Northeast Energy Efficiency Partnerships

EM&V 2.0 – Cutting through the fog: What's Different and What's the Same (relative to traditional EM&V)

June 23, 2016

OVERVIEW

1. Background and Introductions
2. Continuous Measurement for Residential DSM
3. Industry and R&D Needs: Key Issues in M&V 2.0
4. Q&A
5. Next Steps

BACKGROUND

Forum EM&V 2.0 Project

1. March 30 EM&V 2.0 Workshop (Hartford CT)
2. June 23 EM&V 2.0 Webinar (today!)
3. **Sept 22 EM&V 2.0 Workshop (Part 2) Berlin, CT** preceded by Residential lighting and HEMS workshops (Sept 20-21). Registration info forthcoming
4. **November EM&V 2.0 Briefing -** research on developments in industry, vendor profiles, case studies, and update on piloting efforts in region

BACKGROUND

The M&V 2.0 Tool Kit – What's new?

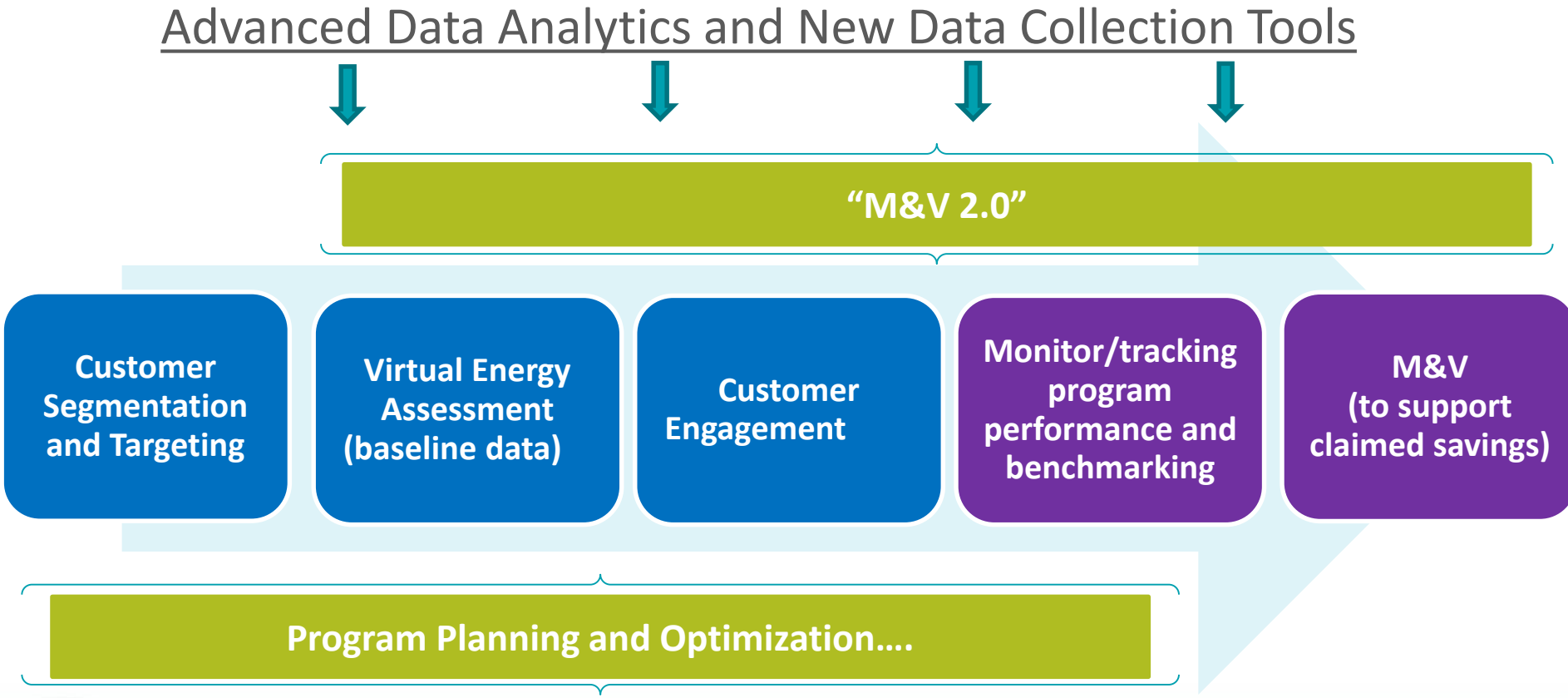
Advanced Data Analytics	**Software:** The growing range of cloud-based software and platforms that process large volumes of data quickly using publically available or proprietary algorithms. These solutions present opportunities for analyzing the increased data collected in order to inform utility programs and EM&V efforts.
Improved Data Collection Tools and Increases in Data Availability	**Hardware:** The use of smart meters and smart thermostats, non-intrusive load monitoring and other tools that both collect energy usage data in new ways and present opportunities to increase the amount and type of data collected.

- In some cases models fall on spectrum between public code and proprietary algorithms i.e., "open methodologies"

See also NEEP's *Changing EM&V Paradigm Report (Dec 2015)*

BACKGROUND

M&V 2.0 'Applications'



M&V 2.0 can apply to measure or whole building level – *our focus today is on whole building…*

BACKGROUND - Definitions

M&V 2.0: the use of more data (interval or volume), analytics, and computation at scale that can help to streamline the M&V process through semi/automation

- Use of (semi) automated software platforms (SaaS, private or public domain models) can support whole-building M&V at scale. *Forum interest is to pilot/test software models and their accuracy in estimating savings relative to traditional M&V, and inform development of software standards that meet specific acceptance criteria*

EM&V 2.0: the integration of M&V 2.0 into program evaluation that evolves traditional EM&V in terms of:

- Evaluation 'preparedness' i.e., real-time access to data and analysis during program implementation.
- Explore whether M&V 2.0 can support claimed savings, leveraging increasing volume and frequency of data to increase sample sizes, reduce on-site visits, and provide more granular results for evaluation. *Forum interest is to test coordination and data sharing between software platforms and evaluators to help streamline evaluation process*

POLL #1-2: Questions on clarify of definitions

Broad Categories of EM&V Approaches

Approach	Meter based?	Net or Gross?	Program/measure sweet spot
Deemed values	Not directly	Gross, Net in some cases	Efficient equipment replacement/installation
Engineering estimates, calculated	Not typically	Gross, Net	Custom industrial and large commercial; new construction
Billing Analysis that can include comparison groups, randomized control trial, or quasi-experimental	Yes, with other data	Usually Net	Programs w large numbers: residential, behavioral, small savings/site
Calibrated simulation modeling (IPMVP Option D)	No (except the calibration)	Gross, Net	Retrofit, large commercial
Whole-building and retrofit isolation M&V (IPMVP Option C, B)	**Yes, with other data**	**Gross**	**Commercial, multi-measure, interactive effects, operational measures**
…	…	…	…

Our focus for today….

SPEAKERS

Diane Duva
Director, Bureau of Energy and Tech Policy
CT Dept of Energy & Environ Protection

Tim Guiterman
EnergySavvy
Director of EM&V Solutions

Dan Zaweski
PSEG Long Island
Manager – Planning & Evaluation

Jake Oster
EnergySavvy
Sr. Director of Regulatory Affairs

Jessica Granderson
Lawrence Berkeley National Lab
Staff Scientist

BACKGROUND

Why the interest in EM&V 2.0?

- 1-2 year lag in impact evaluation results post-program year is inefficient and costly
- As energy efficiency investments scale, EM&V needs to evolve to support growing industry
- Evolving integration of EE, DR and other distributed resources → how can EM&V evolve to address impact of multiple DERs at whole building level
- How can we improve cost-effectiveness of EM&V (not necessarily lower the cost)?

Traditional Evaluation Timeline

Planning & Scoping	Recruiting & Data Collection	Data Analysis	Reporting
Stakeholder engagement	3-6 months of metering	Site data analysis	Stakeholder engagement
Data cleaning & Sampling	Recruitment, installation & retrieval	Site reporting	Reporting & review of results

EM&V 2.0: Continuous Measurement for Residential DSM

Tim Guiterman

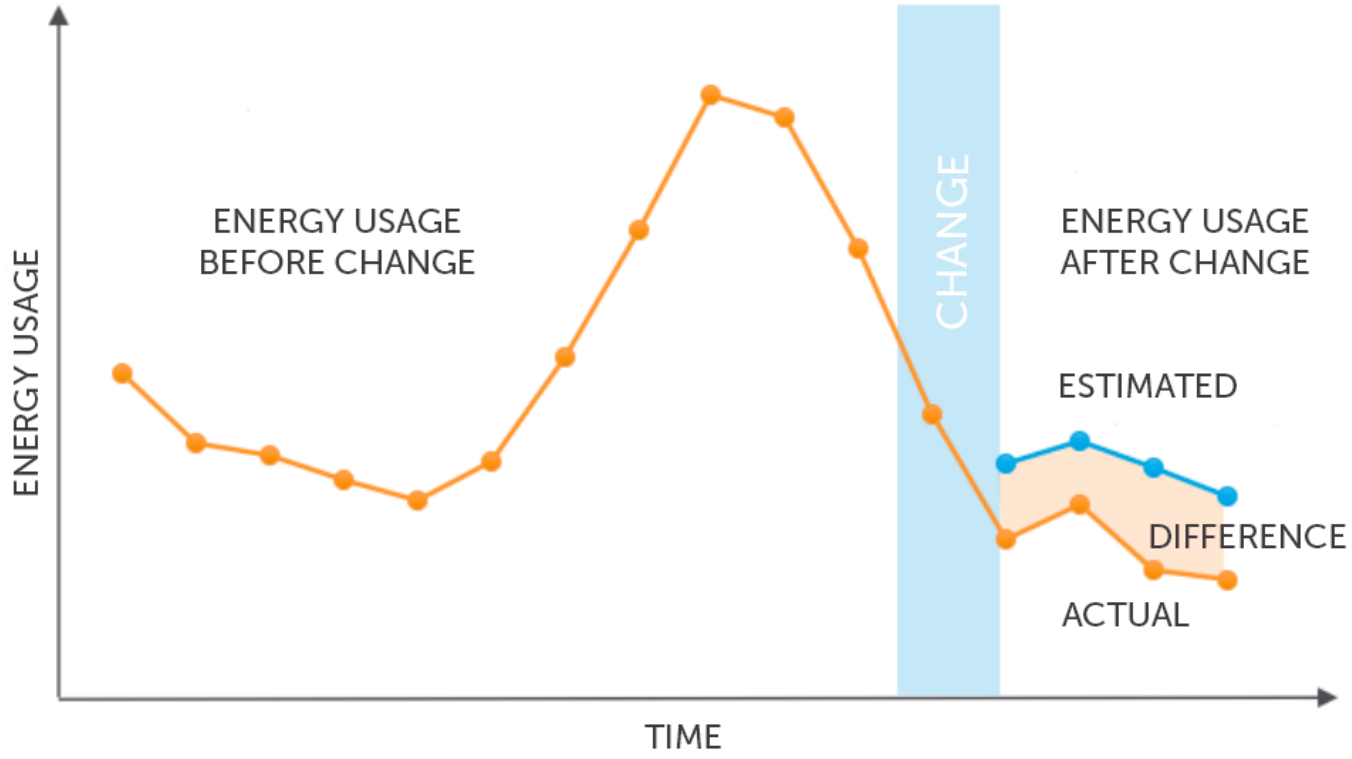
Director of Quantify Solutions

June 23, 2016

REGIONAL EVALUATION,
MEASUREMENT & VERIFICATION FORUM

ENERGYSAVVY

How Does Automated M&V Work?



What Does Automated M&V Look Like?

