Predictive EE Measurement:
Using Data Science to Enable a Reliable Resource

John Backus Mayes

Senior Data Scientist at EnergySavvy

Presented at the 2017 ACEEE National Conference on Energy Efficiency as a Resource

ENERGYSAVVY

Agenda

Topics of discussion

1. Who is EnergySavvy?
2. Continuous EE Measurement
3. Reliable EE Predictions

EnergySavvy at a Glance

Data-driven personalization for the utility customer experience

aps
AEP OHIO
Black Hills Corporation
Columbia Gas of Ohio
A NiSource Company
Columbia Gas of Massachusetts
A NiSource Company
cps ENERGY
EnergyTrust of Oregon
EVERSOURCE
focus on energy
Partnering with Wisconsin utilities
Greater Cincinnati Energy Alliance
1866
Hartford Steam Boiler
INDIANA MICHIGAN POWER
Liberty Utilities
mass save
Savings through energy efficiency
MINNESOTA ENERGY RESOURCES
nationalgrid
New Jersey Natural Gas
New Mexico GAS COMPANY
nyserda
Energy. Innovation. Solutions.
OG&E
PACIFICORP
PNM
PSEG LONG ISLAND
PSE PUGET SOUND ENERGY
SRP
SOUTHERN CALIFORNIA EDISON
An EDISON INTERNATIONAL Company
SOUTHWEST GAS
Unitil
Washington Gas

Quick Facts

- Founded in 2008
- More than 30 utilities and state programs
- Seattle and Boston offices

Setting the Stage

Definitions, context, points of clarification…

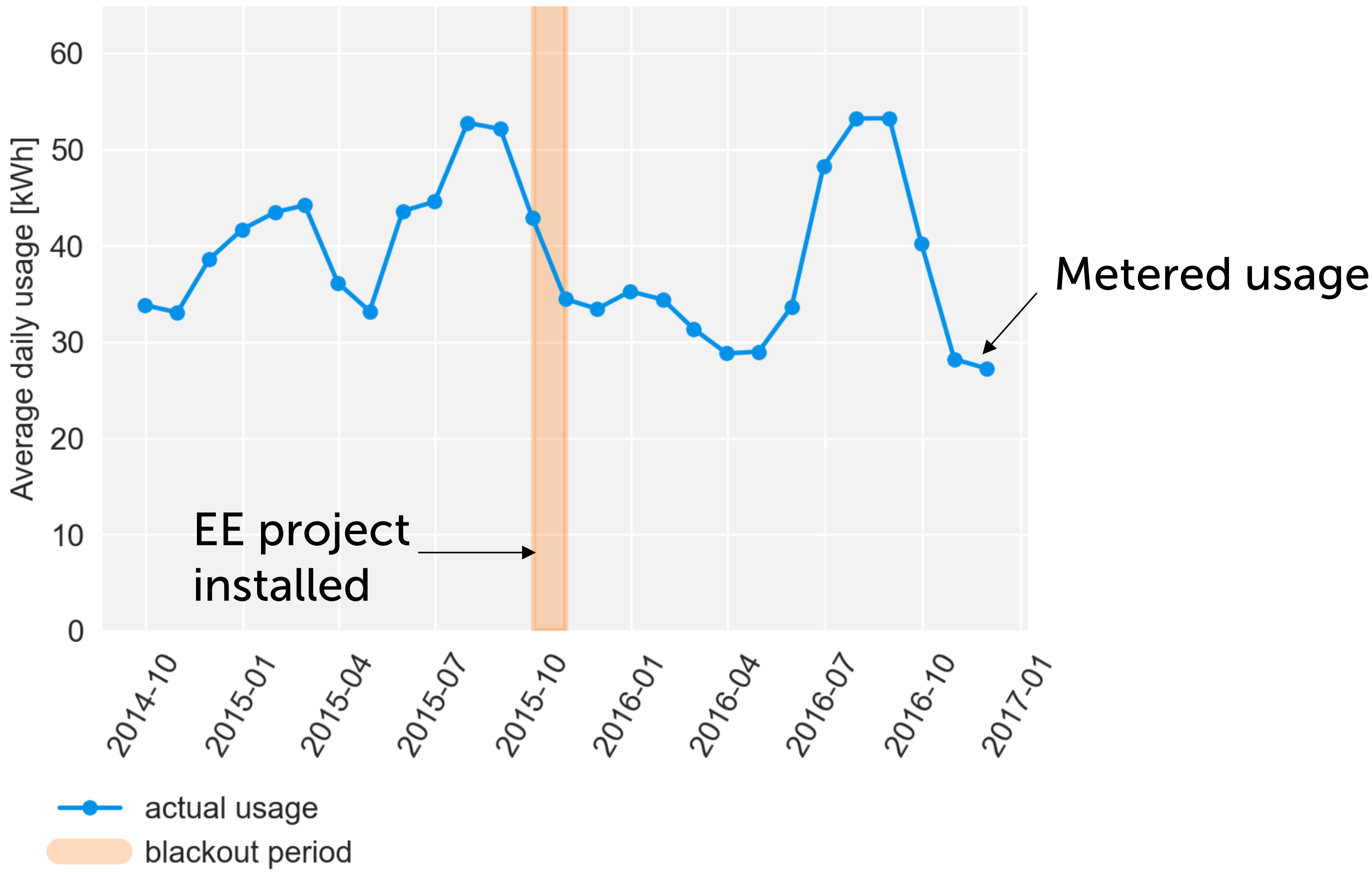
- Our savings measurement software is a tool that enhances EE program management and evaluation—it does not replace formal third-party evaluation.
- Our methods build on existing best practices for whole-building consumption data analysis—see for example the [Uniform Methods Project, Chapter 8].
- This approach is a good fit for many residential and SMB programs; applicability to large C&I is less certain.
- Everything presented here is possible with either monthly billing data or AMI data: "Advanced M&V" does not require AMI.

