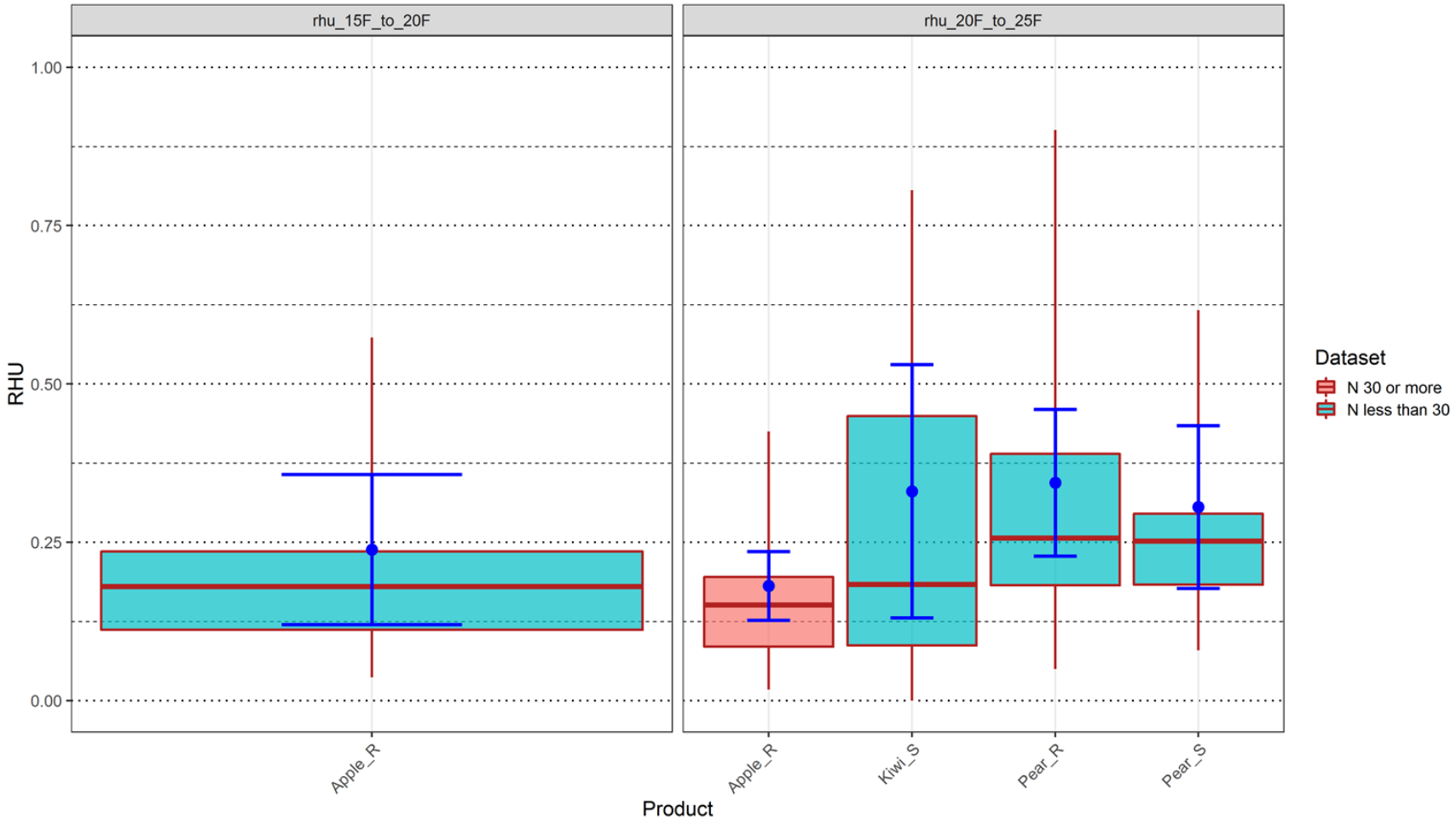
- All
  - Sig CI95s, both oversample and standard. More oversample sig.
- Cold Climate
  - Oversample needed to have enough data
  - Clear differences, statistical sig. CI95s
  - Distribution shifts, median passes q75 of other product
- Mixed Humid
  - Oversample needed for statistical sig. CI95s
  - Distribution shifts, median passes q75 of other product
- Hot Humid: Not much data to draw conclusions in this zone


Product Performance
for All Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_10F_to_15F
rhu_15F_to_20F
rhu_20F_to_25F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Kiwi_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30


Product Performance
for Cold Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_10F_to_15F
rhu_15F_to_20F
rhu_20F_to_25F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Pear_R
Product
Dataset
N 30 or more


Product Performance
for Mixed Humid Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds
rhu_10F_to_15F
rhu_15F_to_20F
rhu_20F_to_25F
RHU
1.00
0.75
0.50
0.25
0.00
Apple_R
Apple_S
Banana_S
Kiwi_S
Pear_R
Pear_S
Product
Dataset
N 30 or more
N less than 30