Continuous Measurement & Monitoring

Feedback on program performance during implementation

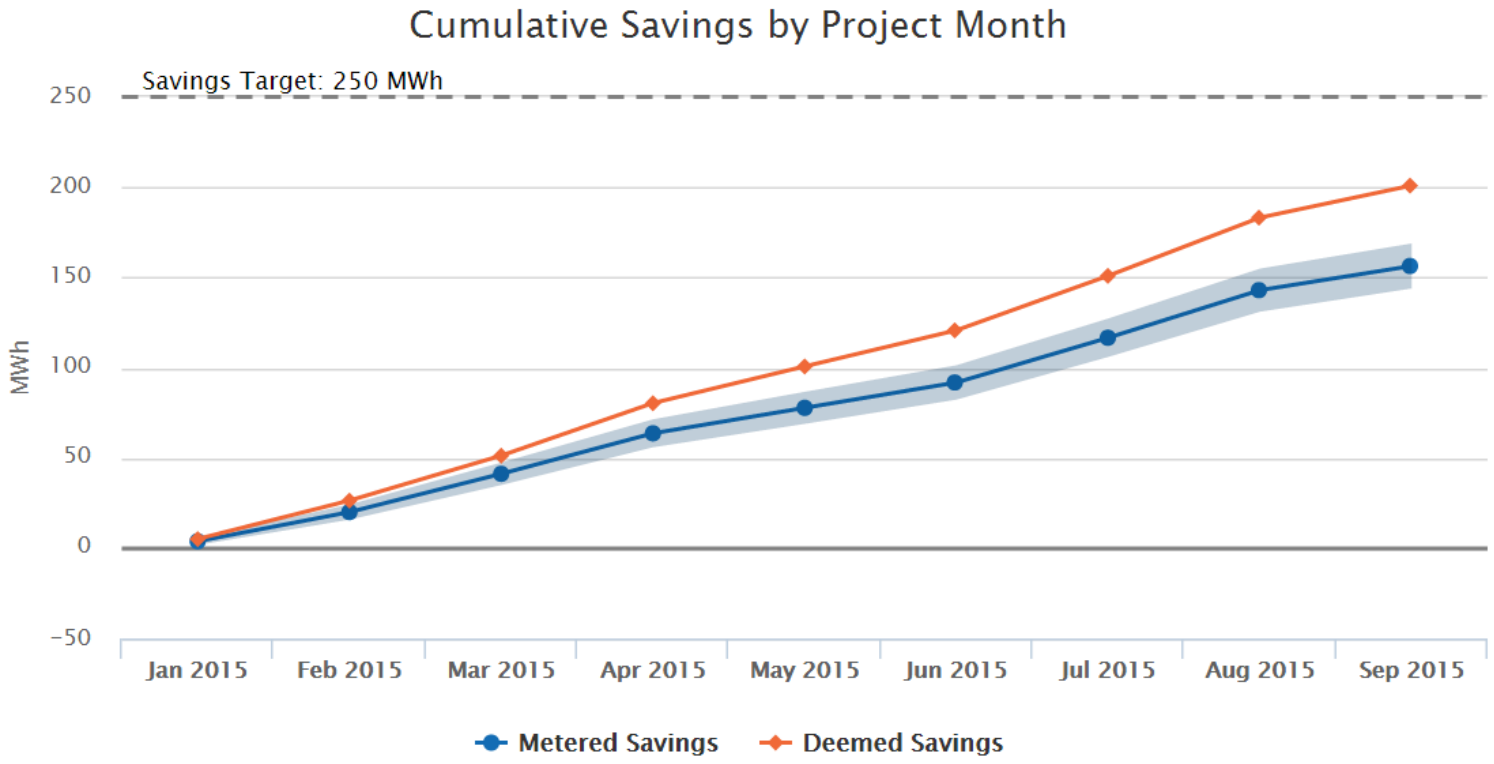
PROGRAM SAVINGS

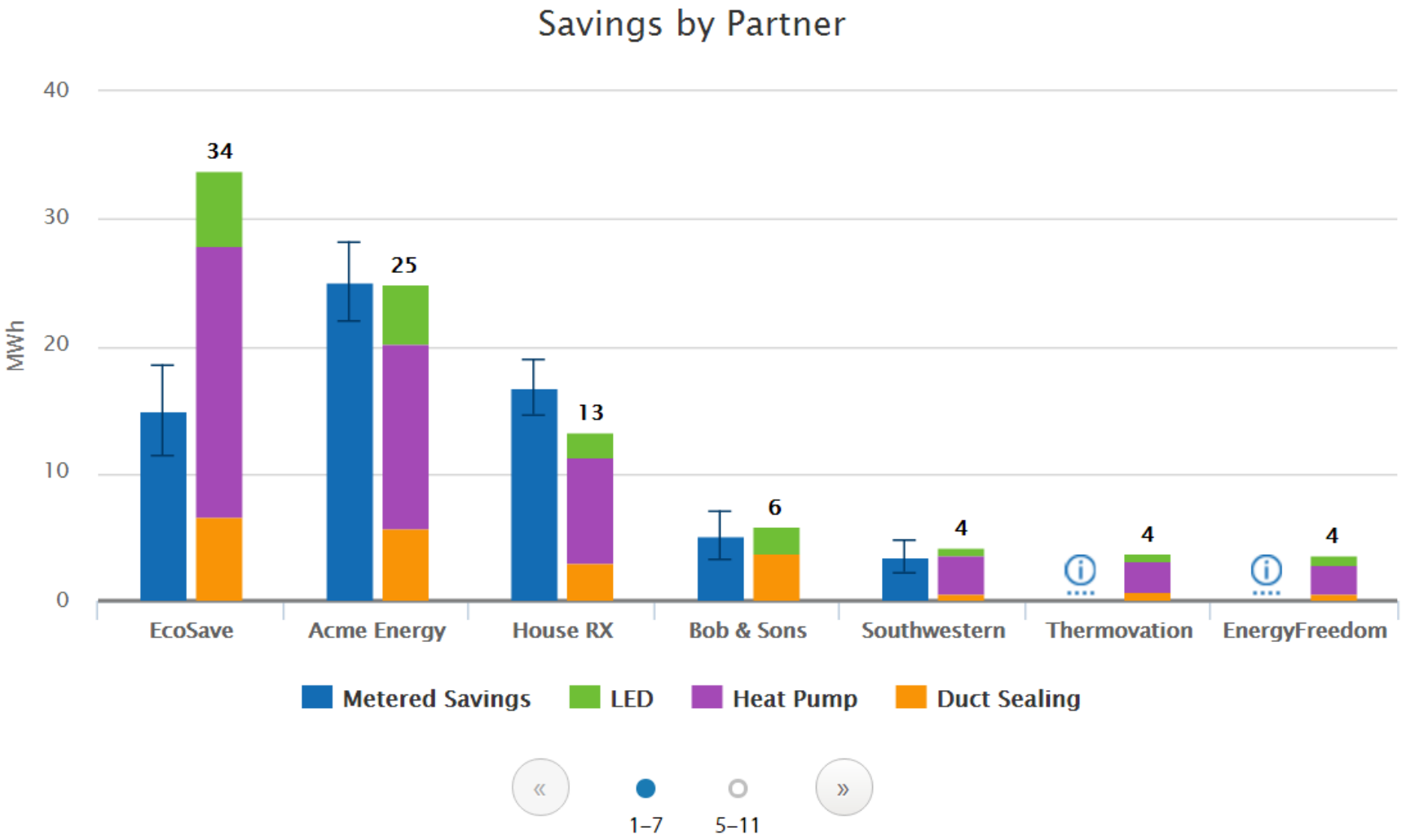
156 ± 12.5 / 250
METERED (MWh)
63% ± 5% OF GOAL

78% ± 6%
REALIZATION RATE

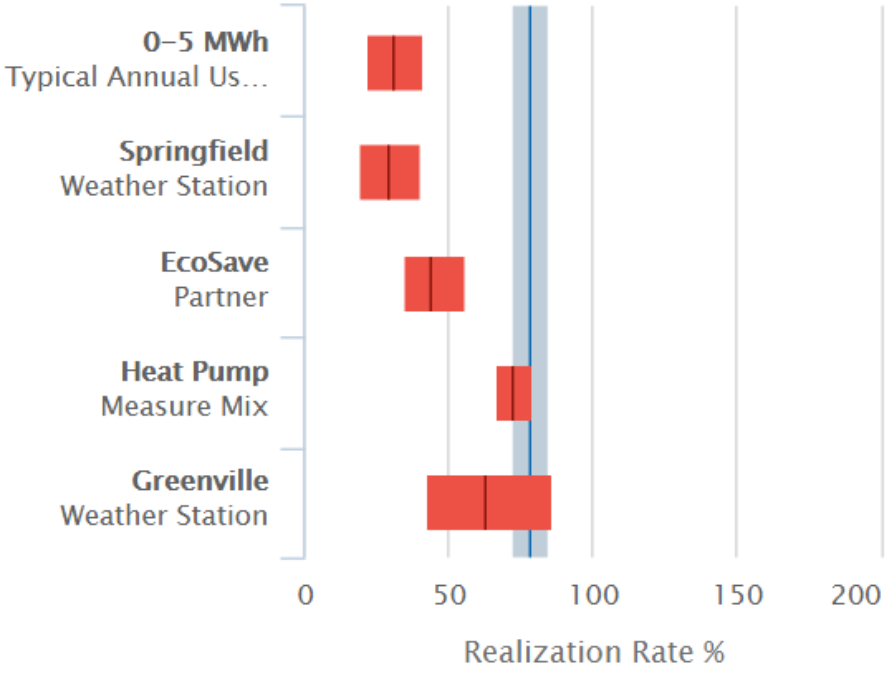
PROGRAM PROGRESS

1,485 / 2,100
PREMISES TREATED
71% OF GOAL




Savings by Partner
40
30
20
10
0
MWh
34
25
13
6
4
4
4
EcoSave
Acme Energy
House RX
Bob & Sons
Southwestern
Thermovation
EnergyFreedom
Metered Savings
LED
Heat Pump
Duct Sealing
«
»
1–7
5–11

