

Recent Developments in Energy Efficiency Evaluation, Measurement, and Verification

Seth Nowak, Maggie Molina, and Martin Kushler

October 2017

Report U1712

© American Council for an Energy-Efficient Economy
529 14th Street NW, Suite 600, Washington, DC 20045
Phone: (202) 507-4000 • Twitter: @ACEEEDC
Facebook.com/myACEEE • aceee.org

Contents

About the Authors.....ii

Acknowledgments.....iii

Executive Summaryiv

Introduction..... 1

Methodology 2

State and Regional Leadership 3

 State Leadership Examples 4

Deemed Savings and Technical Reference Manuals 7

 Deemed Savings Estimates: Definitions and Best Practices..... 9

 TRM Trends and Directions 9

 State and Utility Examples of TRM Evolution..... 11

Common Practice Baselines 15

 Background 16

 CPB for Estimating Net Savings: Context and Discussion..... 17

 Constructing CPBs 19

 CPB Summary and Trends 21

M&V 2.0 23

 Recent Trends 24

 State and Utility Examples of M&V 2.0 25

Challenges Going Forward 28

Conclusions 30

References 32

Appendix A. Interview Responses..... 37

About the Authors

Seth Nowak conducts analysis and writes about energy efficiency programs and policies in the electric and natural gas utility sector. Focus areas of his research include exemplary programs, best practices, and program evaluation, measurement, and verification. Seth earned a master of business administration from the Wisconsin School of Business and a master of public affairs from the La Follette School of Public Affairs at the University of Wisconsin.

Maggie Molina is senior director of policy at ACEEE. She oversees federal, state, and local policy work as well as utility regulation. She conducts energy efficiency program and policy research and provides technical assistance to a variety of groups. She has authored numerous reports on state policy and utility-sector energy efficiency topics, including the first editions of the *State Energy Efficiency Scorecard*, energy efficiency potential studies, utility business models, the cost of saved energy, and efficiency programs. Maggie completed a Master of Engineering and Public Policy (MEPP) at the University of Maryland and a bachelor of arts at Colgate University.

Martin Kushler directs numerous widely acclaimed national studies of utility-sector energy efficiency policies and programs, and provides technical assistance to help advance energy efficiency policies in many states. He has been directing research and evaluation regarding energy efficiency and utilities for three decades, has been widely published, and has provided consultation to numerous states and the federal government. He previously directed ACEEE's Utilities program for 10 years. Prior to joining ACEEE, Marty was supervisor of evaluation at the Michigan Public Service Commission (the utility regulatory commission in Michigan) for nearly 10 years. Duties there included directing all evaluation activities relating to electric and natural gas energy efficiency programs in the state. Marty earned a bachelor of arts, a master of arts, and a PhD from Michigan State University.

Acknowledgments

This report was made possible through the generous support of The Energy Foundation, National Grid, Pacific Gas & Electric Company (PG&E), Eversource, and the Ontario Independent Electricity System Operator (IESO).

We acknowledge Dan Violette from Navigant for his contributions to the research for and writing of the common practice baseline section.

We also gratefully acknowledge the external reviewers, internal reviewers, colleagues, and sponsors who supported this report. External expert reviewers included Beth Delahaij, National Grid; Rafael Friedman and Brian Smith, PG&E; Tracy Dyke-Redmond, Eversource Energy; Bryan Ward, Grid Solutions; Michael Li, US Department of Energy; Niko Dietsch, US Environmental Protection Agency; Cheryl Jenkins and Robert Stephenson, Vermont Energy Investment Corporation (VEIC); Tim Guiterman, Energy Savvy; and Steve Schiller, Rick Ridge, Ralph Prahl, Katherine Johnson, and Gil Peach. External review and support do not imply affiliation or endorsement.

Internal reviewers included Steve Nadel, Ethan Rogers, Dan York, and Brendon Baatz. We would also like to thank Fred Grossberg for managing the editorial process; Keri Schreiner, Sean O'Brien, and Roxanna User for copy editing; Eric Schwass for assistance with the graphics; and Wendy Koch and Maxine Chikumbo for their help in launching this report.

Executive Summary

Evaluation, measurement, and verification (EM&V) plays many vital roles in the planning, development, and deployment of energy efficiency utility system resources – and those resources are growing and changing. As energy efficiency comes to greater scale in the electric and natural gas sectors, EM&V is evolving along with it. Energy efficiency has grown tremendously as an inexpensive utility system resource to meet long-term energy needs; it is also emerging as an important distributed energy resource that provides both time and locational value to meet the short-term energy needs of the distribution system. The multiple benefits of energy efficiency as a low-cost, reliable, and clean resource are, to some extent, dependent on EM&V to quantify them and make them visible.

In this report, we first identify states that exhibit policy and institutional leadership. We then discuss three topics that have recently received significant attention in the industry: deemed savings and technical reference manuals (TRMs), common practice baselines (CPBs), and advanced metering-based M&V enabled by greater data availability and improved data analytics (*M&V 2.0* has become the most commonly used term for this set of tools and approaches). We present specific case studies for each topic, followed by a discussion of major challenges facing the EM&V field.

The three specific topic areas in this report represent only a fraction of the many important developments and ongoing challenges in the energy efficiency evaluation field. Others that are significant, but beyond this report's scope, include the interactions of energy efficiency and other distributed energy resources, such as demand response; the evaluation of market effects and market transformation strategies; the role of efficiency evaluation for air regulators; measurement of nonenergy benefits; and the increased attention – especially from economists – on randomized control trials.

METHODOLOGY

In the first phase of our research, we reviewed literature and conducted interviews with 25 experts in evaluation firms, utilities, government, and nonprofit organizations. In addition to asking these experts specific questions on the selected topics, we asked several open-ended questions about their views on the field's most important challenges and developments. Their responses represent a snapshot of expert views from around the country on the field's salient issues. In the second phase, we focused on state and utility case studies recommended in the first round of interviews and gathered more in-depth information on each of the three topic areas.

STATE AND REGIONAL LEADERSHIP

Experts interviewed agreed that state policy and regulatory frameworks are primary drivers of excellence in EM&V, as well as of the magnitude and quality of energy efficiency programs. When we asked which states have been leaders in terms of evaluation framework, methods, scope, and the issues examined, California and Massachusetts were mentioned most often (11 and 8 times, respectively), followed by Connecticut, which was mentioned four times. States mentioned at least twice include Arkansas, Michigan, New York, Illinois, and Washington, along with the overall Northwest region. These findings should be viewed as a snapshot of expert views, rather than a comprehensive assessment of

state leaders on EM&V. We recognize that the makeup of our interview pool influenced the results and that some states may lead in certain areas but not others.

DEEMED SAVINGS AND TECHNICAL REFERENCE MANUALS

TRMs are critical resources for many EM&V functions. A TRM is an organized system of energy efficiency measures, deemed savings values, algorithms, and related parameters. Currently 28 state or regional TRMs have been adopted, compared to 17 in 2012.¹ The best TRMs thoroughly document each measure, have a transparent public review process, and are updated regularly with information from program evaluations and other primary sources that have been transparently mapped to individual measures.

According to the experts we interviewed, several trends are contributing to the improvement of TRMs and the accuracy of the deemed savings values within them. First, states are demonstrating how the processes for updating TRMs can be clear and transparent, building in balance among competing interests. In Massachusetts, for example, the program administrators maintain the TRM, while the statewide advisory council and other stakeholders act as reviewers. Second, regional collaboration among states helps to reduce costs and increase the accuracy of savings values among similar states in the same region. For example, the Mid-Atlantic TRM is a shared resource among Maryland, Delaware, and Washington, DC. Finally, in another step toward transparency and real-time (or at least rolling) updates, states are engaged in processes to move TRMs online.

COMMON PRACTICE BASELINES

Market data on energy-using equipment and operations are routinely used in program design and have a significant role in assessing markets for energy efficiency. Benchmarks such as common practice or industry standard practice are often used in this context. In addition to this important ongoing role in program design, there are proposals and approaches that use these data to provide ex ante estimates of programmatic energy savings.

CPBs are estimates of what a typical consumer would have done at the time of the project implementation, which becomes the basis for estimating baseline energy consumption. The CPB approach for estimating programmatic energy savings received heightened attention in the evaluation community when it was discussed in the US Environmental Protection Agency's draft Clean Power Plan guidance on energy efficiency.² The CPB approach has

¹ S. Schiller, G. Leventis, T. Eckman, and S. Murphy, *Guidance on Establishing and Maintaining Technical Reference Manuals for Energy Efficiency Measures* (Washington, DC: SEE Action, 2017).
www4.eere.energy.gov/seeaction/publication/see-action-guide-states-guidance-establishing-and-maintaining-technical-reference.

² Environmental Protection Agency, "Draft Evaluation, Measurement and Verification Guidance for Demand-Side Energy Efficiency," Draft for Public Input, August 3, 2015.
www.eebco.org/resources/Documents/CleanPowerPlan/cpp_emv_guidance_for_demand-side_ee_-_080315.pdf. An updated version of the draft is also available at blog.epa.gov/blog/wp-

been used for more than a decade in the Northwest, where it is integrated into the regional planning process and is designed to help produce regional power plans that are consistent over time. However, beyond the Northwest, it does not appear that the CPB approach has been applied widely as a method for estimating attributable program savings.

Massachusetts³ and California⁴ have been experimenting with a CPB variation that they call the industry standard practice (ISP) method. While these ISP approaches are under consideration, efforts are in the preliminary stages and consensus has yet to emerge on how ISPs might be applied to estimate net savings.⁵ Indiana and Delaware included the CPB approach as one possibility for estimating net savings in their state EM&V frameworks, but it is not clear whether they have used it in actual evaluation efforts. In Wisconsin, Focus on Energy (FOE) has used the CPB approach to estimate savings from certain residential programs.

Every evaluation approach has pros and cons, and it is important to consider the assumptions that underlie a given approach, and how the resulting estimates of savings will be used. The CPB approach has several appealing qualities, but the tradeoffs need to be clarified, both in terms of potential biases and the costs associated with using this evaluation approach. Several issues arise in using CPB evaluations:

- Differences of opinion in how the savings are interpreted, i.e., whether savings estimates should be viewed as gross savings estimates or more of a net savings construct
- Challenges related to the collection of market information on equipment sales
- Concerns about the possibly subjective processes by which CPBs are estimated, updated, and interpreted
- Concerns that factors such as self-selection bias may result in a CPB approach yielding an inflated estimate of net savings in some cases

[content/uploads/2016/12/EMV-Guidance-12192016.pdf](#). The approaches to evaluation outlined in each document are similar.

³ ERS and DNV-GL, *Massachusetts Commercial/Industrial Baseline Framework: Final Report* (Boston: MA EEAC, 2017). ma-eeac.org/wordpress/wp-content/uploads/MA-Commercial-and-Industrial-Baseline-Framework.pdf.

⁴ California Public Utilities Commission, *ISP Guide – Industry Standard Practice Guide, Version 1.2A* (San Francisco: CPUC, 2014). T. Xu and R. Friedmann, *A Summary of High-level Issues on the Existing ISP Guide (V1.2A, April 2014) and Recommendations for Revising the ISP Study Guide* (San Francisco: CPUC, 2016). www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=5315.

⁵ For a comparison of work on common or industry standard baseline in California and Massachusetts, see J. Maxwell, S. Haselhorst, and R. Prah, “Baseline Policy Enhancement in Massachusetts and California,” in *Proceedings of the International Energy Program Evaluation Conference, Baltimore* (Madison, WI: IEPEC, 2017). www.ers-inc.com/wp-content/uploads/2017/08/Baseline-Policy-Enhancement-in-MA-and-CA.pdf.

Also, the CPB approach may be appropriate to some programs but not others. Determining whether a CPB approach provides appropriate savings estimates may depend on a jurisdiction's viewpoint and how the estimates are used within that jurisdiction.

M&V 2.0

M&V 2.0 is the term typically used to describe new measurement and verification techniques that take advantage of enhanced data analytics and increased data availability. Data are becoming more granular and more readily available and with faster turnaround times due to the proliferation of advanced metering infrastructure (AMI) meters, building energy management systems, and smart devices. M&V 2.0 tools such as remote data collection, automation, and non-intrusive load monitoring offer capabilities to use these data in ways that are valuable to program designers and implementers, from customer intelligence to end-use load disaggregation.

M&V 2.0 is still in the early development stages. Our research did not find any state that used M&V 2.0 methods as the principal approach for estimating energy savings for program impact evaluation—that is, they are not being used to replace the “E” in EM&V. States and utilities are using M&V 2.0 techniques for various purposes, however. For example, PG&E's Residential Pay-for-Performance (P4P) program will use M&V 2.0 for new program designs. P4P will use gross savings, calculated as weather-normalized pre-program consumption minus weather-normalized post consumption, as the basis for making incentive payments to the program implementers. As another example, Connecticut is facilitating a three-year pilot project with several partners to assess the value of M&V 2.0 approaches for understanding energy savings.⁶ The study will consider M&V 2.0 in commercial buildings and will compare these results to those from traditional engineering analyses.

Our findings indicate that M&V 2.0 is viewed as a positive trend that can help make M&V results more timely and relevant, but states and utilities do not yet have enough information to know the full range of its impacts and opportunities. While remote collection and analysis of energy consumption data can provide early feedback on program, implementer, and participant performance, this does not take the place of independent third-party evaluation. On the contrary, deep expertise and skill in evaluation methods and practice may be more valuable than ever to the energy efficiency field.

