Technical Working Group: Meeting 5

March 7, 2022

Agenda: 3/7/23 Technical Working Group

- CalTRACK 2.1 Topics and Status
- EEWeather and OEEM Pull Requests
- 2.1 Daily Model: Thermal Lag
- 2.1 Daily Model: Probing seasonal bias and RMSE under Cross Validation
- Discussion

CalTRACK 2.1 Topics and Status

LF ENERGY

CalTRACK Daily

Goals of CalTRACK 2.1

- Specific areas of model bias - can we improve?
- Are computational efficiency gains achievable?
- Improved modeling for Solar PV customers?

CalTRACK Monthly

- Delivered Fuels?

CalTRACK Hourly

- Remedy known (small) bugs?
- Specify Demand Response baselines and appropriate model modifications?
- Incorporate Solar PV modeling?
- Improve functional form?
- Develop specifications for additional use cases?

EE Weather

- Code improvements to enhance usability?
- UK/International data?
- Solar irradiance data?

Pull Requests

LF ENERGY

Carbon Co-op

CalTRACK international

CalTRACK Working Group
07 March 2023

www.carbon.coop / @carboncoop

1. Carbon Co-op introduction

- Carbon Co-op is a membership-based **energy services co-operative.**

- We engage in **digital & social innovation** to enable people to reduce their domestic carbon emissions.