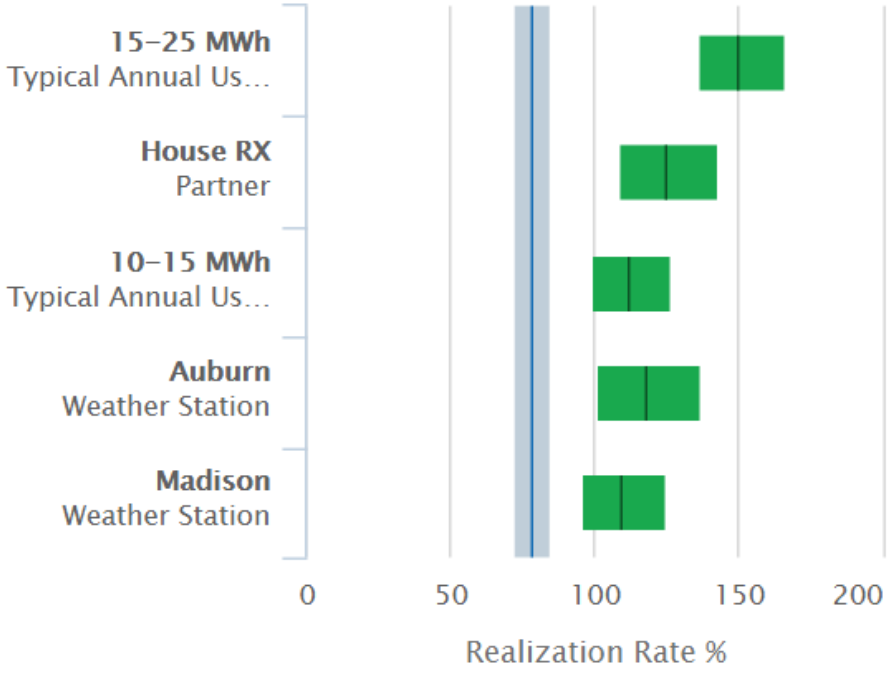
Top Negative Influencers



Realization Rate 90% Confidence Interval

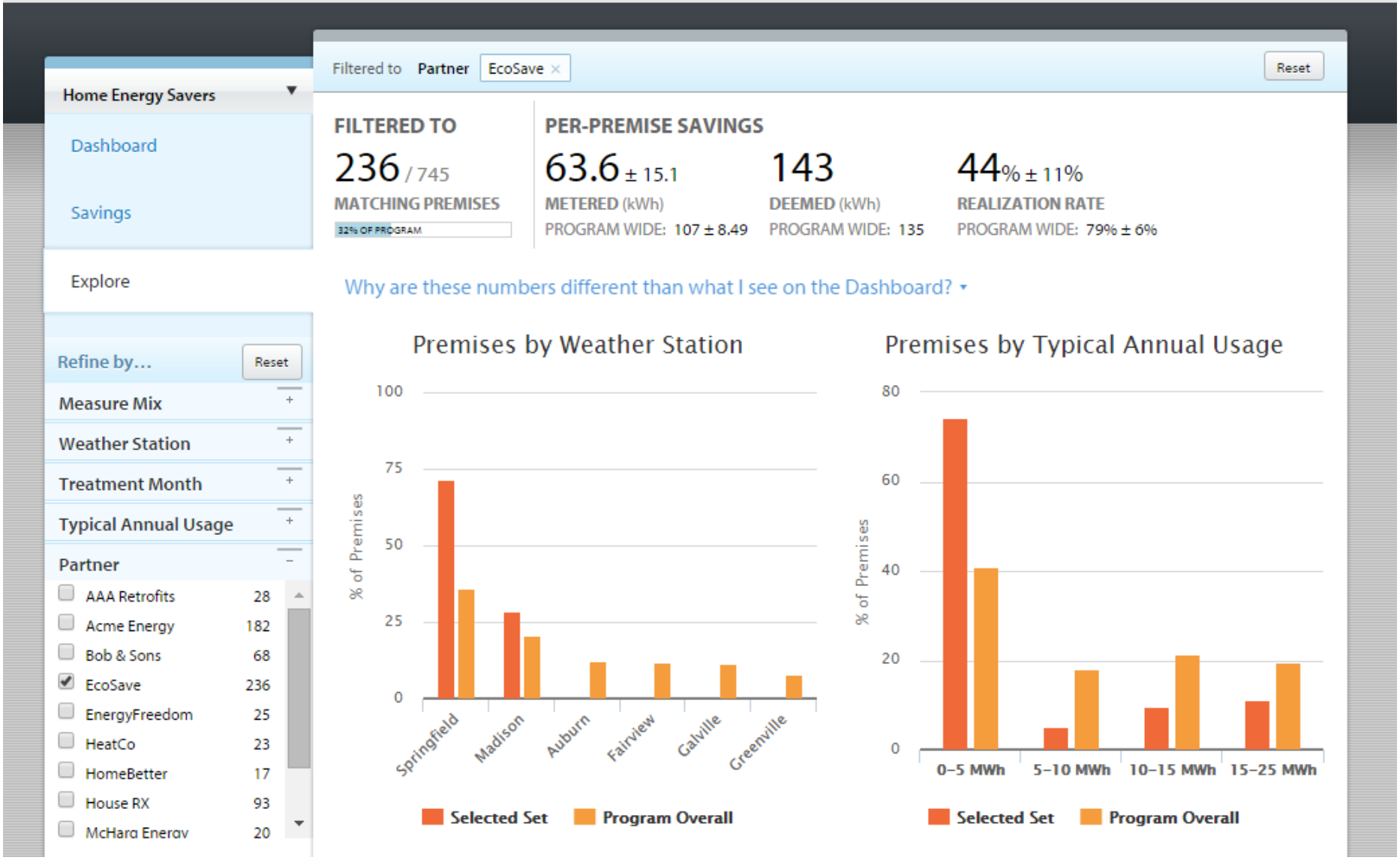
Program-Wide Realization Rate

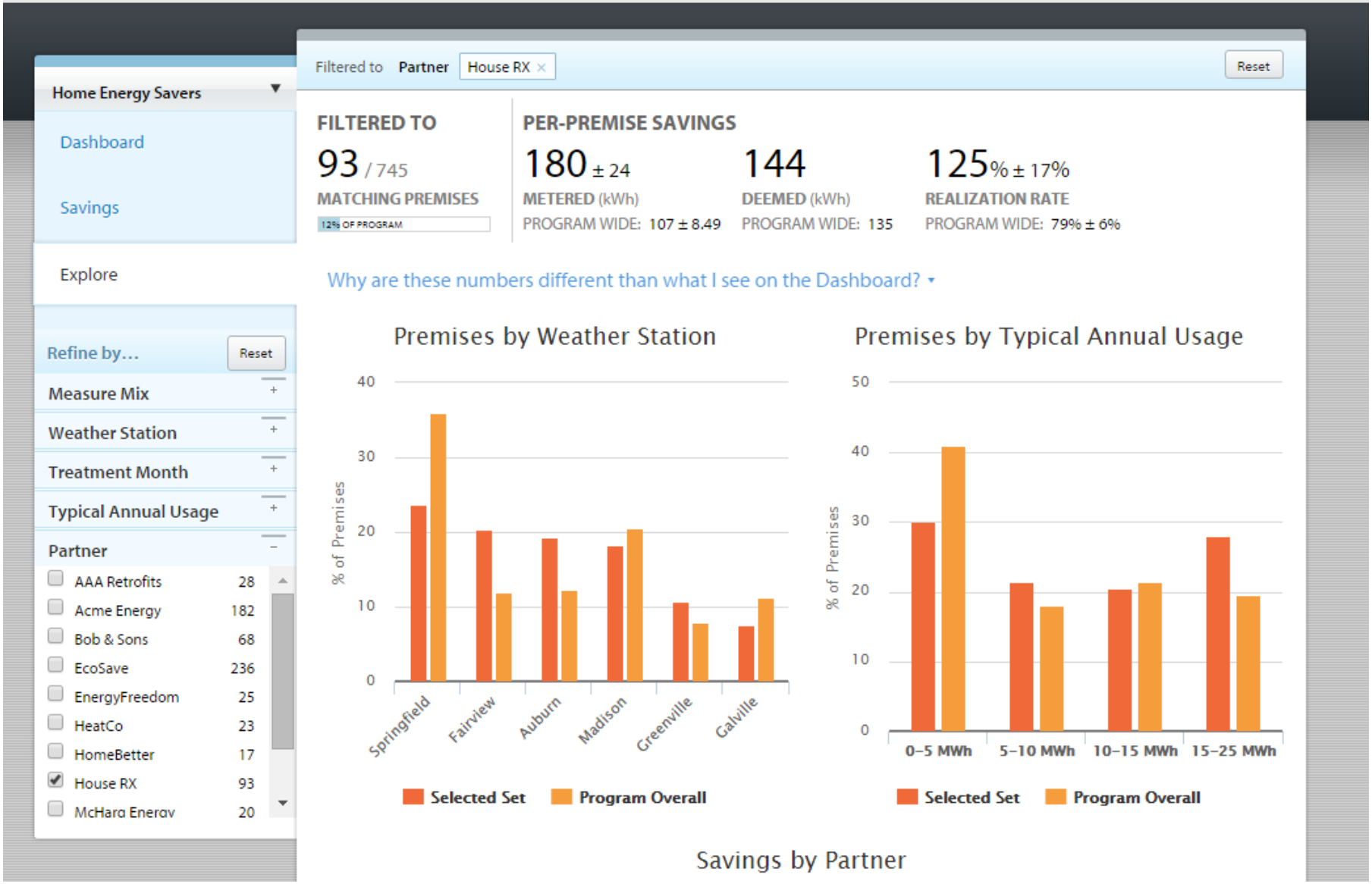
Top Positive Influencers



Realization Rate 90% Confidence Interval

Program-Wide Realization Rate

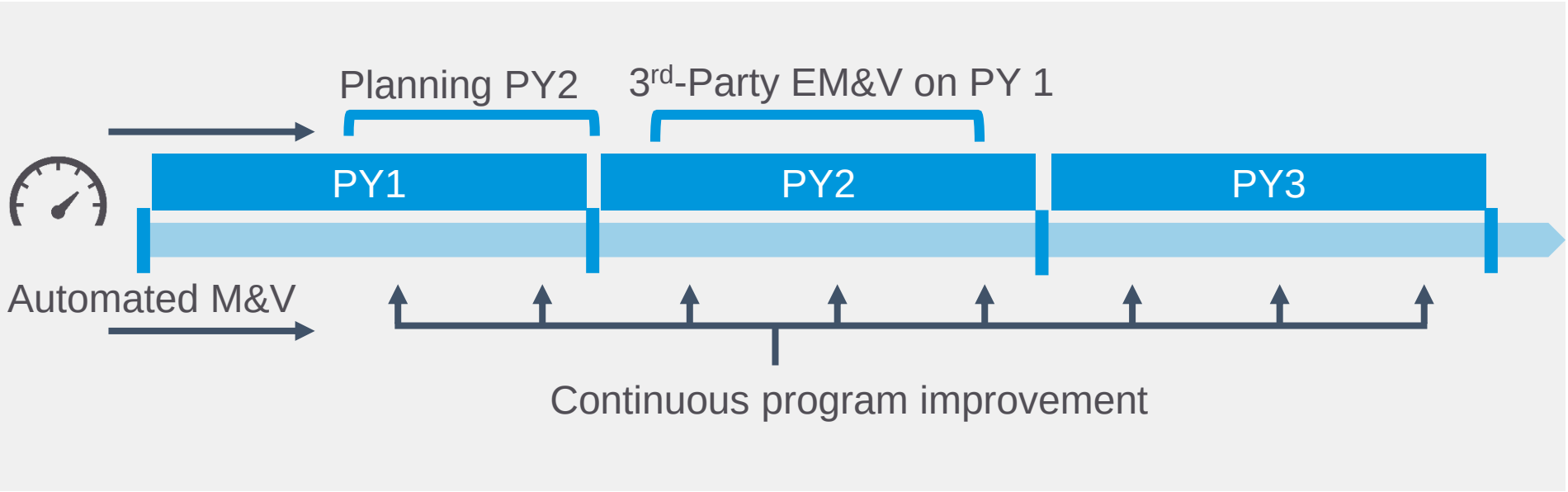

Home Energy Savers
Dashboard
Savings
Explore
Refine by...
Reset
Measure Mix
Weather Station
Treatment Month
Typical Annual Usage
Partner
AAA Retrofits 28
Acme Energy 182
Bob & Sons 68
EcoSave 236
EnergyFreedom 25
HeatCo 23
HomeBetter 17
House RX 93
McHara Energy 20
Filtered to Partner EcoSave
Reset
FILTERED TO
236 / 745
MATCHING PREMISES
32% OF PROGRAM
PER-PREMISE SAVINGS
63.6 ± 15.1
METERED (kWh)
PROGRAM WIDE: 107 ± 8.49
143
DEEMED (kWh)
PROGRAM WIDE: 135
44% ± 11%
REALIZATION RATE
PROGRAM WIDE: 79% ± 6%
Why are these numbers different than what I see on the Dashboard?
Premises by Weather Station
% of Premises
100
75
50
25
0
Springfield
Madison
Auburn
Fairview
Galville
Greenville
Selected Set
Program Overall
Premises by Typical Annual Usage
% of Premises
80
60
40
20
0
0–5 MWh
5–10 MWh
10–15 MWh
15–25 MWh
Selected Set
Program Overall


Home Energy Savers
Dashboard
Savings
Explore
Refine by...
Reset
Measure Mix
Weather Station
Treatment Month
Typical Annual Usage
Partner
AAA Retrofits 28
Acme Energy 182
Bob & Sons 68
EcoSave 236
EnergyFreedom 25
HeatCo 23
HomeBetter 17
House RX 93
McHara Energy 20
Filtered to Partner House RX
Reset
FILTERED TO
93 / 745
MATCHING PREMISES
12% OF PROGRAM
PER-PREMISE SAVINGS
180 ± 24
METERED (kWh)
PROGRAM WIDE: 107 ± 8.49
144
DEEMED (kWh)
PROGRAM WIDE: 135
125% ± 17%
REALIZATION RATE
PROGRAM WIDE: 79% ± 6%
Why are these numbers different than what I see on the Dashboard?
Premises by Weather Station
% of Premises
40
30
20
10
0
Springfield
Fairview
Auburn
Madison
Greenville
Galville
Selected Set
Program Overall
Premises by Typical Annual Usage
% of Premises
50
40
30
20
10
0
0–5 MWh
5–10 MWh
10–15 MWh
15–25 MWh
Selected Set
Program Overall
Savings by Partner