CHALLENGES GOING FORWARD

We asked interviewees what the one or two most important challenges will be going forward. Several noted an overarching theme: the need to keep EM&V relevant in a changing world. Among the greatest specific challenges is trying to separate current energy savings impacts from those of previous program and policy efforts. After decades of policies and programs to promote energy efficiency, it is becoming difficult to isolate net impacts. As

⁶ The Connecticut Department of Energy and Environmental Protection is the lead agency, and is working with Lawrence Berkeley National Laboratory, Eversource, United Illuminating, Northeast Energy Efficiency Partnerships, and the US Department of Energy.

a result, it is becoming even more important to focus on evaluating market effects and moving beyond old ways of assessing attribution.

Other prominent challenges mentioned are those posed by the proliferation of new hardware and software energy data technologies, which supply vast amounts of data. The challenge is to identify the best and most appropriate uses of these data and to decide how to efficiently process and analyze the information. Closely related is the use of M&V 2.0 tools, which have a range of possible applications. Utilities, regulators, and evaluators are faced with questions of how to evaluate those applications and determine which tools can be used to accomplish different objectives.

Some challenges are external to the practice of EM&V, such as the need to better communicate evaluation results to policymakers and regulators, to respond to pressure to reduce EM&V spending, and to determine the impact of political changes that do not favor energy efficiency. We also must find an appropriate balance between independence and oversight. Within EM&V practice, key challenges include addressing self-report evaluation surveys that allow excessive time between participant decisions and data collection, and dealing with ongoing changes in methods and research approaches.

Many more challenges exist as well. A recent overview of EM&V activities identifies several important research gaps, including persistence of energy savings, the need for consistent reporting, and practitioner training and certification.⁷

CONCLUSIONS

The EM&V field is actively growing and changing along with utility energy efficiency, as it must for efficiency to remain a critical resource in the evolving utility industry. An examination of just a few relevant topics reveals a large amount of activity to improve and refine EM&V strategies, processes, and policy frameworks. For example, more states and regions are adopting TRMs and enhancing the accuracy and consistency of the deemed savings values within them. Utilities are examining how they can use M&V 2.0 tools to better understand and engage customers, target program marketing more effectively, and be more nimble and responsive in program implementation using faster feedback on program performance. Many states have been adopting regulatory policies and experimenting with strategies to both seize these new opportunities and tackle challenges to further expand the many benefits of the energy efficiency resource.

⁷ L. Schwartz et al., *Electricity End Uses, Energy Efficiency, and Distributed Energy Resources Baseline. Appendix: Evaluation, Measurement, and Verification of Energy Efficiency and Distributed Energy Resource Activities* (Prepared by Berkeley Lab. Washington, DC: Department of Energy, 2017). emp.lbl.gov/sites/default/files/lbnl-1006983.pdf.

Introduction

The United States has more than three decades of experience implementing energy efficiency evaluation, measurement, and verification (EM&V), and its continued development will be important as energy efficiency as a utility resource continues to expand and evolve. This paper focuses on several recent developments in the evaluation field for utility-sector energy efficiency.

Over the past few decades, energy efficiency has grown tremendously as an inexpensive utility system resource to meet long-term energy needs. More recently, it is emerging as an important distributed energy resource that provides both time and locational value to meet the distribution system's short-term energy needs as well. The multiple benefits of energy efficiency as a low-cost, reliable, and clean resource are all, to some extent, dependent on EM&V to make the impacts visible and persistent, and to demonstrate their reliability. To this end, energy efficiency EM&V is a central and critical factor in the development and growth of utility energy efficiency resources. Meanwhile, innovations in data analytics and data access are setting up new opportunities for EM&V to evolve and improve.

As an example of this greater focus on energy efficiency, 26 states now have efficiency resource standards or similar policies that set long-term energy savings targets for utilities. Also, several states are exploring efficiency's role in distribution system planning, with many already defining distributed energy resources to include energy efficiency. Further, energy efficiency is playing an increasingly important role in regional capacity markets and in planning and forecasting by independent system operators (ISOs). Although this role is uncertain in the near term, its long-term importance as a reliable grid resource (both in resource planning and distribution system planning) will likely grow, as will the importance of EM&V.

Environmental policy has been another recent driver. For example, the resources that went into preparing draft materials for the proposed US Environmental Protection Agency (EPA) Clean Power Plan (CPP) included a significant focus on EM&V guidance (EPA 2015). Despite uncertainty around federal environmental policies, regional environmental policy drivers persist, such as the regional greenhouse gas initiative (RGGI) and an increasing number of states committing to pursue policies that reduce greenhouse gas emissions.

WHAT IS EM&V?

This report is about EM&V of utility customer-funded efficiency, including oversight by regulatory bodies. In this context, energy efficiency EM&V serves three critical objectives: to document program impacts, to identify ways to continuously improve programs, and to support energy resource planning and demand forecasting (SEE Action 2012).

First, EM&V activities document and measure a program's effects and determine its success. This often includes the energy and demand savings, as well as co-benefits such as emissions impacts, transmission and distribution benefits, and water savings.

Second, EM&V evaluates why a program had the effect that it did and how it can be improved, with an eye for both improving existing programs and providing a robust mechanism for estimating savings from planned programs.

Third, EM&V assesses the realization of energy efficiency's expected savings to support energy resource planning and demand forecasting. EM&V methods should be sophisticated enough to provide reasonably high confidence and precision levels, which gives energy efficiency credibility as a viable resource that can be compared to other resources. In the absence of compelling evaluation results, governments may underinvest in cheaper and more beneficial energy efficiency programs, and overinvest in more costly alternatives such as generation, distribution, or pipeline infrastructure. EM&V activities aim to provide highly reliable savings estimates and thereby avoid costly misallocation of public and private resources.

REPORT OVERVIEW

Policymakers and practitioners are constantly improving and adapting EM&V to achieve these objectives. In this report, we give examples of how leading states are improving EM&V and the policy frameworks that enable strong EM&V. We then explore three topics that are attracting significant attention: (1) the use of deemed savings approaches and technical resource manuals (TRMs), (2) the extent to which states are using common practice baselines (CPBs), and (3) how states and utilities are responding to opportunities for improving EM&V through data analytics and data access.

Finally, we summarize some of the most important recent developments and critical future challenges for EM&V that we elicited from a survey of evaluation experts. These represent only a fraction of the many important developments and challenges in the energy efficiency evaluation field. Others that are significant but beyond this report's scope include the interactions of energy efficiency and other distributed energy resources such as demand response, solar, and storage; the evaluation of market effects and market transformation strategies; the role of efficiency evaluation for air regulators; measurement of nonenergy benefits; and the increased attention on randomized control trials (RCTs) and similar comparison group methods. For a discussion of 11 specific research gaps for EM&V activities for energy efficiency and distributed energy resources, see Schwartz et al. (2017).

We organize this report into six sections:

- Methodology
- State and regional leadership in EM&V
- Deemed savings and technical resource manuals (TRMs)
- Common practice baselines
- M&V 2.0 (improved data analytics and data access¹)
- Challenges going forward

Methodology

Our research involved two phases. First, we sought out the most salient issues and recent developments by interviewing experts in the field, reviewing literature, and researching

¹ M&V 2.0 has become the most commonly used term for a set of tools and approaches that can facilitate certain types of advanced metering-based M&V, enabled by greater data availability and improved data analytics.

state-specific examples. We conducted initial interviews with 24 individuals, including various EM&V consultants and vendors, evaluation managers at utilities, and others at NGOs and national labs with energy efficiency evaluation experience.² We first asked about overall EM&V challenges, developments, and specific examples of state leadership. We then delved into specific topics, including M&V 2.0, deemed savings, and CBP and market-based approaches.

Following the initial set of interviews, we began the second research phase, conducting telephone interviews with individuals recommended to us in the first round to collect further information. In particular, we sought additional details on projects mentioned by sources in the first round and on potential candidates for good state- or utility-specific case studies.³

Throughout this paper, we refer to comments made in the interviews, but we do not attribute individual comments to particular respondents. In the first round of interviews, we asked all interviewees to put aside any of their company- or organization-specific interests and give us their objective and candid thoughts as professionals in the field.

In addition to the expert interviews, we also conducted independent research of the literature and on state examples.

State and Regional Leadership

State legislation and regulation determine the framework and process for conducting both EM&V and energy efficiency programs. Although ACEEE has done a national survey of all states on specific EM&V practices (Kushler, Nowak, and Witte 2012) and net savings approaches (Kushler, Nowak, and Witte 2014), systematic ongoing assessment of state policy commitment on EM&V is a current gap in the literature. For example, ACEEE's annual *State Energy Efficiency Scorecard* does not currently include any policy metrics related to the structure or implementation of best practice EM&V policies, which can be difficult to compare across states. To help fill this gap, ACEEE included some elements related to EM&V in its recently published *Utility Energy Efficiency Scorecard* and is exploring new state policy metrics related to EM&V for its next *State Scorecard*.

Our experts concurred that state policies and regulatory requirements are keys to high-quality EM&V. To further explore this issue, we asked interviewees which states have shown leadership on energy efficiency EM&V. We tallied responses to this question and

² The following individuals participated in first-round phone interviews: Kausar Ashraf, Jordana Cammarata, Beth Delahajj, Tracy Dyke-Redmond, Tom Eckman, Ellen Franconi, Rafael Friedmann, Michael Goldman, Steve Grover, Tim Guiterman, Randy Gunn, Nick Hall, Katherine Johnson, Michael Li, Sierra Martinez, Peter Miller, Jake Oster, Gil Peach, Jane Peters, Ralph Prah, Rick Ridge, Steven Schiller, Brian Smith, Liz Titus, and Bryan Ward.

³ The following individuals participated in second-round phone calls or provided information via email: Annette Beitel, Zoe Dawson, Phil Degens, Niko Dietsch, James Fay, Joe Forcillo, Scott Gentry, Ethan Goldman, Jessica Granderson, Vincent Gutierrez, Jennifer Holmes, Cheryl Jenkins, Katherine Johnson, Jason Kupser, Ryan Lee, Ariana Merlino, Rick Morgan, Dean Pollock, Kristen Rowley, and John Ware.

provide the results in table 1. In particular, we asked interviewees to consider each state's framework, methods, and scope of evaluation, as well as the issues the state examines. This should not be viewed as comprehensive assessment of state leaders on EM&V. Rather, it represents a snapshot of expert views from around the country. We recognize that the make-up of our interview pool could affect the results, and that some states may lead in certain areas but not others.

Table 1. Expert interview responses: state leaders on EM&V

State	Number of mentions
California	11
Massachusetts	8
Connecticut	4
Arkansas	3
Michigan	3
New York	3
Illinois	2
Washington	2
Northwest region	2

California and Massachusetts emerged as the overall EM&V leaders in the opinions of our experts. They identified Connecticut, Illinois, and New York as having state policy that drives advanced M&V-type approaches in particular. The new policies in Illinois and New York are emerging in the context of efforts to reform utility regulation and operation—namely, Next Grid in Illinois and Reforming the Energy Vision (REV) in New York. In addition to the states cited in the table, the following jurisdictions received at least one mention: Maryland, New Mexico, Ontario, Oregon, Pennsylvania, Rhode Island, and Wisconsin.

Some experts also noted that examples from states such as California and Massachusetts cannot necessarily be replicated easily in states that are just getting started and lack sufficient budgets. This presents the bigger challenge and question of how advanced EM&V can be deployed in the wider market.

STATE LEADERSHIP EXAMPLES

We now briefly profile the two leading states on EM&V according to our interviewees, along with two regional efforts that they also cited as noteworthy.

California

Our interviewees recognized California as a longtime state leader on EM&V. California has invested heavily in evaluation on multiple fronts for decades, often developing EM&V resources that many other states have later adopted or adapted. Examples include the influential and widely used California Evaluation Framework (CPUC 2004), California Evaluation Protocols (CPUC 2006), the California Standard Practice Manual (CPUC 2001),

and the Database for Energy Efficient Resources (DEER). The California Technical Forum has been moving beyond DEER, developing an electronic TRM (eTRM). California regulators and evaluators have devoted resources to developing and employing methods for estimating net savings in program impact evaluations. This focus on attribution issues—that is, how much of the energy savings were due to program activities, and how much were not—has greater policy and business consequences in California than in any other jurisdiction. This is in part because the scale of energy efficiency portfolios is so much larger, and also because California policymakers have sustained a strong interest in attribution as program portfolios have matured over decades.

Perhaps in part because of the scale of those consequences, California is also known for having a sometimes contentious process surrounding EM&V methods and results, which at times has led to delays in regulatory decision making. Nevertheless, California is widely acknowledged as a leading state for its historical level of support for evaluation and the scope of its EM&V activities.

What the experts we interviewed specifically noted most often, however, were two pieces of recent California state legislation on utility policy, AB 802 and SB 350, which will affect the state's EM&V practices. The first, AB 802, allows for normalized metered energy consumption (NMEC) as well as "EM&V changes from artificial conditions to existing conditions" (California Assembly 2015). The second, SB 350, calls for doubling energy savings. In the words of one expert we interviewed, the question of "How do we measure that?" makes SB 350 one of the top EM&V developments. These changes in state policy, along with others, leave some EM&V activities in flux and somewhat uncertain.