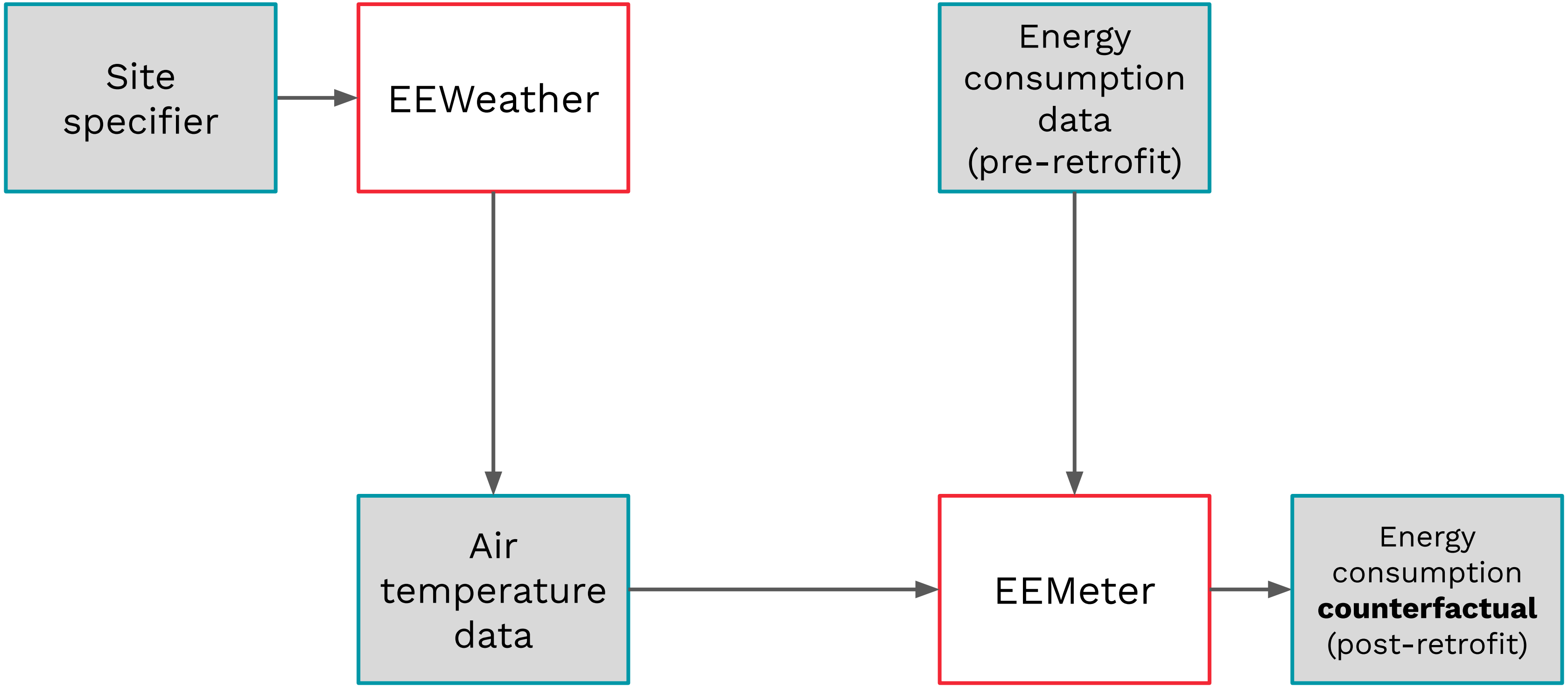
Carbon Co-op

Funded by

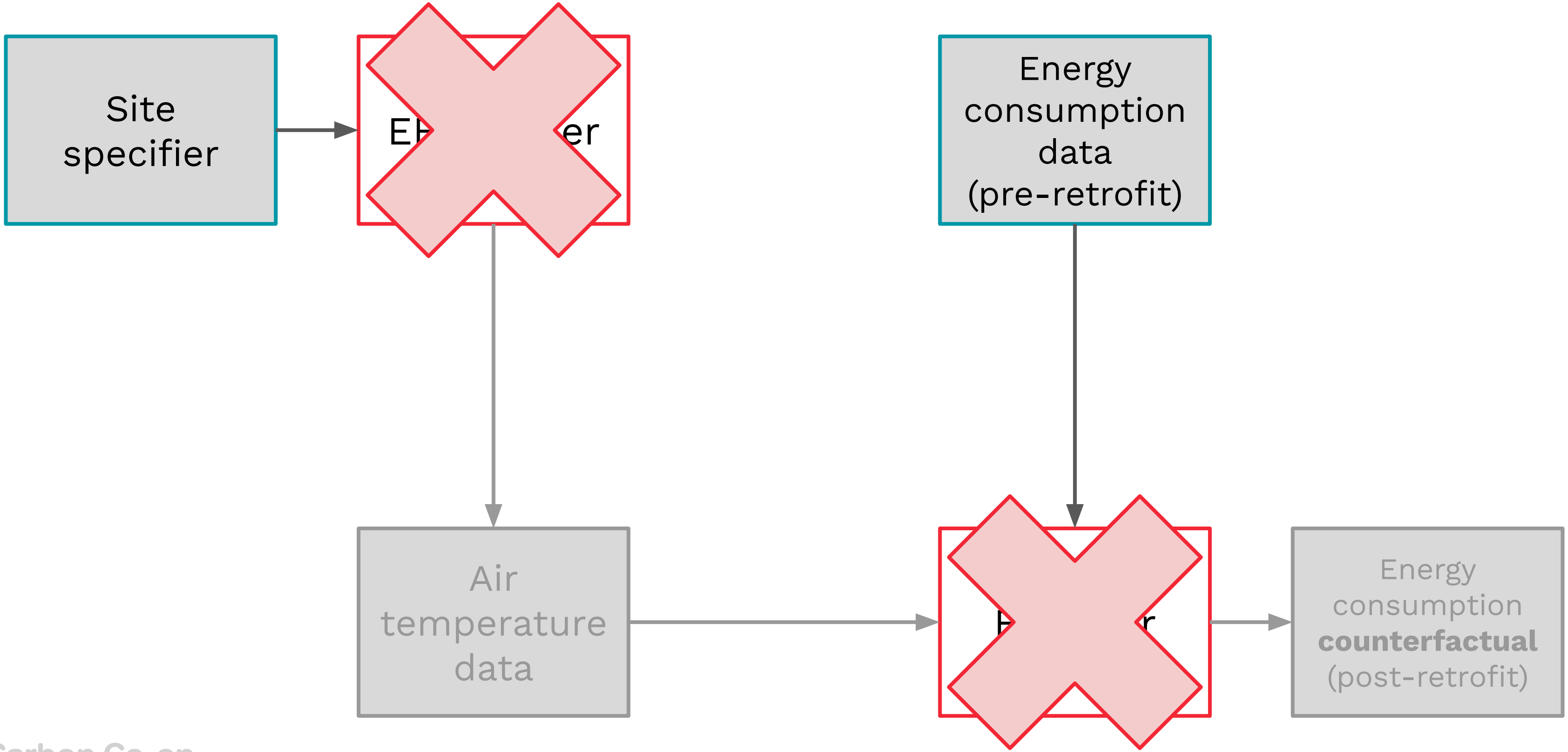
Innovate UK

2. CalTRACK international amendments

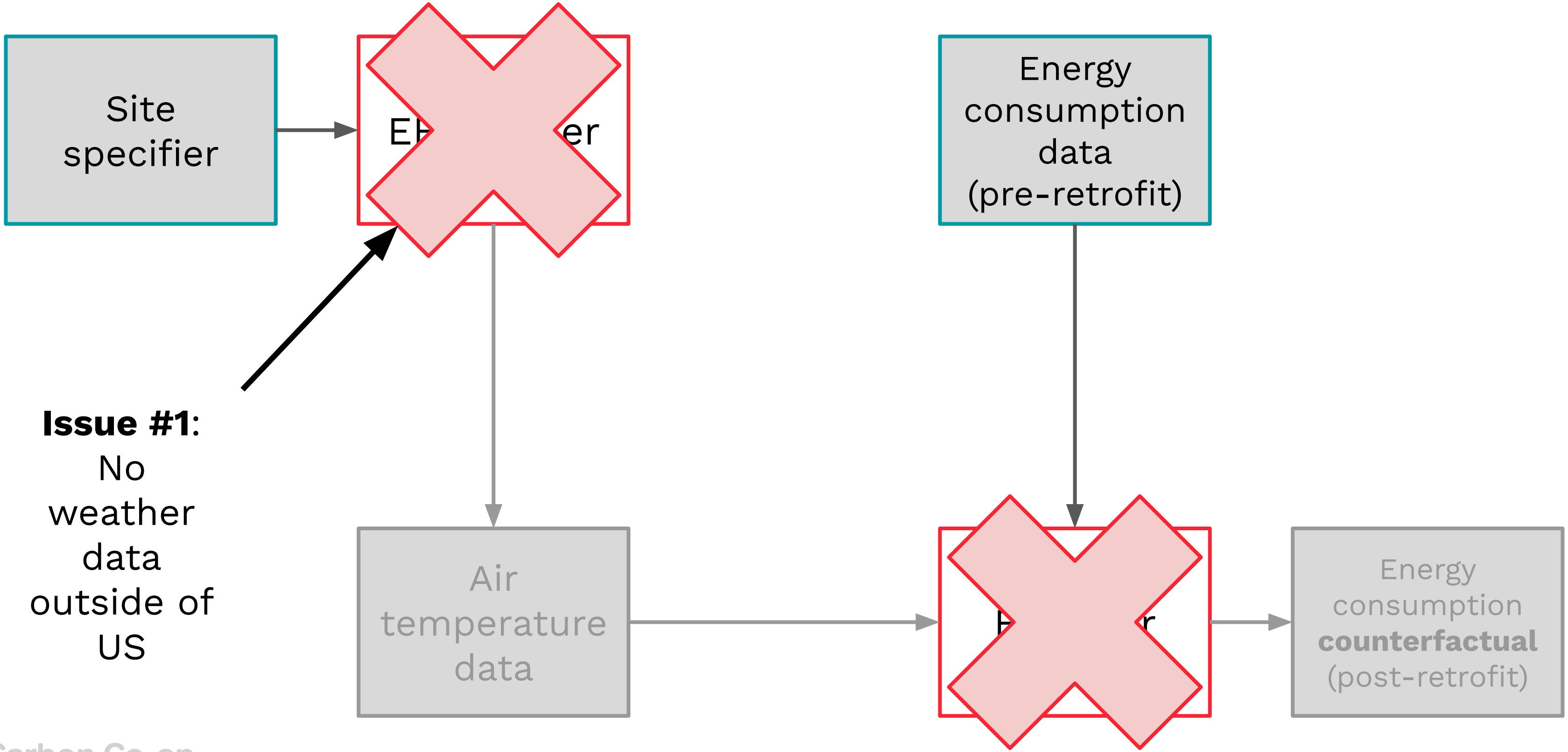
2.1. CalTRACK: USA functionality



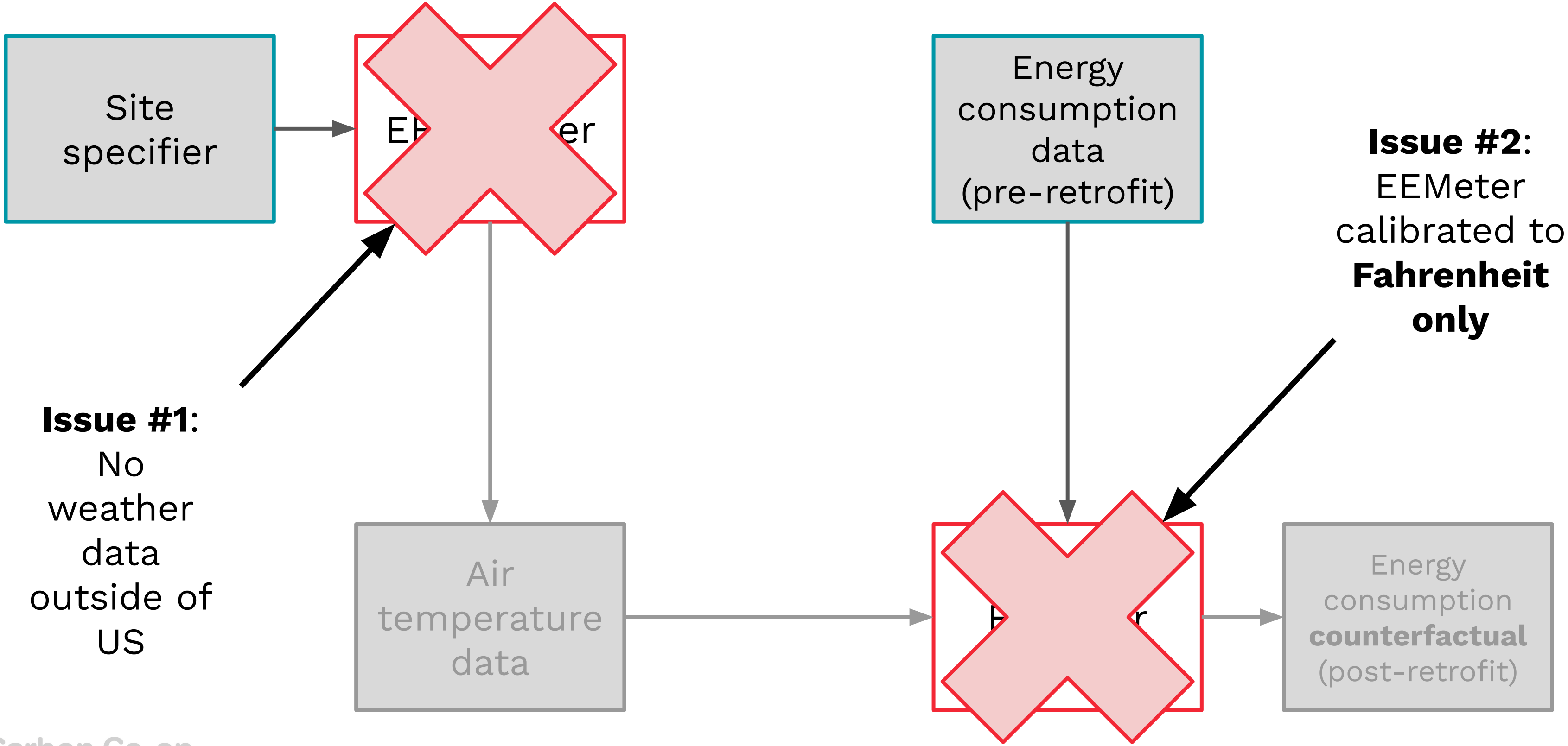
2.2. CalTRACK: UK functionality



2.2. CalTRACK: UK functionality

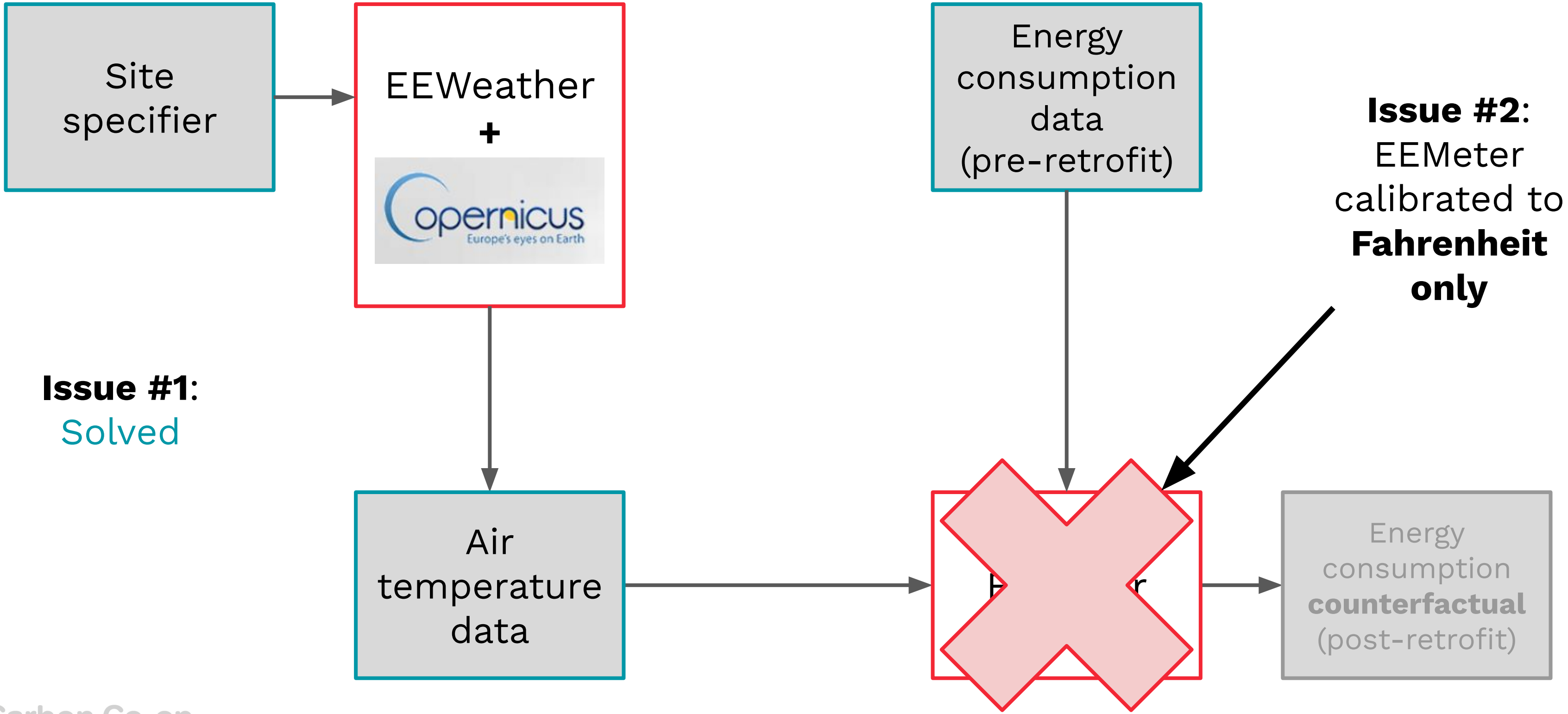


2.2. CalTRACK: UK functionality

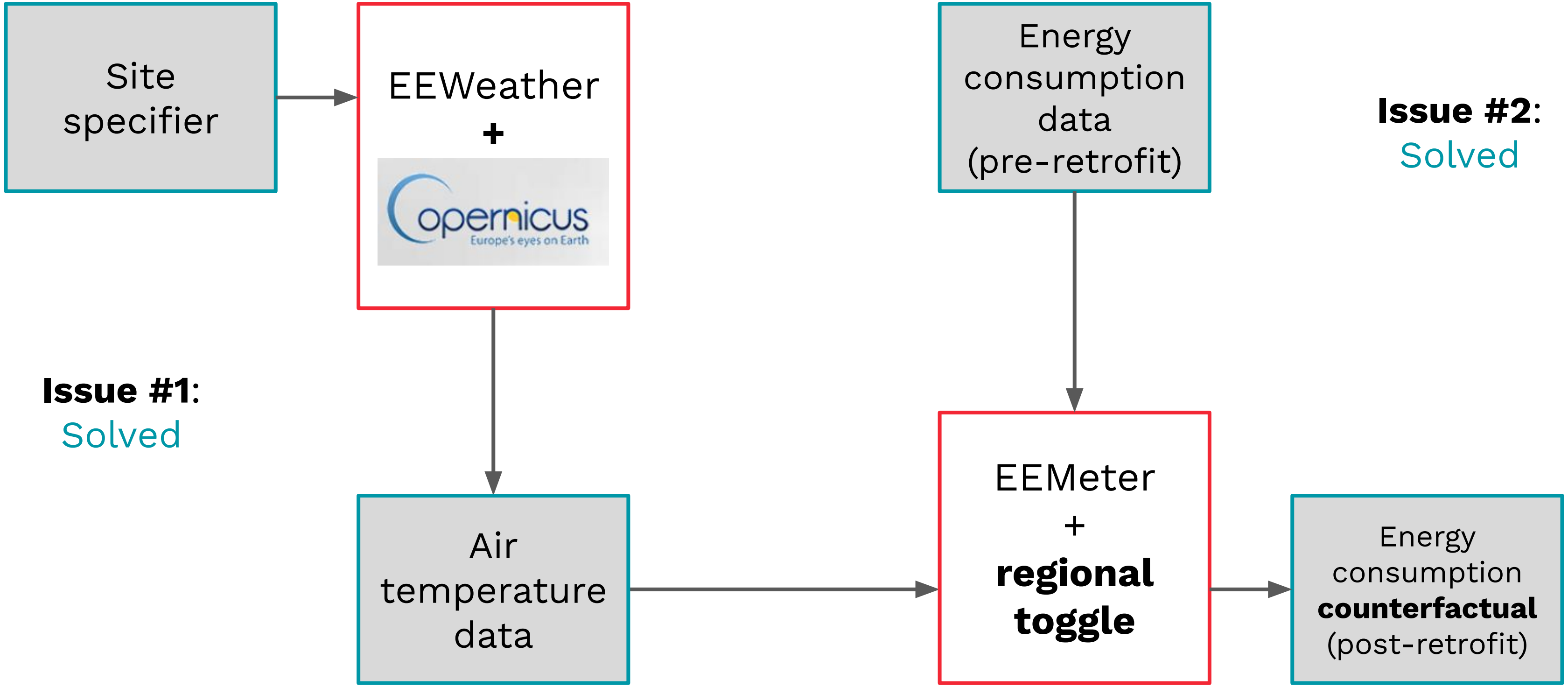


Carbon Co-op

2.2. CalTRACK: UK functionality following amends

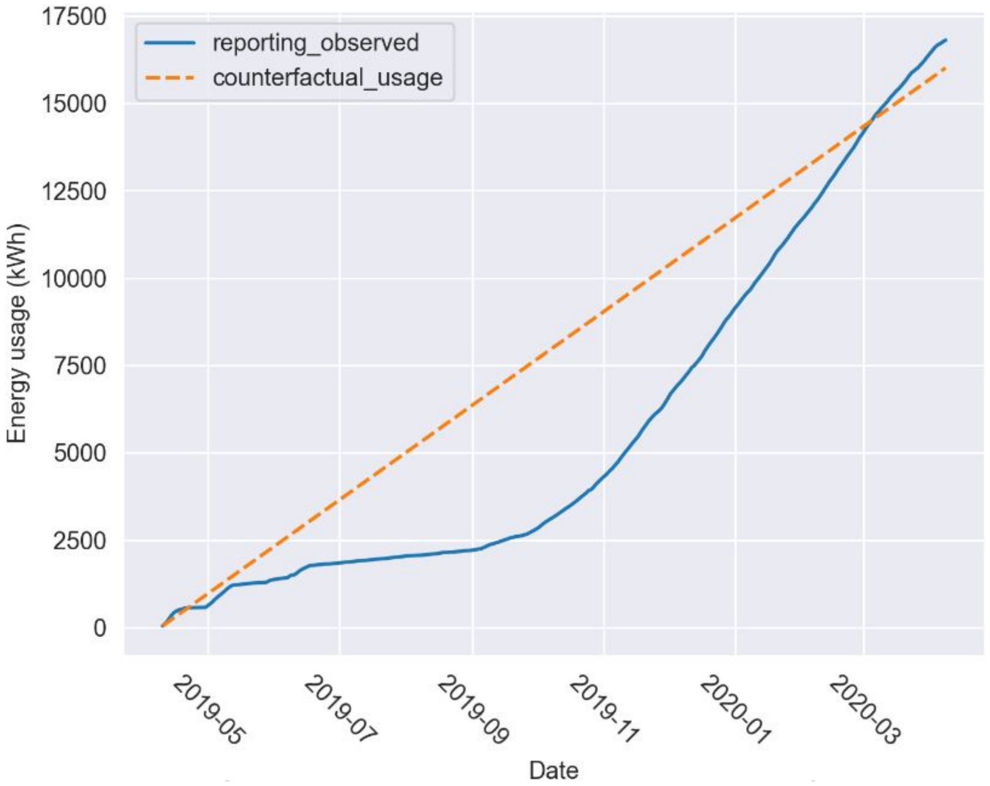


2.2. CalTRACK: UK functionality following amends

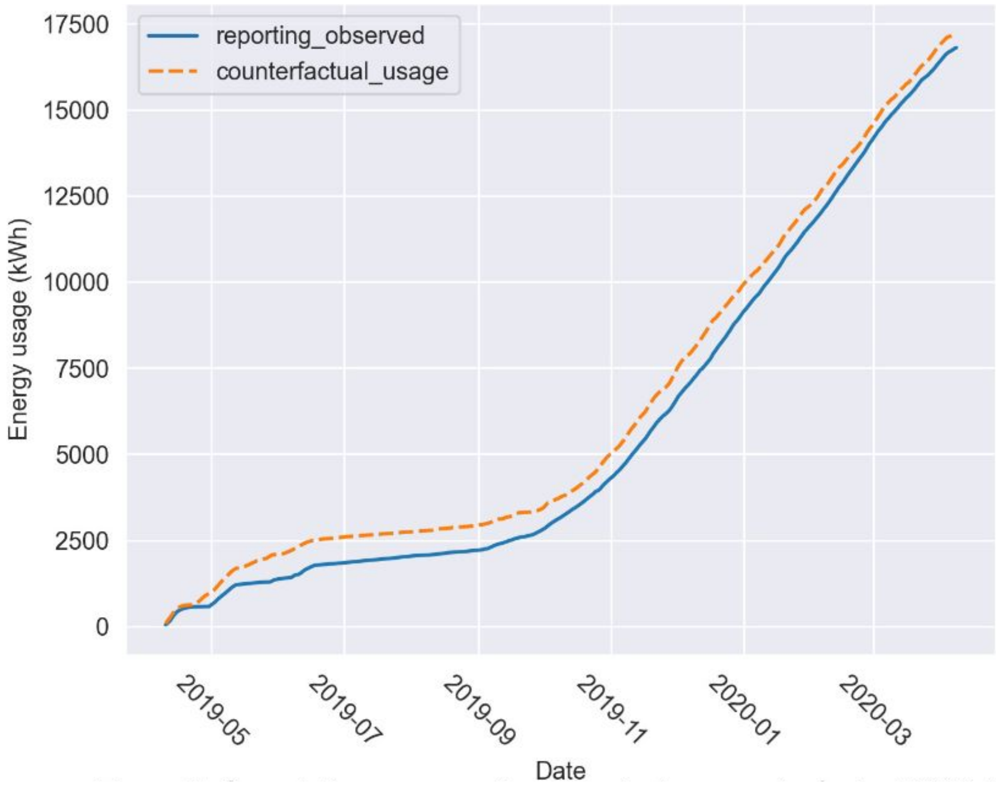


3. Temperature conversion

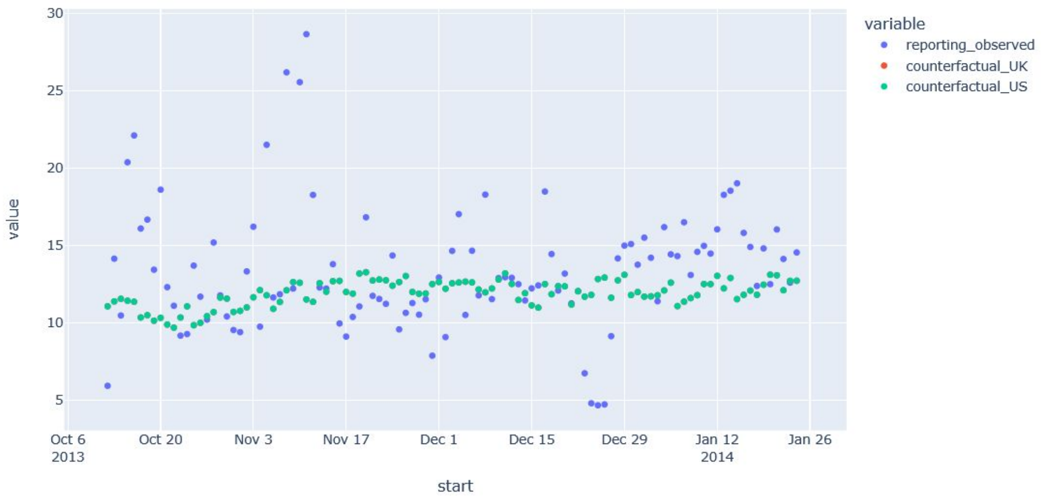
3.1. EEMeter daily: UK analysis before



3.2. EEMeter daily: UK analysis after (i.e. identical to CalTRACK US)