Enhancing EM&V

Integrated Evaluation Support

Supports targeted research and sampling



Enhancing EM&V

An example of work in progress

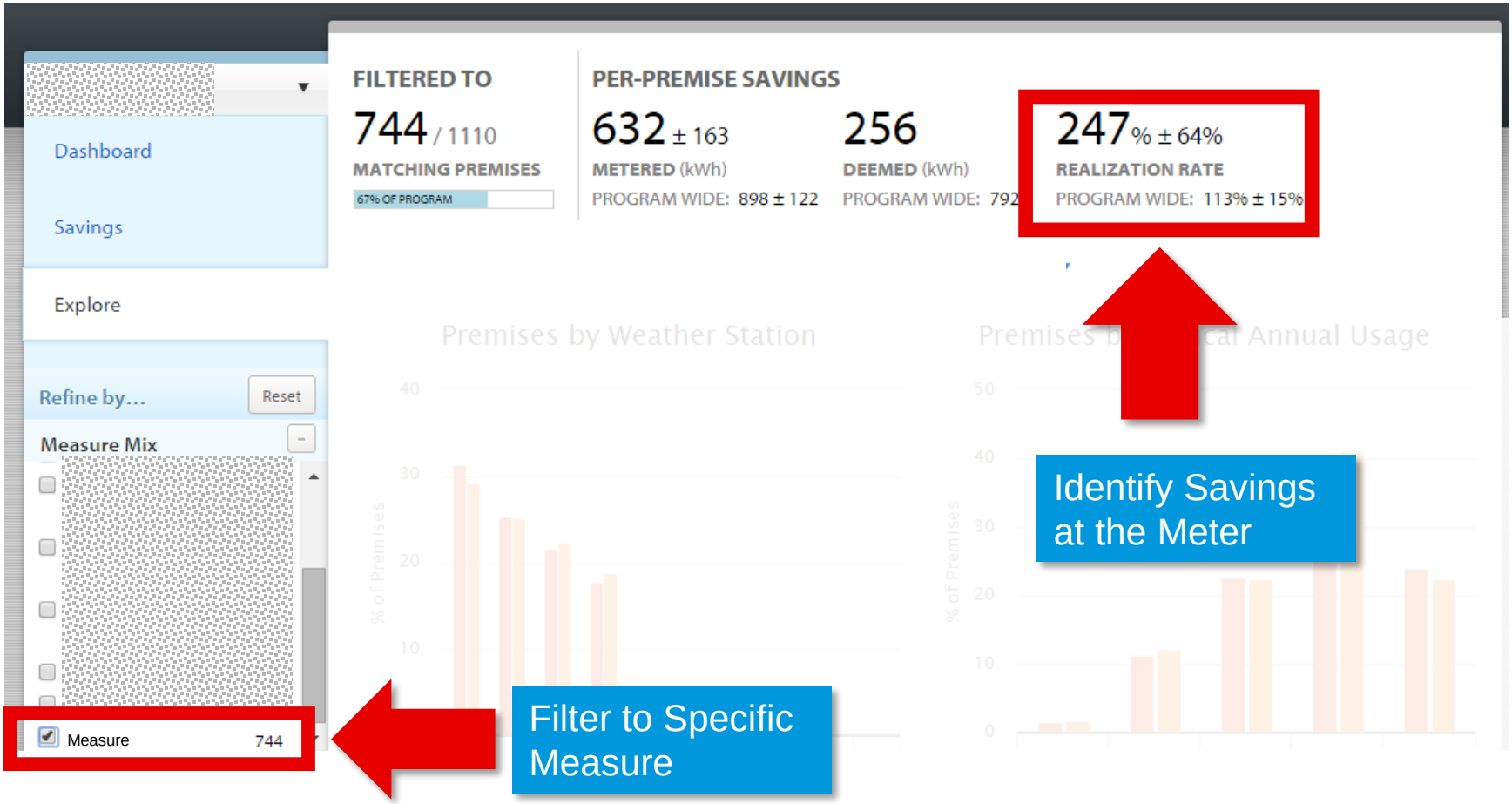
EnergySavvy and an EM&V firm are jointly evaluating a residential HVAC program

→Goal is to better understand how software can enhance and support the evaluation and add value to the PA, to the EM&V team and to the regulators

- Models
  - EM&V firm is validating Quantify's model specifications, inputs, outputs, etc.
  - Tuning Quantify to be the billing analysis of record
  - Applying code baseline adjustments where applicable
- Reporting
  - Quantify serves as a "visual executive summary" and an internal reporting tool
- Evaluation
  - EM&V firm will perform surveys and additional research as needed (e.g., attribution)
- Feedback to programs
  - The team will leverage the software to share findings and explore value of early feedback to program managers

TRM Updates and Calibration

Collecting evidence to refine savings estimates



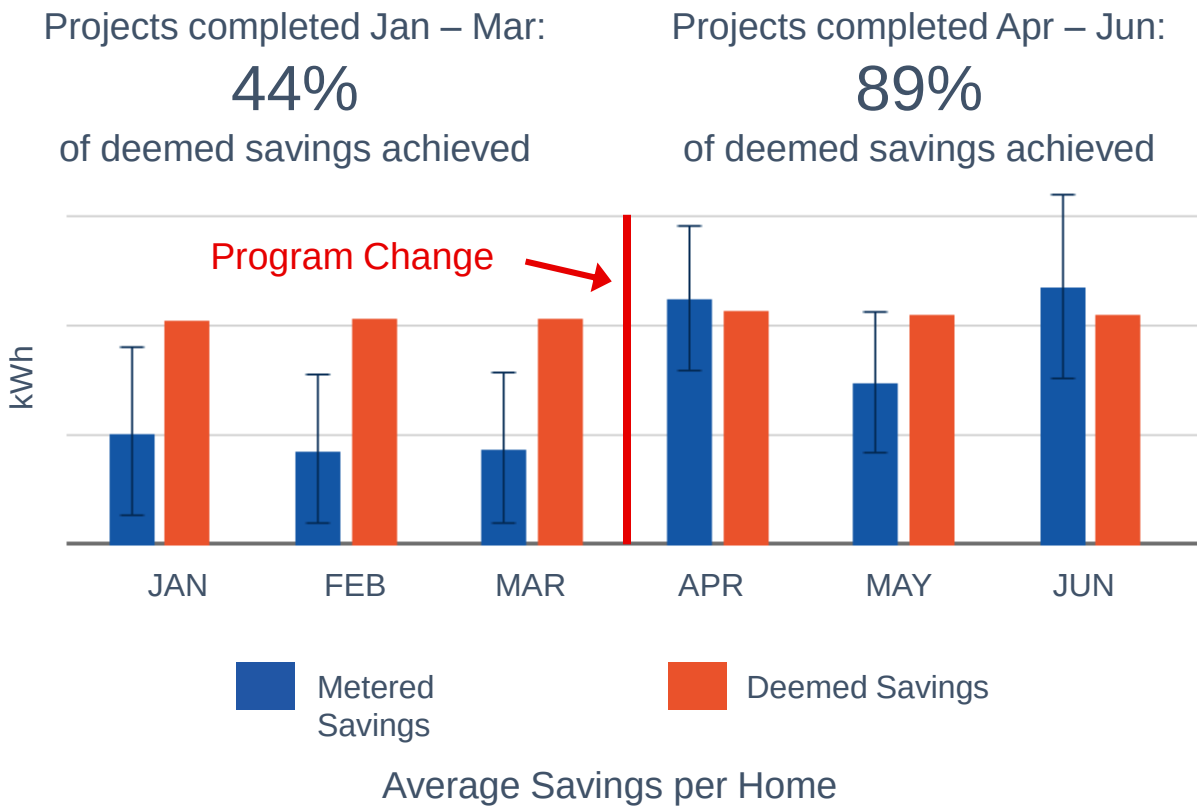
POLL #3: Question on role of EM&V in supporting evaluation

Faster Feedback Loop

Program Optimization

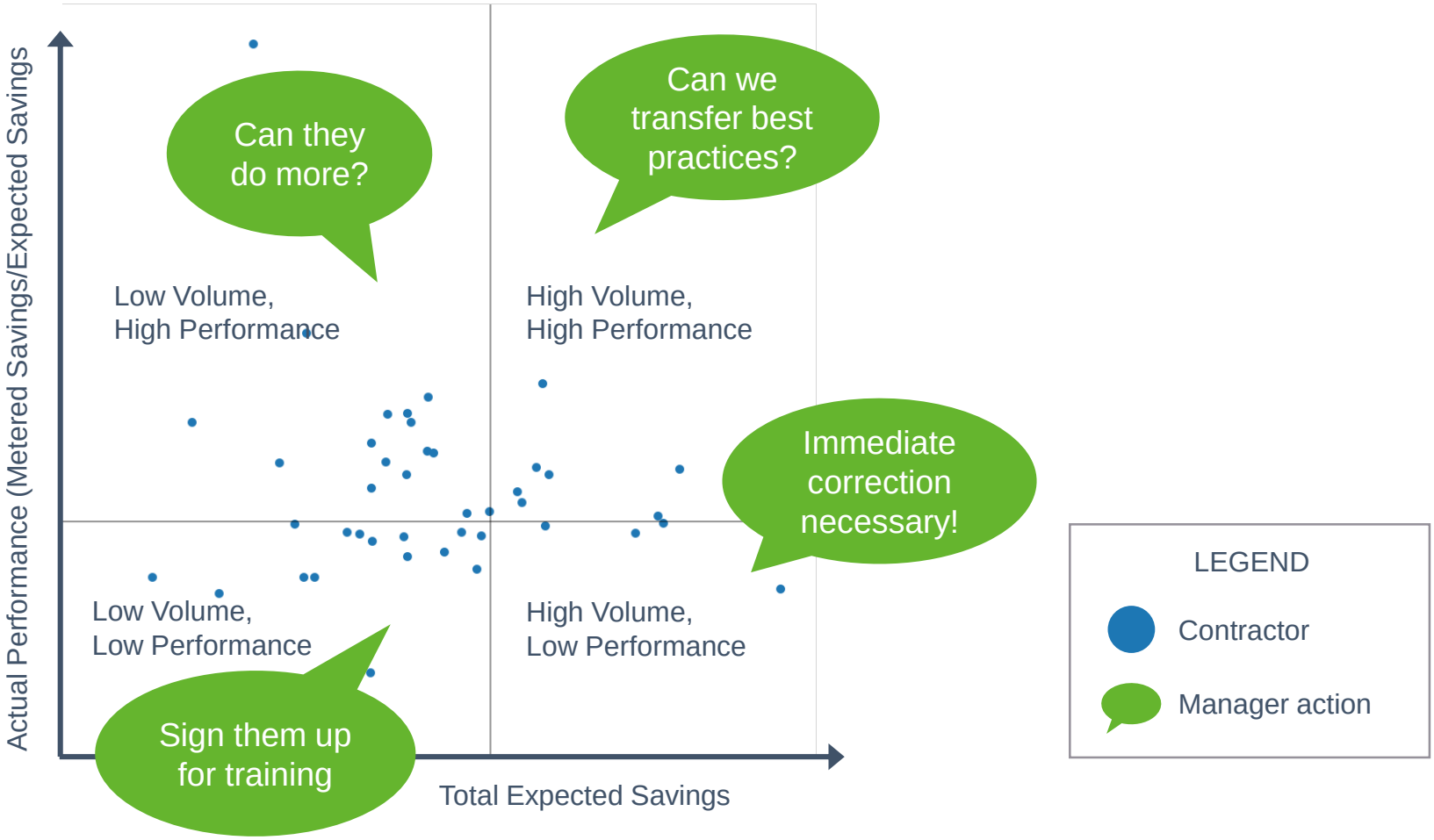
What critical program insights can I discover in time to take action?