Continuous EE Measurement

Building on best practices in billing analysis

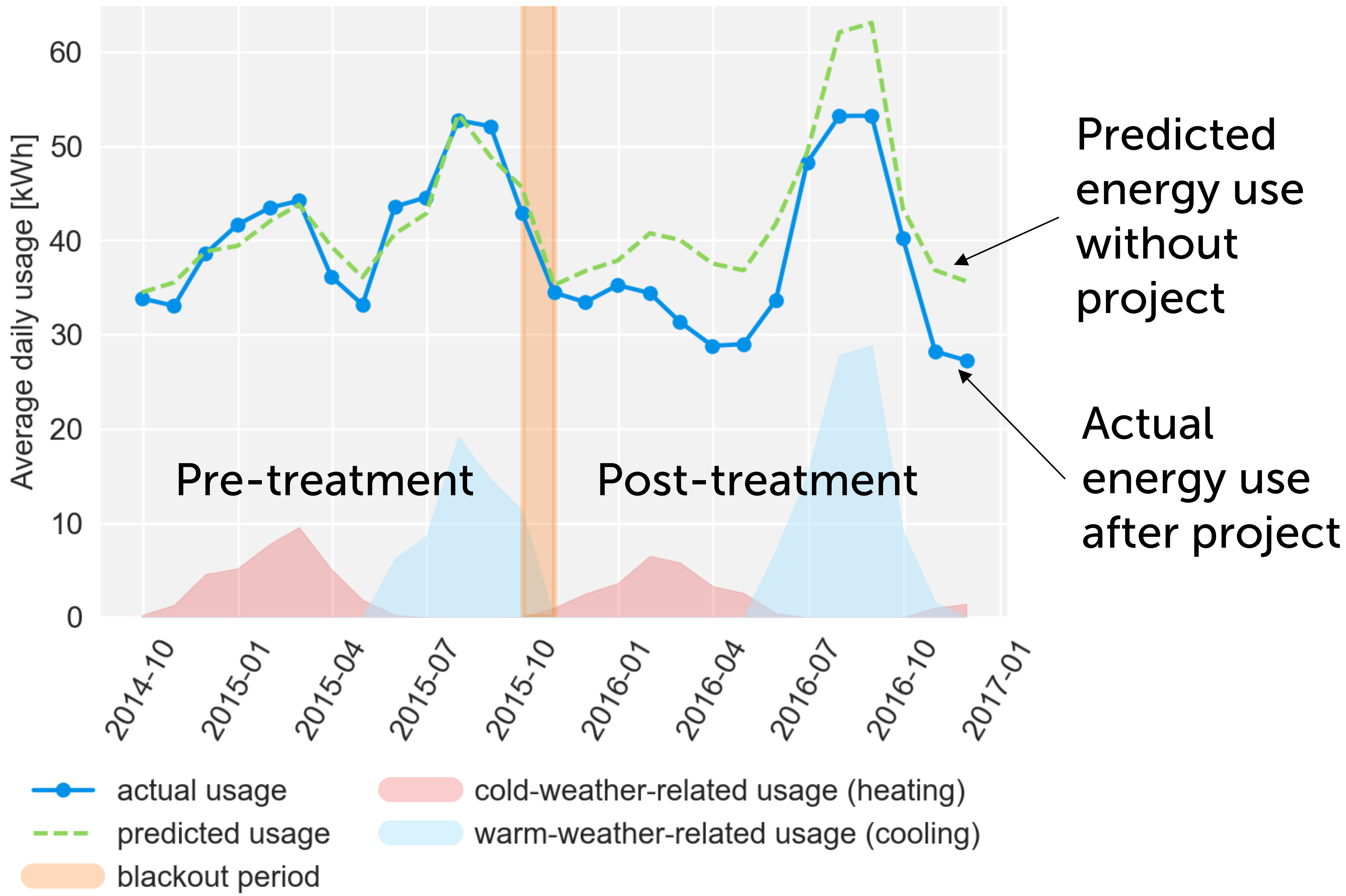
Weather-Normalized Savings

Simple pre-post approach adjusted for changes in temperature



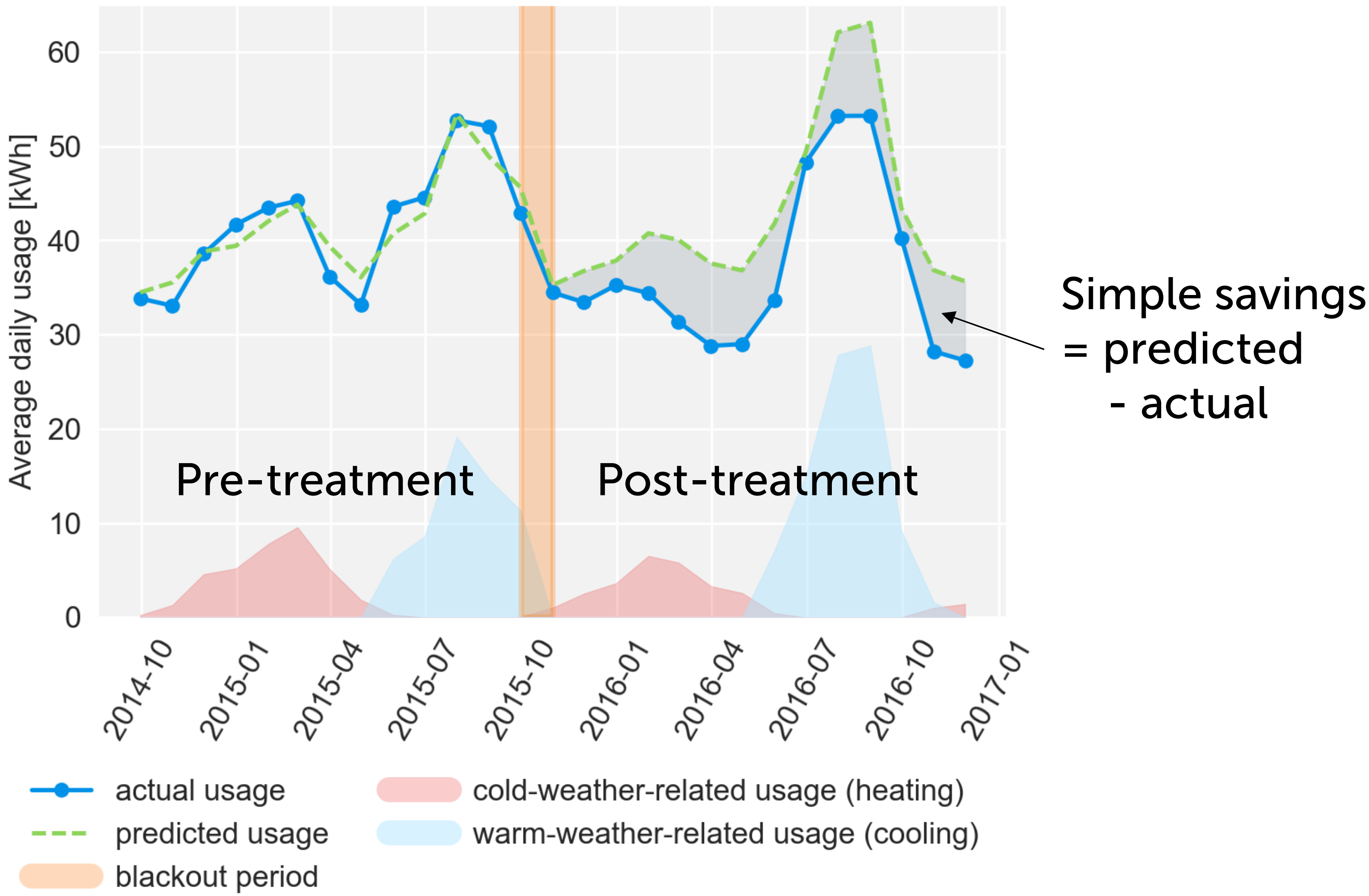
Weather-Normalized Savings

Simple pre-post approach adjusted for changes in temperature