3.3. EEMeter daily: comparison to US outputs (zero)



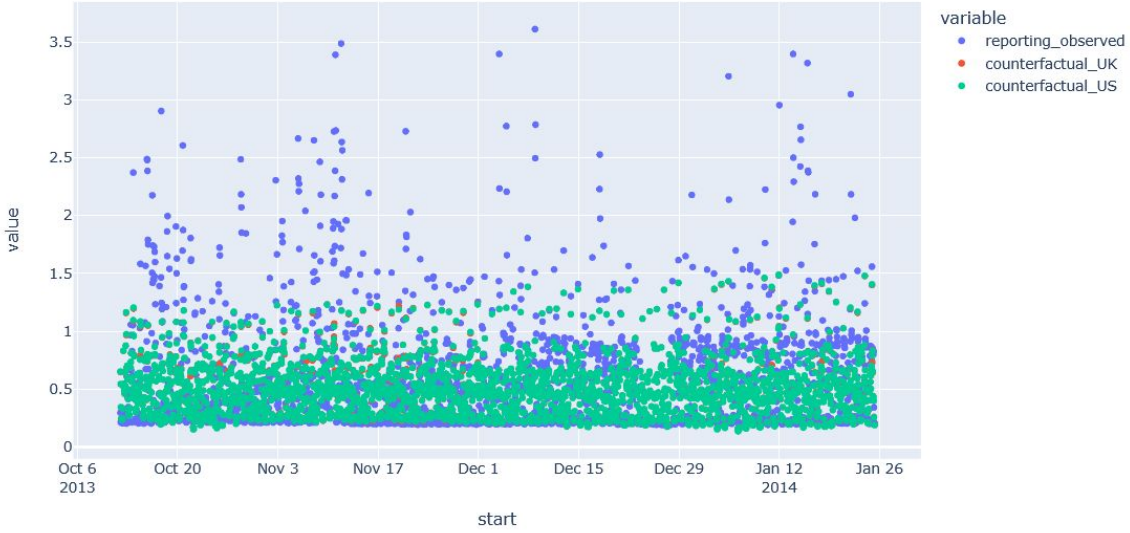

```
The absolute difference between EEMeter daily outputs using the UK and US specifier is: 0.0
```

4. Balance point assumptions (EEMeter hourly)

4.1. Balance point assumptions in EEMeter hourly

	UK	USA	France	Germany	Spain
hbp	60	50.0	64	59	59
cbp	200	65.0	72	72	72
References	https://support.dexma.com/hc/en-gb/articles/36...	NaN	https://support.dexma.com/hc/en-gb/articles/36...	https://journals.ametsoc.org/view/journals/apm...	https://support.dexma.com/hc/en-gb/articles/36...
References	https://www.google.com/url?sa=t&rct=j&q=&esrc=...	NaN	https://rmets.onlinelibrary.wiley.com/doi/full...	https://rmets.onlinelibrary.wiley.com/doi/full...	https://rmets.onlinelibrary.wiley.com/doi/full...

