CalBEM 2025

November 19-20

Sacramento, CA

From Data to Action: Equitable Home Energy Labeling at Scale

Nov 19th, 2025

Mudit Saxena, CEO, XeroHome™ | MSaxena@xerohome.com

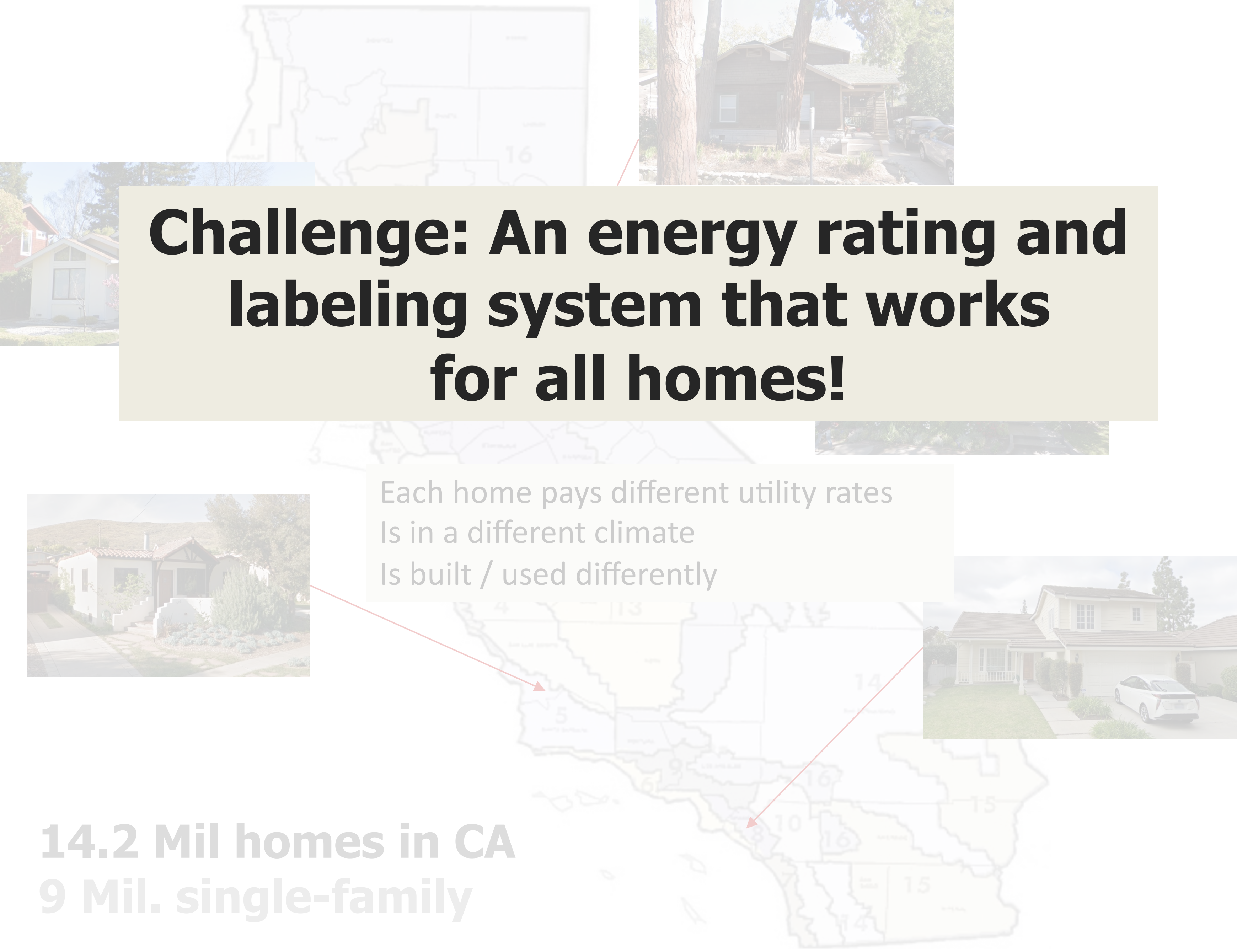
XeroHome™

Acknowledgement

Analysis was funded by **Southern California Edison's Codes & Standards Program**, on request of the California Energy Commission. Analysis was led by XeroHome™ (Vistar Energy) with technical support from 2050 Partners.

XeroHome™

Challenge: An energy rating and labeling system that works for all homes!



Each home pays different utility rates
Is in a different climate
Is built / used differently

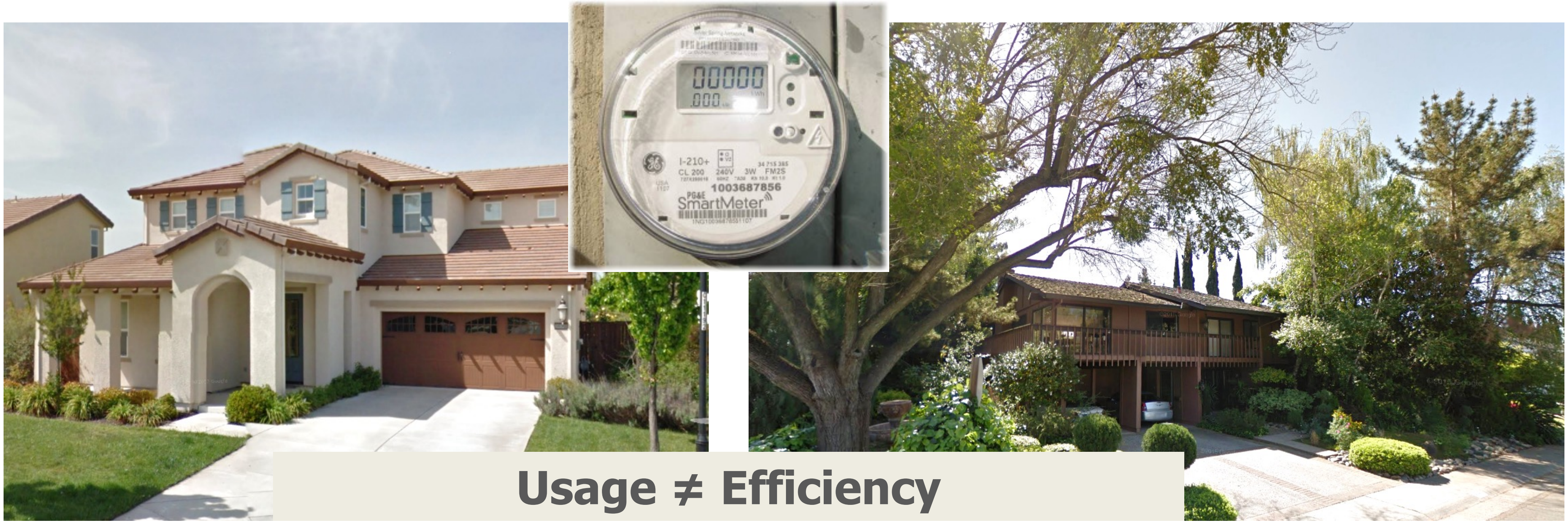
14.2 Mil homes in CA
9 Mil. single-family