Weather-Normalized Savings

Simple pre-post approach adjusted for changes in temperature

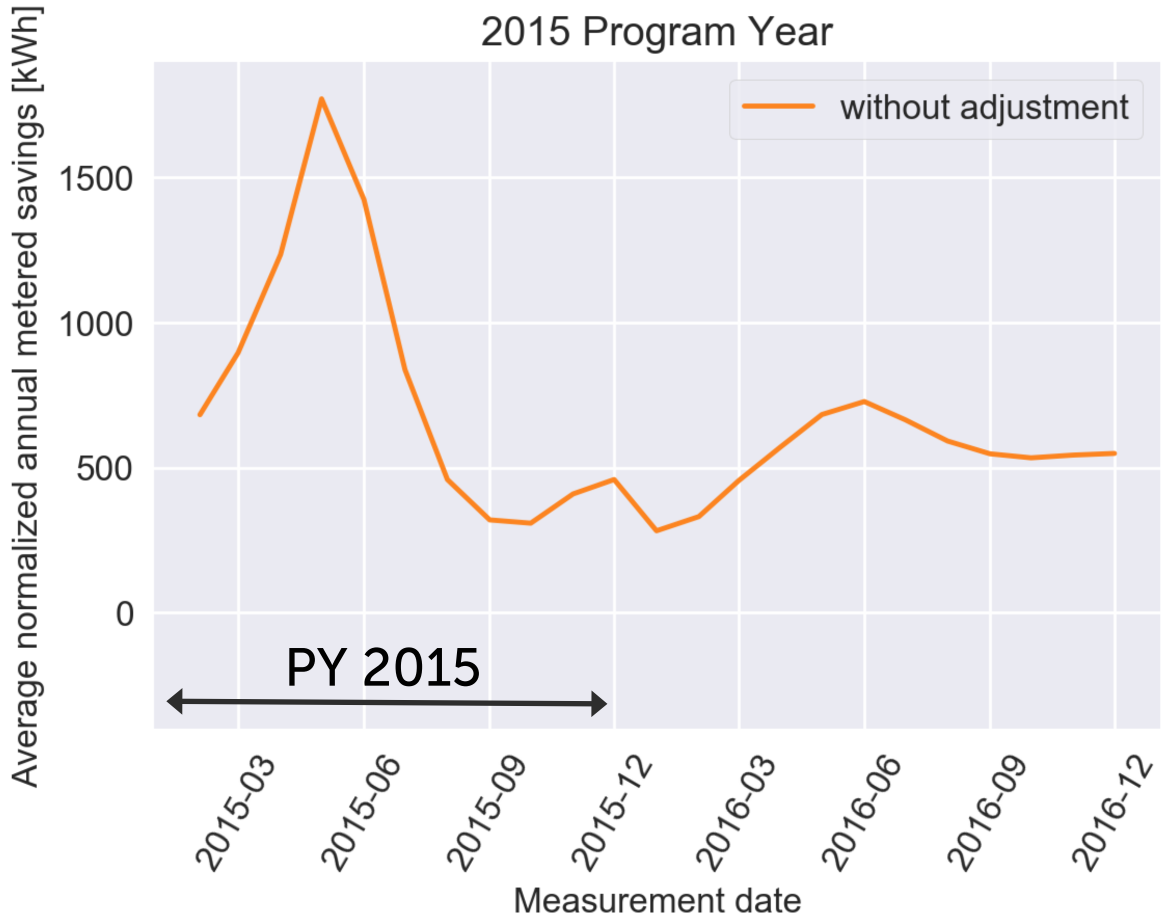


Aggregate Savings

Tracking program performance through savings measurement

Aggregate weather-normalized savings show large variability over time.