Product Performance
for Hot Humid Climate Zone, by temp bin: N>=10
Boxplot Upper and Lower Bounds q90 and q10, solid line mean and conf 95 bounds



RHU Discussion

- What about dual fuel? In theory, dual fuel are excluded from the dataset, so CT service providers need to know which installations are dual fuel. Under the assumption that everyone is doing this right, this is all resistive back up.
- Does it make sense to look at confidence intervals with outlier-contaminated data and skewed distribution?
  - Could be over 5% of installations have problems, so that the RHU is very high because the heat pump is broken, or was installed improperly.
  - Fairly well behaved ¾ of the data, and then 10%, 20%, etc.
  - Can we filter in the metrics software?
- Is sizing also a consideration, e.g. undersized heat pumps in cold climates, that use a lot of strip heat?
  - Can we distinguish between an undercharged and an undersized systems?
  - Also, the aux heat sizing also introduces variation
- Instead of filtering outliers, could we asses the actual distribution?

RHU Discussion

- Partly this is a small sample problem. OK, how large a sample would you need?
  - Really need to look at data to know?
  - You would need a really huge sample if you want to include outliers
- Avenues to redress long resistance heat run times:
  - Ramping set points carefully to avoid RHU, nudging, etc.
  - Message customers who have high RHU, attempt to get them to fix their systems. To reward this, we would need to keep outliers in the data set.
- Another noise source is weather – if a polar vortex comes in, and the system time in a temperature bin is well outside its design temperature, well there will be a lot of aux heat and that doesn't necessarily indicate a problem.
- Defining outliers: 2-3 interquartile rangers beyond the median.
- More detail on the distribution? (Could be a fork in the code)
- Widen temperature bins for more thermostats per bin? Proposal: <10F bin, 10-20, 20-30, 30-35, 35-40, 40-45,

RHU Discussion

- Most of the information we really want is visible in the 30 degrees and below, where compressor lockouts happen.
- Could we modify the bins to take the design temperature into account, e.g. 5F below design temp, etc.
- Another way is to set up so that the bins are 20% of hours in each bin. This means that the temperature edges of the bins for each thermostat would be different, which would make looking at them together iffy. We could do something similar for the entire climate zone and use different bins for each zone.
- Is a thermostat with 1 hour in a bin weighted the same as a thermostat that has 100 hours in the bin? Yes. We might be able to weight more heavily those thermostats with many hours in each bin. To do this right, we need the compressor run time hourly, not just daily.
- Ask for total res heat hours and compressor hours? Others say not so useful.
- Process for how to decide what programming to ask for in the next week?

RHU Discussion

- Filter so that you only include thermostats with a minimum number of hour in the temperature bin?
- Weight by number of hours that thermostat had in the bin? Or, average thermostat-hours in the bin, instead of thermostats.
- At least know the average thermostat hours in the bin? That would let us know if we want to ignore the bin. Separate step would be how we weight or roll up to get a meaningful conclusion.