4.2. EEMeter hourly: comparison to US outputs (negligible, down to balance point difference)




```
The absolute difference between EEMeter hourly outputs using the UK and US specifier is: 4.964859235637663
```

5. EEWeather international & reanalysis data

5.1. CDS reanalysis vs. NOAA data - benefits

- Accurate weather data for anywhere around the world, regardless of proximity to weather stations
- 2m temperature data included, also solar irradiance, wind speeds, precipitation
- NOAA data may be of limited quality outside of US

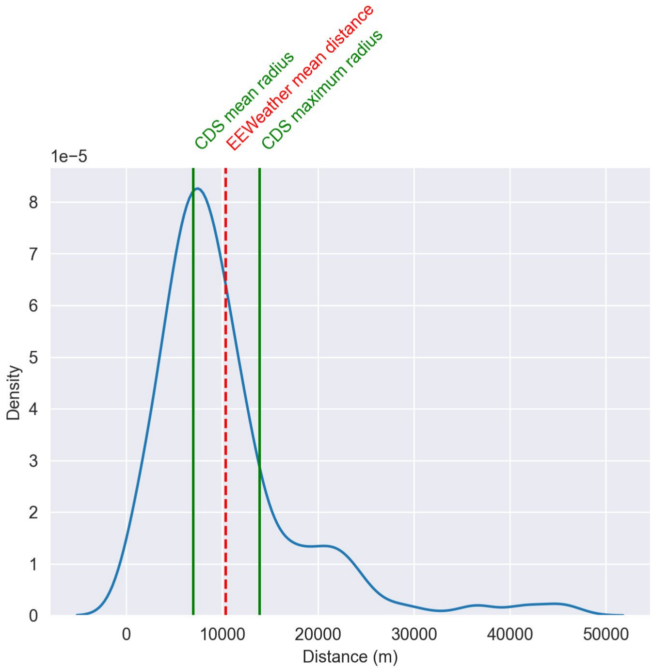


Figure 1: EEWeather minimum distance distribution vs. CDS maximum radius

5.2. CDS reanalysis vs. NOAA data

- Reanalysis data compares well against NOAA data in terms of accuracy

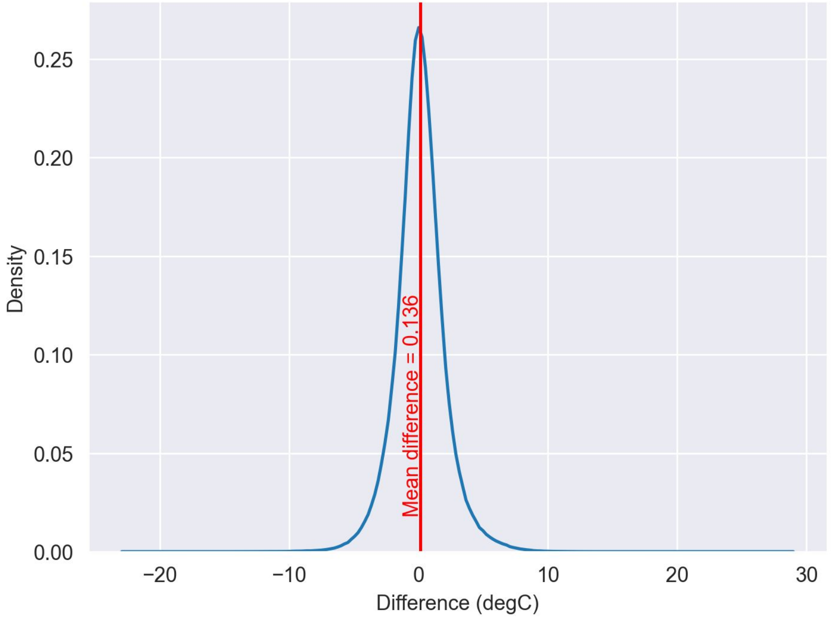


Figure 2: difference distribution between CDS and 'original' EEWeather across 100 US sites

5.3. CDS reanalysis vs. NOAA data - issues

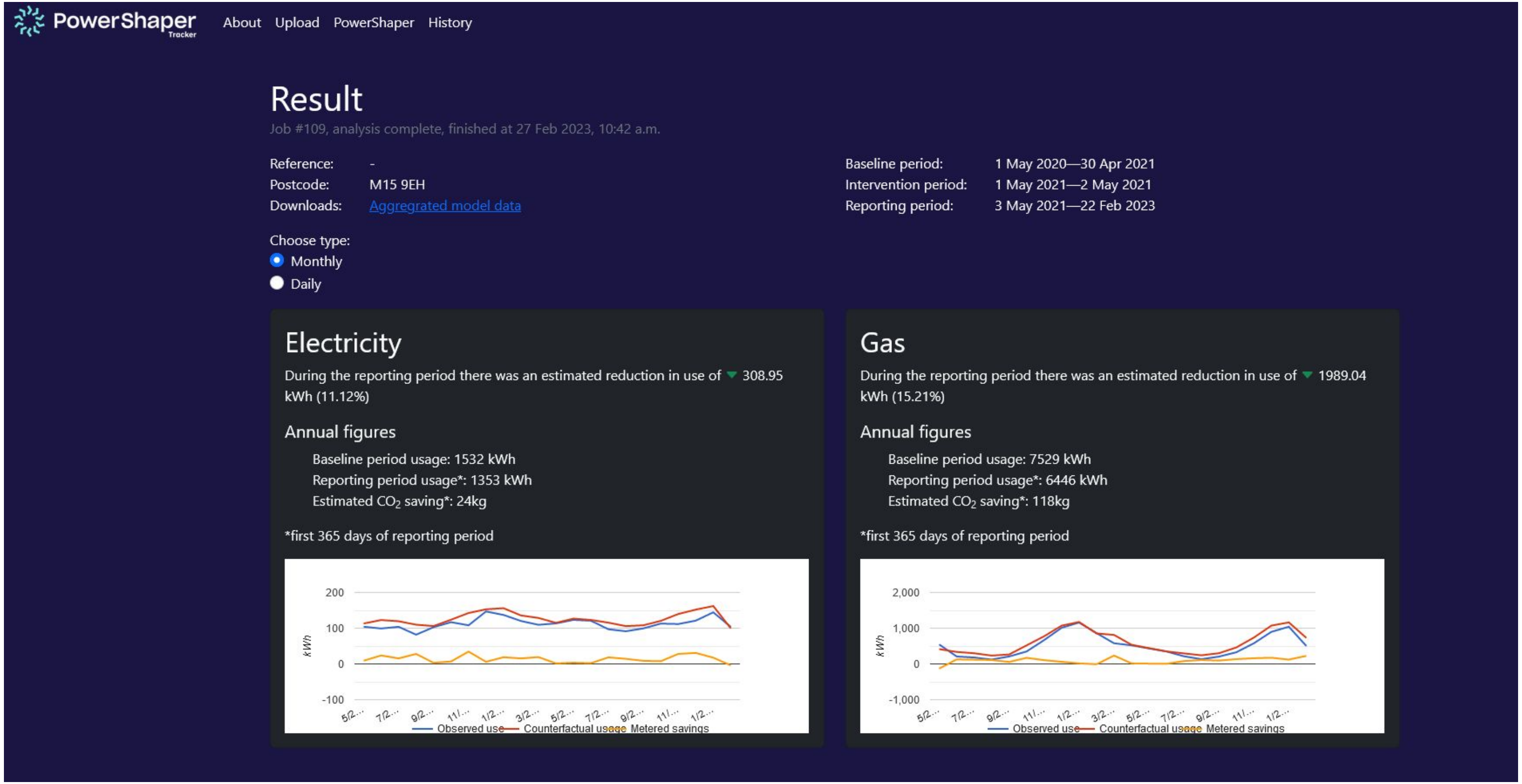
3. Speed

While outputs from the EEWeather international package (via CDS) are accurate, API traffic on the CDS means that download speeds are poor compared to 'original' EEWeather. A typical 2-year interval call for US sites using the 'original' package can be expected to return temperature data at a mean of 0.12 seconds; the same call using EEWeather international might typically take **552 seconds**.

Users in the United States should therefore continue to use 'original' EEWeather.

6. Web app (PowerShaper Tracker)

6.1. CalTRACK UK web app



7. Round up

7.1. Key questions

- Is there a benefit to a future EEWeather international in pursuing a reanalysis approach?
- Should balance points be hard coded for international/site-specific usage of EEMeter?

- Any more?

PEOPLE POWERED NOT FOSSIL FUELLED

Carbon Co-op

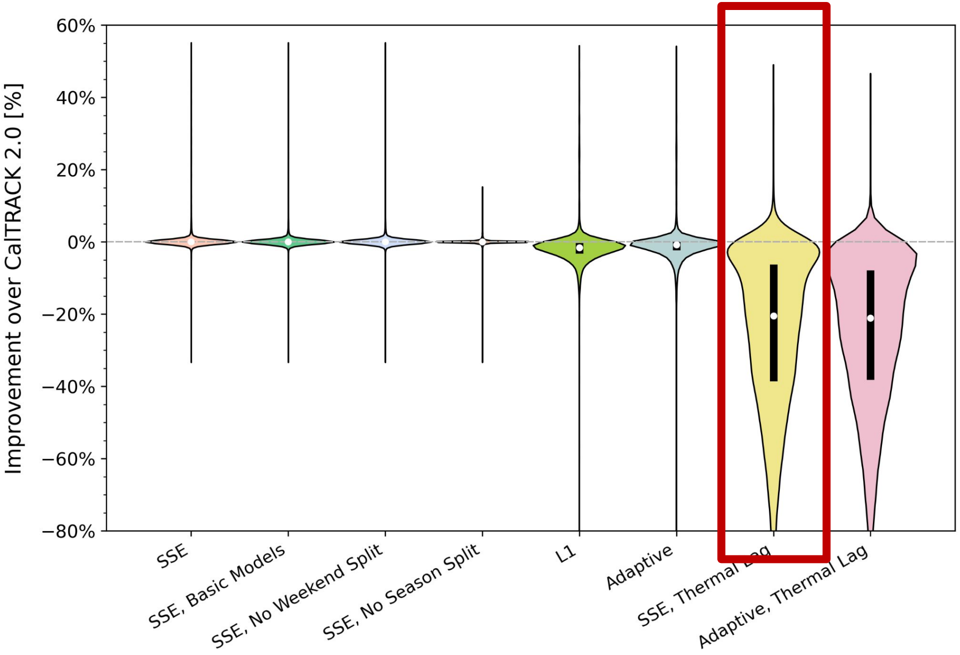
James Fenna: james@carbon.coop
Matt Fawcett: matt@carbon.coop