XeroHome™

If Only Some Homes Get Labeled …
We Risk Leaving Others Behind

A statewide home energy label must be universal and free, to avoid **inequities**

Can Energy Usage Data Alone Be Used To Develop a Home Rating?

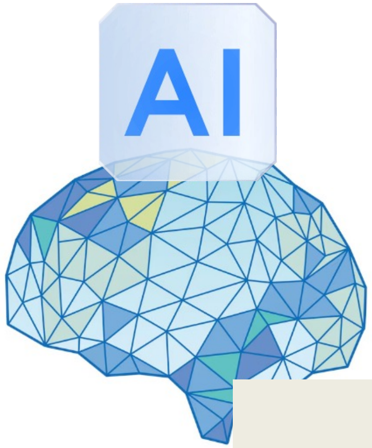


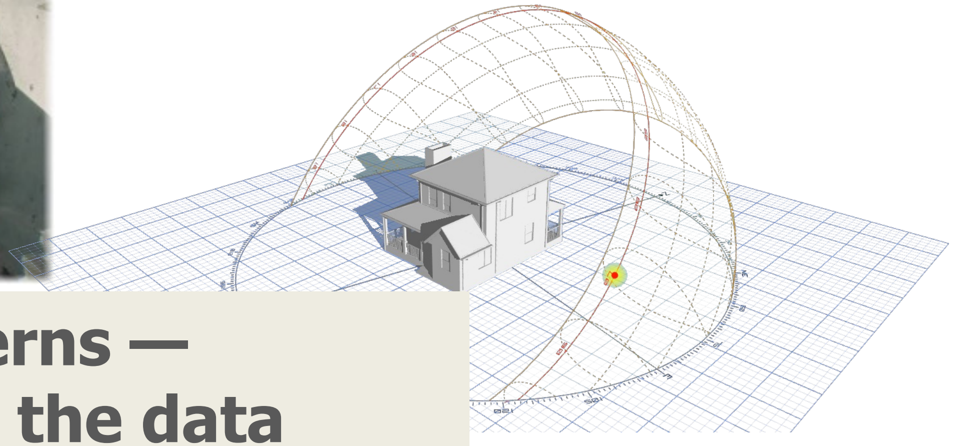
Usage ≠ Efficiency

- 🏠 **Size Effect**: Small inefficient homes, same usage as large efficient homes.
- ⚡ **Behavior Bias**: Changes with occupants / occupant lifestyle.
- 🌦️ **Weather Variability**: Year-to-year weather shifts distort results.

XeroHome™

What About Machine Learning?
LLMs, Neural Networks, AI





AI can find patterns —
Physics can explain the data

ML Model:
Usage Data In → Rating Out
(No explanation)

Physics-Based Model:
Understands how the
home works

- **Black Box**: Hard to explain why two homes get different ratings.
- **Data Burden**: Reliable ML needs tagged, standardized data at scale for millions of homes.
- **Missing Physics**: Without a building model, AI can't predict upgrade impacts

XeroHome™

A Scalable, Equitable Path:
'Building Energy Modeling (BEM) + Energy Usage Data'

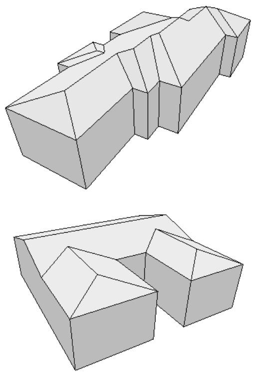
PUBLIC DATASETS

- Parcel Data, GIS Shapes Files
- Tax Assessor Data
- Building Permit Data
- MLS Data
- NREL ResStock Data
- Homeowner Input (via XeroHome)