- Monitor 100% of premises (treated and untreated) in real time
- Timeliness of data makes it actionable
- Quick feedback loop means lots of chances to iterate



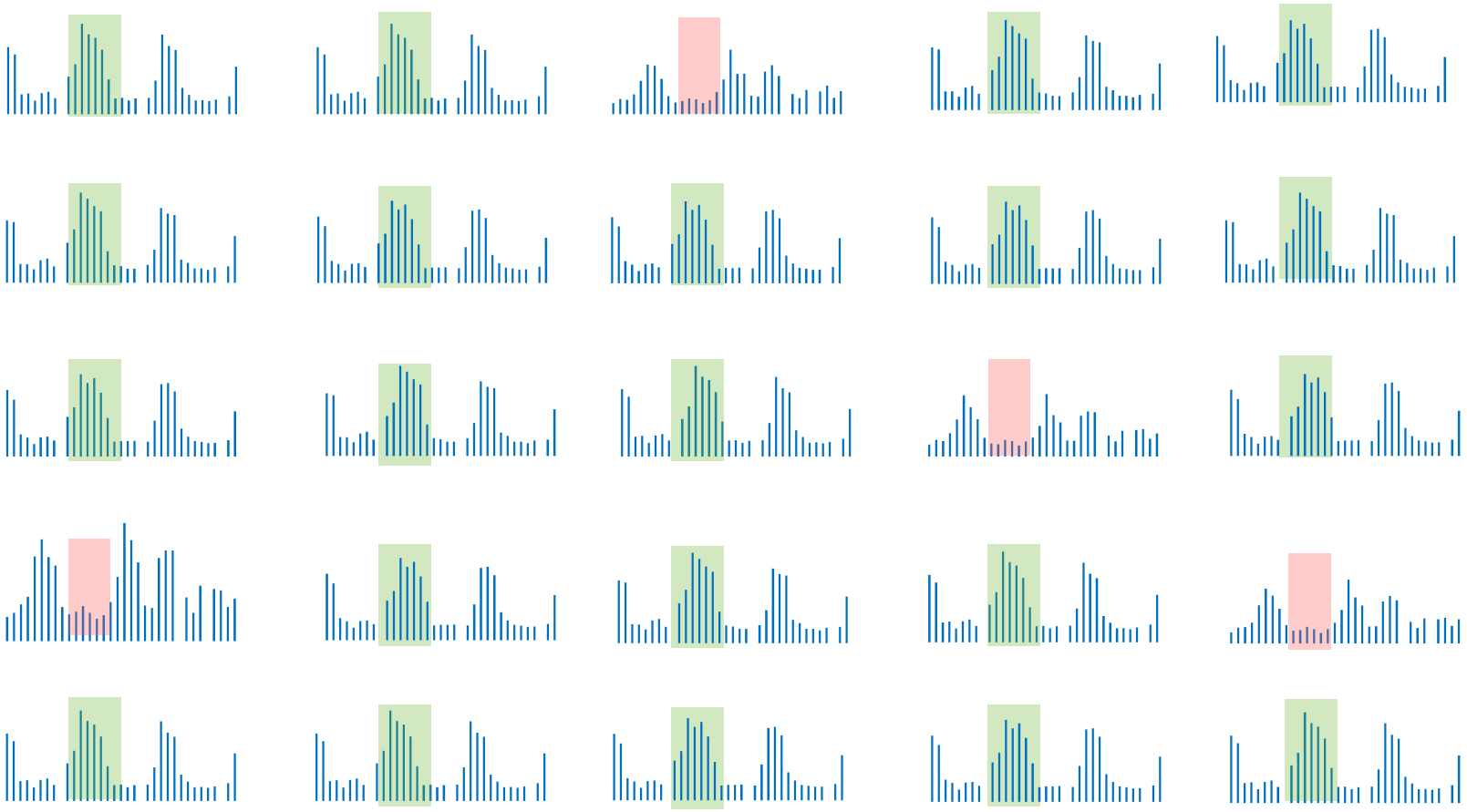
Contractor Performance Management

Comparing savings at the meter to expected savings reveals good, bad and ugly.



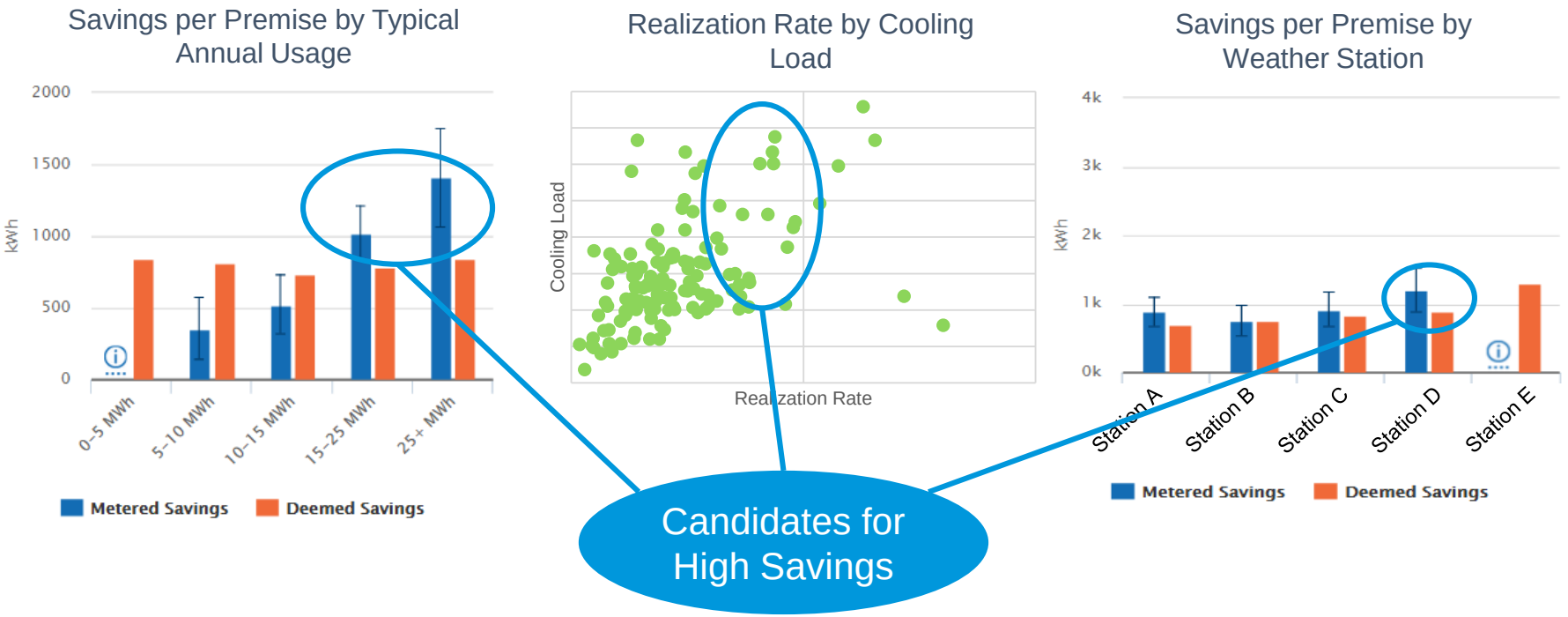
Intelligent QA/QC Targeting

Target in-home inspections and reduce QA/QC costs



Customer Segmentation

Case Study: Identify future program candidates based on performance of others



POLL #4 – Question on Feedback Loop

PSEG Long Island – Pilot & Vision

DAN ZAWESKI

MANAGER – PLANNING AND EVALUATION

PSEG LONG ISLAND
We make things work for you.

POLL #5 – Question on benefits of EM&V 2.0

Questions and FAQs

- Do you need AMI data for EnergySavvy's EM&V 2.0 tools?
  - No – this tool has been put to use with AMI, Monthly and Bi-Monthly data
- Will this approach only work for programs that expect savings above 10% per project?
  - No – with monthly data and enough projects, savings are detectable down to 2-3% per project. With AMI data it may be possible to go lower.
- Is this intended to replace formal evaluation?
  - No – this tool is intended to enhance and support evaluation while providing feedback information to PA's

Thank You

ENERGYSAVVY

Tim Guiterman
Director of Quantify Solutions
tim@energysavvy.com
802-557-4755

Jake Oster
Sr. Director of Regulatory Affairs
jake@energysavvy.com
802-598-1175

Industry and R&D Needs: Defogging Key Issues in M&V2.0

Jessica Granderson

Team:
Samuel Fernandes, Samir Touzani
Lawrence Berkeley National Laboratory

David Jump
KW Engineering

BERKELEY LAB
Lawrence Berkeley National Laboratory

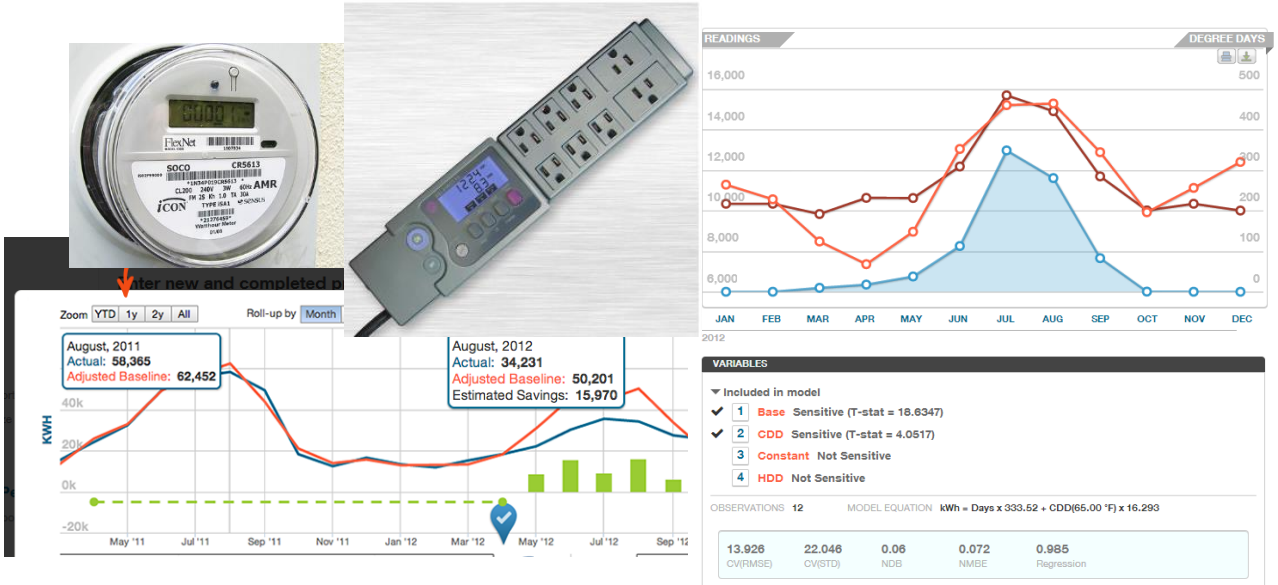
Outline

- Intro to M&V2.0 and webinar focus
- Who is doing what on the general topic
- What do we know about commercial M&V2.0, what is next, and where are we going?

What is M&V2.0?

- Generally understood as: use of more data (interval or volume), analytics, computation at scale
  - to streamline the M&V process through semi/automation

- Delivered in proprietary tools, 'open' algorithms



BERKELEY LAB
Lawrence Berkeley National Laboratory

What are the potential benefits of M&V2.0?

- Increase visibility, quickly obtain ongoing and interim results feedback
  - Increase savings and enhance customer experience?

- Automate parts of the process that computers do well, streamline data acquisition and processing
  - Reduce time and cost?
  - Increase throughput, number of projects going through the pipeline?

What is the vision for where we might end up?