The California Public Utilities Commission (CPUC) and investor-owned utilities (IOUs) are currently evaluating whether to update the evaluation framework, in part because of these two new laws. These stakeholders' preliminary findings are that updates are warranted, but that other gaps and challenges remain (for example, a changing policy landscape means the document should be developed in a way that is easy to update); they also found that a framework update should not be conducted independently from an evaluation protocols update because the two are linked (California Public Utilities Commission 2017). AB 802 directs the increased use of existing condition baselines as a default assumption for program implementation.⁴ SB 350, which includes a goal of doubling California's energy efficiency savings by 2030, emphasizes approaches such as market transformation and pay-for-performance.

Changes in California's electricity supply, including plant closures, increased solar penetration, and the duck curve phenomenon have been key drivers for policy discussion around energy efficiency.⁵ There is now an increasing need for energy efficiency to replace some of the lost baseload and to reduce demand at certain times of day. The CPUC is

⁴ Certain justified exceptions are allowed in cases where a baseline determined by codes and standards and/or a dual baseline would be appropriate, as determined by the commission.

⁵ The state used to see peak demand in the hottest afternoons of summer, which are now times of oversupply due to solar generation; peak demand has now shifted into the evening. The demand curve for electricity now has the shape of a duck's back.

approaching efficiency resource planning through multiple channels. One of these is phase III of the CPUC's energy efficiency rulemaking (R. 13-11-005), also known as the *rolling portfolio proceeding*, which is expected to remain open into mid-2018.

These broader policy and market trends are driving some of the state's new EM&V trends, and utilities are responding in part by tapping into opportunities through M&V 2.0 tools. As we discuss below, some Pacific Gas & Electric (PG&E) energy efficiency programs are increasingly looking at whole-premise evaluation, ongoing engagement, and multiyear commitments. As we also discuss below, another utility, Southern California Edison (SCE), has been working on a Preferred Resources Pilot (PRP) to examine its energy efficiency efforts on the grid and understand how the measures perform.

Massachusetts

Several experts named Massachusetts as a leader in EM&V. They noted that Massachusetts has a strong EM&V policy framework, uses empirically driven results to update the TRM, supports a more extensive research effort than other states, has dedicated EM&V funding, and is a leader in examining market effects. The state views EM&V as a critical engine for providing objective, fact-driven results to ensure reliable savings claims.

The Massachusetts Energy Efficiency Advisory Council (EEAC) oversees all evaluation activity, while the utility program administrators (PAs) oversee the evaluations through an Evaluation Management Committee that includes a representative from each PA and from the EEAC consultants.⁶ The EEAC, which was established by state law and has council members appointed by the Department of Public Utilities (DPU), is the state's energy efficiency collaborative.⁷ In 2009, the EEAC adopted a formal resolution establishing an EM&V framework, which was then submitted to the DPU. The EEAC has a dedicated EM&V consulting team; its purpose is to enhance and ensure the objectivity, accuracy, timeliness, and usefulness of utility EM&V activities on behalf of the EEAC. The team must oversee all aspects of the EM&V process, including evaluation planning, contractor procurement, study implementation, reporting and outreach, and application of study results.⁸

Before 2009, EM&V was done at the utility level. Under the 2009 resolution, EM&V was made primarily a statewide function and formal decision making authority was given to the EEAC or its agents. Aside from potential studies and pilots developed by utilities, most studies are statewide. Multiple contracting teams collectively produce approximately 40–50 studies per year. Most Massachusetts programs have consistent designs and marketing under Mass Save, which enables a statewide approach to EM&V.

⁶ Massachusetts EM&V studies completed over the last several years can be found here: ma-eeac.org/studies/.

⁷ Massachusetts is one of only a few states that have an “enhanced” energy efficiency collaborative (the Energy Efficiency Advisory Council), characterized by a significant operating budget, statutory permanence, and a broad array of specific tasks and responsibilities (SEE Action 2015).

⁸ For additional information on the 2009 resolution, see ma-eeac.org/wordpress/wp-content/uploads/Consultant-Team-Year-2014-Approved-Workplan1.pdf.

According to the 2009 resolution, the EEAC can overrule decisions by individual PAs. However the goal was for that authority not to be needed, and to instead establish a collaborative framework. From the start, the EEAC and PAs set out to reach negotiated solutions to every issue, and thus EEAC has never in eight years had to overrule any PA decision. Still, the underlying authority provides leverage. The resolution enabled this collaborative framework, which is a delicate balancing act, but ultimately yields trust and respect. This collaboration and overall framework have been important elements of the effort's success. They lead to an atmosphere of stability and deal with conflict of interest by having an outside party with authority to oversee the process.

Regional Efforts

Two regional EM&V efforts were mentioned by our interviewees as particularly noteworthy. The Northwest and the Northwest Regional Technical Forum (RTF) in particular were mentioned for a very effective EM&V collaboration. (The states of Washington and Oregon were also cited individually.) The RTF was established in 1999 as a technical advisory committee and standard-setting body on evaluation issues to the Northwest Power and Conservation Council (NW Council). The RTF's goal is to support consistent and reliable quantification of energy savings. It operates as a regional network of efficiency experts who collaborate (online and in person) on a dynamic catalog of data for a comprehensive set of efficiency measures. An evaluation expert we interviewed described the RTF as having an "effective, transparent, collaborative, productive" process that leads to savings estimates that are "accurate, timely, and useful."

Another leading regional organization is the Regional Evaluation, Measurement and Verification Forum (EM&V Forum), which serves the northeastern and mid-Atlantic states. The EM&V Forum, established in 2009, is an initiative of the Northeast Energy Efficiency Partnerships (NEEP). Our experts commended NEEP for the following:

- Advancing evaluation frameworks set up for net savings, not gross
- Cost-effectiveness measures that quantify nonenergy benefits
- A transparent and accessible analysis process, including a public website showing how it derives its savings results
- Encouraging movement toward more real-time evaluation results rather than results that come two years or more after the program year began

Deemed Savings and Technical Reference Manuals

The first of the three topics we consider as a means to describe recent developments in EM&V is the use of deemed savings and TRMs. An essential resource for EM&V, TRMs are used in deemed savings EM&V methods to determine savings of specific efficiency measures, and in the M&V method for deemed savings calculations. For program planning and implementation, TRMs can serve as a single, consistent source to be used as the basis for calculations of per-measure impacts (Schiller et al. 2017).

TRMs contain energy efficiency measure deemed savings values, algorithms, and related parameters, compiled and maintained in a unified system.⁹ TRMs can take the form of a database, website, document, or set of spreadsheets, and they are usually operative at the state level. There are also utility-specific TRMs and regional TRMs. TRMs provide consistent and accurate prescribed energy savings values and algorithms for planning and forecasting, reporting, cost-effectiveness analysis, and evaluation. Many stakeholders, including utilities, program implementers, evaluators, and regulators rely on TRMs.

A recent review found that 28 state or regional TRMs have now been adopted, compared to 17 in 2012 (Schiller et al. 2017). As more jurisdictions adopt and improve their TRMs, they are gaining more attention in the industry. The Midwest Energy Efficiency Alliance (MEEA) has cataloged the attributes of TRMs in its region to provide a comprehensive side-by-side look at their features and applications (MEEA 2017).¹⁰ The SEE Action Network has also published a national guide to TRMs (Schiller et al. 2017).

Best practices for TRMs include clear, well-documented information about the measures, a public measure review process (to add, change, or remove measures), regulatory commission final approval, and regular (often annual) updates of values based on program evaluation (Beitel et al. 2016). Updating deemed savings values often involves prioritizing measures based on the value of energy savings to the portfolio and the length of time elapsed since values were last updated.¹¹ Reviews and updates may be done less frequently for measures that provide a relatively small share of energy savings.

If any of these best practices are not followed to keep a TRM up to date, with a current set of properly documented measures, its accuracy and effectiveness will be limited. Five common TRM weaknesses illustrate this. If a TRM's values are based on data from another state, the values may not be accurate and must be tested against primary data collected in-state. Similarly, savings values are sometimes deemed based on engineering simulations instead of primary data collection from buildings in the field (however this approach can be improved by ensuring that engineering models are calibrated to actual usage).¹²

A third issue is that the deemed effective useful life (EUL) of a TRM's measures might be outdated; many states, for example, use EUL numbers from the California DEER, which was started in the early 1990s. A fourth potential problem is that operational assumptions regarding how customers use measures might be inaccurate; these assumptions need regular testing to validate their accuracy. For example, it might make a big difference in estimated savings if it was assumed that, on average, a screw-in light bulb was used for 150

¹⁰ See <http://www.mwalliance.org/sites/default/files/media/Midwest-TRM-Inventory-Feb2017.pdf>.

¹¹ Deemed savings values are “estimates for the energy and/or demand savings for a single unit of an installed energy efficiency measure that (1) have been developed from data sources (such as prior metering studies) and analytical methods that are widely considered acceptable for the measure and purpose, and (2) are applicable to the situation being evaluated.” For a more complete description, see aceee.org/sector/state-policy/toolkit/emv.

¹² For example, see psdconsulting.com/nyserda-hpwes-rr-study/.

hours per year, instead of 300 hours per year. A fifth concern is that, for some measures, customer behavior can be a factor in the savings and other parameter values. For example, programmable thermostats result in energy savings beyond traditional thermostats only if they are programmed properly (or programmed at all). In part because of this uncertainty in behavioral response, some jurisdictions no longer include programmable thermostats as a measure.

In addition to concerns regarding content, it is important that states use an impartial and unbiased update process. Because measure technologies, products, and cost-effectiveness levels are constantly evolving, TRMs will become inaccurate over time unless processes are in place to calibrate deemed savings values, add new measures, and remove measures that no longer qualify. Utilities, program implementers, and manufacturers all have an interest in higher energy savings. Having a rigorous and transparent update process can help add a balance of interests, in part by engaging multiple stakeholders such as consumer advocates, large energy users, evaluators, utility regulators, and nonprofits.

DEEMED SAVINGS ESTIMATES: DEFINITIONS AND BEST PRACTICES

Deemed savings estimates are one of the many components of a TRM. The definition of deemed savings estimates falls along a spectrum of approaches—some savings estimates are based entirely or partially on previous EM&V studies, while others are based on engineering analyses but perhaps make too small a contribution to savings to warrant detailed studies. Along this spectrum, certain assumptions and algorithms may be deemed based on previous studies, while the actual savings estimates are not. Deemed savings estimates based on prior EM&V results and calibrated engineering estimates can be highly accurate. Also, deemed savings estimates developed based on results from multiple studies can be more comprehensive and accurate than any single evaluation for a specific program.

It is also important to include protections against possible inappropriate use of deemed savings values. For example, if measures have large variations in savings, they are not good candidates for deemed savings estimates. Also, estimates should specify the measures and conditions to which the savings estimates apply to ensure that the estimates are used only in proper situations.

TRM TRENDS AND DIRECTIONS

We asked the experts what the leading states have been doing to solidify or improve the use of deemed savings. We found that the use of TRMs and deemed savings¹³ is widely regarded as an important tool for saving time and reducing EM&V costs. However the experts note important caveats regarding TRM use, including that they should be constructed in an independent and unbiased manner, transparent, and regularly updated as better and newer information becomes available.

¹³ Energy savings are only one type of parameter that can be deemed; other parameters include methods, calculations, factors, and other values.

The experts also said that evaluation results were being used to improve the quality of deemed savings and other values like measure lives within the TRMs. Some of the trends and directions they observed are as follows.

Clear and transparent processes. TRMs should be, and increasingly are, updated, timely, clear, and well documented. There should be an “ongoing effort to update the extensive technical manual through a clear annual process.” Leading states using TRMs have an advantage because they include “key parameters, source of data, and [a] clear path to update those parameters—[they are] clear and transparent.” A key design feature of these processes is collaboration. Arkansas is a prime example. The Arkansas Parties Working Collaboratively (PWC) comprises the investor-owned utilities, EM&V contractors, program implementers, and intervener groups. One of the PWC’s most important responsibilities is to update the TRM. In Massachusetts, the TRM is viewed as a mechanism for applying all impact evaluation study results. For each impact study, stakeholders assess how it will affect the TRM. While PAs administer and maintain the TRM, the EEAC and other stakeholders act as the reviewers.

Increased regional collaboration. Geographically adjacent states sometimes share deemed values or entire TRMs. This reduces costs and increases consistency of savings values. Examples provided during interviews were the Massachusetts and Rhode Island collaboration; Arkansas as a “*de facto* regional TRM”; the Mid-Atlantic TRM serving Maryland, Delaware, and Washington, DC (NEEP 2017); and the Northwest RTF library of unit energy savings measures (RTF 2017).

Online TRMs. California is in the process of developing an eTRM through the California Technical Forum. In Massachusetts, the TRM has been moved online to serve as a technical resource library (Mass Save 2017). One objective is to make the information more transparent and accessible. The Northwest’s RTF has many online resources, including a library of current unit energy savings measures and a deemed savings database (RTF 2017). In Vermont, Washington DC, and Ohio, TRMs are accessible to program implementers, but not to the public.

More-frequent or rolling updates. Many jurisdictions are updating parts of their TRMs on a more frequent periodic basis, such as annually. Massachusetts is a good example here, too, as the state has had a clear process in place for updating the TRM annually for many years. Once the TRM was moved online, decision makers began discussing the capability of having values updated in near-real time for some purposes, rather than having only annual updates.