+

BEM & COMPUTE

CLOUD PLATFORM



Develops Individual EnergyPlus™ Building Energy Models (BEM) At Scale

+

ENERGY USAGE DATA

UTILITY DATA API

Performs Data Analysis With Speed, Accuracy And Scale

=

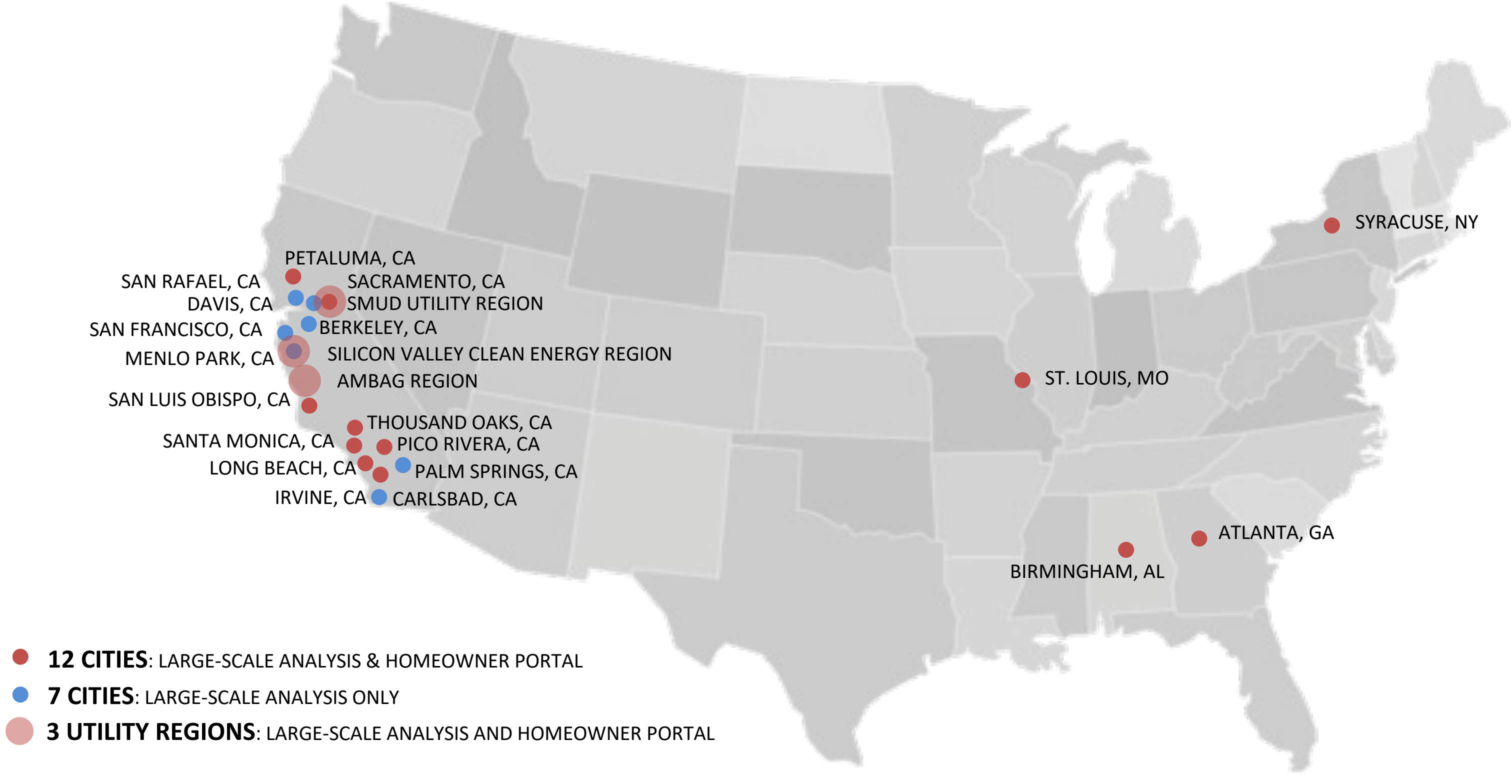
WEB-PORTAL

XeroHome™

Provides Home Upgrade Insights & Recommendations

XeroHome™

Piloting 'BEM + Energy Usage Data' Approach
XeroHome™ Deployments 2018 - 2025



Deployments funded by: Southern California Edison (SCE), Sacramento Municipal Utility District (SMUD), Pacific Gas & Electric (PG&E), Silicon Valley Clean Energy (SVCE), Association of Monterey Bay Govts. (AMBAG), City of Sacramento, New York State Energy & Research Development Auth. (NYSERDA), Ameren Missouri, National Grid New York, Alabama Power, Georgia Power, Electric Power Research Institute (EPRI)

XeroHome™

Analysis of Data from XeroHome™ Deployments

Analysis of XeroHome™ deployment data was done to demonstrate the feasibility of a ‘no cost to participant’ statewide home energy rating that is:

- ☑ **Feasible** to implement
- ☐ **Accurate** in its results ⇦ Today’s presentation
- ☑ **Scalable** across millions of homes

Acknowledgement: Analysis was funded by Southern California Edison, led by XeroHome™ (Vistar Energy) with technical support from 2050 Partners.

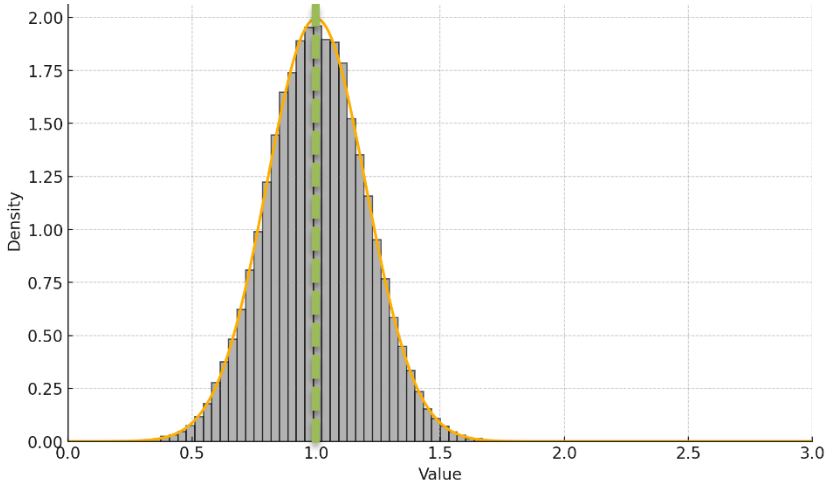
XeroHome™

Visualizing Model Accuracy

- Modeled electric energy use **compared to** metered electric energy use for each home