CT Metric Discussion

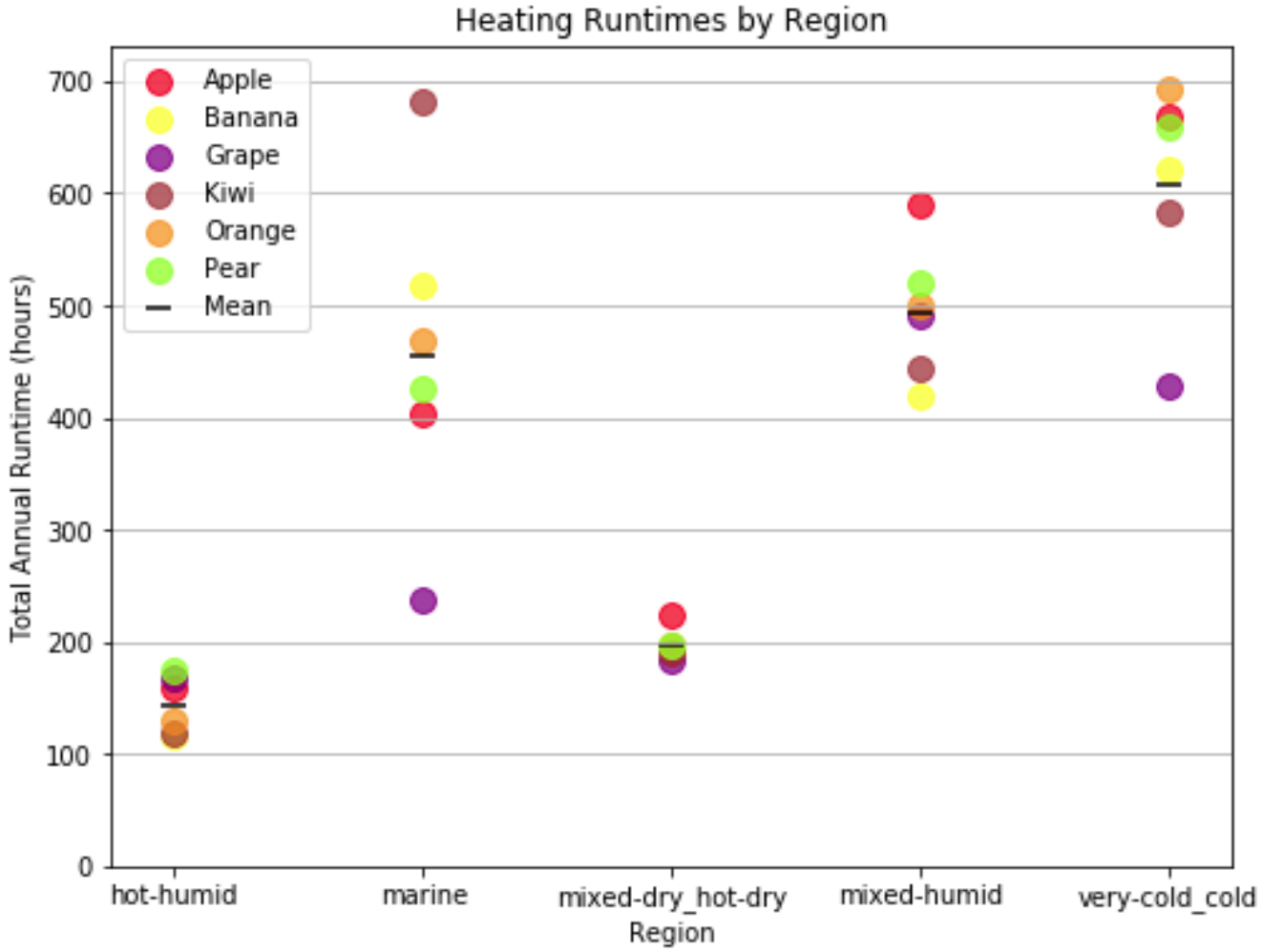
Leo Rainer and Alan Meier, LBNL
March 8, 2019