Thermal Lag: Brief Update

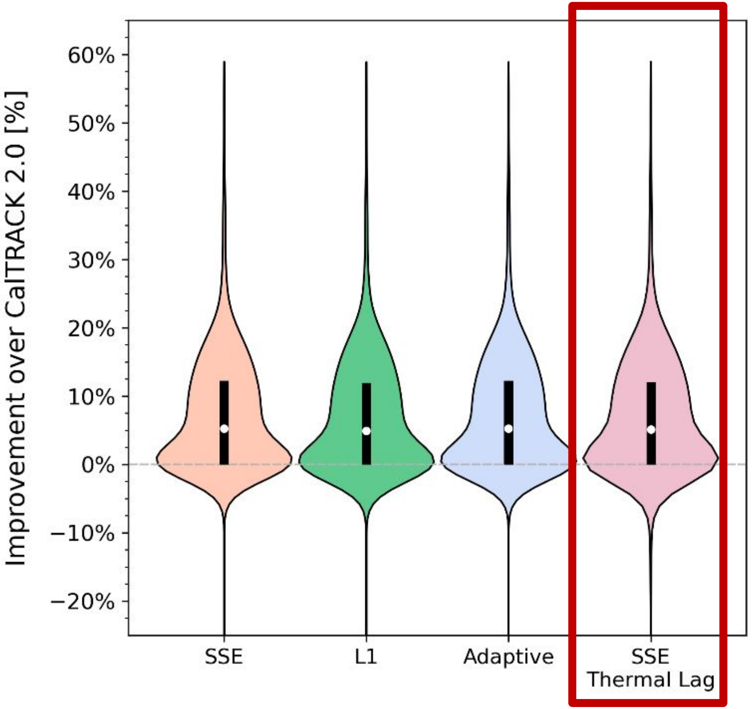
LF ENERGY

Fixing Thermal Lag Bug

Previous



Current



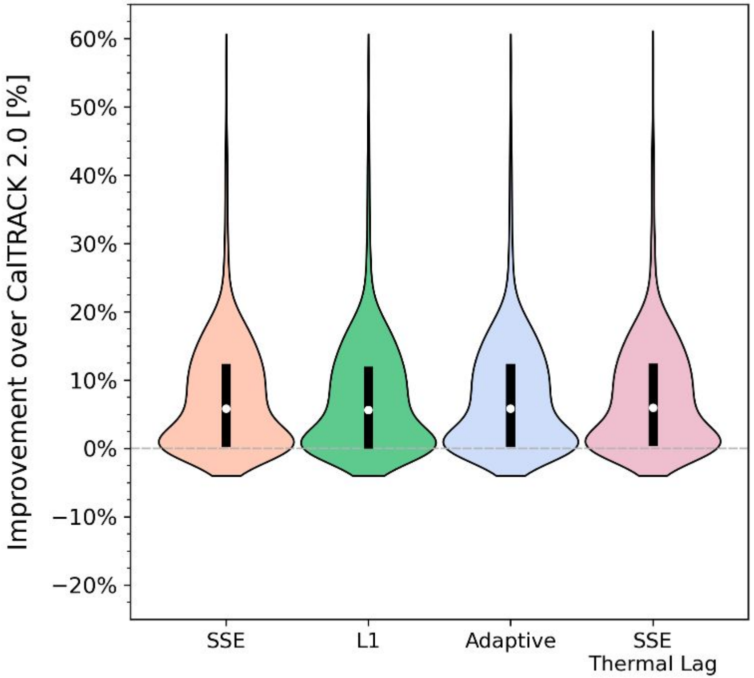
Previously thermal lag parameters were incorporated into all models. Now those parameters must overcome penalization

CalTRACK 2.1 Daily: Preliminary Full Model Results

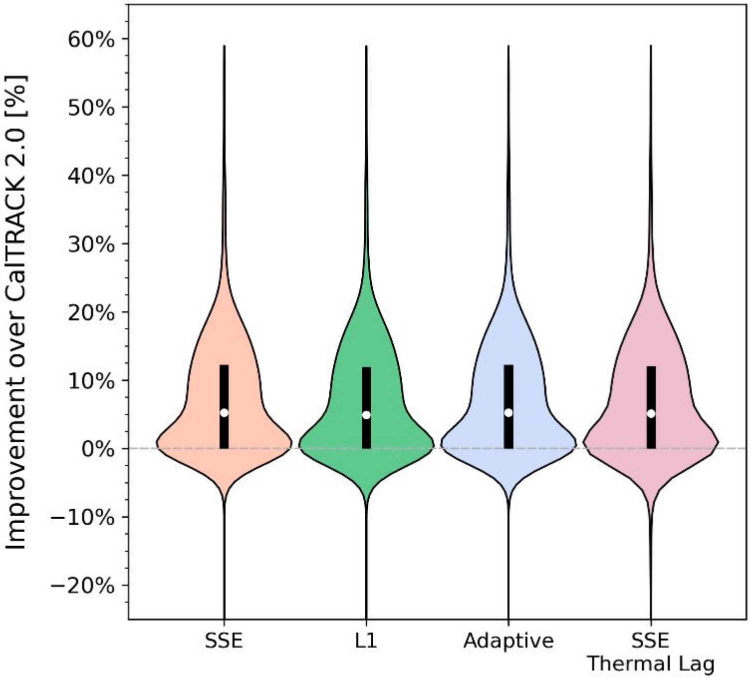
RMSE: Gas-Residential (All Meters)

Using Cross Validation to Generate Train and Test Samples

Train



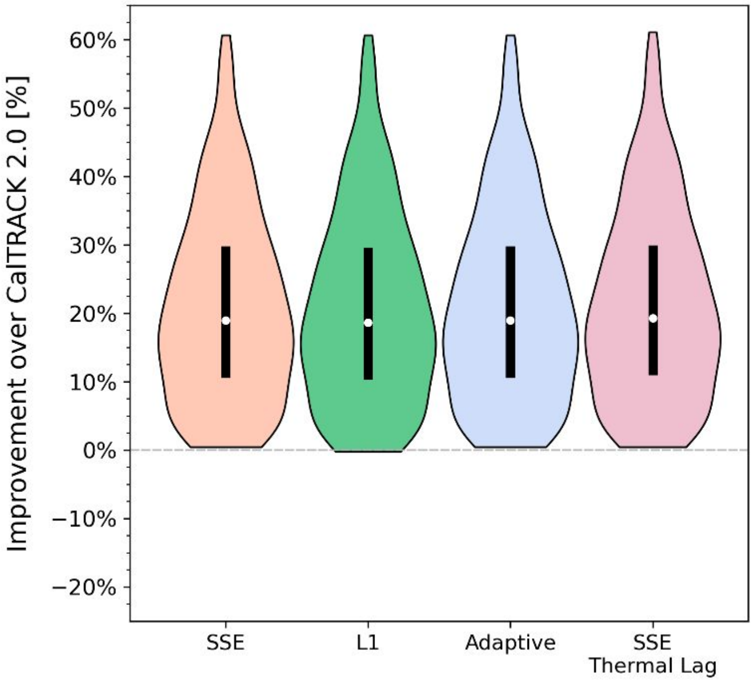
Test



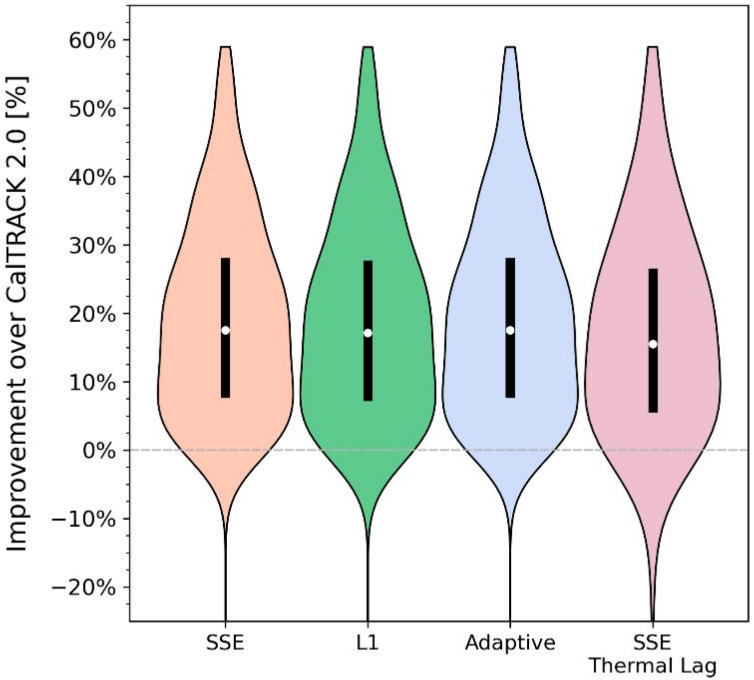
RMSE: Gas-Residential (Split Models Only)

Huge improvement…But only 8% of meters currently fall into this category.

Train

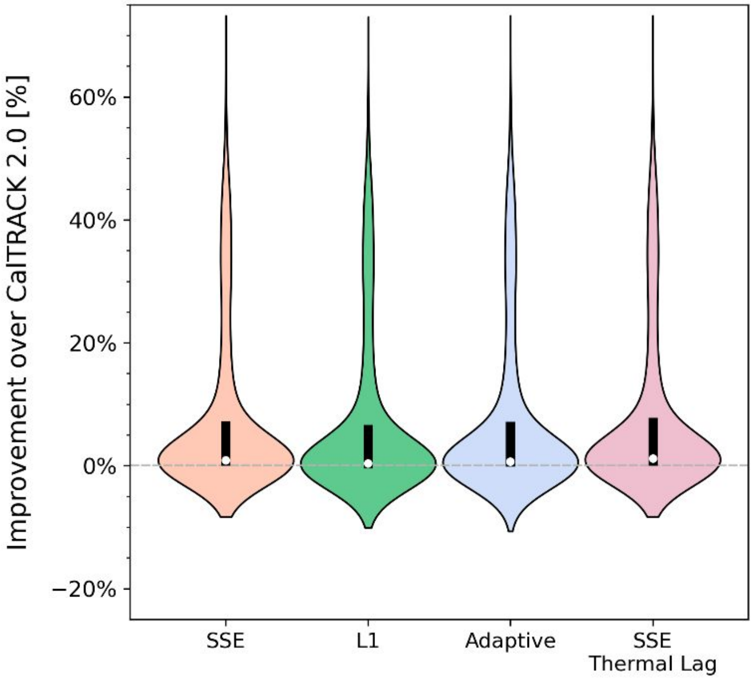


Test

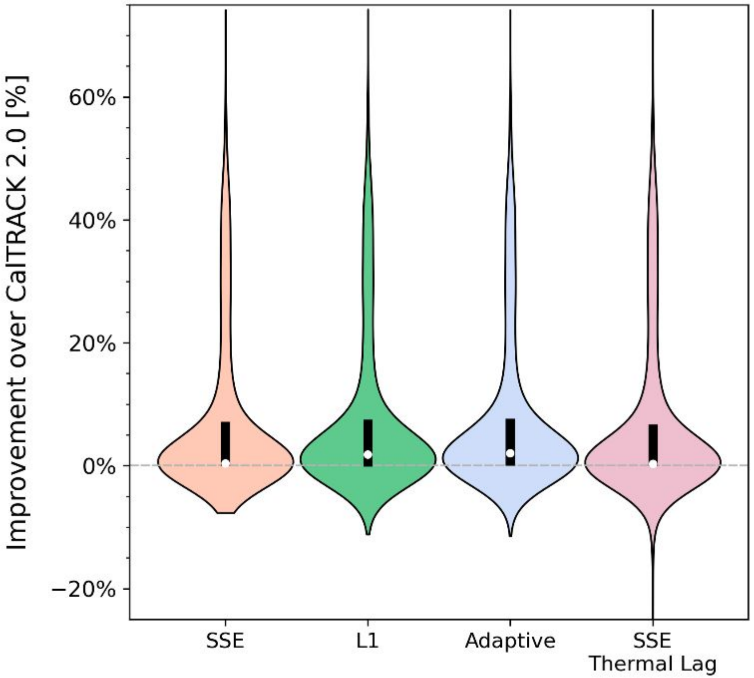


RMSE: Electric-Commercial (All Meters)