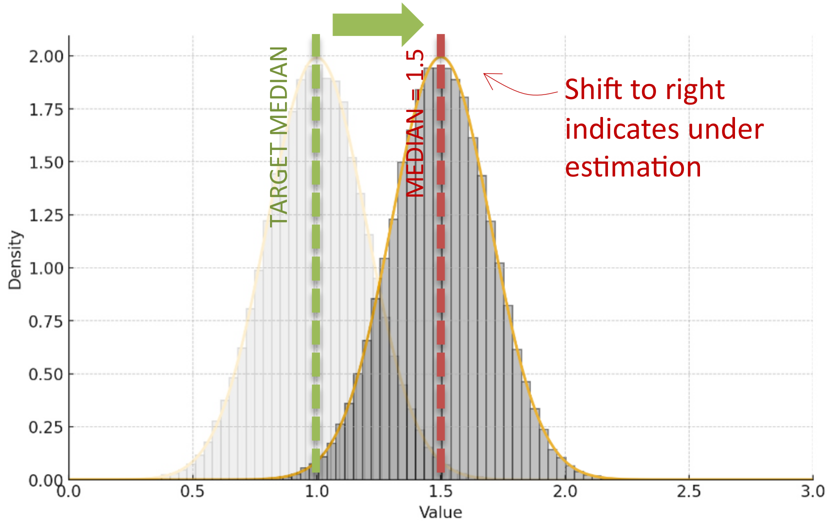
$$Calibration\ Coefficient = \frac{Energy\ Use_{Actual}}{Energy\ Use_{Modeled}}$$

- <1 Calib. Coef. means actual energy use was **less** than modeled
- 1 Calib. Coef. means an **exact match**
- >1 Calib. Coef. means actual energy use was **more** than modeled

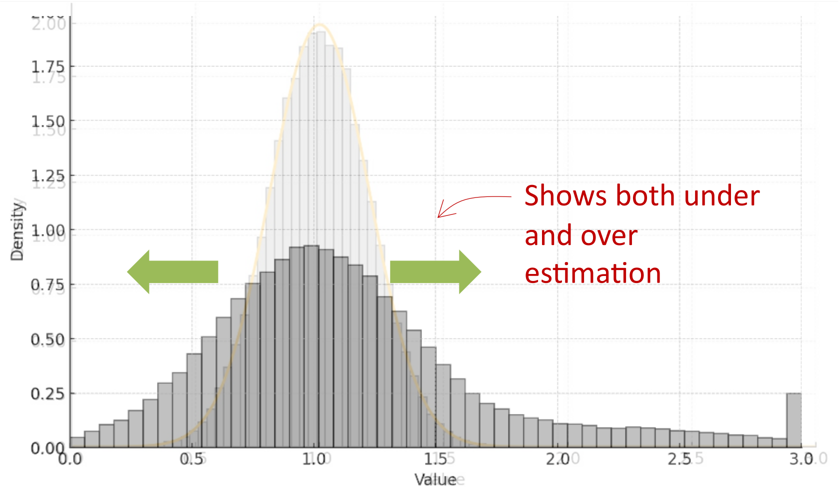


XeroHome™

Bias and Uncertainty



- When most predictions cluster to one side of the true value, the models have a consistent **bias** - systematically under or over estimation.



- When predictions spread widely around the true value, the models have high **uncertainty** - sometimes under, sometimes over estimation, but no consistent pattern.

XeroHome™

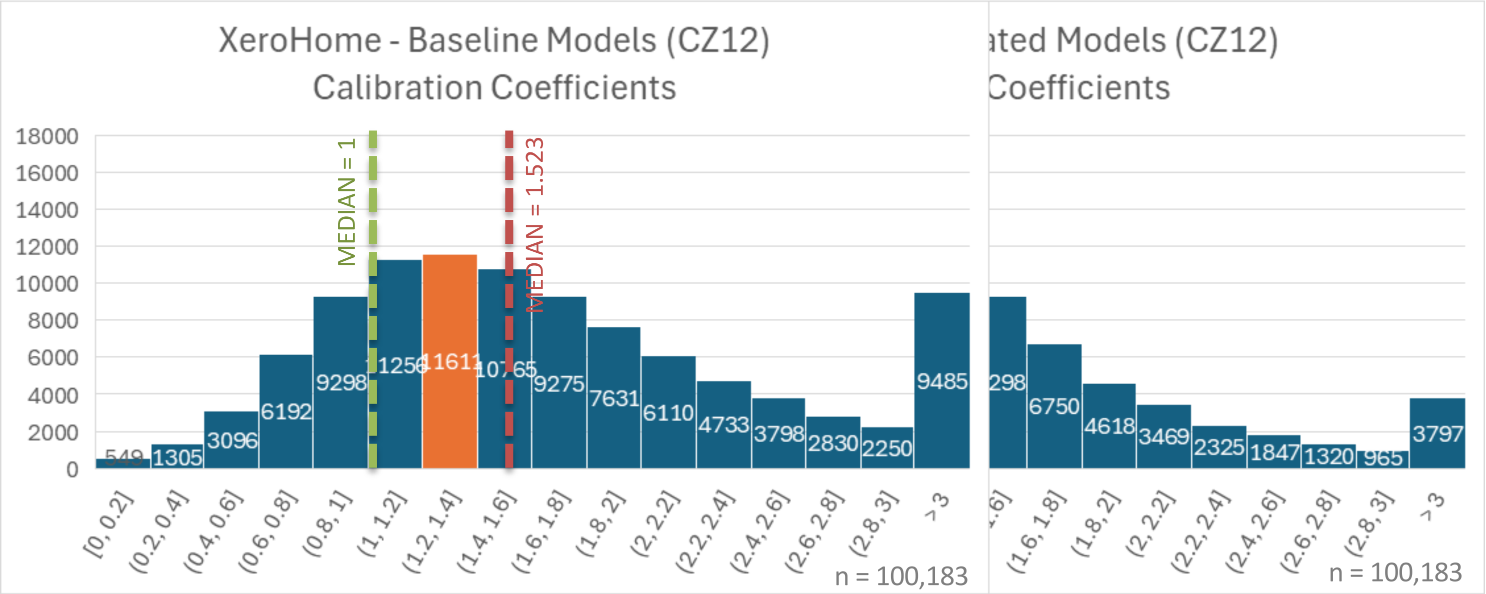
Hypothesis #1 – Adding Energy Use Data Improves Accuracy of the Home Energy Models

- Quantify the improvement in **Bias** and reduction in **Uncertainty** when energy use data is used to calibrate the home energy models.
- Is there a difference between mild (coastal) vs extreme (inland) climates?