Reliance on deemed savings values. Some experts said that jurisdictions are moving toward using deemed savings as their accuracy and consistency improves, while others noted that a few states are trying to move beyond past reliance on deemed savings methods. Overall, most experts noted that deemed savings make program implementation easier, and that they are incredibly useful, as long as they are not overused or improperly used. Along these lines, even as more states and regions adopt or improve their TRMs, some experts asked, “Given more powerful technologies, why deem it when you can measure it?” While a full answer is beyond this report’s scope, we can say that certain programs and approaches are

better suited to project-level M&V and technologies that can measure energy consumption with far more precision than was previously possible. In some cases, data from M&V 2.0 can be used to update deemed savings values. Some experts highlighted California, New York, Illinois, and Maryland for taking steps to reduce over-reliance on deemed savings values. Others cite the value to utilities of having savings values set in advance (i.e., deemed) to reduce risk and uncertainty for program planning and investment decisions. Deemed savings approaches are not going away, but perhaps we will see more segmentation going forward, with deemed approaches used for some program segments.

Increasingly adaptive and adaptable. One expert aptly summarized the evolution of TRMs as moving from “static repositories to evolving databases” and from compilations of deemed values to “system[s] of protocols and practices, tied in with, and responsive to, programs.”

STATE AND UTILITY EXAMPLES OF TRM EVOLUTION

Commonwealth Edison Smart Thermostats

Experience with smart thermostats in Illinois illustrates aspects of how TRMs relate to the evolution of the energy efficiency industry and particularly to measure technologies. Smart thermostats are a noteworthy example because they are at the intersection of several trends in utility energy efficiency and EM&V. First, sales have been projected to double in 2017 (Statista 2017). Additionally, smart thermostats are a dual-fuel measure, they impact both energy savings (kWh) and demand (kW), and they have technological capabilities (such as real-time remote data collection and analysis) that could potentially make them good candidates for M&V 2.0.

Commonwealth Edison, in collaboration with Nicor Gas, North Shore Gas, Peoples Gas, smart thermostat manufacturers, and advocacy groups, launched an effort to accelerate smart thermostat adoption in 2015 with the five-year goal of installing one million smart thermostats (Wood 2015). Among the initial goals the group set were to include deemed energy savings values into the TRM and to develop common definitions and standards for smart thermostats (Thornburg and Jewell 2015).

In Illinois, smart thermostats are now defined as having three features that distinguish them from programmable thermostats: they let customers control home heating and cooling remotely via smart phones and computers using Wi-Fi, and they have both occupancy sensing and learning capabilities. Having learning capabilities makes a difference for the estimated useful life, and therefore, the deemed energy savings of the measure, because otherwise energy savings would depend on customer behavior to keep the thermostat programmed. In the Illinois TRM, the estimated useful life for smart thermostats is the full physical life (that is, the duration until it physically stops working). This is longer than for programmable thermostats, which at some point will predictably be left unprogrammed. Questions concerning the net energy savings for programmable thermostats led to ENERGY STAR suspending the product category in 2009 (EPA 2017).

Manufacturers of smart thermostats can collect and analyze granular data on internal temperature and on whether the thermostat is signaling the HVAC system to heat or cool the house. Using these data, energy consumption can be modeled. On the surface, this might imply that the energy savings could be measured at some point, rather than deemed ex ante,

but there are barriers to doing so. Measured energy savings data for individual residential customers' homes are not always easily available to utilities, program evaluators, or regulators for legal and data security reasons. Smart thermostat manufacturers have a competitive business interest in protecting their proprietary data, while customers have rights to keep their individual energy consumption data confidential. In any case customer data must be made anonymous before it is shared. One manufacturer, ecobee, has a data sharing initiative called *Donate Your Data* that uses anonymization. Utility customers are invited to do as the name says and provide granular data to be analyzed in order to find energy savings opportunities (ecobee 2017).

As a result, despite the product's high profile and its technical potential to more precisely measure energy consumption at the individual home level, Illinois currently uses single deemed savings values for smart thermostats, where the deemed values are based on impact evaluation studies. It uses one set of deemed values for homes with heating and cooling, and individual electric and gas values for homes in which smart thermostats control only heating. That is, a prescribed amount of energy savings is in the TRM for natural gas (from furnaces) and another set value is used for electricity (from air conditioners and furnace fans.)

Illinois is moving toward more precise and accurate energy savings estimates for smart thermostats in the TRM through the Illinois Stakeholder Advisory Committee and a subcommittee on smart thermostats. Commonwealth Edison (ComEd) is engaged in the process as well and has made a data request to the manufacturers. Another possible source of data that is more granular than the monthly billing is whole-house meter data collected in 30-minute intervals from advanced metering infrastructure (AMI) meters, also known as smart meters. These much more detailed data hold greater potential explanatory power for program optimization as well as impact evaluation. ComEd's evaluator, Navigant, may consider examining AMI data in more depth to assess energy savings for smart thermostats in the future.

Michigan

Michigan is a noteworthy example of a pragmatic EM&V approach, particularly in relation to deemed savings. Michigan evaluators, regulators, and utilities have focused on increasing the accuracy and consistency of deemed savings values. According to some of the experts we interviewed, Michigan relies more on deemed savings than other states. The state's mechanism, comparable to a TRM, is the Michigan Energy Measures Database (MEMD). The MEMD is updated and enhanced through an annual collaborative process with input from utilities, Michigan Public Service Commission (PSC), and others.

An important feature of Michigan's EM&V approach is the use of a deemed net-to-gross (NTG) ratio. The state has used 0.9 widely for many programs, but not for compact fluorescent lamps (CFLs), for which it uses 0.82. The initial deemed ratio was established via a review of available data from other states, with values that were to be updated over time as more data became available. Deeming the NTG ratio reduces uncertainty for utilities by eliminating the risk of a retroactive application of a different NTG ratio than that assumed in program planning, and avoids the controversy and arguments over attribution issues that have occurred in other jurisdictions. A deemed NTG is another reason that utilities and

other stakeholders pay attention to the accuracy and consistency of deemed savings measure values in the MEMD.

Michigan's overall approach has several advantages. It is statewide, which provides consistency of measure savings values across utilities, programs, and contractors. It is updated annually through a collaborative of utility and non-utility members, which builds consensus and buy-in among stakeholders. The savings values themselves are enhanced through ongoing calibration research. When the system was originally established, Michigan initially borrowed and adapted many values from other states. Over time, the state conducted more primary research, including conducting evaluations and building modeling simulations, to refine and improve the accuracy of its database.

Another advantage is that administering the MEMD update process costs less than \$100,000 per year, a relatively small part of EM&V spending in Michigan, which is several million dollars per year. For context, its overall utility program spending was more than \$250 million in 2016.

Pilot programs are an important aspect of the MEMD update process. Michigan utilities can use up to 5% of their efficiency budget for pilot programs. They then get credit for an equivalent percentage of their annual energy savings goal based on that spending. For example, if their energy savings goal is 1% and they spend 5% of their budget on pilots, they would get credit for 1/20th of their goal by virtue of the pilot. Having such a strong incentive for utilities to run pilot programs creates a pipeline of measures that have been field tested and may ultimately be added to the deemed savings database.

One weakness of the MEMD is that it is based on average savings values for each measure as calibrated in a variety of commercial and residential building types. That is, it is a set of statewide average values rather than custom or utility-by-utility values. While deemed savings values are applied using weather factors for different parts of the state (Northern Michigan has a different climate, for example), there is still some loss of precision.

The MEMD administrator has worked with the commission and utilities to refine deemed values over time to take advantage of smart meter data and analytic capabilities for better data gathering and better billing analysis. More data and sub-segments to sample yield higher-quality savings numbers. In the past, this level of analysis would have required sub-metering of buildings.

Arkansas

Arkansas has become a leading state in the Southeast in utility energy efficiency evaluation. The state uses a multiple stakeholder collaborative process to oversee and provide input on evaluations. A key to its success is that stakeholders consider EM&V and program planning, design, and implementation as an integrated whole.

Starting in 2012, the state's independent evaluation monitor (IEM) worked with the public service commission to develop an extensive evaluation framework alongside the creation and expansion of energy efficiency programs. They did this through a stakeholder group called the *Parties Working Collaboratively* (PWC), comprising investor-owned gas and electric utilities, EM&V contractors, program implementers, and intervener groups. Having the IEM

as an independent coordinator facilitated effective collaboration. Together, they spelled out the framework in a multi-volume TRM that includes savings and parameter values as well as evaluation protocols.

Arkansas was the first southeastern state to develop a TRM, even before Texas, which was the first state to establish an energy efficiency resource standard (EERS).¹⁴ Arkansas Commission staff members have consulted frequently with staff from other states, such as Louisiana and Mississippi, on the TRM development process. By using primary data to provide data calibrated to the South, the Arkansas TRM has become to some extent a de facto regional TRM. It is used as a reference not only by Louisiana and Mississippi, which are just beginning to establish their energy efficiency programs, but also states with a longer history, including Georgia and Texas.

The Consistent Weatherization Approach (CWA) is another example of Arkansas leadership in EM&V. As the state began expanding utility energy efficiency, the Oklahoma Gas and Electric (OG&E) and Arkansas Oklahoma Gas (AOG) utilities came together to propose a new weatherization program to the commission, because the “quick start”¹⁵ program was not meeting its performance goals. The utilities share service territory and together run the joint program, which addresses both gas and electric end uses. They created the AOG/OG&E Weatherization program in response to the needs of hard-to-reach customers in their service territory who were unable to be served effectively through the traditional CAP agencies. This joint-utility model was the basis for the CWA, which was developed and later approved by the Arkansas Public Service Commission after it saw the approach’s effectiveness.

Today, the CWA has consistent electric and gas program offerings statewide across all utilities. As evaluation results come in, they are used to update the TRM’s deemed savings values. Because programs are consistent statewide, the TRM numbers are more reliable for all utilities for program planning and implementation, in turn leading to better evaluation quality and fewer surprises. Utility program managers say this helps them all year, since they can see more accurately where they are and where they are likely to end up for savings, costs, participation, and other measures. Contractors and trade allies use an online portal that gives them an order of measures based on cost effectiveness derived from the TRM’s deemed values.

To enhance the TRM’s accuracy, Arkansas is taking the critical step of moving toward greater use of prospective deemed savings values. *Prospective deemed savings values* mean that impact evaluation results from Year X programs are used to calibrate the deemed savings values in the TRM in Year X + 1, increasing accuracy and reliability. The energy savings estimate is updated if the evaluated savings estimate for a measure represents a 10% difference from an existing value. When the state began to develop EM&V, regulators based

¹⁴ Energy efficiency resource standards are specific, long-term energy savings targets required by state law or regulation, usually expressed as a percentage of annual retail sales, which gas or electric utilities must achieve through their energy efficiency program portfolios.

¹⁵ Arkansas had utilities implement several quick start programs initially, laying the foundation for longer-term efforts.

the initial savings values on those from other states, and then adjusted them to make them more Arkansas-specific with respect to variables such as hours of use for each measure, heating and cooling temperatures for weather adjustment, and other parameter values. With prospective deemed savings, continuous improvement is built into the process.

Common Practice Baselines

We now examine CPBs, an evaluation approach that received increased attention across the evaluation community when it was discussed prominently in the EPA's "Draft EM&V Guidance for Demand-Side Energy Efficiency" for the CPP (EPA 2015 and 2016).¹⁶

All evaluation approaches have pros and cons, and CPB approaches are no different. Decisions regarding the application of any evaluation methods require an understanding of the assumptions that underlie different approaches, as well as the evaluation's goals and potential uses of the estimates. In this context, SEE Action (2012, 7-2) defines CBPs as "estimates of what a typical consumer would have done at the time of the project implementation. Essentially, what is 'commonly done' becomes the basis for baseline energy consumption."¹⁷

However the CPB approach has not been widely addressed in the broader evaluation literature, which leads to confusion in its use. In general, this approach uses available information to develop an ex ante estimate of baseline consumption. This information can then be used, with limited adjustments based on ex post data and analysis, to develop savings estimates.

Common practice baseline is the term used in the CPP's "Draft EM&V Guidance." The term used in the Pacific Northwest is slightly different, as the Northwest Power and Conservation Council (NW Council) calls the approach *current practice baselines* (also abbreviated CPB), and other applications use the term *industry standard practice* (ISP). These are the terms most commonly used in North America; in Europe, the same methods are called *business as usual* (BAU). Here, we use the term *common practice baseline*, or CPB.

A number of the experts we interviewed indicated that they were first introduced to this evaluation method in the CPP's "Draft EM&V Guidance." Appendix A offers some details about their responses to the CPB method. Here, we describe the underlying concepts of the CPB method, discuss how the method can be applied, and outline recent trends with several jurisdictions that are experimenting with CPB approaches.

16 This section benefitted by contributions from Daniel Violette, who was an advisor/contractor to ACEEE for this study. Some of the material in this section also appears in Violette and Rathbun 2014.

17 SEE Action illustrates this "commonly done" baseline using an appliance example. "For example, if the program involves incenting consumers to buy high-efficiency refrigerators that use 20% less energy than the minimum requirements for ENERGY STAR® refrigerators, the common practice baseline would be refrigerators that consumers typically buy. This might be non-ENERGY STAR refrigerators, or ENERGY STAR refrigerators, or, on average, something in between."

BACKGROUND

In general, the CPB approach uses available market information on common practices to develop an estimate of baseline consumption. This information is then used, with limited adjustments, based on ex post data and analysis, to develop savings estimates. Part of the confusion likely stems from the use of conceptually similar approaches known by different names—including current practice baseline (also CPB), ISP, and BAU.