Are these variations related to EE program activity?



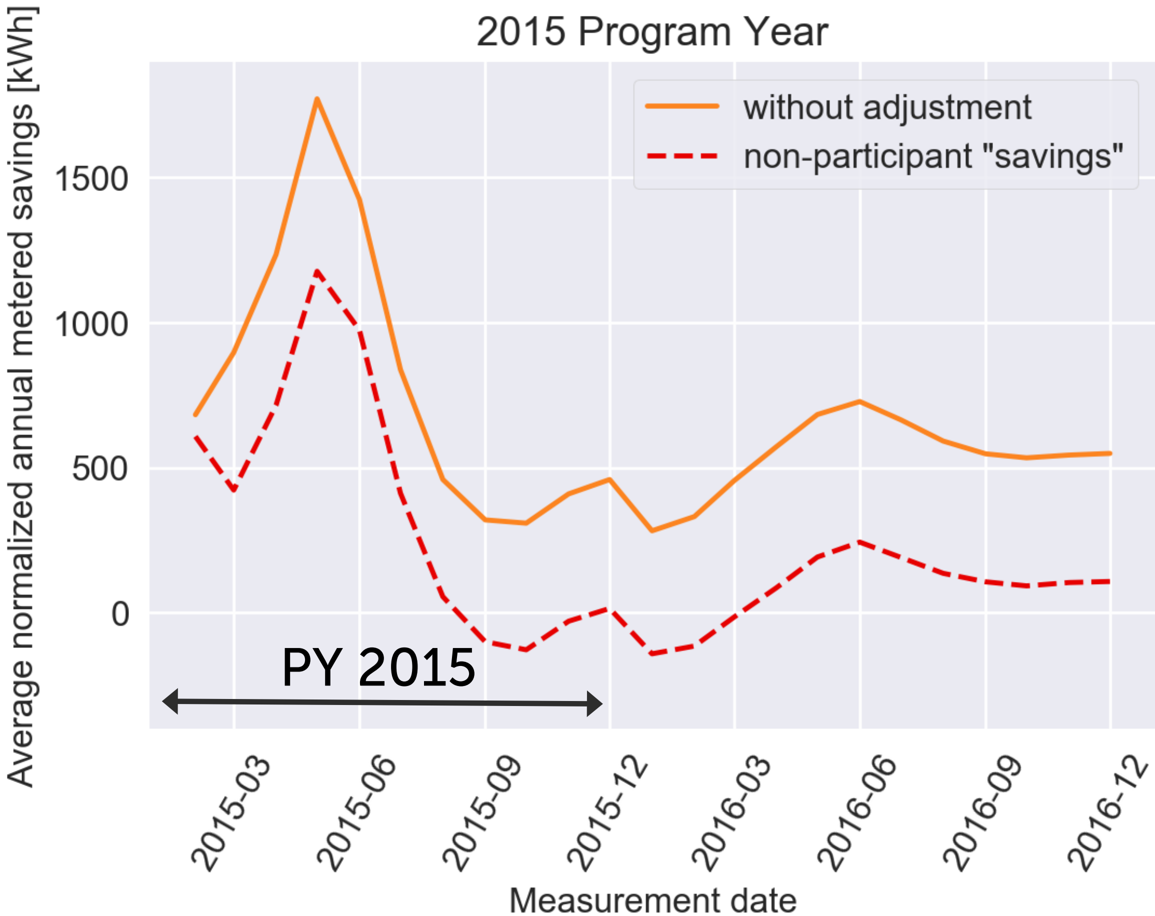
Exogenous Effects

Large confounding factors not captured by weather normalization

What if we do the same analysis for similar homes that did not participate in the program?

Non-participant "savings" may arise from many sources:

- Economic trends
- Extreme weather
- Modeling error
- Etc...