Model Accuracy: Without / With Energy Data
Calibration Coefficients Histogram

Example 1 – **CZ12** XeroHome™ home energy models

Without Energy Data   With Energy Data



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
1.086	1.523	2.129	1.043	1.049

1st Quartile	Median	3rd Quartile	IQR	Std Dev.
		1.567	0.778	0.793

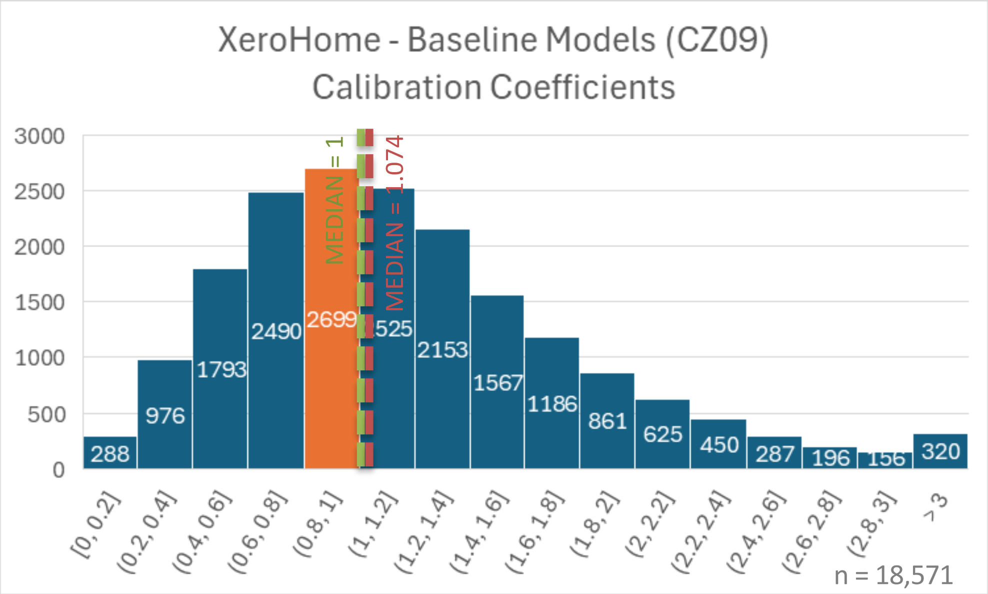
- **Bias**: **78% reduction in systematic bias**. Median moved closer to 1: from 1.523 → 1.117
- **Uncertainty**: **24% reduction in uncertainty**. Std Dev moved closer to 0: from 1.049 → 0.793

XeroHome™

Model Accuracy: Without / With Energy Data
Calibration Coefficients Histogram

Example 2 – **CZ9** XeroHome™ home energy models

Without Energy Data



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
0.734	1.074	1.521	0.787	0.734

- **Climate Zone 9** is a coastal (mild) climate and the home energy models showed high accuracy even without adding energy data.

XeroHome™

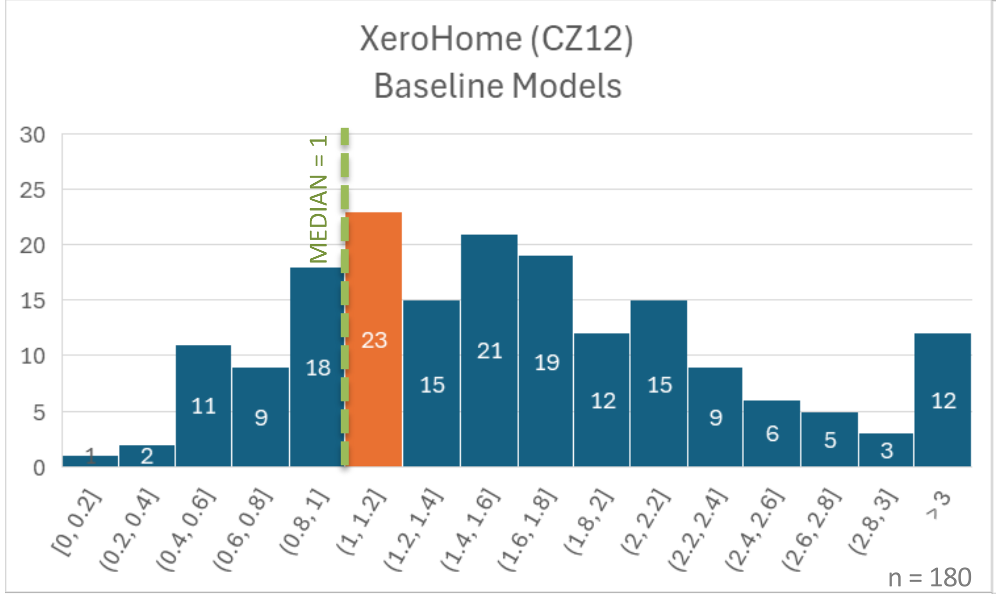
Hypothesis #2 – Adding Inputs Collected by a Homeowner Improves Accuracy of the Home Energy Models

- Quantify the improvement in Bias and reduction in Uncertainty when data collected by a homeowner is used to improve the assumptions in the home energy models.