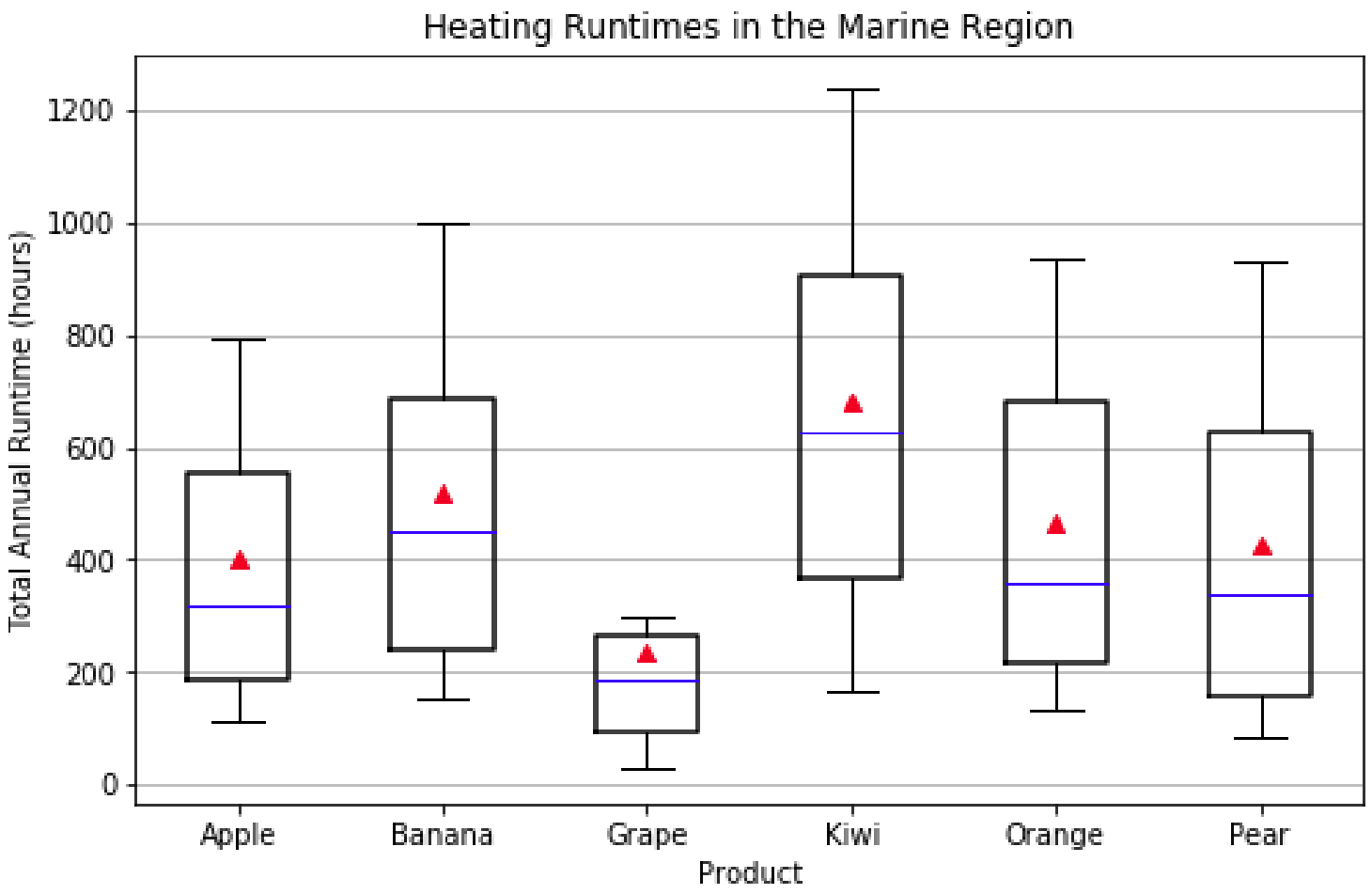
Metric Options

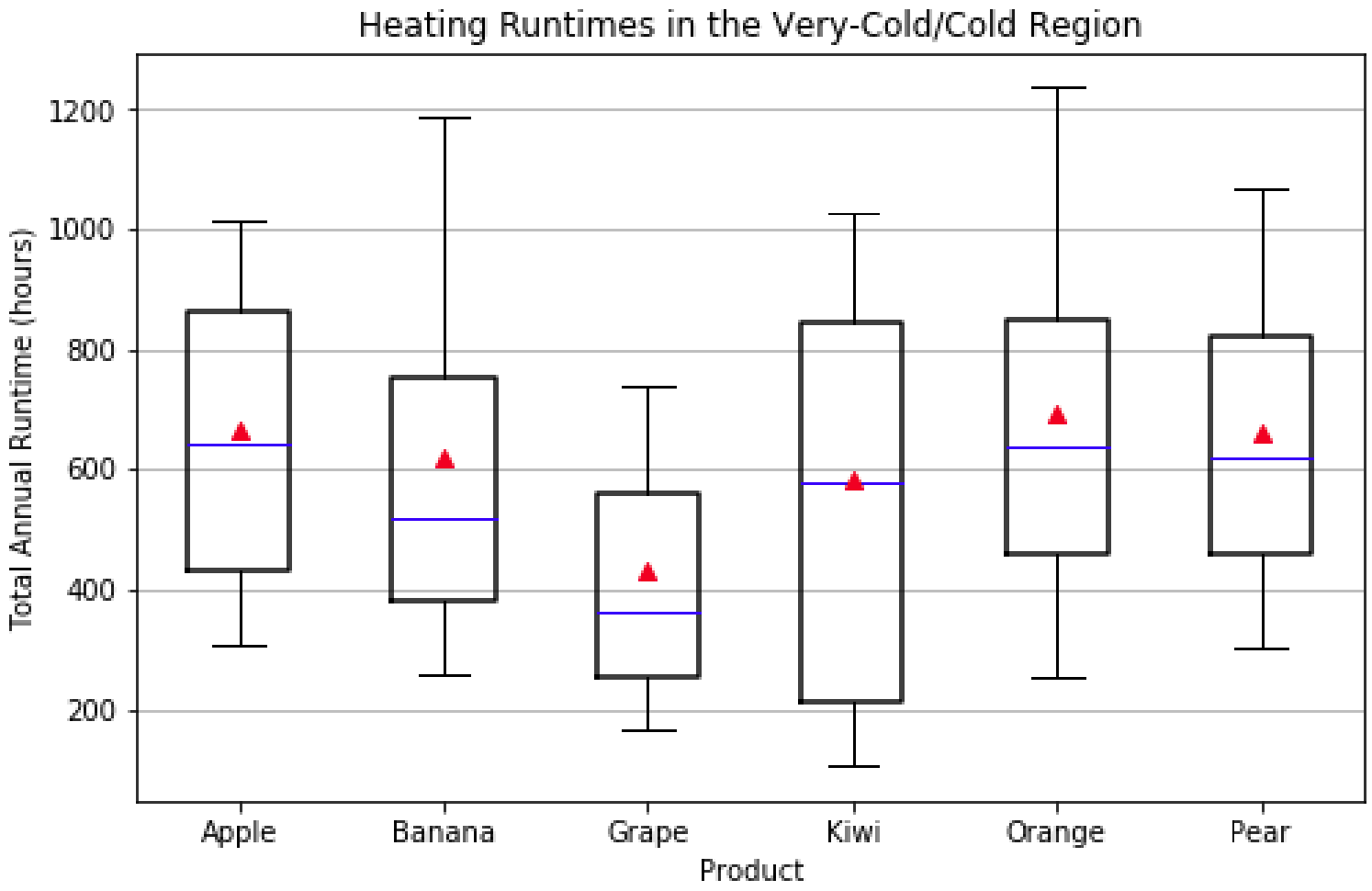
Metric	Description
Current	Runtime reduction calculated using self-referential (90/10) comfort temperatures
Regional Baseline	Runtime reduction calculated using regional baseline temperatures
Indoor Temperatures	Maintained indoor temperatures during core or operating hours
Equipment Runtimes	Gross or core cooling and heating equipment runtimes
Hybrid	A weighted combination of the above four metrics