- New M&V2.0 methods can be objectively tested as industry continues to innovate and new data source become available
- Multiple real-world pilots are used to assess M&V 2.0 value proposition
  - Cost, accuracy, time, tradeoffs vs. traditional M&V
  - Value of continuous feedback in increasing savings as well as customer value and experience – for both residential as well as commercial
- Processes/work flows are established to leverage automation while using engineering expertise where needed to maintain a quality result
- Analytical solutions to flag the non-routine adjustments are developed and tested for effectiveness
- Industry establishes acceptable levels of uncertainty and confidence, and documentation requirements for transparent evaluation

What is new about M&V2.0? What is not new?

- M&V2.0 tools are built upon savings estimation techniques that have been used for decades
  - Comparison group analyses,
  - IPMVP Options B&C, whole-building and submeter-based
  - IPMVP Option D, calibrated simulation modeling

- What's new is:
  - Degree of automation in data acquisition, and model creation
  - Granularity and volume of data can improve quality of result
  - Potential for continuous feedback
  - Integration of M&V capability with other analyses for operational efficiency
    - eg load visualization, portfolio tracking, end-use monitoring, etc.
  - Software as a service offerings for owners, managers, program administrators

Two examples

ENERGYSAVVY

- Customer engagement
- Program administrators
- Continuous savings feedback
- Net savings, comparison group billing analysis
- Residential

lucid

- Operational efficiency, SEM, MBCx
- Owners and operators
- Continuous savings feedback
- Gross savings, pre/post whole-building or submeter Option B or C
- Commercial

Screen shots of M&V2.0 capability

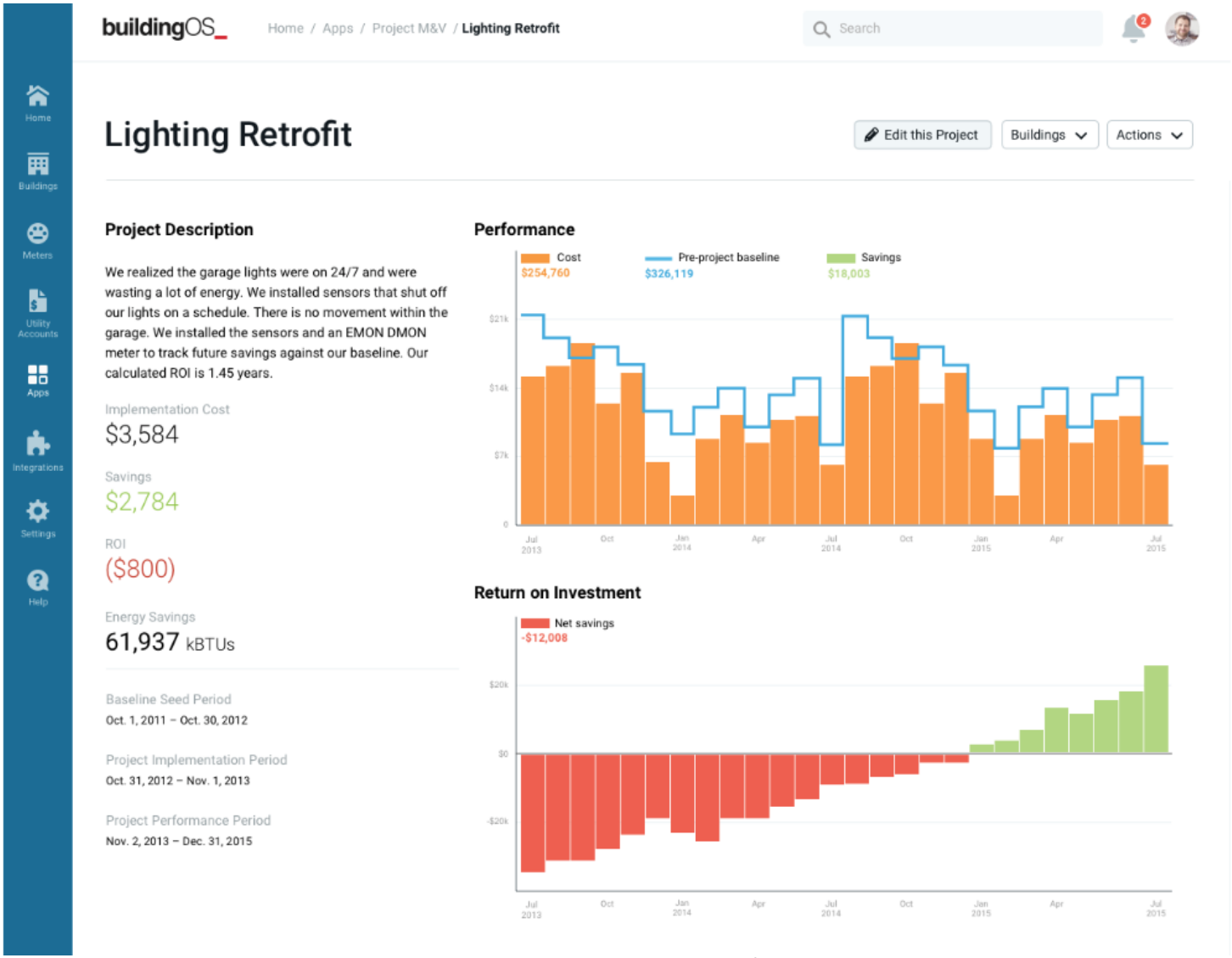


Image Source: Lucid

Screen shots of M&V 2.0 capability

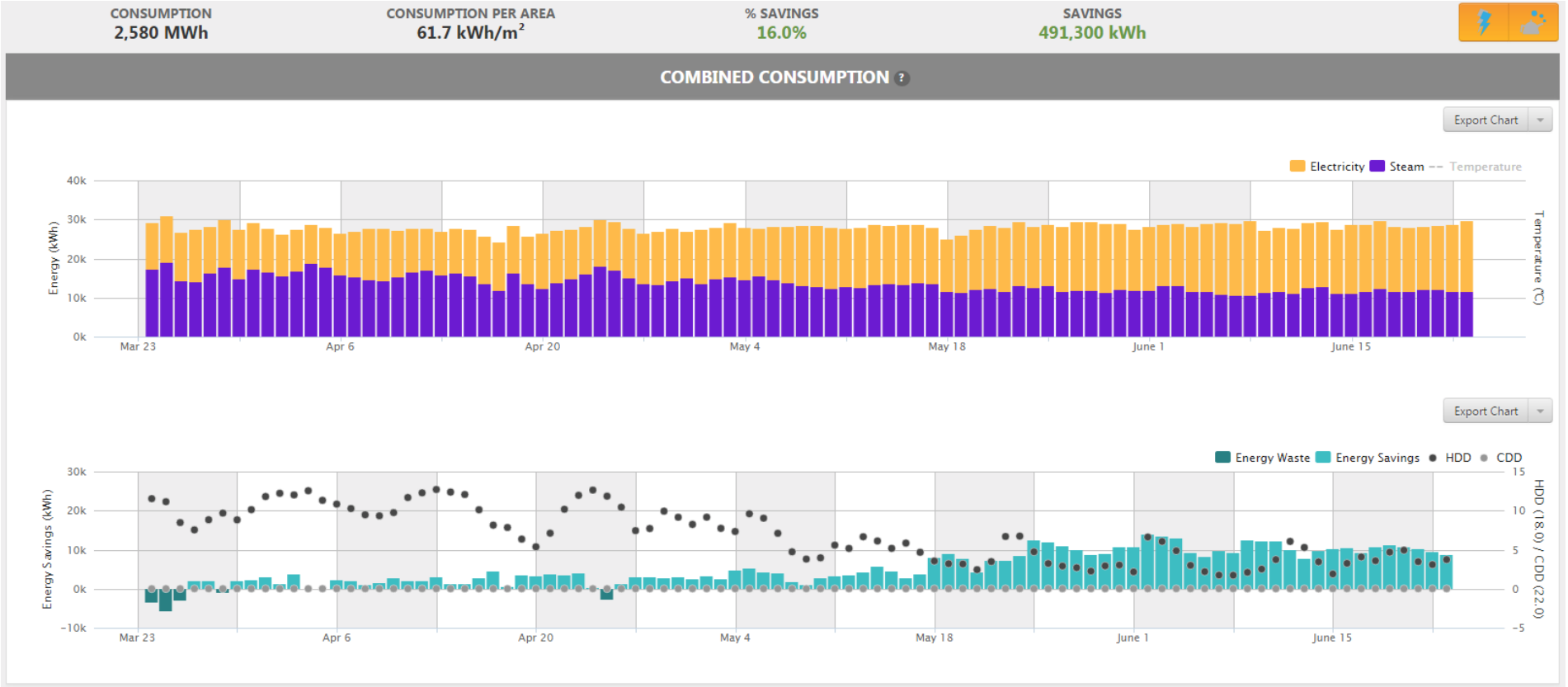


Image Source: EnerNOC

Screen shots of M&V 2.0 capability

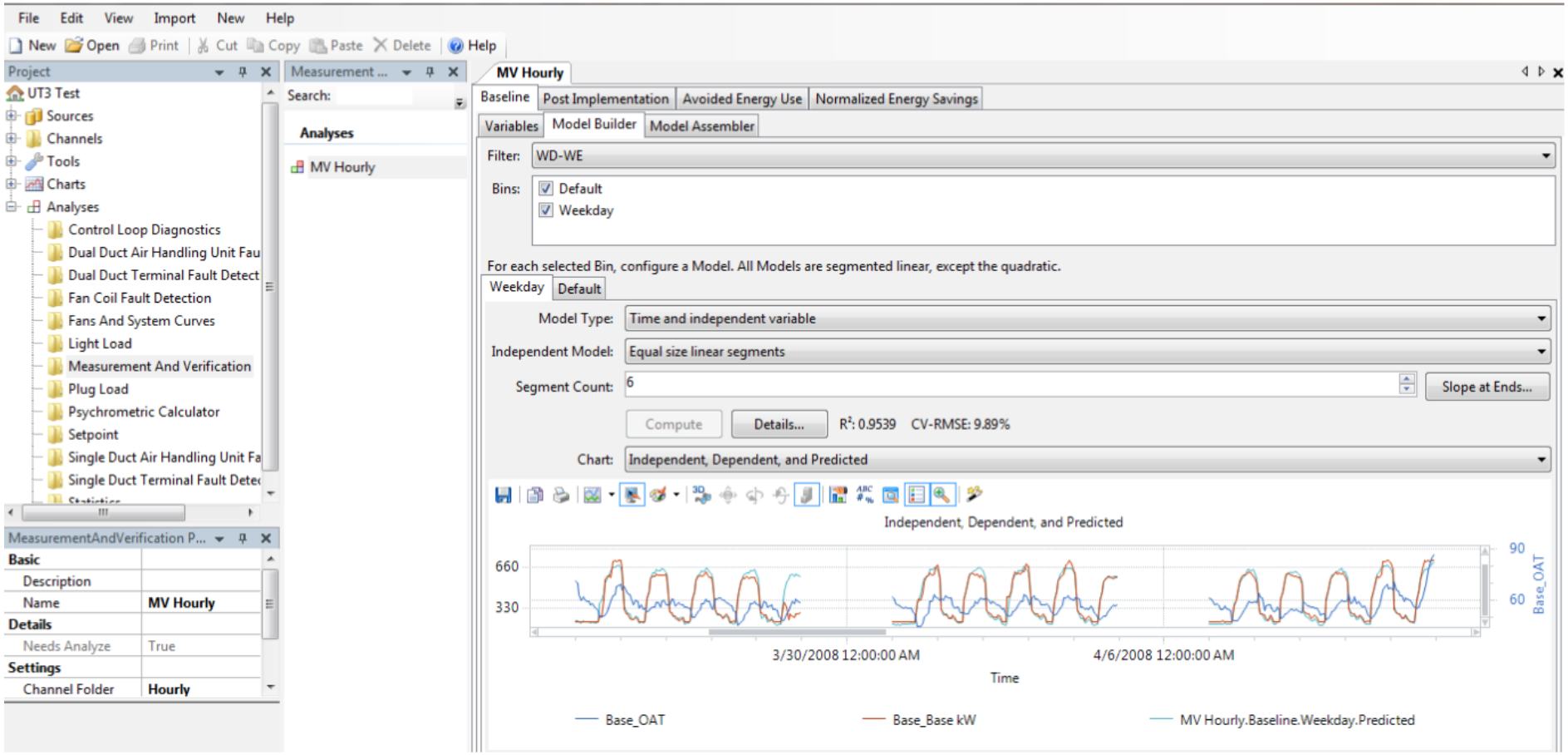


Image Source: Universal Translator 3

A diversity of savings estimation approaches is used today

Approach	Meter based?	Net or Gross?	Program/measure sweet spot
Deemed values	Not directly	Gross, Net in some cases	Efficient equipment replacement/installation
Engineering estimates, calculated	Not typically	Gross, Net	Custom industrial and large commercial; new construction
Billing Analysis that can include comparison groups, randomized control trial, or quasi-experimental	Yes, with other data	Usually Net	Programs w large numbers: residential, behavioral, small savings/site
Calibrated simulation modeling (IPMVP Option D)	No (except the calibration)	Gross, Net	Retrofit, large commercial
***Whole-building and retrofit isolation M&V (IPMVP Option C, B)**	**Yes, with other data**	**Gross**	**Commercial, multi-measure, interactive effects, operational measures**
…	…	…	…

***This is the focus of this presentation**

Where are meter-based approaches most appropriate?