Train

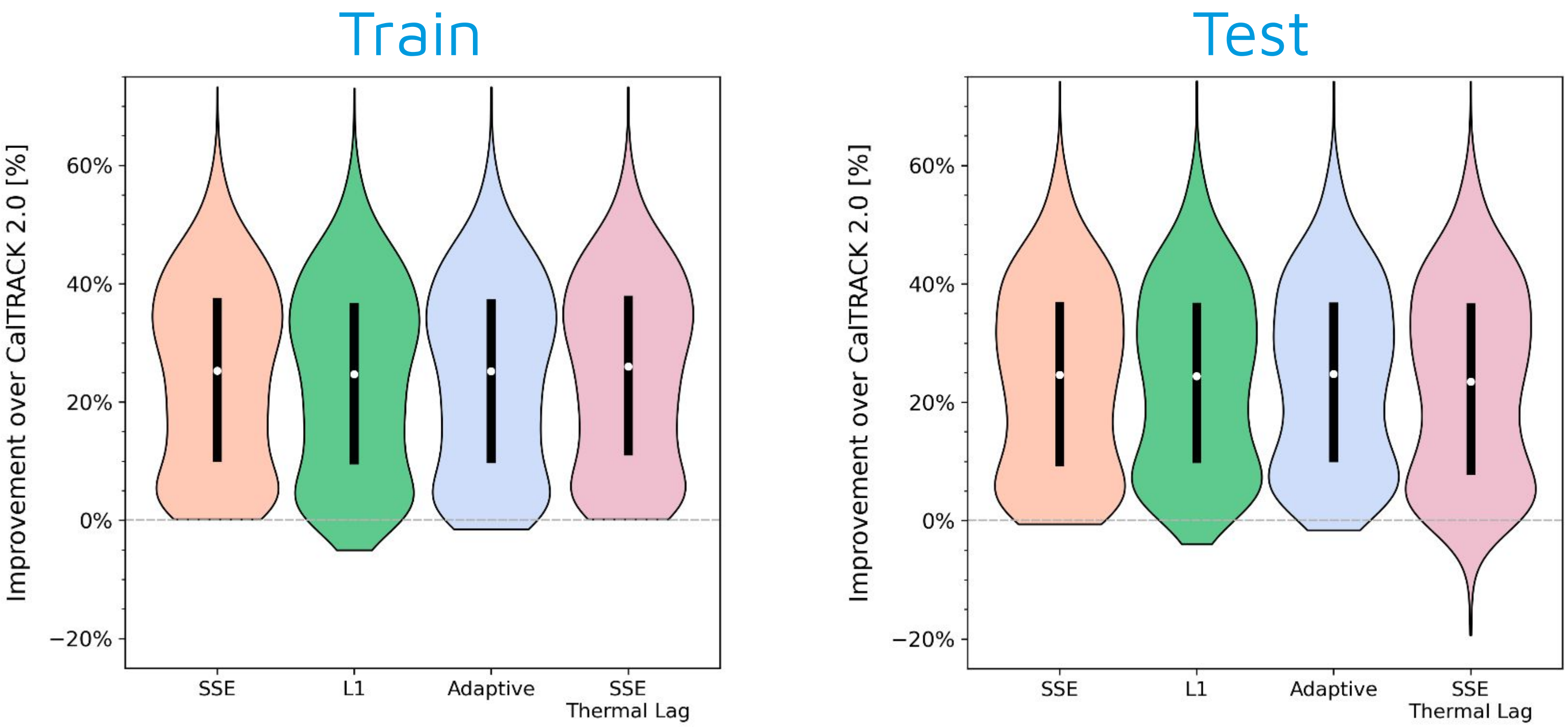


Test



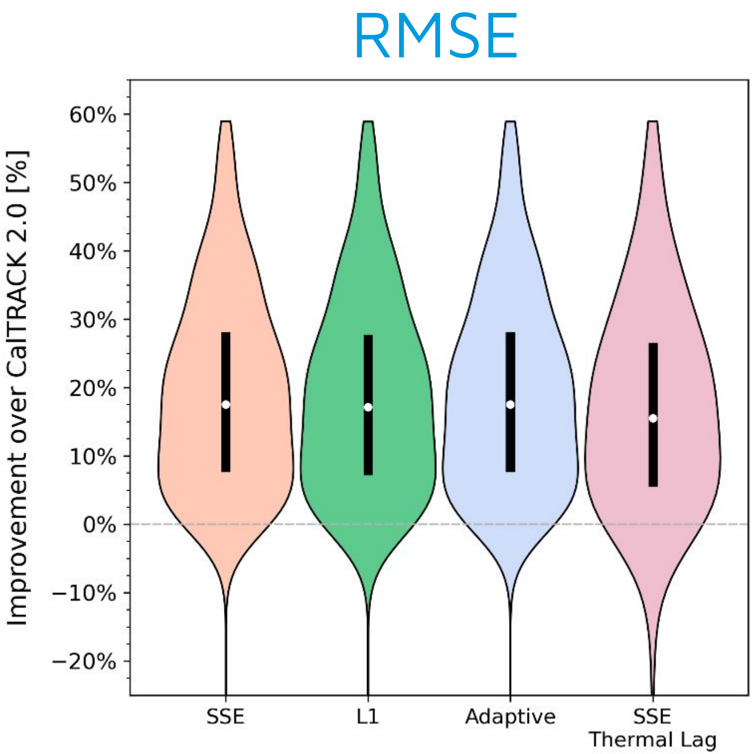
RMSE: Electric-Commercial (Split Models Only)

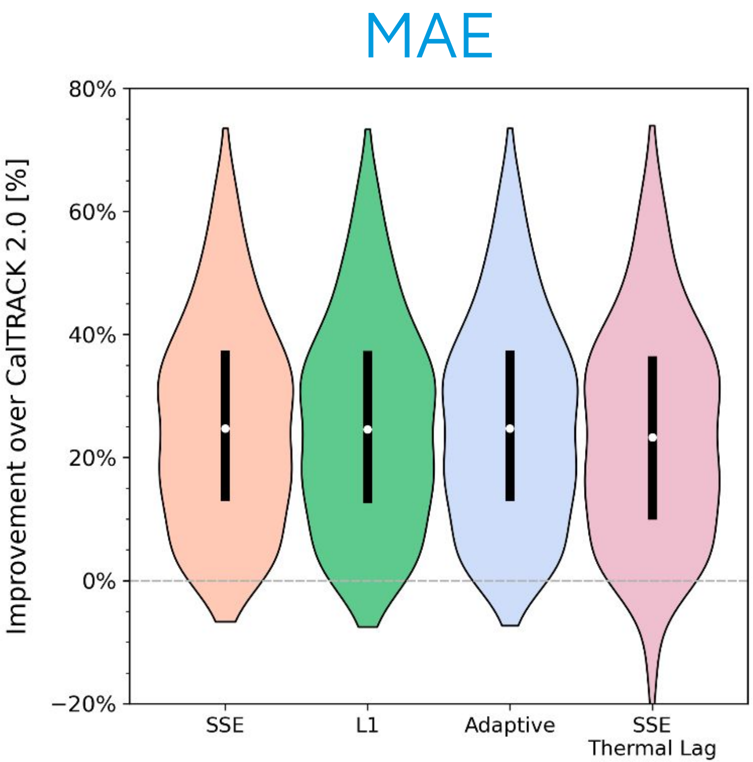
Huge improvement…25% of meters currently fall into this category.



Gas-Residential (Split Models Only)

Small effect but thermal lag appears to worsen average model performance



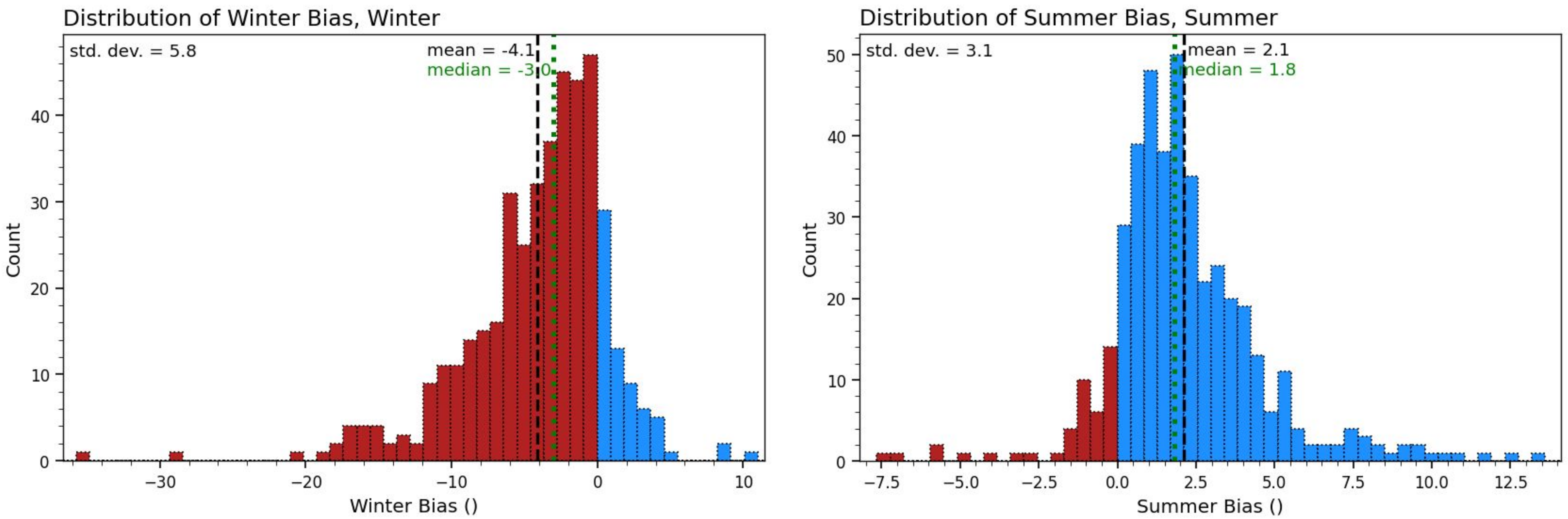


Probing Seasonal Bias and RMSE Under Cross Validation

LF ENERGY

CalTRACK 2.0 Winter, Summer Bias (Res Gas)

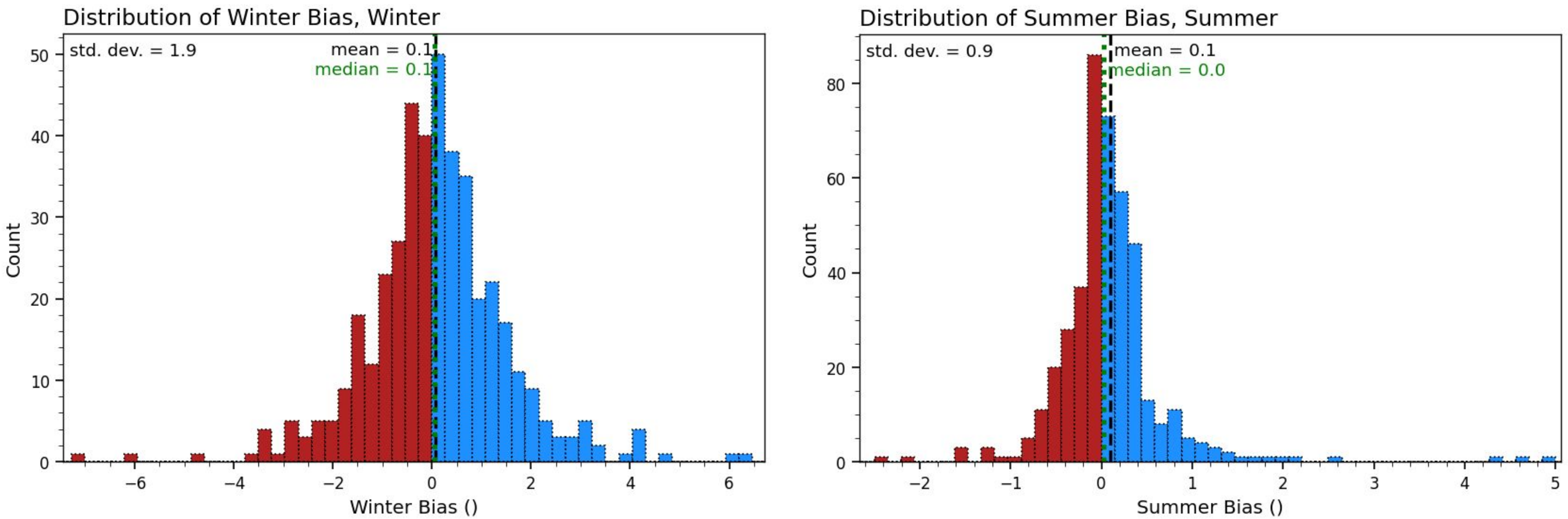
No seasonal splitting



Avg. Annual Test SSE = 511.6

CalTRACK 2.0 Winter, Summer Bias (Res Gas)