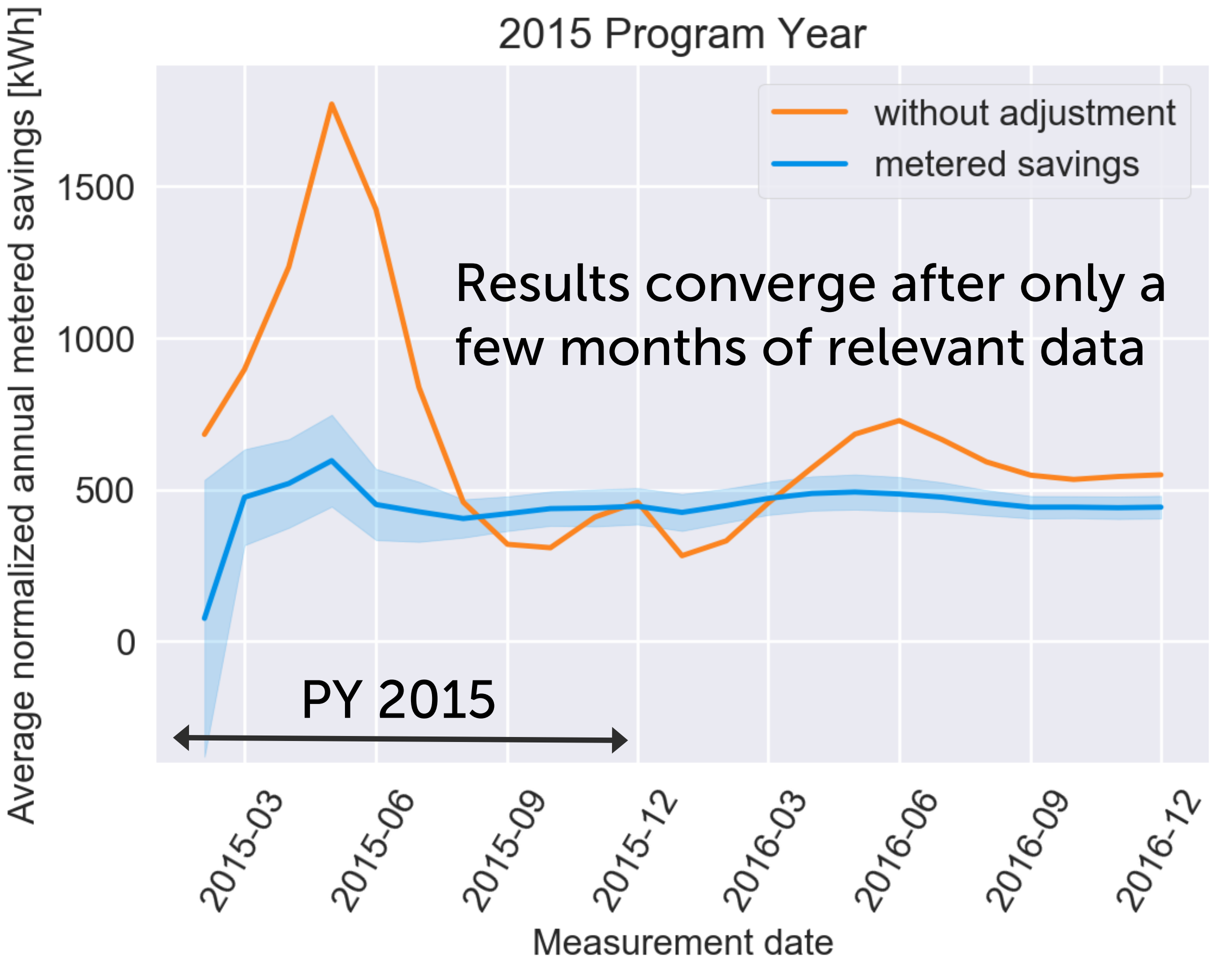
Predicting Exogenous Effects

Moving beyond the matched comparison group

- The goal: predict the exogenous change in usage at a home, using all the relevant data we can obtain about that home.
- Machine learning (ML) is a great fit for this problem: approximate unknown predictive relationships in a large, diverse data set.
- We've used a standard ML algorithm (Random Forests) for this purpose, training the model with a large sample of representative non-participants.

"Metered" EE Savings

Subtract the exogenous effects to obtain consistent EE measurements



Even with 12 post-treatment months across all projects, exogenous effects can amount to 20-50% of average savings.