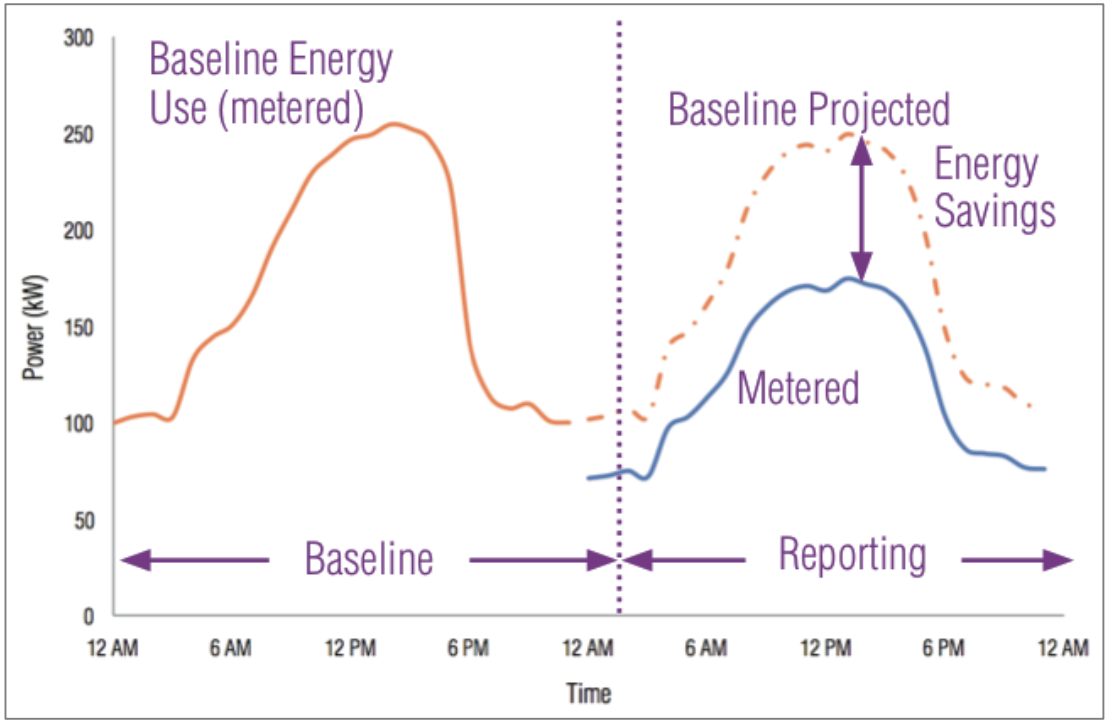
- 'Predictable' buildings
  - Weather sensitive, regularly scheduled
- Projects with multiple and interactive measures
  - Affecting several building systems (HVAC, lighting, etc.)
- Difficult to quantify measures

  Duct sealing, envelope upgrades, etc.
- Projects with larger savings, 'above the noise'
- **Measures using existing condition as baseline**
  - **Retrocommissioning, behavioral, operational**

Promising opportunities associated with meter-based M&V approaches

- Enabling delivery of whole-building programs that combine strategies for deep savings
- Enabling pay-for-performance programs
- Scalability and streamlining
  - Reduce labor time and costs
  - Maintain an accurate result
  - Quickly obtain ongoing and interim results
  - Increase throughput, number of projects

How are meter-based site savings quantified?



Metering at whole-building (Option C), or submetered measure isolation level (Option B)

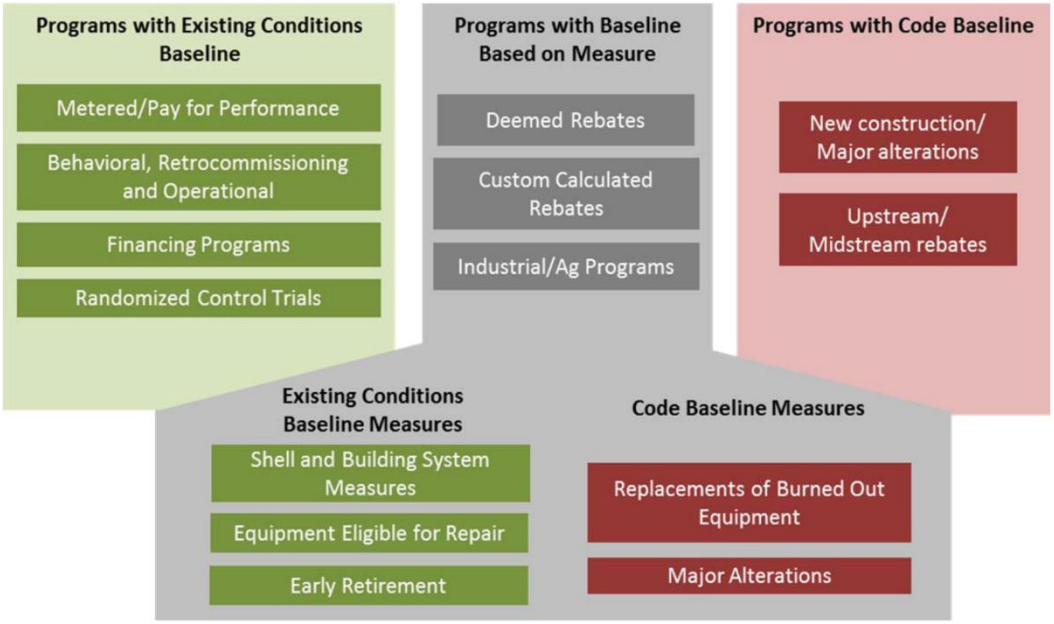
In M&V2.0 tools baselines are automatically created with meter and weather data feeds

User enters date of measure implementation, savings are calculated by the tool

Relevant California Activities

- CA AB802: CPUC to authorize programs using "normalized meter-based" energy savings (existing conditions baseline) for:
  - To- and beyond-code savings, and retrocommissioning, operational, behavioral programs
  - Counting savings towards goals when feasible and cost-effective
- CPUC providing guidance on where existing use baselines are/not appropriate, EM&V plans

Figure 3: Proposed Baseline Framework



Other Relevant Activities

- RMI e-Lab Accelerator cross-stakeholder group group, more detailed articulation of M&V2.0, potential benefits, outstanding industry needs
- CEE Guidebook resource to understand uncertainty principles for whole-building M&V approaches, in context of whole-building program deign and delivery
- EVO has started an M&V2.0 group to determine how IMPVP will address the topic
- ASHRAE technical committees discussing ‘standard methods of test’

POLL # 6-7 Questions

Are you involved in, or going to be involved in any programs that rely on an existing use baseline?

Are you interested in exploring 2.0 tools and methods in your work?

Motivating Industry Questions, R&D Approach, and Highlights

Industry questions motivate LBNL's R&D

- Are these proprietary tools reliable?
- How can I verify their accuracy and compare them?
- Are proprietary tools any better or worse than standard regressions?
- Even if a tool is generally robust, how do I know that it will work for my specific projects or program?
- How "big" do my savings have to be to use these approaches?
- How do I know that a robust tool was applied to generate a quality result?

Four-step R&D approach to answer these questions

1. **Population-level (many buildings) M&V2.0 testing to verify general, overall robustness, compare and contrast tools**

2. **'Off-line' demonstration of promising models with historic utility program data**

3. Identification of reporting requirements and quantitative acceptance criteria for savings claims (in progress)

4. Larger pilots, demonstrations on 'live' programs (future)

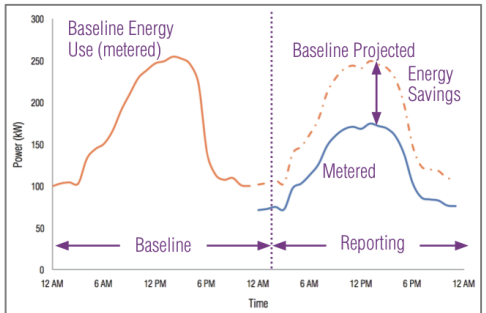
1. Population-level general testing and tool comparison

- Tested accuracy of baseline projections in proprietary tools and open standard models against data set from 500-600 untreated buildings

- Given 12mo whole building interval data, predicted 12mo of energy use
  - Within {-4, 5}% error for a full half of the buildings, CV(RMSE) well within industry guidelines, errors even smaller when aggregating buildings into portfolio
  - No clear 'winner' across 10 models

- No attempts to refine models based on expertise, knowledge of buildings, additional variables
  - Floor of predictive accuracy

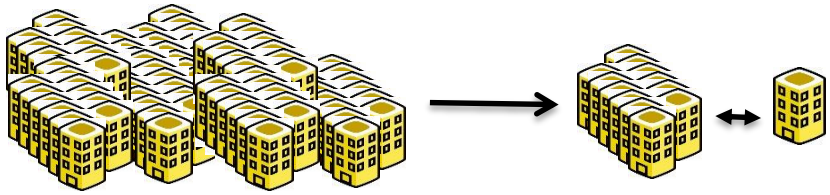
- Test procedure is published, was used by PG&E to prequalify tools for pilot, is available for use by others



2. Demonstrating 2.0 tools with historic program data

- Given tools that generally predict energy well, use them to automatically quantify savings

- Develop practitioner workflows to leverage automation while retaining accuracy of the savings result
  - Many, but not all buildings are 'predictable'
  - Gross savings at the meter may not be gross savings due to the measure, i.e., non-routine adjustments may be needed
  - Use uncertainty analysis to quantify *accuracy of the savings results when applied to specific projects/buildings/programs*



We draw from ASHRAE Guideline 14

- Provides **recommendations for accuracy in M&V**
- Covers **'goodness of fit' between the model and the baseline period data**, with suggested thresholds for bias (NMBE) and CV(RMSE)
- Covers suggested formulae to quantify **uncertainty due to error in the baseline model**
- Suggests **that fractional uncertainty be no more than 50% with at least 68% confidence** (what will EE programs require?)