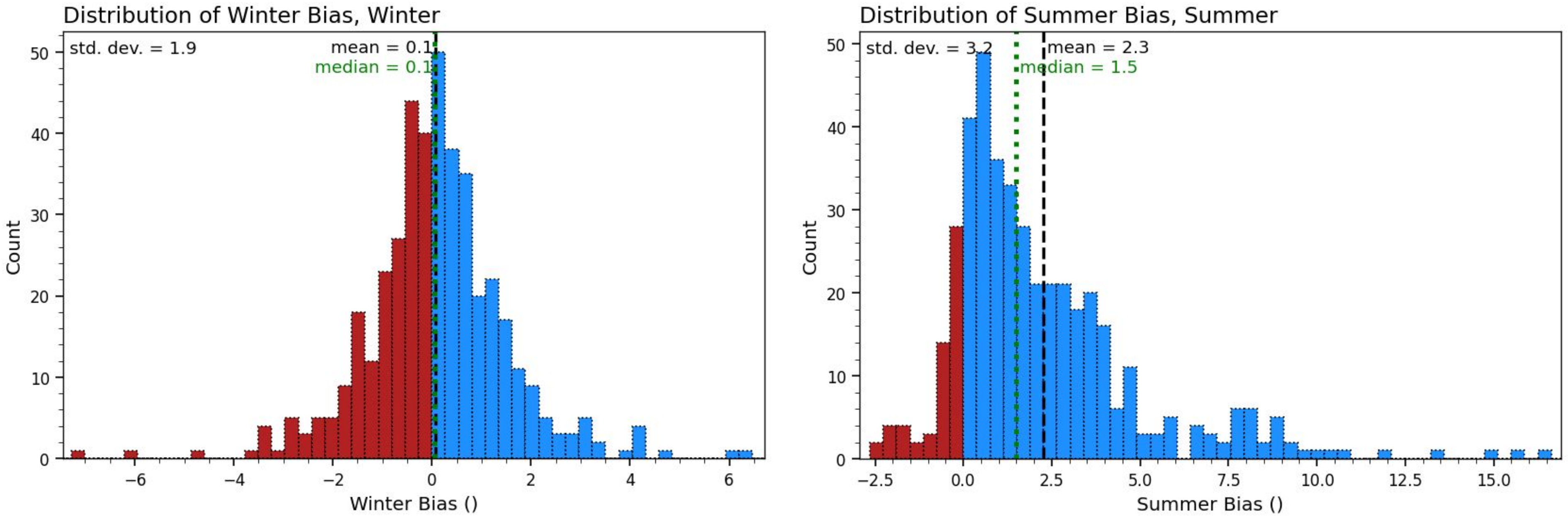
Split along all seasons (summer, shoulder, winter)



Avg. Annual Test SSE = 482.0

CalTRACK 2.0 Winter, Summer Bias (Res Gas)

Split winter vs summer+shoulder



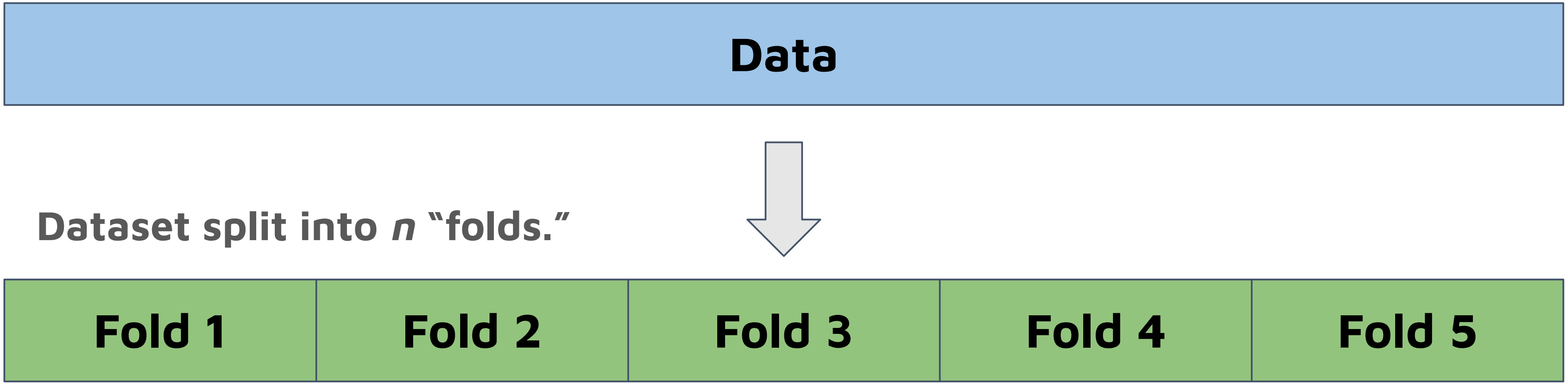
Avg. Annual Test SSE = 503.0

Open Questions

1. What is the best way to penalize additional parameters?
2. Is Cross Validation a viable penalization option itself or is it too computationally intensive?
3. Is there a better formulation for thermal lag?
4. What is ultimately the balance we should be striving for between **solving seasonal bias** vs. **remedying other sources of bias** vs. **overfitting** vs. **model complexity** vs. **computational cost**?

Appendix

Cross Validation: A Rigorous Approach to Test/Train Splitting



Each fold takes a turn as the test data with remaining folds serving as training data. Each iteration is a "split."

n models developed. Model parameters and performance determined by averaging results on test samples.

How to Avoid Overfitting: Penalization

Balancing model error/performance with number of parameters

How (Preliminary)?

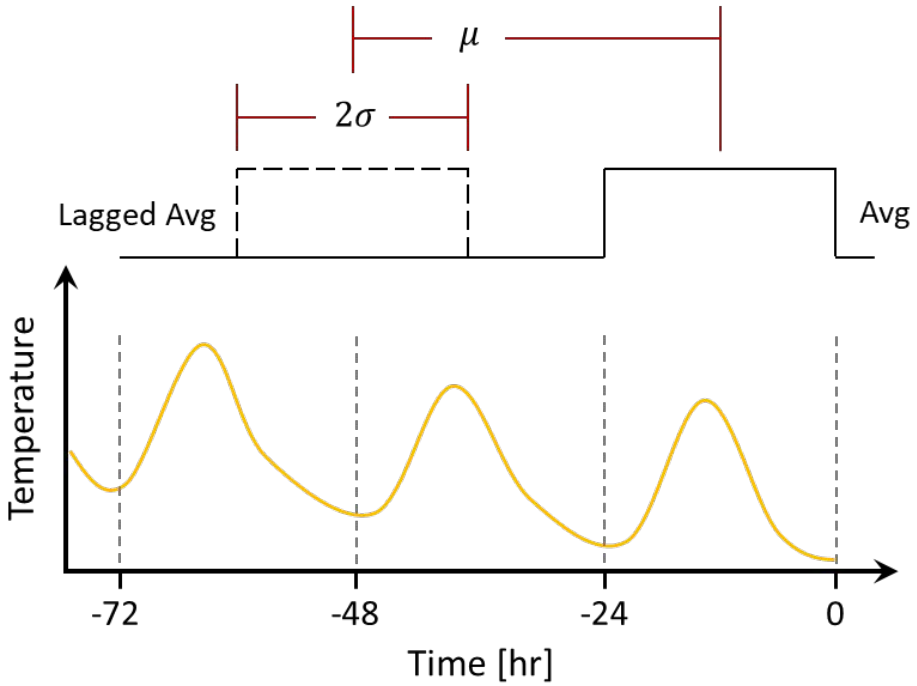
- Selection Criterion (AIC for example): Introduce penalty that increases with model complexity
- Additional penalty term
- Cost/Benefit test on adding parameters
- Empirical estimate of Cross Validation

Review: Thermal Lag

Possible solution: Lag all days' temperature

-3 < μ < 3 days

0.5 < σ < 3 days



Open Questions

What is the best penalization/selection criterion for splits

- Few models are being split in residential gas
- Use existing formula but add % effectiveness multiplier to penalty term

$$log\ likelihood = -\frac{\mathrm{N}}{2}\left(\ln(2\pi) + \ln\left(\frac{loss}{N}\right) + 1\right)$$

$$\mathrm{AIC} = -2\ log\ likelihood + eff_{\%}\ (2\ K)$$

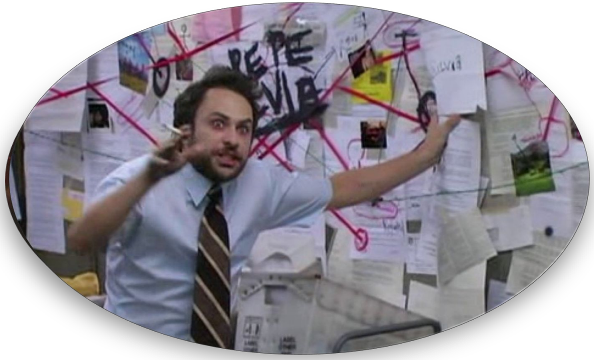
$$\mathrm{BIC} = -2\ log\ likelihood + eff_{\%}\ (K \log(N))$$

$$\mathrm{SABIC} = -2\ log\ likelihood + eff_{\%}\ \left(K \log\left(\frac{N+2}{24}\right)\right)$$

$$N = Number\ of\ data\ points\ (days)$$

$$loss = combined\ error$$

$$K = number\ of\ predictors + split\ penalty$$



Open Questions

What metric to use?

- RMSE/weighted RMSE/MAE
- Savings uncertainty of non-participants in reporting year

Thermal lag seems largely detrimental

- Is this agnostic to implementation?
- Bugs?