Model Accuracy: Without / With Homeowner Inputs
Calibration Coefficients Histogram

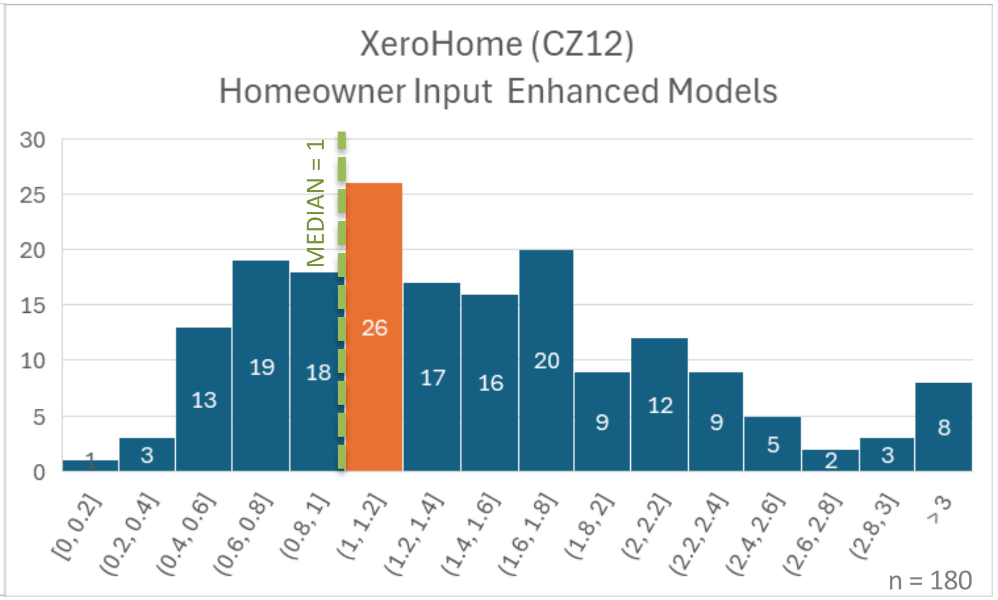
Example 1 – **CZ12** XeroHome™ Models built without and with access to homeowner data

Without Homeowner Input



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
1.048	1.510	2.066	1.018	0.983

With Homeowner Input



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
0.906	1.310	1.849	0.943	0.951

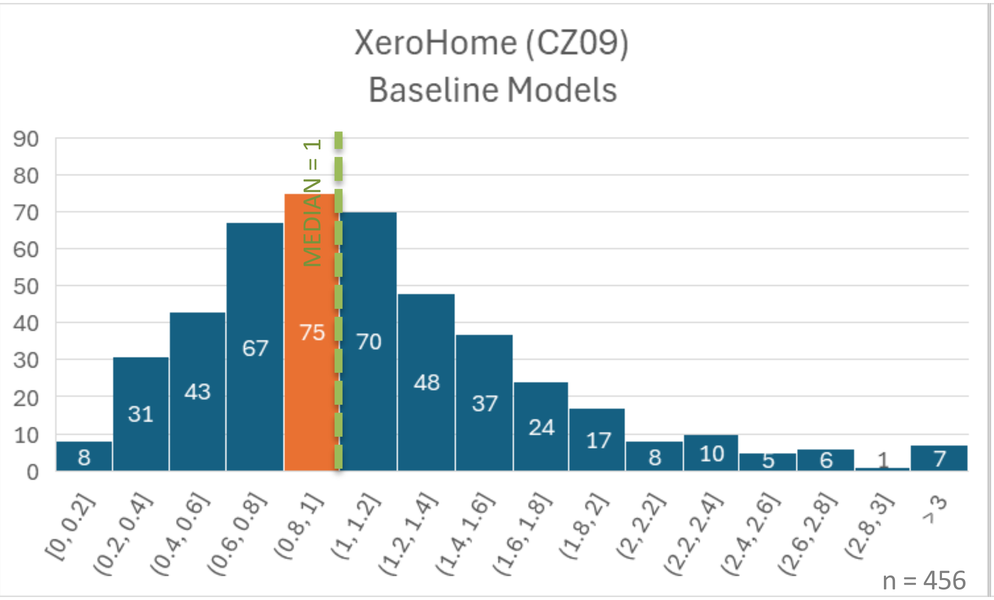
- **Bias**: **39% reduction in systematic bias**. Median moved closer to 1: from 1.510 → 1.310
- **Uncertainty**: **3% reduction in uncertainty**. Std Dev moved closer to 0: from 0.983 → 0.951

XeroHome™

Model Accuracy: Without / With Homeowner Inputs
Calibration Coefficients Histogram

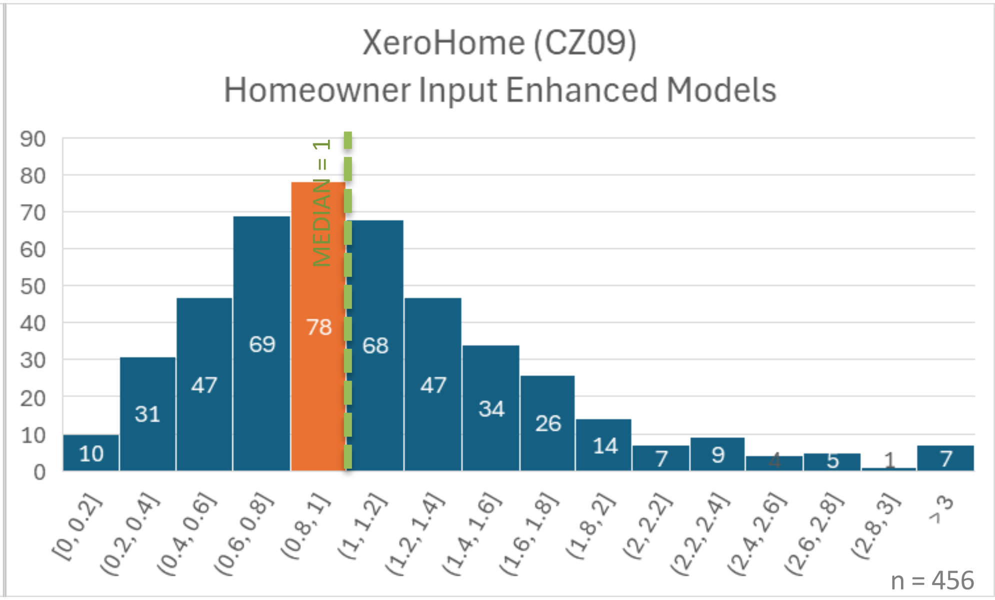
Example 2 – **CZ9** XeroHome™ Models built without and with access to homeowner data