Model demonstration with historic program data

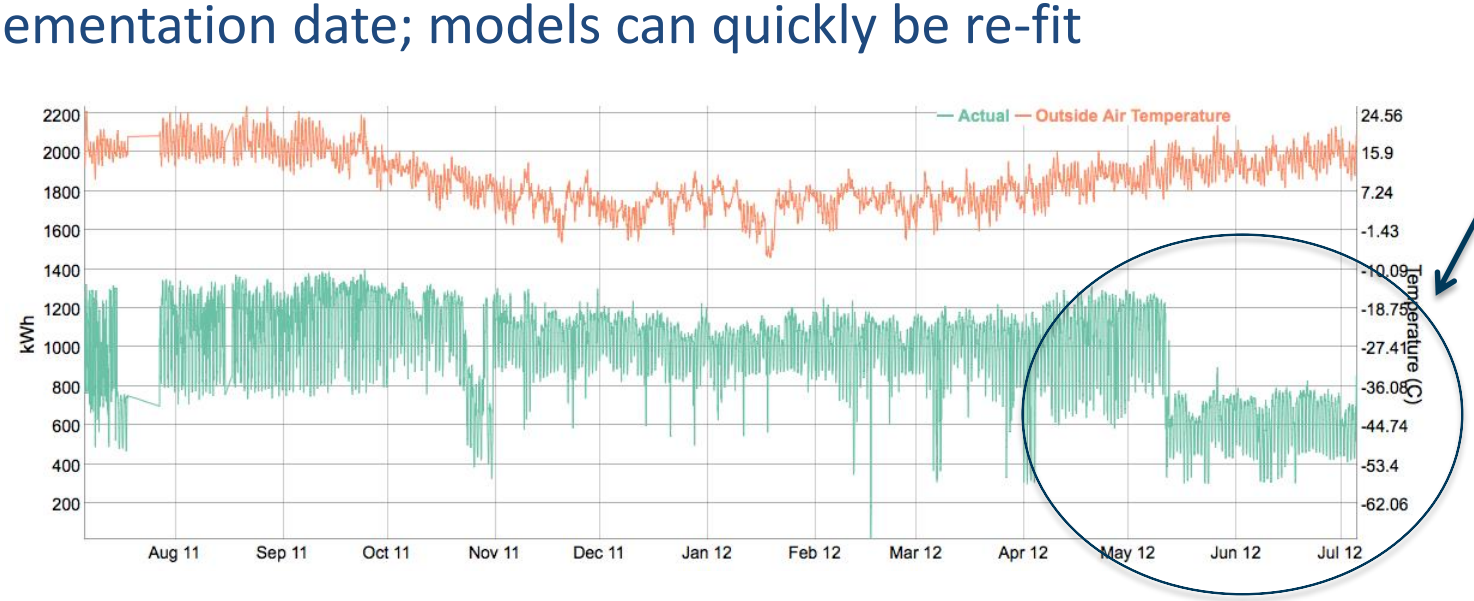
- Data from 51 buildings that underwent RCx and in some cases retrofits

- Preliminary workflow, drawing from ASHRAE Guideline 14
  - Auto fit the model to data from baseline period, and compute goodness of fit metrics
  - Set aside buildings that do not meet suggested fitness thresholds - these will require further investigation
  - For 'good' buildings auto compute savings and uncertainty using M&V 2.0 tool
  - Aggregate savings and uncertainties for each building to determine portfolio-level results

Findings from applying this workflow to historic program data

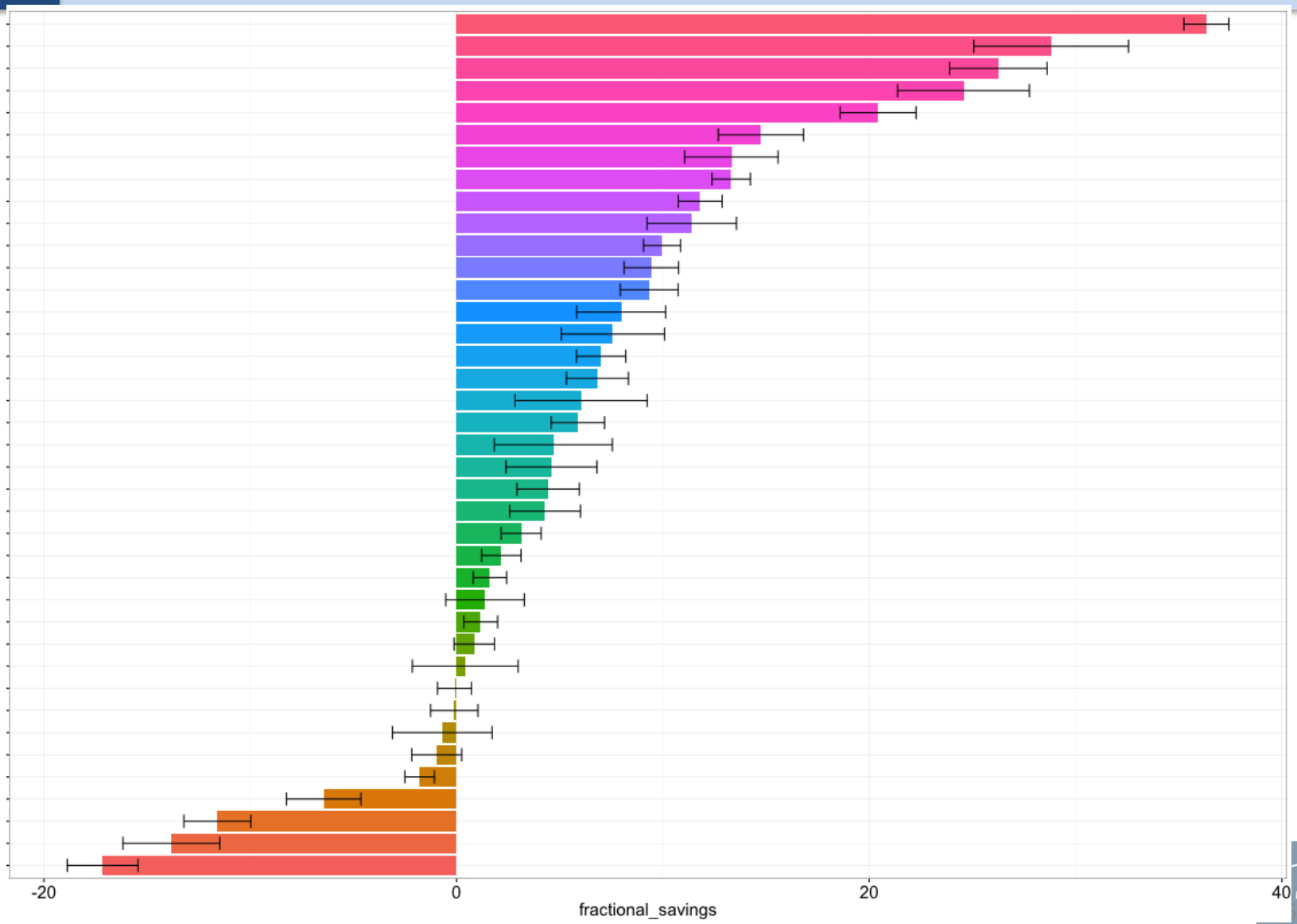
- Of the 51 buildings, 39 'passed' the goodness of fit tests using ASHRAE guidance

- Of the 12 that did not 'pass', 5 had incorrect documentation of measure implementation date; models can quickly be re-fit



- **For this data set, 44 of 51 buildings look to be well-suited to automated analysis; 7 may require more manual investigation**

Findings from uncertainty analysis with historic program data



Savings uncertainty ranges for each of 39 buildings, at 95% confidence level

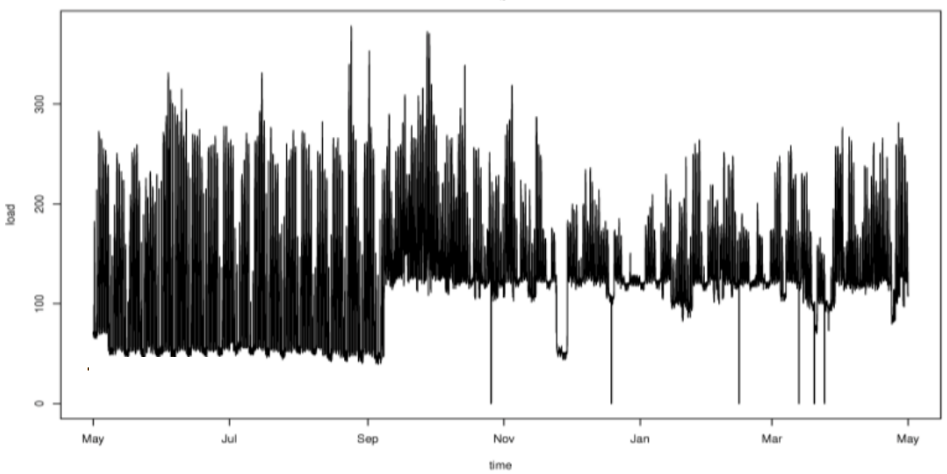
Summary of uncertainty findings in the demonstration on historic program data

- **32 of 39 individual buildings satisfied** or exceeded **ASHRAE** uncertainty requirements

- **At portfolio-level** for the aggregate of the 39 buildings, at 95% confidence level
  - Savings = 3.96% =/-.3, that is within confidence interval of [3.66%; 4.26%]
  - Aggregate ***far*** **exceeds ASHRAE guidance for sufficiency**

Some comments on non-routine adjustments

- Gross metered savings may not reflect gross program/measure savings
  - E.g. Occupancy or schedules may change or loads may be added/removed



- By definition, these Option-C compliant M&V2.0 baseline models do *not* handle NR Adj.

- It is possible that 2.0 analytics can flag cases where savings drop or increase unexpectedly, so that implementers can make timely inquiries of the site

Some comments on uncertainty, confidence, and documentation requirements

- General tool testing can tell us that we have good well-made hammers

- If we have well-made hammers, uncertainty and confidence can verify that we've driven our nails straight and true
  - But how straight do we need to be?
  - An how do we prove it to 3rd parties?
  - What documentation will we need?

80-20? 90-10? And how do we set these values?

Some comments on net, gross, other baselines and methods

- Even with deemed savings you commonly need to layer additional analysis to determine a net
- Existing conditions baselines are critical to less common programs that
  - Promise deep savings, offer opportunity beyond equipment-based measures
  - Focus on operational, retro-commissioning, behavioral, multi-measure, whole-building
- Calibrated simulation can be complex, costly and difficult to scale
- Comparison groups may not always be possible to establish for commercial buildings

POLL #8: Is there value in both population-level testing for tool comparison, *and* assessment of tool accuracy for specific buildings/projects or programs

Where Have We Gotten and Where are We Going?

Where have we gotten?

- Appreciation of the potential benefits of M&V2.0
- Replicable test procedures to assess overall robustness of M&V 2.0 tools for commercial buildings
  - *Many* models predict within a few percent for *many* buildings using commonly available data
  - Use by large utility to pre-vet vendors for pilot, published for ongoing use
- Initial exploration (ongoing beyond the 51 projects shown here)
  - High confidence and low uncertainty when applying M&V2.0 tools
  - Start on defining practitioner workflows to retain a quality result
  - Indication that with interval data savings may not have to be as big as 10% to 'see' at the whole-building level

Returning to the the vision for where we might end up

- **New M&V2.0 methods can be objectively tested** as industry continues to innovate and new data source become available
- Multiple **real-world pilots are used to assess M&V 2.0 value** proposition
  - Cost, accuracy, time, tradeoffs vs. traditional M&V
  - Value of continuous feedback in increasing savings as well as customer value and experience – for both residential as well as commercial
- **Processes/work flows are established to** leverage automation while using engineering expertise where needed to **maintain a quality result**
- Analytical **solutions to flag the non-routine adjustments** are developed and tested for effectiveness
- Industry establishes **acceptable levels of uncertainty and confidence, and documentation requirements** for transparent evaluation

What do we still need to know or do?

- Demonstrate 2.0 tools on more historic program data (ongoing)
- What do we do for buildings that don't have a good fit, and aren't well suited to the meter-based approach?
  - How can we leverage targeting and pre-screening
- How does M&V2.0 compare to traditional approaches, 'in the field'?
  - **Can we conduct a sufficient number of pilots and what 'proof' points should they be designed to produce?**
- How do we handle non-routine adjustments?
- What uncertainty, confidence, and documentation requirements are needed for evaluation?
- What group might serve as a testing body for new M&V2.0 tools to verify their general robustness?

POLL #9-10 Questions on Piloting M&V 2.0

POLL #11

Efforts are underway to test tools, including uncertainty analysis, for the acceptance of M&V 2.0 tools and methods….

Are you interested in working on topics of tool testing and acceptance criteria?

Questions and Discussion

Thank You!

For more information please contact Jessica Granderson
JGranderson@lbl.gov, 510.486.6792

For more detailed reports and presentations: eis.lbl.gov