The energy efficiency field has a long history of looking at market information on equipment purchases and operations. This information provides insights into which efficiency measures a program should include and helps assess appropriate programmatic incentives. In addition, the information helps in the development of both initial savings estimates found in TRMs and program implementation tracking software. Much of the literature on common and industry standard practices addresses these applications, and this research will continue to be important in developing cost-effective programs regardless of whether CPBs are used directly in ex ante program evaluation.

The CPB method is relatively new in the broader evaluation literature as an approach for estimating net savings,¹⁸ and its application has been somewhat limited; however the NW Council has applied a variant of this method for several years in estimating ex ante net savings.¹⁹ The NW Council continues to evolve this approach with new protocols developed by the RTF (RTF 2015).

Ridge et al. (2013) indicate that, in addition to the NW Council, Indiana and Delaware have included CPB in their state evaluation guidelines as approaches that could be used to address net savings. For example, the Evaluation Framework for Delaware (2015) states: "Net savings can also be calculated directly without Free Ridership adjustments by using experimental design evaluations, quasi-experimental design evaluations, or by setting the energy impact baselines to include Free Ridership considerations, such as when the baseline is set at market standard practice."

In this framework, market standard practice is viewed as conceptually the same as CBPs and is one among several approaches for estimating net savings. While the State Evaluation Frameworks for Indiana and Delaware include CPB approaches as one estimation option among others, it is not clear how often these have actually been used.

¹⁸ One of the evaluation experts indicated that "everyone uses common practice baseline" concepts as they go through the process of evaluating appropriate baselines for TRMs, deemed savings, and in program design in terms screening of measures that are viewed as cost-effective and eligible for financial incentives.

¹⁹ Tom Eckman of the NW Council indicated that this general approach has been applied in setting deemed savings since the 1980s, and it was designed to fit with the NW Council's integrated planning process; that is, it is meant to provide an estimate of the increment of savings beyond what system planners assume for naturally (or currently) occurring efficiency in their demand models. Additional information can be found at the RTF website of the NW Council and in RTF (2012), as well as in the roadmap for assessing energy efficiency measures (RTF 2015).

CPB FOR ESTIMATING NET SAVINGS: CONTEXT AND DISCUSSION

In the context of estimating net savings, the CBP approach is designed to directly assess the savings attributable to energy efficiency program activities.²⁰ One advantage claimed for CPB approaches is that they can help avoid the double counting of free riders. The concern is that the two-step approach—in which (1) gross savings is estimated using a baseline that may be similar to common practice, and (2) an NTG ratio is applied to this savings estimate—can double count at least some free riders (Ridge et al. 2013; Hall, Ladd, and Khawaja 2013). The argument presented in these papers is that the process of producing ex ante estimates of “gross” savings may end up being closer to net savings without any adjustment for NTG factors such as free ridership, spillover, and market effects. This view assumes that some NTG factors are already accounted for by the process used to produce the “gross” savings estimates. This emphasizes the need to (1) understand the derivation of gross estimates as part of the energy efficiency evaluation process and (2) explicitly set out the assumed counterfactual scenario in both the gross savings and net savings methods used.²¹ Taking these two steps helps prevent the double counting that results in higher-than-appropriate free ridership estimates.²²

The NW Council, through its RTF, has the longest history of using a CBP approach. The RTF assumes that its current practice baseline directly defines the conditions that would prevail in the program’s absence (the counterfactual scenario) as dictated by the “typical choices of eligible end users, as dictated by codes and standards and the current practices of the market.” (RTF 2015, 3) The RTF estimates this baseline based on recent choices of eligible end users in purchasing new equipment and services. These choices may be inferred from data on shipments, stocking of equipment,²³ purchases (equipment or services), or selected design/construction features.

For example, the baseline for more-efficient televisions is the average efficiency of recent television shipments. These baselines, along with the measure unit energy savings, are subject to a sunset date. The sunset date is “shortened as needed to reliably estimate savings for a measure whose baseline is rapidly changing” (p. 10). The RTF sets out indicators used to determine if current practice is the appropriate baseline. However, “as a general rule, the RTF will use a baseline that is characterized by current market practice or the minimum requirements of applicable codes or standards, whichever is more efficient” (p. 10).

²⁰ Violette and Rathbun (2014) present definitions of gross and net savings, as well as a review of industry practices for estimating net savings.

²¹ It is important to remember that gross savings and net savings are difference estimates and both need a baseline for estimation (see NEEP, 2016).

²² Some experts indicated that this double-counting problem may be the result of inconsistent program rules as set out by the program administrators and regulators, and is not an estimation issue. If this is the case, evaluators still must decide whether the ex ante savings are net, gross, or somewhere between, because the ex post estimates must be used in an internally consistent way to adjust the claimed ex ante savings.

²³ Searching the websites of equipment providers to see which equipment is available and stocked by suppliers is an increasingly common process.

The case for using CBPs appears to derive from two issues. First, the definition of gross savings may include factors that are more appropriately viewed as components of net savings, and additional adjustments to these original estimates are not needed.²⁴ Net savings are considered to be the reduction in energy use resulting from the program-induced change to more efficient technologies.²⁵

Second, program evaluations that report net savings may do so inconsistently. Unfortunately, the components of the net savings calculation differ among jurisdictions and are often based on what the jurisdiction's stakeholders view as appropriate and measurable (see NEEP 2012). Although spillover is widely recognized and can be significant, a number of jurisdictions resist estimating spillover values and including them in net savings calculations. Market effects values have faced similar challenges.

Through its RTF, the NW Council has used CBPs for energy savings more consistently and longer than any other region or jurisdiction. Much of the RTF work is regional, which can help define appropriate markets for both residential and nonresidential appliances and equipment. In addition, regional organizations in the Northwest (e.g., the Bonneville Power Authority and NEEA) conduct market characterization studies for important energy-using equipment, providing information that the RTF can use to develop these baselines. Finally, the RTF supports the NW Council in developing a regional power plan every five years. The RTF's use of the energy efficiency baselines is designed to be consistent with the assumptions used in the most recent power plan. The RTF has designed these savings estimation methods to take into account the regional planning context and regional needs.²⁶

Some entities in the Pacific Northwest use other methods discussed in this section to estimate net savings. For example, the Energy Trust of Oregon (ETO), a voting member of RTF, evaluated a smart thermostat energy efficiency program using a regression-based random encouragement design (RED) approach, along with matching designs to estimate net savings (Energy Trust of Oregon 2016). In addition, the ETO has contributed to the literature and the development and application of self-report survey approaches for estimating net savings (see Castor 2012).

²⁴ This CPB method uses an ex ante approach to estimating net savings as the current practice baseline is typically developed before a program implementation and before the program influences current practices.

²⁵ Tom Eckman of the NW Council expands on this point (See Violette and Rathbun 2014), noting that this approach assumes that "What is occurring prior to program launch is a better measure of what would have occurred absent the program (that is, the counterfactual scenario) than a determination made after the program has influenced the market." Essentially, the NW Council performed an ex ante net analysis when it developed deemed savings estimates that are by design viewed as net savings. For the NW Council's purposes, this is viewed as being as accurate as performing complex studies after the program has been implemented. More information on the NW Council approach can be found in RTF (2012) and at the RTF website at rtf.nwcouncil.org.

²⁶ RTF guidelines (2015) state "The terms 'net' or 'gross' are intentionally not used to modify the term 'savings,' as they may conflict with the definition of baseline" in these guidelines. A presentation by Jennifer Light, RTF Chair, at the April 2017 Forum meeting offers a context for the current practice baselines (Light 2017).

Determining whether a CBP approach provides appropriate savings estimates may depend on a jurisdiction's viewpoint and how the estimates are used within that jurisdiction. When used as part of a five-year regional planning process, as with the NW Council, one viewpoint might emphasize the estimation of energy savings across a five- or ten-year period of program activity. With this perspective, CBPs that are reestimated periodically (as in the RTF) may reflect broad market changes over time. CBPs change over time and are, in part, influenced by the ongoing energy efficiency efforts across all the years of program activity. An alternate view might be applied when looking at incremental resource investments. Energy efficiency investments designed to offset other transmission, distribution, and generation investments focus on incremental or attributable energy savings.²⁷ In this case, the fact that past energy efficiency programs may have changed the current baseline represents sunk costs and should not be considered in economic assessments of new programs; only the savings that are incremental and attributable to that year's energy efficiency investments should be used.²⁸ This illustrates how different jurisdictions may have different needs and uses for savings estimates and how these needs and uses can drive the choice of approach (See NEEP 2016).

Self-selection bias is a significant concern with CBPs.²⁹ The average action taken in a current market may not be representative of the customers who participate in a specific program. A CBP will include a range of equipment with different levels of efficiency. A program that allows consumers to select themselves in may attract customers who would have been predisposed to select the promoted high-efficiency equipment; in that case, applying a CBP might overestimate net savings by not accounting for the unique characteristics of those customers.

Conversely, if customers predisposed to implement the high-efficiency measure don't bother with the program, and those who participate are truly motivated by the incentives and/or other program features, then CPB may underestimate the program's true savings. Specific elements of program design can heavily influence which of those tendencies prevail. Additionally, if the program results in nonparticipant spillover, it is not clear how the CBP approach would capture those savings.³⁰

CONSTRUCTING CPBS

We can describe the theory underlying the definition and the pros/cons of CBPs; however there still is the task of developing these baselines. Developing CBPs for all of the individual

²⁷ Non-wire alternatives are an example of other T&D investments.

²⁸ While this perspective may be appropriate for estimating incremental impacts for near-term resource planning purposes, the value of market efficiency improvements over time produced by the ongoing energy efficiency program effects should be recognized in policy and regulatory decision making.

²⁹ Reviews and discussion of procedures for addressing self-selection in energy efficiency evaluation can be found in Goldberg et al. 2017 and Randazzo et al. 2017.

³⁰ This will not be an issue in applications where market-wide sales data are available on standard and energy-efficient equipment, but these data are unavailable in many markets targeted by energy efficiency programs.

measures included in a portfolio of residential and nonresidential programs can be a daunting assignment. The NW Council's RTF built up its library of measure protocols over several years.³¹ In addition, the data and information needed for these multiple CBPs can be challenging to develop if a framework and approach have not been developed over time.

SEE Action 2012 indicates that appropriate CBPs can be estimated through surveys of participants and nonparticipants, as well as analysis of market data. Discussions with the RTF indicate that it often scans the websites of equipment providers to see what types of equipment are currently for sale online. In addition, supporting studies characterize the markets for energy-using equipment undertaken by other regional entities. A CBP should be based on current equipment sales and not on the stock of equipment installed. Equipment sales represent the current choices of equipment for customers. Sales data can be tough to come by and, even if available, may reflect only parts of the market. Access to sales data also varies by jurisdiction; those jurisdictions that have developed strong connections to equipment manufacturers, suppliers, and trade allies through long-standing energy efficiency programs are likely to have better access.

When possible, the baseline should be substantiated by actual sales data from retailers and installers, rather than from surveys and anecdotal information.³² Considerations should include the following:

1. How much data is required to set a current practice baseline?
 - a. Will additional market research and/or studies be required to set current practice baselines?
 - b. How will minimum required confidence levels be determined?
2. Do the CBPs need market segments – by business type, application, or region?
Regional variation was a significant issue for one IOU in the Northwest, as pricing and availability were viewed as varying across the service territory.

Interviewees made comments on issues related to collecting market information, indicating that developing CBP estimates may be easier for some markets than others. For some markets, the diversity of available equipment and end-uses can pose challenges, e.g., the commercial and industrial (C&I) lighting market is diverse in terms of equipment choices, controls, and end-uses served. On the other hand, rebate programs for specified equipment may allow CPBs to be more easily defined. Program complexity may also be an issue. A custom C&I program can span several end-use applications, technology choices, and changes in operations to achieve efficiencies, which can make defining CPB baselines challenging.

³¹ The list of measures currently addressed by the RTF is available at rtf.nwcouncil.org/measures.

³² Comment from Puget Sound Energy on the proposed baselines for nonresidential lighting applications (2016).

Experts in states actively looking at CPB approaches mentioned other challenges, including:

- Setting up the processes by which CPBs are estimated, updated, and interpreted
- Some subjectivity in determining what is, in fact, a common business practice at a given point of time.³³

Although similar issues exist in other evaluation approaches for estimating programmatic energy savings, here the issues must be addressed to assess how much information is needed to characterize a CBP for a complex market or program.

CPB SUMMARY AND TRENDS

CPBs have pros and cons. The decision to use this approach for certain measure categories and programs will depend in part on the jurisdiction's view of its evaluation needs and the expected uses of the energy savings estimates.

Ridge et al. (2013) make the point that previous energy efficiency programs have affected current markets for energy efficiency equipment through spillover and market effects. This results in current CBPs that are more efficient than they would have been if these past energy efficiency programs were not offered. The effect of these past programs is to reduce the annual energy use of the measures that constitute the current practice. This results in an estimate of current programs' energy savings that is lower than would have been the case without programs in prior years.

This challenge seems to be partly analytical and partly a policy consideration. Ideally, past evaluations of energy efficiency programs should have included all of the impacts attributable to the programs, but because spillover and market effects were generally omitted from past evaluations, they have not been counted. The annual energy savings from today's programs are less than would have been the case if these past programs had not been offered. From this perspective, the use of CPBs as estimates of net savings seems to be an effort to make up for the omission of spillover and market effects in past evaluations that impact the overall market. As a result, savings that are accruing today resulting from previous programs may need to be accounted for when looking at overall energy savings over a period of time. In this context, savings from current programs combined with market effects from past programs that are embedded in the CBP comprise a reasonable estimate of energy efficiency program impacts over the long term; this approach best represents the estimate of the overall return on investments in energy efficiency.