Without Homeowner Input



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
0.706	0.982	1.350	0.643	0.650

With Homeowner Input



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
0.718	1.009	1.407	0.690	0.643

- Bias: **47% reduction in systematic bias**. Median moved closer to 1: from 0.982 → 1.009
- Uncertainty: **1% reduction in uncertainty**. Std Dev moved closer to 0: from 0.650 → 0.643

XeroHome™

Key Takeaways

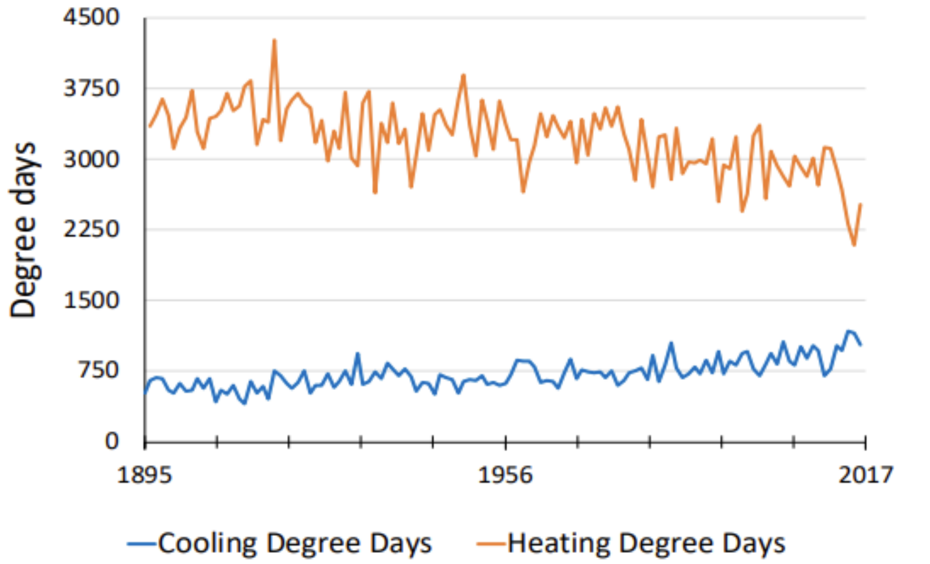
- **Hypothesis #1 – Energy Usage Data**: With energy data, modeling predictions can become both more accurate (less biased) and more consistent (less uncertainty), creating a stronger foundation for statewide ratings.
- **Hypothesis #2 – Homeowner Input**: Using homeowner inputs to replace key initial assumptions about the home can improve accuracy mainly making the models less biased.

XeroHome™

Discussion on Variance

- **Weather Impacts**: Energy use can vary year over year due to weather changes – Hotter years drive more cooling, less heating and vice versa

Figure 1. Cooling and Heating Degree Days



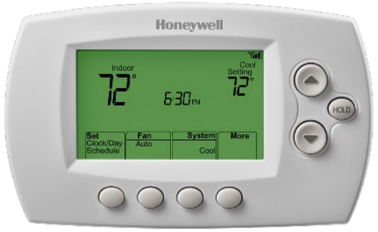
Note: Degree days measure the difference between the average daily temperature and a reference temperature of 65 degrees Fahrenheit (°F). Cooling degree days measure how much the average daily temperature is higher than 65°F; heating degree days, how much it is lower than 65°F. For example, an average daily temperature of 75°F corresponds to a cooling degree day value of 10. Each value shown in the graph is the sum of degree days for that year.