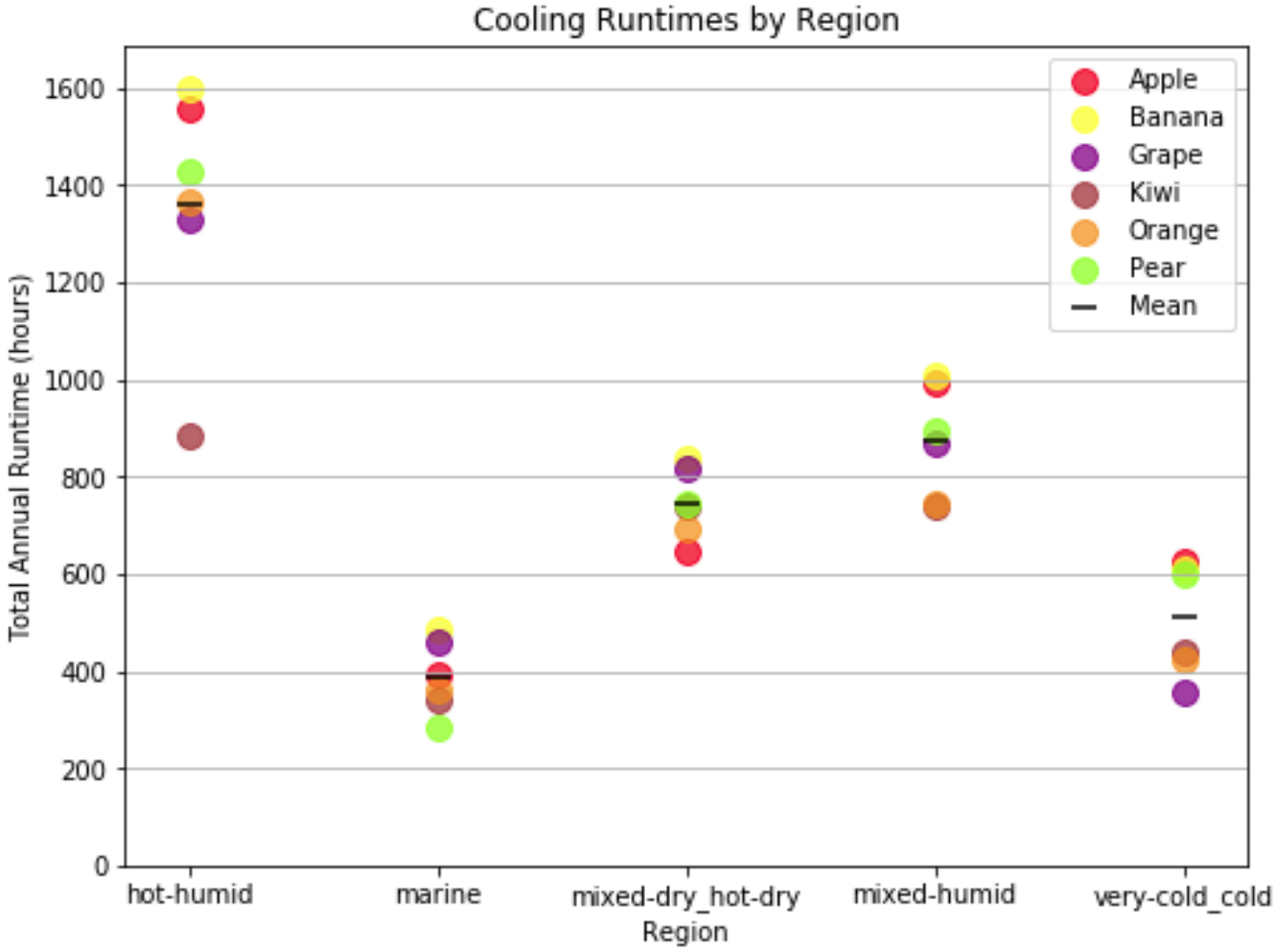
Metric Advantages and Disadvantages

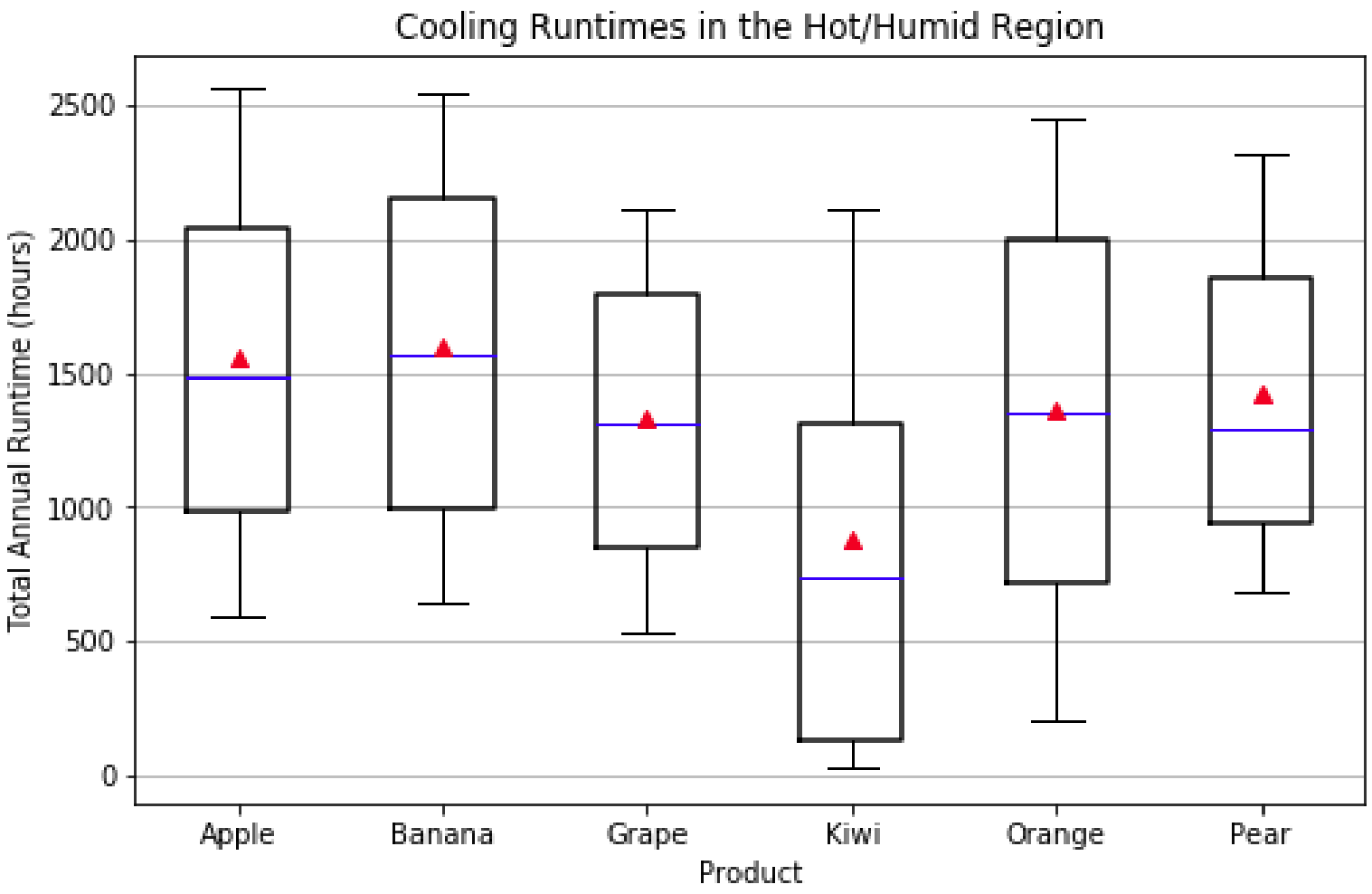
Metric	Advantages	Disadvantages
Current	Not affected by differences in customer base. Separates equipment choice from equipment operation.	Captures savings only from temperature choices. Only rewards setback savings.
Regional Baseline	Fixed and regionally responsive	No clear relationship between regional data set and vendor submitted sample data
Indoor Temperatures	Independent of house characteristics. Valid for all system types.	Does not capture savings from better HVAC control. Does not directly estimate energy savings.
Equipment Runtimes	Captures savings from HVAC control. Directly related to energy use.	Hard to separate equipment choice from equipment operation. No good choice of baseline.