Reliable EE Predictions

Moving beyond deemed savings

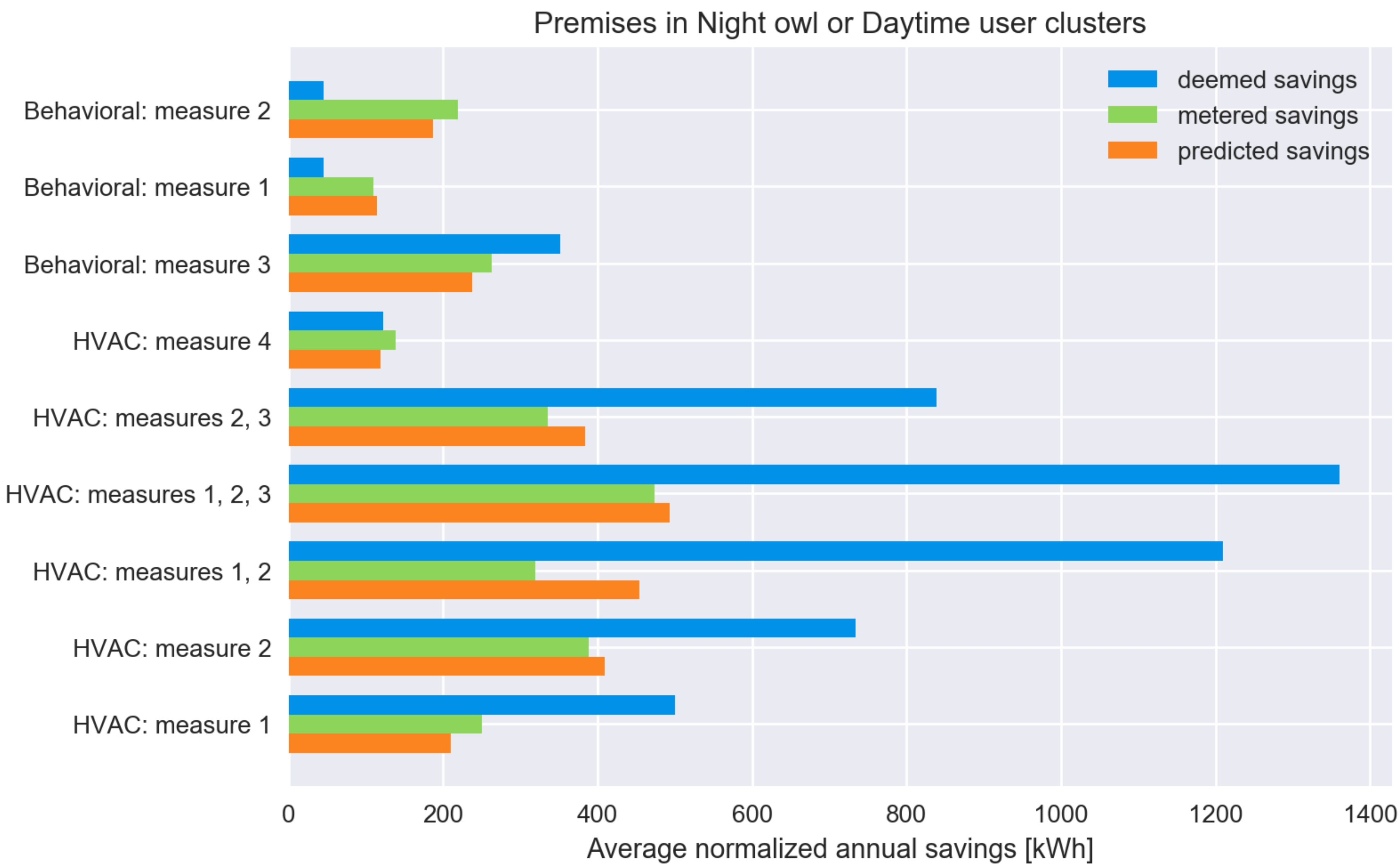
Predictive Models of Metered Savings

How do EE savings vary between homes, from project to project?

- Deemed savings tend to be optimistic, and they offer little or no granular insight into program or measure performance.
- With metered savings and machine learning, we can predict how much specific customers are likely to save from specific EE measures.
- This provides realistic, granular predictions for future EE planning and targeted marketing.