Source: NOAA, 2017a

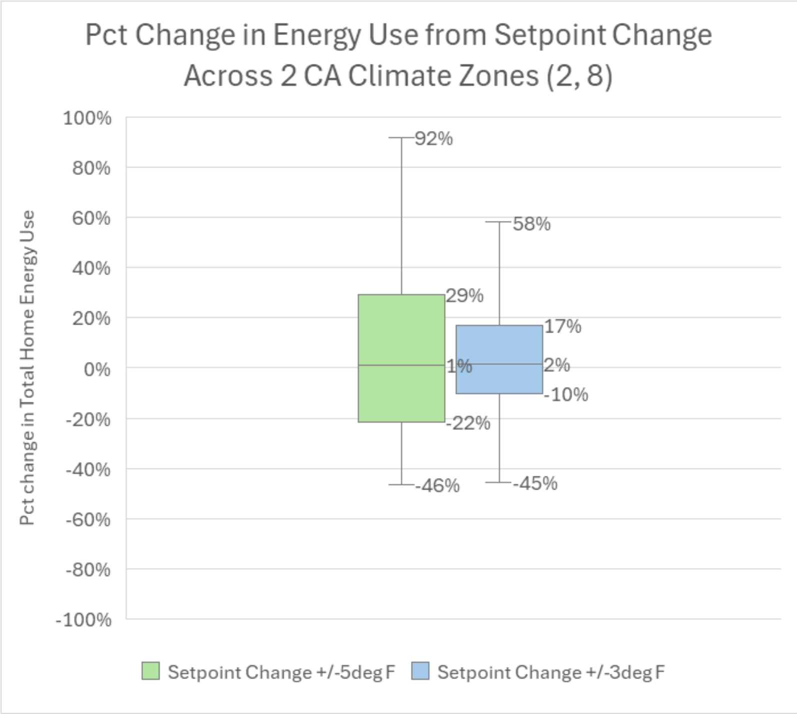
XeroHome™

Discussion on Variance

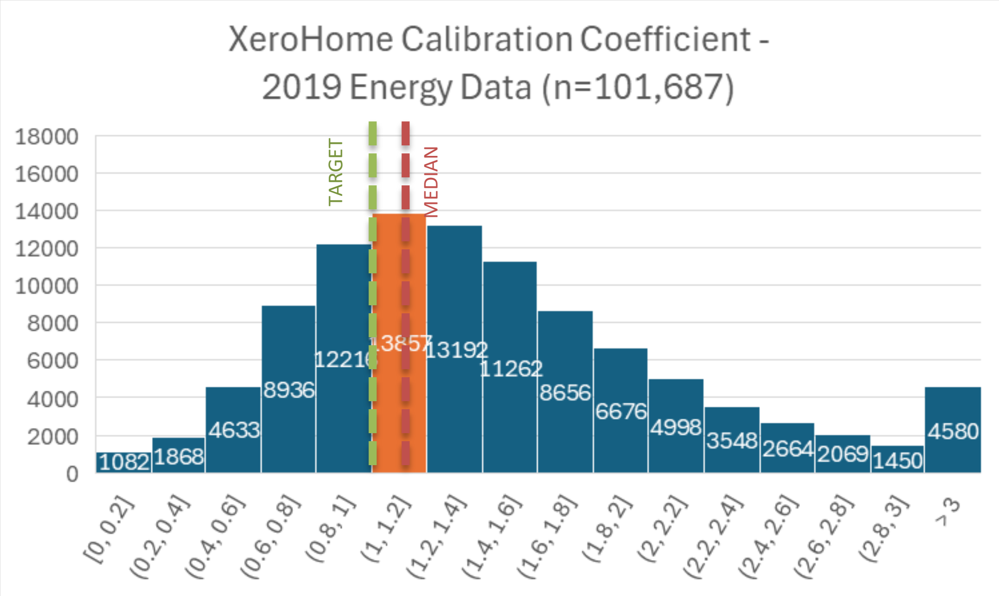
- **Behavioral Impacts**: Energy use can vary with operational changes – Thermostat settings, time spent at home (e.g., working from home), window operation, and number of occupants, etc.



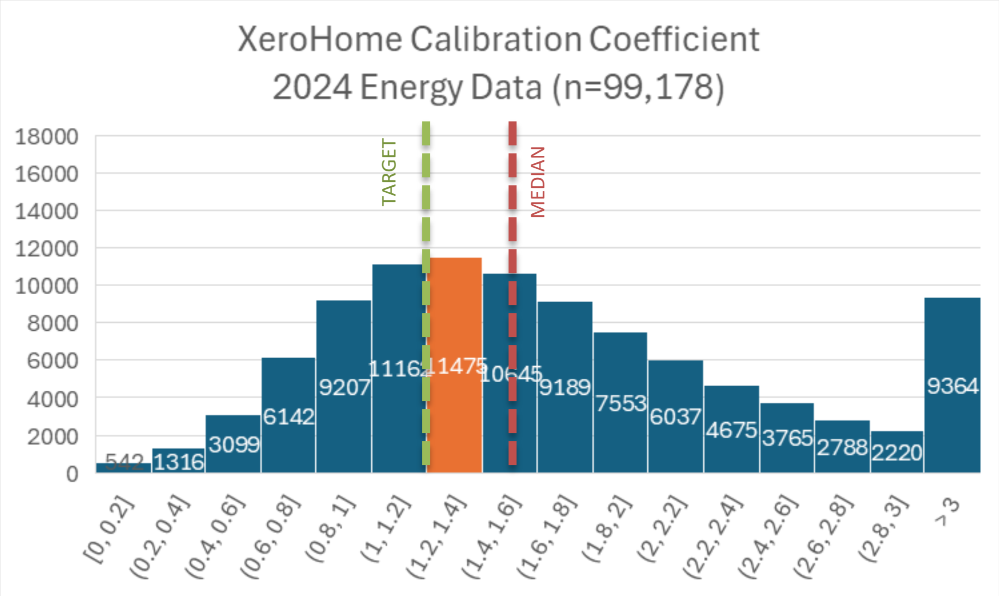
Setpoint changes:



Natural Energy Usage Variance
Due to Weather + Behavior



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
0.949	1.321	1.815	0.866	0.839



1st Quartile	Median	3rd Quartile	IQR	Std Dev.
1.085	1.521	2.127	1.043	1.048

Comparing the same home energy models against 2019 and 2024 energy data shows:

- **±7% variance in systematic bias** due to factors like weather, occupant behavior etc.
  Median varies: 1.321 – 1.521
- **±11% variance in uncertainty** due to factors like weather, occupant behavior etc.
  Std Dev varies: 0.839 – 1.048

These differences highlight the role of external factors – rather than model error – in year-to-year energy use variance.

Conclusions

- **'BEM + Energy Usage Data'** offers a scalable, no-cost path to statewide home energy labeling that is credible, equitable and cost-effective.
- Publicly available data can form the foundation of an energy model, integrating actual energy-use data and homeowner-provided details enhances accuracy.
- Even a well-calibrated home energy model may not perfectly align with measured energy use, due to inherent variability in a home's consumption patterns.

Mudit Saxena
CEO & Founder, XeroHome™
MSaxena@xerohome.com
https://about.xerohome.com

XeroHome™