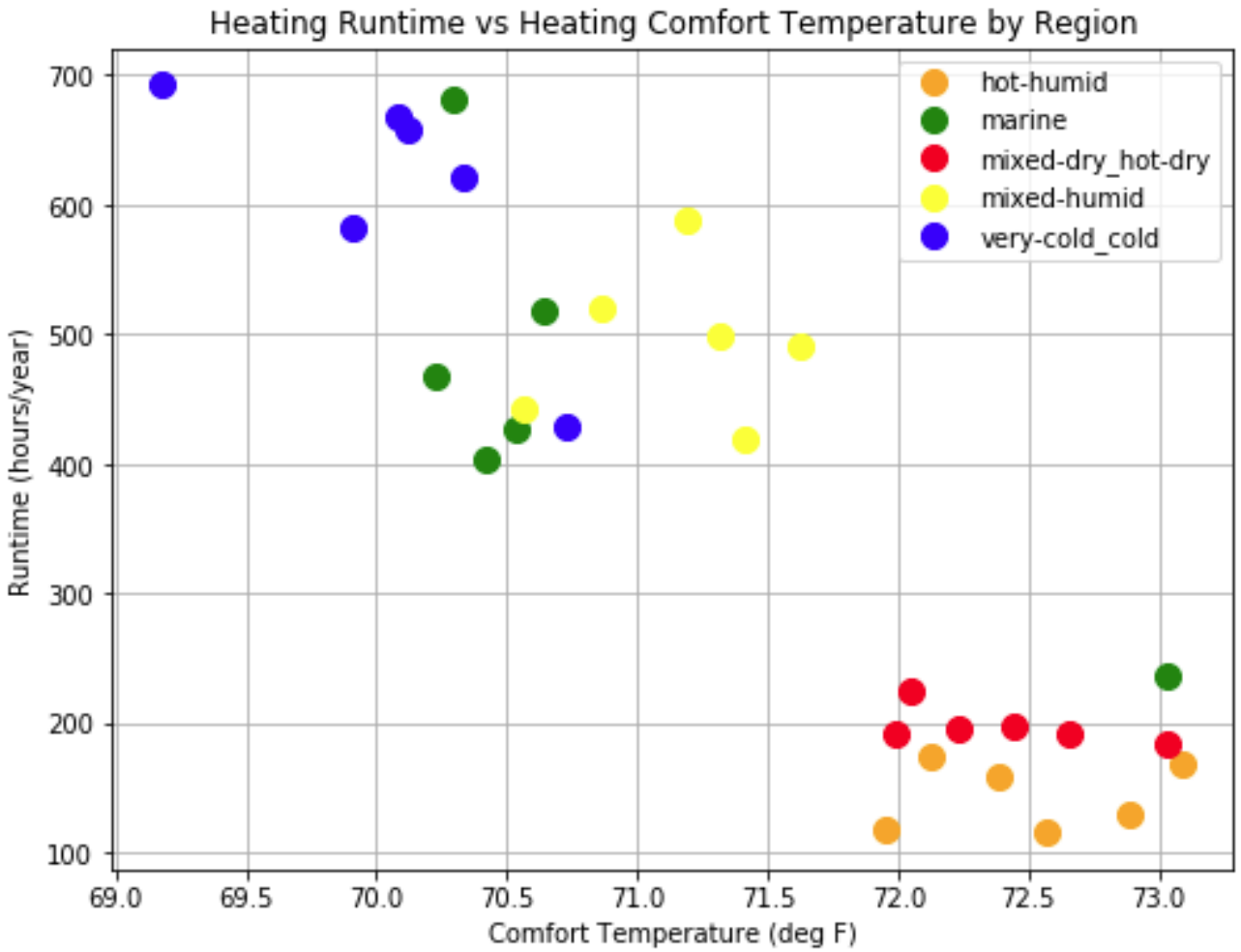

Heating Runtimes by Region
Apple
Banana
Grape
Kiwi
Orange
Pear
Mean
Total Annual Runtime (hours)
0
100
200
300
400
500
600
700
hot-humid
marine
mixed-dry_hot-dry
mixed-humid
very-cold_cold
Region