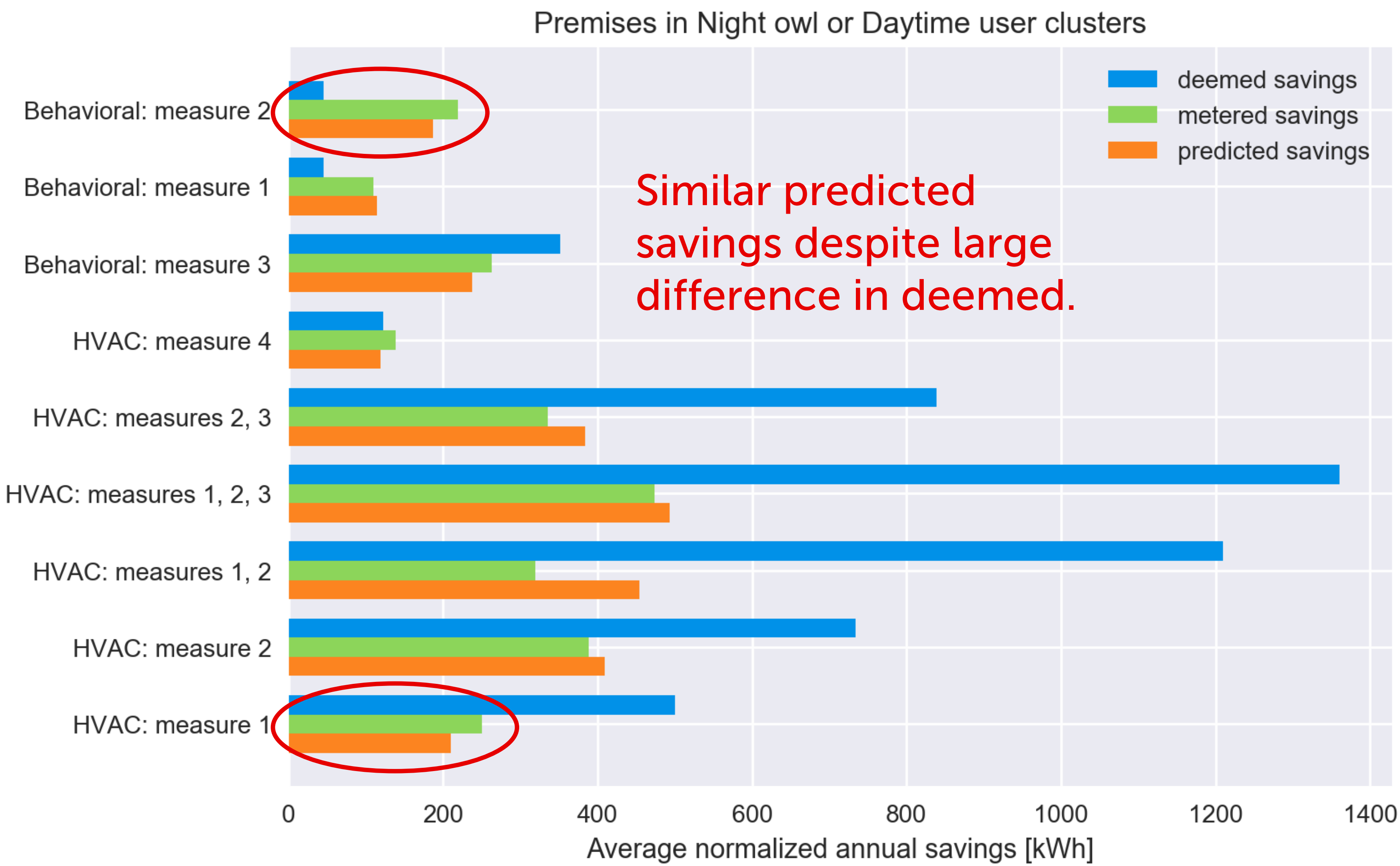
EE Savings Predictions

Reliable expectations for any customer and measure mix



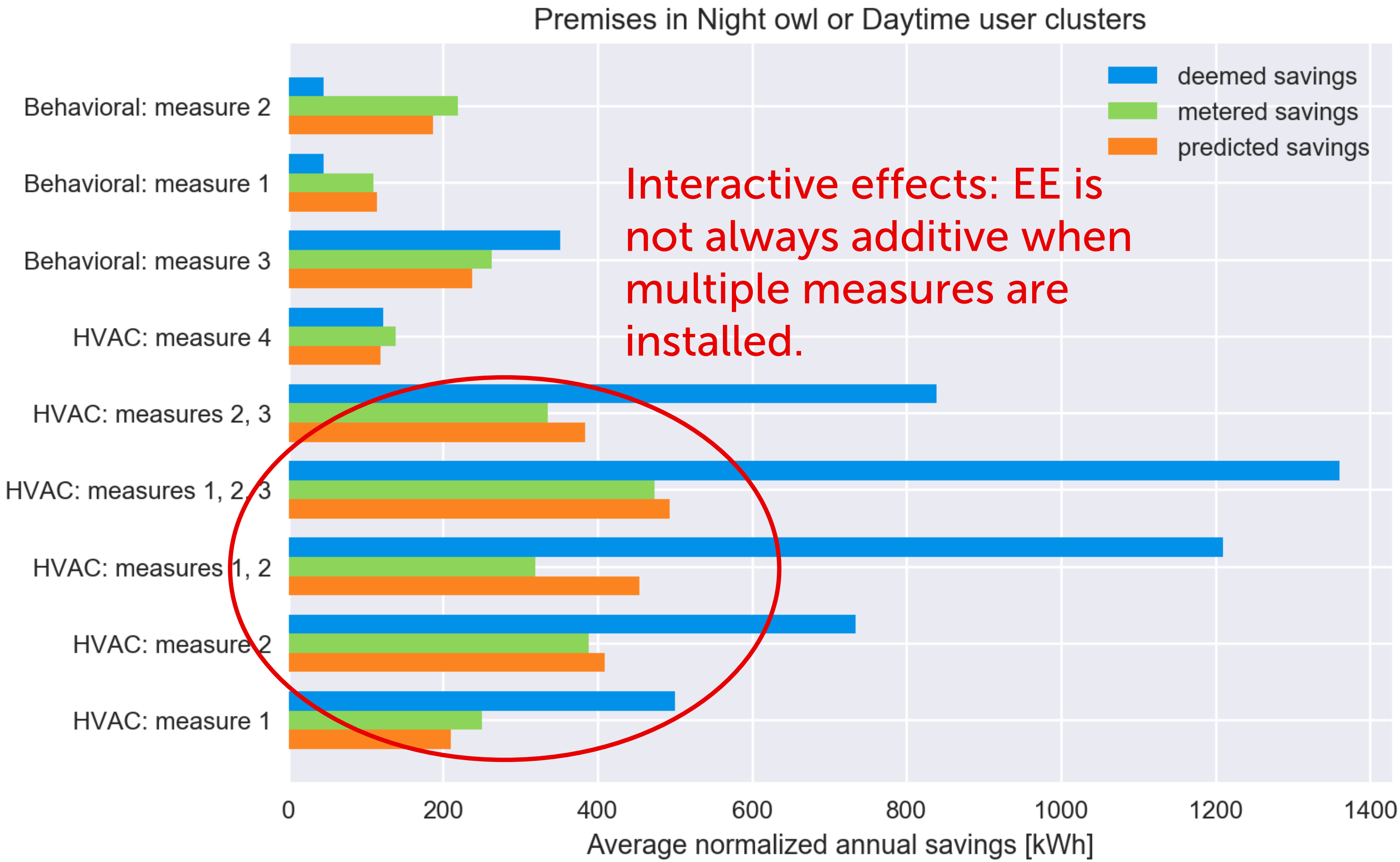
EE Savings Predictions

Reliable expectations for any customer and measure mix



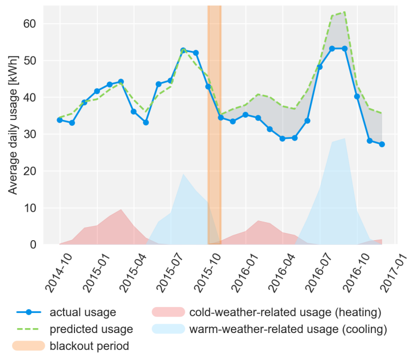
EE Savings Predictions

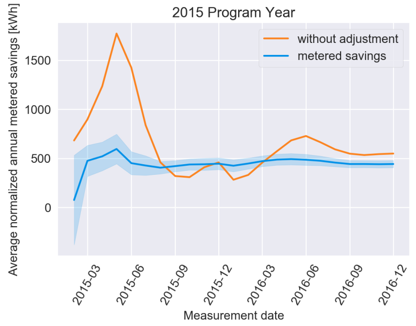
Reliable expectations for any customer and measure mix

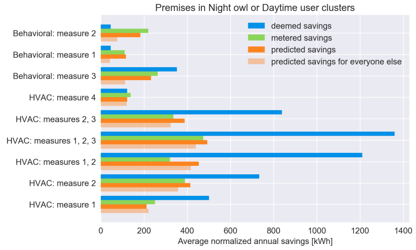


In summary…

- Weather-normalized savings is a good starting point, but it's not sufficient—we can do better.
- For continuous and granular EE measurement, it's crucial to understand exogenous effects.
- With machine learning, we can leverage all the available data to predict exogenous effects and uncover real insight into program performance.
- With reliable savings predictions, EE can be confidently deployed where it matters most.