However, if an investment or resource acquisition view is more appropriate, each efficiency program should be evaluated as an incremental investment. That is, a program implemented in 2017 should be evaluated against what is attributable to that investment

³³ Examples of where subjectivity might be increased included (1) industries with just a few large customers and/or suppliers within a geographic area, where those few customers may not share any true common practice, and (2) a market with many customers and many suppliers, along with a diverse set of end-use applications, that challenges the definition of common practice(s) across the market.

only – all impacts from prior years’ programs are essentially sunk costs and should not be considered. In such cases, policy and analytic views of net savings estimation are linked.

Clearly, opinions about CPB use vary widely among evaluation professionals.

Appropriately understanding and accounting for the cumulative effect of sustained energy efficiency programs over years is a substantial, ongoing challenge for the evaluation industry.

The bottom line for assessing the CBP approach is the same process that is used in all other methods: (1) understand the construction of the baseline used in the evaluation; and (2) analyze the implications of this baseline against an appropriate counterfactual scenario for that program. The potential uncertainty and magnitude of bias need to be at least subjectively assessed. Based on this standard approach, decisions can be made about the estimation methods most appropriate for evaluating an energy efficiency program, taking into account jurisdictional priorities and needs.

If an evaluator encounters a jurisdiction that uses a CPB method and refers to the savings as “net savings,” that evaluator should proceed in an internally consistent manner. For example, the evaluator should explain how the utility, agency, or regional body is defining attributable or net savings. In addition, if there is a gross savings construct in the evaluation approach for any programs, the evaluator should define the differences between the gross and net savings values.

In summary, the Pacific Northwest has been using the CPB approach for more than a decade. In addition, several jurisdictions are assessing the approach to see if it might make sense for some of their programs or in certain contexts, such as overall return from energy efficiency investments across a number of years.

Evaluation experts interviewed indicated that a potential advantage of CPB is its clarity, particularly for certain program types. The CPB approach’s objective is to develop a CBP against which participants’ energy use is measured. This clearly sets out the goals of the evaluation process. The questions are how this estimate of savings should be interpreted and how the CBP should be developed. Developing CBPs must be internally consistent across a defined market, which may have a distribution of different types of technologies applicable to different uses (e.g., consider the lighting sector and the diverse uses of lighting equipment).

A potential strength of the CBP approach is its use in upstream and market transformation energy efficiency programs. It can be applied market wide and, unlike randomized trials and quasi-experimental designs, it does not require participants to be identified (assuming appropriate sales data are available).

A limitation of CPB approaches is that the method is susceptible to self-selection bias (that is, the average consumer may not be the type of consumer who participates in the program). It is not clear how this can be addressed, other than by conducting surveys to determine specific characteristics of purchasers of efficient equipment relative to the CBP. However this survey effort would negate the unique aspects claimed for the CBP approach – i.e.,

specific consumers who have and have not purchased the high-efficiency equipment would need to be identified.

In the last few years, Massachusetts (ERS and DNV-GL, 2017) and California (CPUC 2014; Xu and Friedmann, 2016) have initiated discussions on the role of common practice or industry standard baselines in evaluation, but these are initial efforts with no consensus yet on how they might be applied to estimate net savings.

Based on our review, it appears that the CPB approach is most appropriate for programs promoting specific types of equipment for which sales data are available. At the other end of the spectrum, it will be very difficult to apply the CPB to comprehensive retrofit programs that install many different measures at individual sites.

M&V 2.0

M&V 2.0 is the term used to describe the opportunity for new M&V techniques through enhanced data analytics and increased data availability. Hardware technology drives this increased data availability through new advances and innovations, and increased adoption of existing hardware technologies such as smart meters, smart thermostats, and home and building energy management systems. Smart meters generate energy consumption data at more frequent intervals (e.g., every 5 or 15 minutes) than traditional monthly meter readings. Another example is nonintrusive load monitoring (NILM), which, when combined with software, can disaggregate loads and provide more end-use data (Eckman and Sylvia 2014). The result of these evolving data tools is the increased access, availability, timeliness, and granularity of energy consumption data.

Increased available data drive demand for new software to analyze them, as well as business opportunities for data analytics vendors to enter the market. Cloud computing enables these larger volumes of data to be processed and turned into useful information faster. M&V can use one or more methods and can involve metering measurements in combination with engineering calculations, statistical analysis, and/or computer simulation modeling. It is a way to look at data more closely and uncover previously hidden details. As one expert put it, “M&V 2.0 is billing analysis at scale, using a software solution.”

A growing body of research explores M&V 2.0. For example, in its *The Changing EM&V Paradigm* report, NEEP outlines M&V 2.0 trends across two major areas: (1) advanced data analytics and program enhancements (enabled by new software) and (2) advanced data availability (enabled by new hardware) (NEEP 2015). ACEEE also began to explore this topic in 2015 (Rogers et al. 2015). Recently, Rocky Mountain Institute (RMI) and Lawrence Berkeley National Laboratory (LBNL) provided a helpful overview of M&V 2.0 tools, methods, and applications (Franconi et al. 2017). An ongoing LBNL stakeholder working group also focuses on these topics.

When describing M&V 2.0 opportunities, stakeholders are increasingly acknowledging the importance of distinguishing its role; it is not the “E” in EM&V (evaluation), but rather one type of savings estimation approach for M&V applications. The evaluation part includes a broader set of activities, including attribution and baselines, as well as process evaluations, which are not directly obtainable through software solutions. M&V is narrower and is

typically applied at the project level to determine gross energy savings at individual sites or projects.

M&V 2.0 tools offer a range of capabilities valuable to program designers and implementers, from customer intelligence to end-use load disaggregation. M&V 2.0 is one tool in the toolbox within savings estimation approaches, but it may be more often used and valuable on the program side, e.g., for use in program monitoring. Timelier data and increased granularity mean that M&V results can surface more quickly, and this faster feedback means more-responsive program implementation. The increasing availability of more-granular data could also have significant implications for calculation of demand impacts, based on multiple definitions of peak.

M&V 2.0 is not without challenges. Two key challenges are dealing with non-routine adjustments, and addressing situations in which past performance is not the appropriate baseline (Gruendling et al. 2017). Another set of issues concerns the privacy of individual customer data: as more and more detailed data on energy use in homes and businesses become available, the considerations for how and when data may be accessed become increasingly important (ACEEE 2014).

RECENT TRENDS

Over the past year, the initial reaction to M&V 2.0 by the broader EM&V community has evolved from skepticism and a feeling of being threatened to new partnerships on EM&V. Utilities are increasingly learning how to use these tools to better understand customer energy usage and program participation and to obtain more timely information on how to better serve customers.

We find that M&V 2.0 is viewed as a positive trend, but that states and utilities do not have enough information yet to know the full range of its impacts and opportunities. Also, definitional uncertainty remains, with various stakeholders interpreting the phrase differently and in some cases preferring not to use it at all. Although new data analytics and data sources offer significant opportunity, we are still in the very earliest stages and there is still regulatory caution. Most of our respondents noted that states and utilities are responding with interest *and* caution. For example, we have yet to see any jurisdiction that has relied on M&V 2.0 to estimate savings claimed by a utility or a third party.

Some also question whether M&V 2.0 is truly a new technique, suggesting that the industry has had similar approaches to monthly billing analysis for years. Many view billing analysis as a blunt tool, with the same challenges that have always complicated it persisting today. What is new and helpful, however, is the precision that comes from collecting and handling more data faster and at more frequent intervals.

Some experts are skeptical that we can ever automate savings estimates. This is because analysts must make numerous individual decisions to assess gross or net savings, including decisions driven by the target data, such as how to treat outliers (data points that are outside the usual deviation that would be statistically predictable and controlling for exogenous factors other than weather). If we remove the human component and automate, we will get an answer – but it may not be the best answer.

Increased available data are driving business opportunities for data analytics vendors to enter the market. While some utilities may be considering in-house M&V 2.0 solutions, they recognize that contracted vendor solutions may be more advantageous because they have server capacity and other technical capabilities to store, process, and write code for increased data volumes. Internal IT departments at utilities handle revenue collection that gets preferential treatment, and energy efficiency departments have to compete for IT attention. Specialist vendors often gain experience from multiple utilities and implementers and can apply this learning.

In summary, M&V 2.0 is still at an early stage, but initial pilots now underway will provide much useful information on paths forward. Still, based on the information that is available, it is likely that M&V 2.0 will not be appropriate for all programs and needs, lending itself more to some applications than others. With experimentation, we can better understand these appropriate and inappropriate applications.

STATE AND UTILITY EXAMPLES OF M&V 2.0

We identified a few specific examples of M&V 2.0, all of which are in the early stages of development.

Connecticut Pilot

Many interviewees pointed to Connecticut as a leader in this area, notably for embarking on a new three-year pilot on commercial building M&V 2.0. The Connecticut Department of Energy and Environmental Protection (DEEP) is facilitating the pilot project in partnership with LBNL, Eversource, United Illuminating, NEEP, and the Department of Energy (DOE), with the latter providing financial support. The goal is to assess the value of M&V 2.0 approaches for understanding energy savings and for use as evaluation tools, starting with commercial buildings. An additional goal is to analyze the savings estimates derived from the tool compared to results from traditional engineering analyses. With data from the utilities, the project team is now trying to select buildings and measures to test an LBNL modeling tool. As of spring 2017, it was in the early stages of identifying specific projects.

Pacific Gas & Electric's Pay for Performance Program

PG&E will be launching a new pay for performance (P4P) program by late 2017. The utility's current home upgrade program (HUP) and Advanced HUP have seen high costs and low cost-effectiveness, which are an impetus for the utility to move to a P4P approach.³⁴ The goals with the new approach are to encourage contractors to go deeper into each home to find additional savings (i.e., look for opportunities to install cost-effective measures beyond the standard program offerings). This program is an outcome of the previously mentioned AB 802 legislation. IOUs were invited to submit proposals for expedited review for the High Opportunity Projects and Programs. The utility submitted an Advice Letter to the CPUC for

³⁴ For a review of P4P programs and business models, see NRDC 2017. The report finds that P4P has important opportunities for increasing energy savings, but also key limitations that will need to be better understood through piloting and experimentation.

consideration, providing information about the budget, program plan, and detailed EM&V plan.³⁵

To date, PG&E has completed a competitive bidding process for implementers, vetted multiple companies that which submitted proposals, and selected a couple of firms to move forward. A key element of the bids was the proposed levelized \$/kWh and \$/therm payments potential aggregators would need. Billing analysis of AMI data will be used with the weather-normalized pre/post difference serving as a basis for making incentive payments. PG&E is committing to provide multiple performance-based payments beginning one year after program intervention. Given that weather-normalized pre/post billing analysis without control or comparison groups does not address exogenous factors, there is a risk to PG&E that the evaluation results will differ from the paid savings. However PG&E believes that providing the market with consistent and straightforward pre/post billing analysis methods for determining incentive payments is essential to give future savings aggregators sufficient certainty.

The savings on which incentive payments are based will be calculated using CalTRACK, a protocol being developed by a technical working group to standardize a set of billing analysis methods.³⁶ Calculations in CalTRACK are powered by OpenEEMeter, an open-source software platform. The objective is to provide transparency in the M&V savings estimation process by having open-source codes that anyone can use and, in theory, arrive at similar estimates of savings. If CalTRACK can meet that objective, the goal is for energy efficiency savings to be perceived as less risky by both system operators and investors.

Southern California Edison's Preferred Resources Pilot

SCE's Preferred Resources Pilot (PRP) is a multiyear study that will assess whether clean energy resources can offset increasing customer electricity demand in the central region of Orange County, California, which is home to approximately 204,000 residential customers and 30,000 commercial and industrial customers (SCE 2017). The region faces growing demand due in part to the closure of the San Onofre Nuclear Generating Station and the impending retirement of other nearby plants. Depending on the pilot's results, the utility may be able to avoid or defer the need for new gas-powered plants.

SCE is considering clean energy resources such as solar, wind, energy efficiency, and energy conservation. The PRP project creates opportunities to consider new energy efficiency EM&V approaches to estimate savings performance and persistence. SCE is interested in estimating energy efficiency savings using a bottom-up approach—that is, at the individual customer level—that could also be discernable at the grid level within the PRP area.

The utility contracted with an M&V 2.0 firm to analyze energy usage and to detect significant and sustained energy use changes in a sample of 62 commercial building customers that installed energy efficiency measures (EEMs) (SCE 2016). The efficiency

³⁵ See March 25, 2016, "Submission of High Opportunity Projects and Programs (HOPPs) Proposal: Residential Pay-for-Performance Program" at www.pge.com/tariffs/tm2/pdf/GAS_3698-G.pdf.

³⁶ www.caltrack.org/about.html

impact study classified building energy use changes into two categories: (1) *recorded changes*, which refer to the known date of installation of an EEM and the estimated ex ante savings of the measures, and (2) *detected changes*, which are sharp, sustained changes identified through algorithms and caused by other factors such as occupancy or business activity changes within the building. (Detected changes are sometimes referred to as *non-routine events* or *anomalies*.) Detected changes can potentially mask or amplify energy efficiency program impacts.

The analysis found that 54 of the customers (87%) had energy use patterns or baselines that could be modeled with statistical confidence using machine learning algorithms (computer programs that use iterative calculations to learn how to make more accurate predictions). Of these customers, offices, retail stores, and schools appear to be least affected by behavior and other building use changes aside from anticipated EEMs. These customers showed the highest level of savings at the meter.