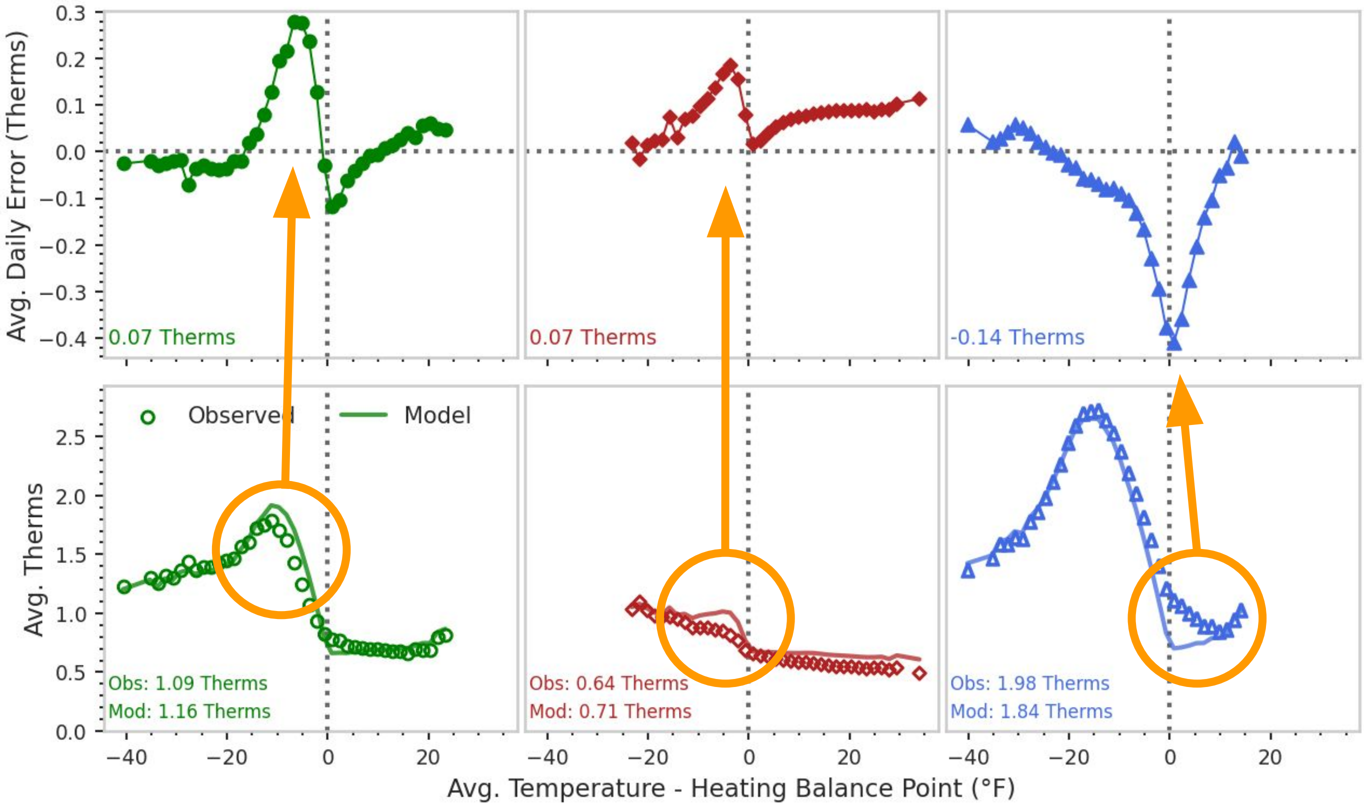
Optimization!

- Many parameters here:
    - Optimization Algorithm (Subplex, SLSQP, COBYLA, ...)
    - Selection criteria (AIC, BIC, ...)
    - Segmentation penalty amount
    - SSE, L1, Adaptive

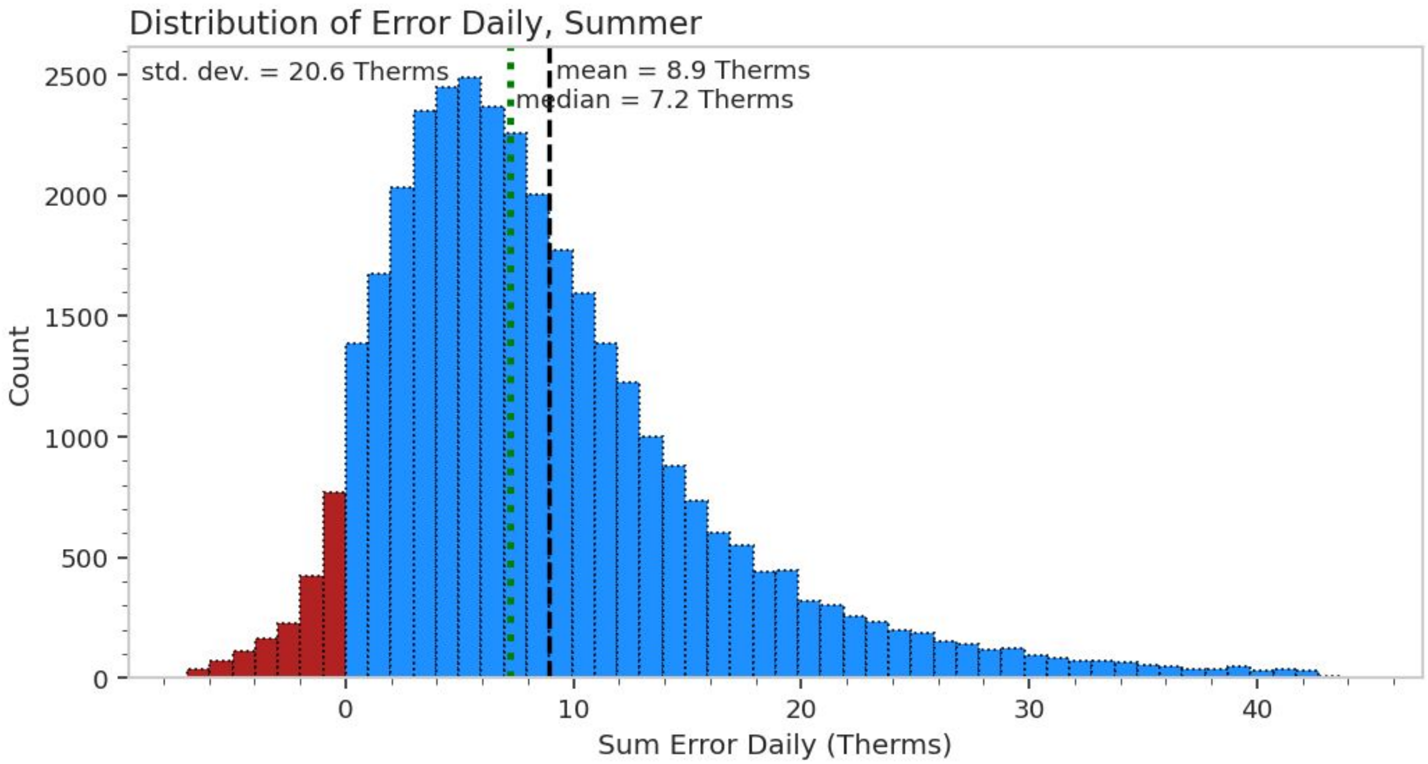
Why is there Seasonal Bias?

Error profiles differ by season with sharp features centered around the balance point temperature

Space heating tends initiated at warmer outside temperature in the winter compared to the rest of the year.

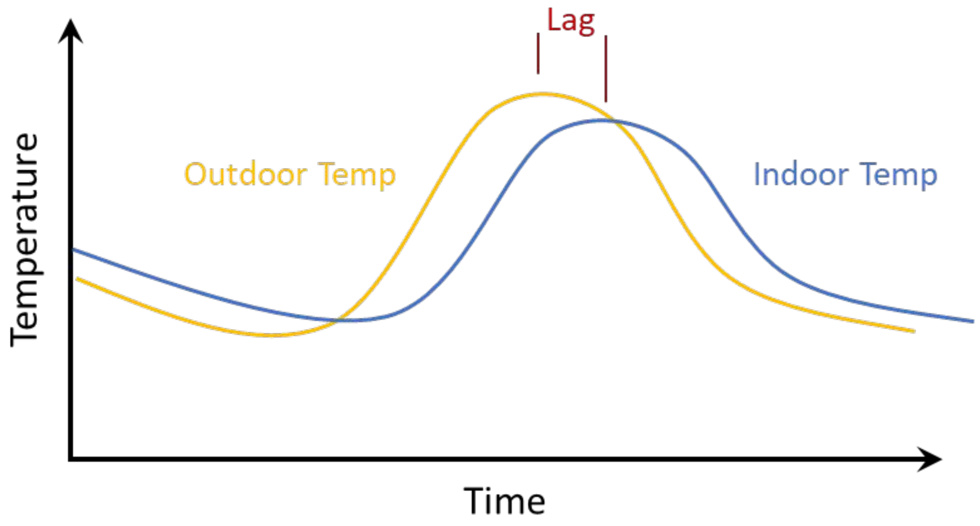


Seasonal Bias Distribution



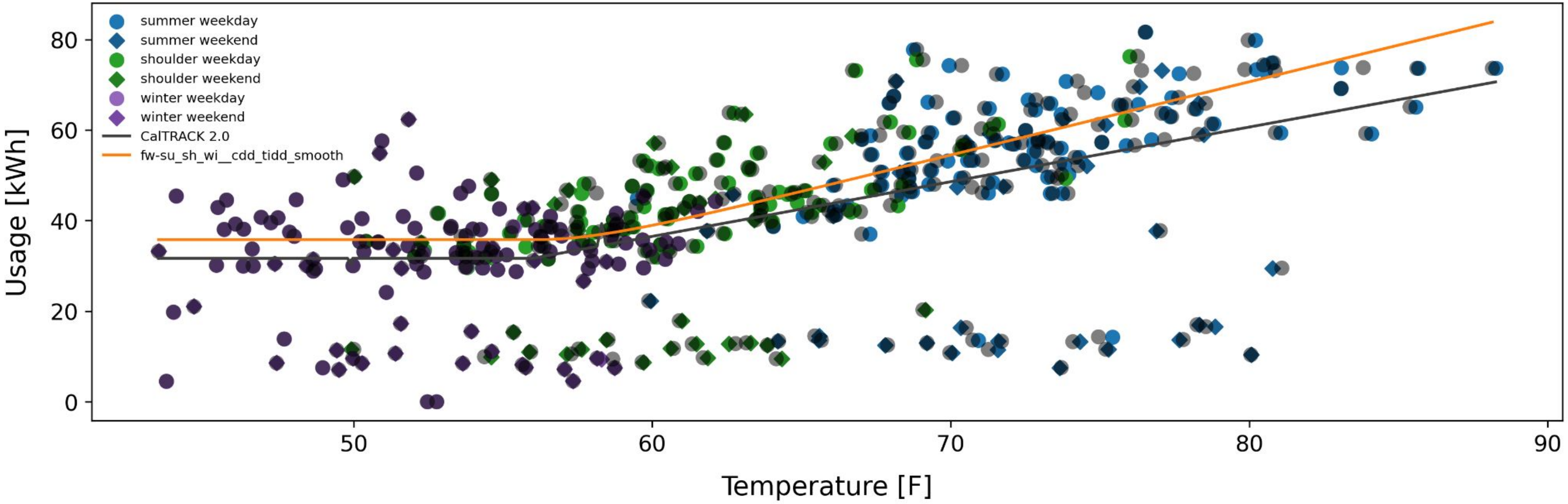
Review: Thermal Lag

Buildings and behavior change based on prior temperature



Review: Thermal Lag

Example of thermal lag



RMSE: 14.8 ± 0.8
mu: -2
sigma: 12