Thank You

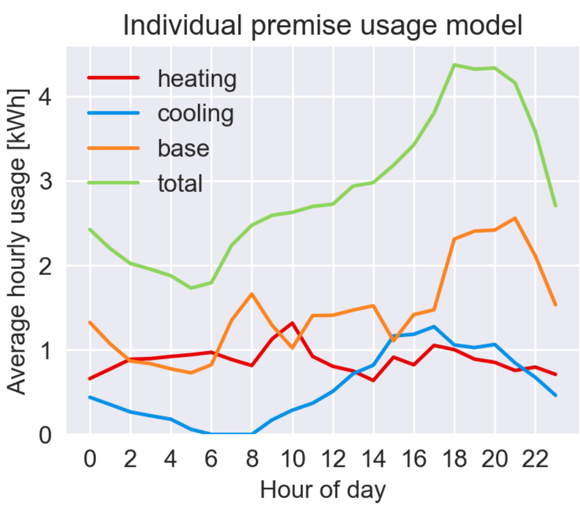
ENERGYSAVVY

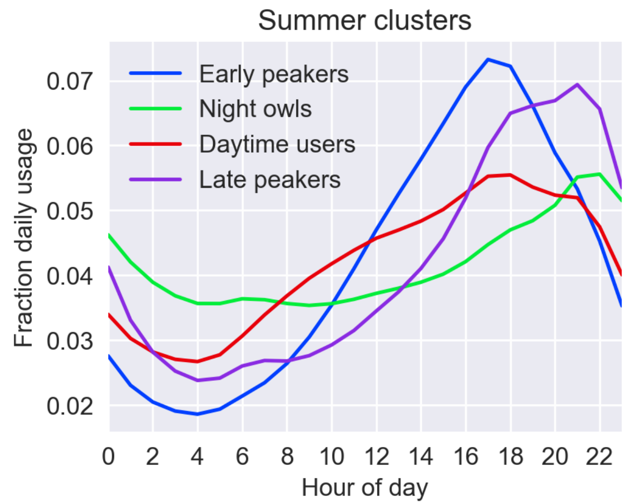
John Backus Mayes
Senior Data Scientist
john@energysavvy.com

BACKUP

Training the Model

Insights derived from AMI data are useful predictive features





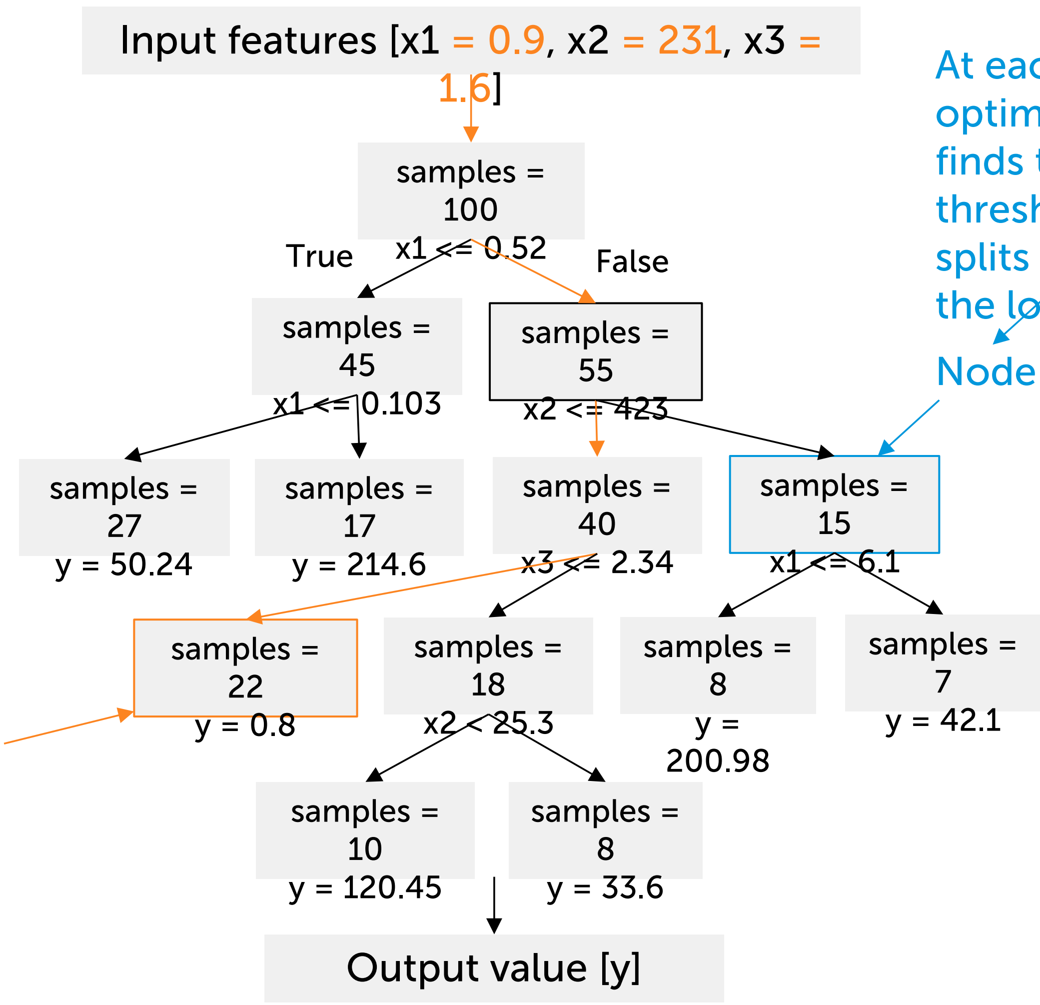
<u>Additional information</u>

- Premise
- Location
- Demographics

Metered savings model
(variant of Random Forest)

Predicted savings

Decision trees: the building block of random forests



At each node an optimization algorithm finds the feature and threshold value that best splits the data (results in the lowest MSE)

Node

Each sample in the data follows a path determined by a series of true/false statements based on its feature values to reach a predicted output value

Combine decision trees to make a random forest

Many weak learners come together to make a strong learner

- Individual decision trees are sensitive to noise in the training data
- Averaging the predictions from a forest combats this tendency to overfit as long as trees are not correlated
  - Reduce correlation with random subsets of the training data, and optimization of random sets of features at nodes