For example, “Building 1” had LED lighting installed and showed a reduction of approximately 42% compared to the business-as-usual prediction, with savings sustained throughout the evaluation period (as shown in figure 1). On the other hand, 30 of the 54 customers appear to have behavior and other usage changes that impact their net energy savings calculated from the customer’s meter (in addition to the anticipated EEM changes). This does not mean that the buildings had lower savings from energy efficiency. However it shows that occupancy and other changes can mask the savings from energy efficiency when looking for impacts at the meter alone.

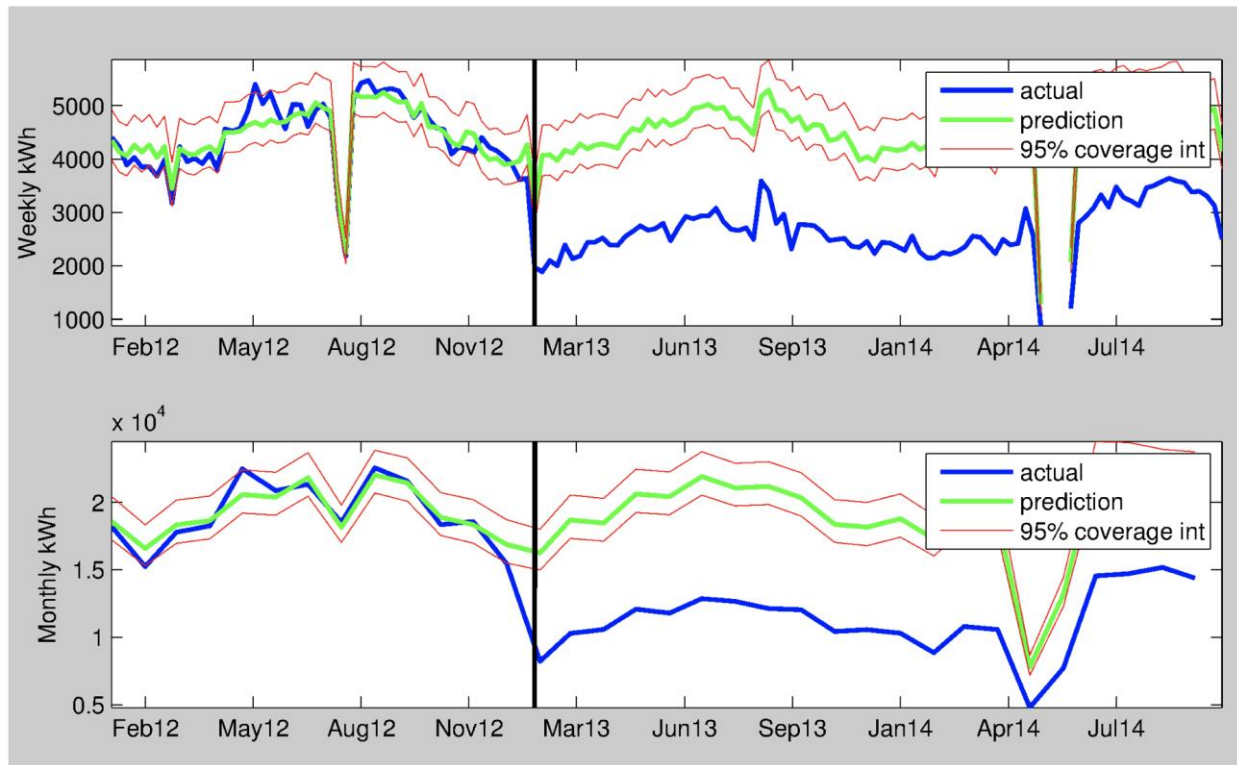


Figure 1. PRP “Building 1” weekly (top) and monthly (bottom) actual use (blue), along with BAU predictions (green), confidence intervals, and change events (pink). *Source:* SCE 2016.

The utility also contracted with a traditional EM&V firm which found similar savings estimate results, where a comparison was possible (Shahana Samiullah, senior manager, Customer Insights and Measurement & Evaluation, Southern California Edison, pers. comm., May 9, 2017). Although the utility remains interested in using M&V 2.0 tools, it is not yet sure if it is a permanent third-party solution. This is partially due to the costs and partially because California stakeholders are still in learning mode regarding the new state policy of going to meter-based savings.

PG&E Commercial Whole Building Performance Demonstration Project

Another example is PG&E's Commercial Whole Building (CWB) Performance project, which is a performance-based approach designed to deliver at least 15% post-installation energy savings in existing commercial buildings. PG&E's CWB demonstration project is extending the use of meter-based measurement for creating projections of energy use to establish a counterfactual baseline for use in determining performance incentives.

Given the more heterogeneous nature of energy consumption in the commercial and industrial spheres, CWB relies on documenting routine and non-routine events for adjusting observed performance for changes in premise use that affect energy consumption. The International Performance Measurement and Verification Protocol (IPMVP) Option C protocol is used as the basis for this approach.

PG&E is targeting buildings without significant process loads and with stable end uses, such as large and medium-sized commercial office buildings, supermarkets, and public buildings such as schools and libraries. The first year of performance has been completed and the impact evaluation is underway. As of this writing, the first of 12 buildings in the demonstration project has completed the post-reporting period (Gruendling et al. 2017). The project is using multiple models to project counterfactual baselines, including two open-source IPMVP Option C models, two proprietary Option C models, and an ensemble model developed by PG&E data scientists that blends estimates of the Option C models. The project is also using IPMVP Option D (building simulation) to ascertain when less-expensive Option C-based estimates yield comparable estimates. Similar to the residential example, an overall goal is to move away from program designs that require state regulators to pre-approve the unit energy savings (gross) for every measure. In addition, PG&E is determining how the program could be expanded to serve more customers.

Challenges Going Forward

Challenges inevitably arise as EM&V evolves. When asked about the one or two biggest challenges facing EM&V today, our interviewees identified several key issues. (See table A2 in Appendix A for the compiled results of these issue areas.)

Keeping EM&V relevant in a changing world. Efficiency may be becoming a victim of its own success. After decades of policies and programs to promote energy efficiency, it becomes difficult to isolate net impacts. As a result, it becomes even more important looking ahead to focus on metrics such as market impacts and, in some cases, to move beyond old ways of assessing attribution. For example, we need to weigh the value of attribution studies with their costs.

Using new data sources effectively. AMI meters, smart thermostats, nonobtrusive load monitoring, and building energy management systems can supply vast amounts of data. The challenge is to identify the most appropriate and best uses of these data and to decide how to efficiently process and analyze this information.

Using M&V 2.0 tools. M&V 2.0 gives rise to a range of possible applications. How do programs assess their value and select those tools that best accomplish different objectives?

Making evaluation results and language more accessible, impactful, timely, and relevant. We need to more clearly communicate high-level M&V and EM&V results to policymakers and stakeholders without using too much technocratic jargon. This extends not only to EM&V results, but also to the need for robust evaluation itself. We need faster program performance results, rather than results 1–2 years after a program ends. As one interviewee put it, “the current status of performance results is too late, too little, too expensive.”

Budget challenges. Among these challenges are insufficient EM&V budgets and the constant pressure to reduce EM&V spending. Research has shown that utilities spend about 3–5% of efficiency portfolios on EM&V.³⁷ However it is worth noting that, while the percentages are within a similar range, the absolute amounts vary significantly due to the wide differences in portfolio budgets. One expert commented that, with economies of scale, larger states and utilities should be able to handle a percentage at the lower end of the 1–6% range, while states just getting started should have a much higher percentage, perhaps 10%.

Regulatory perceptions. Some regulators still suspect that energy efficiency does not produce real savings, while other regulators suspect that utilities inflate their savings estimates.

Political changes. Conservative policy trends in particular pose a challenge for EM&V institutions, focus, and expenditures at the state level and potentially at the federal level.

Attribution issues. Among the many issues raised in this area was a concern that utility customers might not accurately recall their decisions because too much time may elapse between their making decisions and their taking the self-report evaluation surveys. Various protocols govern the use of self-reports to estimate NTG ratios (e.g., Ridge et al. 2007). A key challenge is ensuring that evaluators comply with these protocols.

Insufficient EM&V independence or oversight. In past research, we have found that most jurisdictions required independent evaluation of some type, but the particular structures and approaches to achieve this varied greatly (Kushler, Nowak, and Witte 2012). A related challenge is how to properly define the role of evaluator. For example, how should independence be defined? What level of oversight by regulators and/or advisory councils should be required?

Increasing consolidation. The evaluation industry has become increasingly concentrated into fewer larger firms, which could negatively impact work quality. When more entry-level

³⁷ Based on data from the [Consortium for Energy Efficiency](#).

employees take on increasing responsibilities, the potential for lower-quality work increases as well. This industry concentration also creates a challenge for some utilities to find a vendor without conflicting interests. PUCs require utilities or program administrators to hire independent evaluators. Evaluators are also hired by administrators and implementers for internal quality control functions. On the positive side, larger evaluation contractors can develop experience and expertise beyond the capacity of smaller firms.

This is far from an exhaustive list of challenges in the EM&V field. LBNL recently prepared an overview of EM&V activities relevant to energy efficiency and distributed energy resources (Schwartz et al. 2017). In that overview, LBNL identifies 11 specific research gaps that require more readily available data to both assess impacts of energy efficiency and support investments. As listed in the report, these gaps include the following general categories:

- Reliability and certainty of evaluated impacts
- Input data access and availability needs
- Consistent reporting and program typologies
- Timeliness of EM&V reporting and utilization
- EM&V factors: attribution of savings, measure lifetime and persistence of savings, and rebound
- EM&V practitioner training, certification, and independence
- Opportunities for further development of EM&V methods: deemed savings, randomized control trials, EM&V 2.0, and top-down evaluation
- EM&V for transmission and distribution (T&D) system efficiency
- EM&V for codes and standards
- EM&V for financing programs
- EM&V for nonenergy impacts

This list of gaps provides a roadmap for the evaluation field as it continues to evolve.

Conclusions

As policymakers and utility system planners continue to see energy efficiency's importance as a resource, the value of—and demand for—accurate, reliable, and consistent EM&V must grow as well. We are seeing an increase in efficiency's eligibility for use as an energy resource comparable to other resources, as well as a growing recognition of its value by parties not historically familiar with efficiency. These developments require rigorous EM&V to assure the resource is valid and the impacts are appropriately valued.

Many jurisdictions have been passing legislation and adopting regulatory policies to both seize these new opportunities and tackle challenges to further expand the efficiency resource. The leaders invest heavily in program evaluation, producing dozens of evaluation studies each year.

For example, more states and regions are adopting TRMs and enhancing the accuracy and consistency of the deemed savings values within them. Transparent, collaborative multi-stakeholder processes for updating efficiency measures and calibrating savings values

contribute to these improvements. Jurisdictions are increasingly using the results from program evaluations to update TRMs in a cycle of continuous improvement.

Baseline methods and their application have implications for how EM&V meets the overarching challenge of staying relevant in a changing world, particularly in states that have been pursuing efficiency the longest. Although the CBP approach gained heightened attention in the evaluation community when it was discussed in the EPA's draft CPP guidance on energy efficiency and has become well-established in the Northwest by the NW Council, it is not widely used elsewhere. CPB is not equivalent to M&V 2.0 and TRMs as a dynamic area of innovation and rapid change in the evaluation landscape. More recently, Massachusetts and California have been experimenting with a variation that they call the *industry standard practice* (ISP) method, though they have yet to reach consensus on how ISPs might be applied to estimate net savings, and these efforts are in the preliminary stages.

The concurrent proliferation of advanced hardware and software technologies for collecting and analyzing energy consumption data has led to innovations designed to meet some of the demands on EM&V. New energy data analytics start-up companies, energy efficiency program designs, evaluation research approaches, and policy processes are emerging in response to these challenges.

Utilities and other stakeholders are examining how they can use advanced M&V tools to better understand and engage customers, target program marketing more effectively, and be more nimble and responsive in program implementation using faster feedback on program performance. The evaluation community has a growing set of tools in its toolbox and utilities have new partners for program implementation. As regulators receive evaluation results, it is increasingly important that the evaluation community present a unified message on the reliability of energy efficiency impacts.

For states that are in the early stages of developing energy efficiency programs, or for those that are ramping up, it will be important to ensure comprehensive policy frameworks and allocate sufficient resources for EM&V. Tapping into emerging national and regional resources can help, but will not be sufficient. Emerging states with smaller budgets may need to allocate higher percentages of energy efficiency portfolio costs to EM&V to keep up with best practices.

Our review finds that many states have been adopting regulatory policies and experimenting with strategies to both seize these new opportunities and tackle challenges to further expand the many benefits of the energy efficiency resource.