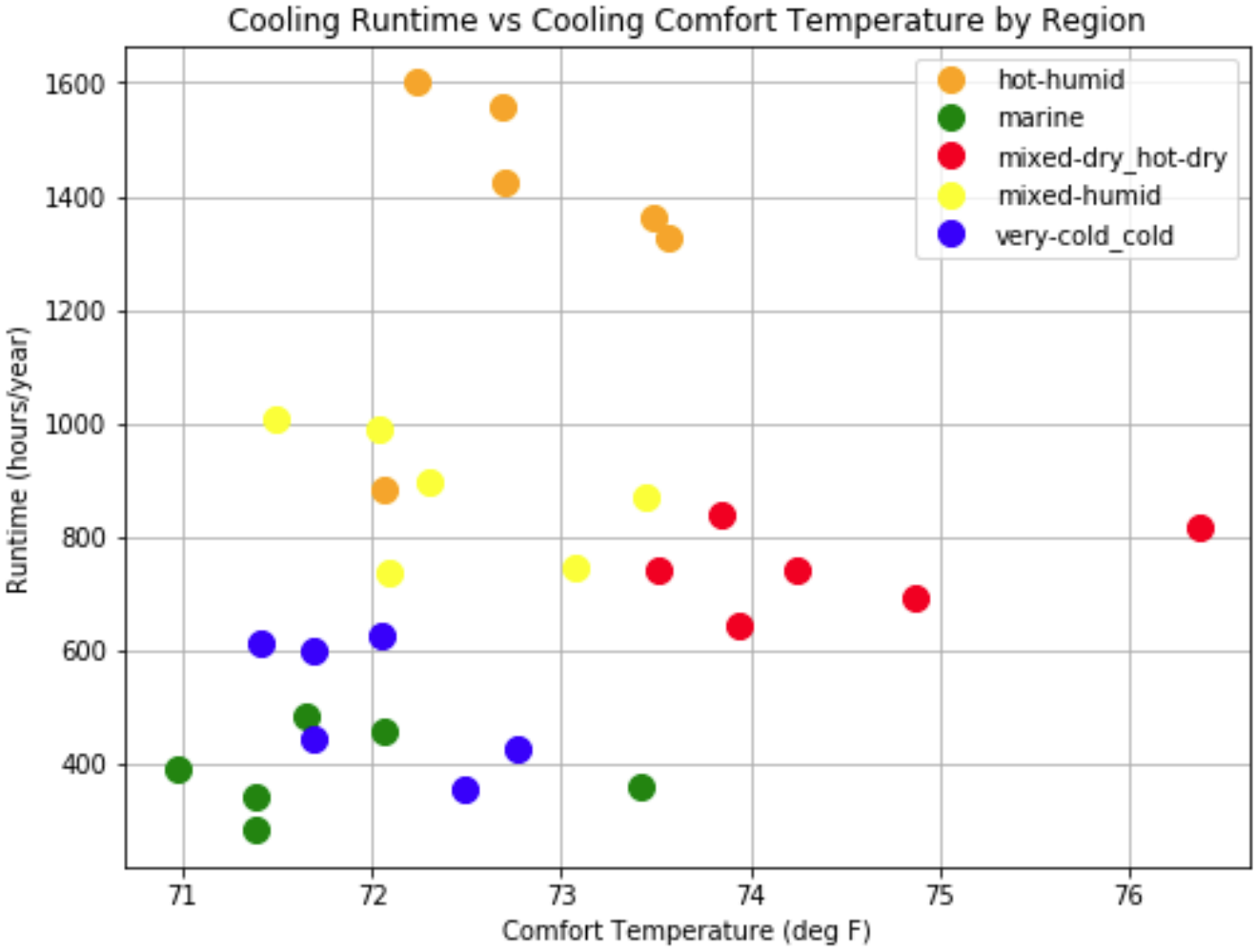

Heating Runtimes in the Marine Region
Total Annual Runtime (hours)
0
200
400
600
800
1000
1200
Apple
Banana
Grape
Kiwi
Orange
Pear
Product


Heating Runtimes in the Very-Cold/Cold Region
Total Annual Runtime (hours)
1200
1000
800
600
400
200
Apple
Banana
Grape
Kiwi
Orange
Pear
Product


Cooling Runtimes by Region
Total Annual Runtime (hours)
0
200
400
600
800
1000
1200
1400
1600
hot-humid
marine
mixed-dry_hot-dry
mixed-humid
very-cold_cold
Region
Apple
Banana
Grape
Kiwi
Orange
Pear
Mean


Cooling Runtimes in the Hot/Humid Region
Total Annual Runtime (hours)
2500
2000
1500
1000
500
0
Apple
Banana
Grape
Kiwi
Orange
Pear
Product


Heating Runtime vs Heating Comfort Temperature by Region
Runtime (hours/year)
Comfort Temperature (deg F)
700
600
500
400
300
200
100
69.0
69.5
70.0
70.5
71.0
71.5
72.0
72.5
73.0
hot-humid
marine
mixed-dry_hot-dry
mixed-humid
very-cold_cold


Cooling Runtime vs Cooling Comfort Temperature by Region
hot-humid
marine
mixed-dry_hot-dry
mixed-humid
very-cold_cold
Runtime (hours/year)
1600
1400
1200
1000
800
600
400
71
72
73
74
75
76
Comfort Temperature (deg F)

Discussion Questions

- Does the metric need modification?
- A hybrid of metrics?
- Data additions
  - Non-core runtimes
  - Indoor temperatures
- How to handle variable speed?
- Add humidity regionally?

Regional Baselines Discussion

- Have you looked at correlation with average outdoor temperature for each data set? It's in the stats file. [Good suggestion, will do]
- Also, mean indoor temperature? Also in stats file.
- Many questions, can we concentrate on one?
- **Is the run time a valid metric of performance for CTs?**
- **Is it reasonable to assume that vendors' customer populations are comparable?**
  - A pretty far leap
- **Is it reasonable to assume that vendors' customer populations have different average temperatures?**
  - Could be – more appeal to elderly (higher set points) or more households with someone home all day
- Could we drill down to a more geographically fine grained baseline?

Regional Baselines Discussion

- Edge cases that are significant for any metric: Can we talk about multiple thermostats in the same home? And vacation homes.
  - Can these cases be detected algorithmically? Vacation homes, yes – either no occupancy or for weeks on end no comfort, or comfort only on weekends.
  - Multiple thermostats are much harder, unless they all belong to the same vendor.
- Another edge: heat pumps w/o backup, esp. variable speed
- Agree that current metric has a significant issue that more efficient set points ding the ENERGY STAR score.
- Some third party products aren't made to work with variable speed systems
  - Some advanced systems have their own staging on board if they are used with a non-proprietary thermostats
  - EPA currently advises consumers to use a proprietary thermostat with variable speed units