References

- ACEEE. 2014. *Best Practices for Working with Utilities to Improve Access to Energy Usage Data*. Washington, DC: ACEEE. aceee.org/files/pdf/toolkit/utility-data-access.pdf.
- Beitel, A., T. Melloch, B. Harley, and A. Mejia. 2016. "Technical Reference Manuals: Best Practices from Across the Nation to Inform the Creation of the California Electronic Technical Reference Manual (eTRM)." In *Proceedings of the ACEEE 2016 Summer Study on Energy Efficiency in Buildings* 6: 1–14. Washington, DC: ACEEE. aceee.org/files/proceedings/2016/data/papers/6_1027.pdf.
- California Assembly. 2015. Assembly Bill No. 802. www.leginfo.ca.gov/pub/15-16/bill/asm/ab_0801-0850/ab_802_bill_20150916_enrolled.pdf.
- California Public Utilities Commission. 2001. *The California Standard Practice Manual: Economic Analysis of Demand-Side Programs and Projects*. San Francisco: California Public Utilities Commission. [www.cpuc.ca.gov/uploadedFiles/CPUC_Public_Website/Content/Utilities_and_Industries/Energy - Electricity and Natural Gas/CPUC_STANDARD_PRACTICE_MANUAL.pdf](http://www.cpuc.ca.gov/uploadedFiles/CPUC_Public_Website/Content/Utilities_and_Industries/Energy_-_Electricity_and_Natural_Gas/CPUC_STANDARD_PRACTICE_MANUAL.pdf).
- . 2004. *The California Evaluation Framework*. Prepared by TecMarket Works. San Francisco: California Public Utilities Commission. www.calmac.org/publications/California_Evaluation_Framework_June_2004.pdf.
- . 2006. *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. Prepared by TecMarket Works. San Francisco: California Public Utilities Commission. www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=5212.
- . 2014. *ISP Guide – Industry Standard Practice Guide, Version 1.2A*. San Francisco: California Public Utilities Commission. www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=5315.
- . 2017. *California EM&V Framework Refresh Needs Assessment: Draft Report*. Prepared by Research Into Action and Navigant Consulting. San Francisco: California Public Utilities Commission. pda.energydataweb.com/-/documents/1881/view.
- Castor, S. 2012. *Fast Feedback Results: 2011 Final Report*. Portland: Energy Trust of Oregon. www.energytrust.org/wp-content/uploads/2016/12/Fast_Feedback_-_20110.pdf.
- Delaware Department of Natural Resources and Environmental Control. 2015. *Delaware Evaluation Framework*. Prepared by Optimal Energy and Wirtshafter Associates. Dover: Department of Natural Resources and Environmental Control. [www.dnrec.delaware.gov/energy/information/otherinfo/Documents/EEAC/DRAFT Delaware Evaluation Framework.pdf](http://www.dnrec.delaware.gov/energy/information/otherinfo/Documents/EEAC/DRAFT_Delaware_Evaluation_Framework.pdf).
- Eckman, T., and M. Sylvia. 2014. "EM&V 2.0 – New Tools for Measuring Energy Efficiency Program Savings." *Electric Light & Power*, February 12. www.elp.com/Electric-Light-

Power-Newsletter/articles/2014/02/em-v-2-0-new-tools-for-measuring-energy-efficiency-program-savings.html.

ecobee. 2017. "Donate Your Data: Advancing Science for a Better Planet."
www.ecobee.com/donateyourdata/.

Energy Trust of Oregon. 2016. *Energy Trust of Oregon Smart Thermostat Pilot Evaluation*. Prepared by Apex Analytics. Portland: Energy Trust of Oregon.
[assets.energytrust.org/api/assets/reports/Smart Thermostat Pilot Evaluation-Final_wSR.pdf](http://assets.energytrust.org/api/assets/reports/Smart%20Thermostat%20Pilot%20Evaluation-Final_wSR.pdf).

EPA (Environmental Protection Agency). 2015. *Evaluation, Measurement and Verification Guidance for Demand-Side Energy Efficiency: Draft for Public Input*. Washington, DC: EPA.
www.eebco.org/resources/Documents/CleanPowerPlan/cpp_emv_guidance_for_demand-side_ee_-_080315.pdf.

—. 2016. *Draft EM&V Guidance for Demand-Side Energy Efficiency*. Washington, DC: EPA.
blog.epa.gov/blog/wp-content/uploads/2016/12/EMV-Guidance-12192016.pdf.

—. 2017. "Programmable Thermostats Specification." Accessed September.
www.energystar.gov/index.cfm?c=archives.thermostats_spec.

ERS and DNV-GL. 2017. *Massachusetts Commercial/Industrial Baseline Framework: Final Report*. Boston: MA EEAC (Massachusetts Energy Efficiency Advisory Council). ma-eeac.org/wordpress/wp-content/uploads/MA-Commercial-and-Industrial-Baseline-Framework.pdf.

Franconi, E., M. Gee, M. Goldberg, J. Granderson, T. Guiterman, M. Li, and B. Smith. 2017. *The Status and Promise of Advanced M&V: An Overview of "M&V 2.0" Methods, Tools, and Applications*. Berkeley: Berkeley Lab. eta.lbl.gov/sites/all/files/publications/lbnl-1007125.pdf.

Goldberg, M., G. Agnew, M. Fowle, K. Train, and B. Smith. 2017. "Not Just Another Pretty Formula: Practical Methods for Mitigating Self-Selection Bias in Billing Analysis Regressions." In *Proceedings of the 2017 International Energy Program Evaluation Conference, Baltimore*. Madison, WI: IEPEC. www.iepec.org/2017-proceedings/65243-iepec-1.3717521/miriam-goldberg-1.3717583.html.

Goldberg, M., M. Marean, C. Puckett, C. Godin, W. Todd, S. Bodmann, and K. Kelly. 2015. *The Changing EM&V Paradigm: A Review of Key Trends and New Industry Developments, and Their Implications on Current and Future EM&V Practices*. Prepared by DNV GL. Lexington, MA: Northeast Energy Efficiency Partnerships. www.neep.org/changing-emv-paradigm.

Gruendling, P., P. Jacobs, A. Kanungo, A. Wortman and D. Jump. 2017. "M&V 2.0 and Commercial Whole Building Programs: The Army We Wish We Had." In *Proceedings of the 2017 International Energy Program Evaluation Conference, Baltimore*. Madison, WI: IEPEC. www.iepec.org/2017-

proceedings/polopoly_fs/1.3718096.1502900960!/fileserver/file/796599/filename/074.pdf.

Hall, N., D. Ladd, and M. Khawaja. 2013. "Setting Net Energy Impact Baselines: Building Reliable Evaluation Approaches." In *Proceedings of the 2013 International Energy Program Evaluation Conference, Chicago*. Madison, WI: IEPEC. www.iepec.org/conf-docs/conf-by-year/2013-Chicago/131.pdf - page=1.

Kushler, M., S. Nowak, and P. Witte. 2012. *A National Survey of State Policies and Practices for the Evaluation of Ratepayer-Funded Energy Efficiency Programs*. Washington, DC: ACEEE. aceee.org/research-report/u122.

—. 2014. *Examining the Net Savings Issue: A National Survey of State Policies and Practices in the Evaluation of Ratepayer-Funded Energy Efficiency Programs*. Washington, DC: ACEEE. aceee.org/research-report/u1401.

Light, J. 2017. *Guidelines Updates*. Portland: Northwest Power and Conservation Council. nwcouncil.app.box.com/v/GuidelinesAprilRTFPres.

Mass Save. "Massachusetts Technical Reference Library." www.masssavedata.com/Public/TechnicalReferenceLibrary.

Maxwell, J., S. Haselhorst, and R. Prahl. 2017. "Baseline Policy Enhancement in Massachusetts and California." In *Proceedings of the International Energy Program Evaluation Conference, Baltimore*. Madison, WI: IEPEC. www.ers-inc.com/wp-content/uploads/2017/08/Baseline-Policy-Enhancement-in-MA-and-CA.pdf.

MEEA (Midwest Energy Efficiency Alliance). 2017. *Midwest TRM Inventory: Statewide Technical Resource Manuals for Energy Efficiency*. Chicago: MEEA. www.mwalliance.org/sites/default/files/media/Midwest-TRM-Inventory-Feb2017.pdf.

NEEP (Northeast Energy Efficiency Partnerships). 2012. *Regional Net Savings Research, Phase 2: Definitions and Treatment of Net and Gross Savings in Energy and Environmental Policy*. Prepared by NMR Group and Research Into Action. Lexington, MA: NEEP. www.neep.org/sites/default/files/resources/NEEP-RegionalNetSavingsReport12-05-12.pdf.

—. 2016. *Gross Savings and Net Savings: Principles and Guidance*. Prepared by Navigant Consulting and Tetra Tech. Lexington, MA: NEEP. www.neep.org/gross-and-net-savings-principles-and-guidance.

—. 2017. *Mid-Atlantic Technical Reference Manual V6*. Lexington, MA: NEEP. neep.org/mid-atlantic-technical-reference-manual-v6.

NRDC (Natural Resources Defense Council). 2017. *Putting Your Money Where Your Meter Is: The Opportunity for Pay-for-Performance Efficiency Programs to Increase Energy Savings*. New

York: NRDC. www.nrdc.org/sites/default/files/programs-increase-energy-savings-ib.pdf.

Puget Sound Energy. 2016. *Comments on Baseline Recommendations*. Presentation to Regional Technical Forum meeting.

Randazzo, K., R. Ridge, S. Wayland, and B. Smith. 2017. "Comparison Groups for Whole Building Program Impact Evaluations: They Are Harder and Easier Than You Think." In *Proceedings of the International Energy Program Evaluation Conference, Baltimore*. Madison, WI: IEPEC. www.iepec.org/2017-proceedings/polopoly_fs/1.3718058.1502900902!/fileserver/file/796583/filename/090.pdf.

RTF (Regional Technical Forum). 2015. *Roadmap for the Assessment of Energy Efficiency Measures*. Portland: Northwest Power and Conservation Council. nwcouncil.app.box.com/s/7kec25hhxpalsccdi6woj1s481e659q.

—. 2017. "RTF Library of Unit Energy Savings Measures." rtf.nwcouncil.org/measures.

Ridge, R., M. Baker, N. Hall, R. Pahl, and W. Saxonis. 2013. "Gross Is Gross and Net Is Net: Simple, Right?" In *Proceedings of the 2013 International Energy Program Evaluation Conference, Chicago*. Madison, WI: IEPEC. www.iepec.org/conf-docs/conf-by-year/2013-Chicago/094.pdf - page=1.

Ridge, R., K. Keating, L. Megdal, and N. Hall. 2007. *Guidelines for Estimating Net-to-Gross Ratios Using the Self-Report Approaches*. San Francisco: California Public Utilities Commission.

Rogers, E. E. Carley, S. Deo, and F. Grossberg. 2015. *How Information and Communications Technologies Will Change the Evaluation, Measurement, and Verification of Energy Efficiency Programs*. Washington, DC: ACEEE. aceee.org/research-report/ie1503.

Schiller, S., G. Leventis, T. Eckman, and S. Murphy. 2017. *SEE Action Guide for States: Guidance on Establishing and Maintaining Technical Reference Manuals for Energy Efficiency Measures*. Prepared by Lawrence Berkeley National Laboratory. Washington DC: SEE Action. www4.eere.energy.gov/seeaction/publication/see-action-guide-states-guidance-establishing-and-maintaining-technical-reference.

Schwartz, L., M. Wei, W. Morrow, J. Deason, S. Schiller, G. Leventis, S. Smith, W. Ling Leow, T. Levin, S. Plotkin, Y. Zhou, and J. Teng. 2017. *Electricity End Uses, Energy Efficiency, and Distributed Energy Resources Baseline. Appendix: Evaluation, Measurement, and Verification of Energy Efficiency and Distributed Energy Resource Activities*. Prepared by Berkeley Lab. Washington, DC: Department of Energy. emp.lbl.gov/sites/default/files/lbnl-1006983.pdf.

SEE Action (State and Local Energy Efficiency Action Network). 2012. *Energy Efficiency Program Impact Evaluation Guide*. Prepared by S. Schiller, Schiller Consulting, Inc. Washington, DC: SEE Action.

www4.eere.energy.gov/seeaction/system/files/documents/emv_ee_program_impact_guide_0.pdf.

—. 2015. *Energy Efficiency Collaboratives*. Prepared by M. Li and J. Bryson. Washington, DC: SEE Action.

www4.eere.energy.gov/seeaction/system/files/documents/EECollaboratives-0925final.pdf.

SCE (Southern California Edison). 2016. *Energy Efficiency Impact Study for the Preferred Resources Pilot*. Rosemead: Southern California Edison.

www.sce.com/wps/wcm/connect/5b0de293-4a61-472b-a32b-ed9c2cd6aea2/EEImpactStudy_SCEWhitePaper.pdf?MOD=AJPERES.

—. 2017. "Our Preferred Resources Pilot." [www.sce.com/wps/portal/home/about-us/reliability/meeting-demand/our-preferred-resources-pilot/!ut/p/b1/hc-xDoIwGATgR-](http://www.sce.com/wps/portal/home/about-us/reliability/meeting-demand/our-preferred-resources-pilot/!ut/p/b1/hc-xDoIwGATgR-qVFijTyDlr4lKaiJ2MZ1ME0UH4_MLhIW47ZLvhhNBDCKM8ZNu8Z2eY7zPPRRXxw1IqzO23rSgXX7MqkbCOE7gMgH8CWFrfxbhR6Sx1LEH25POwXVWkluzKn2-)

[qVFijTyDlr4lKaiJ2MZ1ME0UH4_MLhIW47ZLvhhNBDCKM8ZNu8Z2eY7zPPRRXxw1IqzO23rSgXX7MqkbCOE7gMgH8CWFrfxbhR6Sx1LEH25POwXVWkluzKn2-](http://www.sce.com/wps/portal/home/about-us/reliability/meeting-demand/our-preferred-resources-pilot/!ut/p/b1/hc-xDoIwGATgR-qVFijTyDlr4lKaiJ2MZ1ME0UH4_MLhIW47ZLvhhNBDCKM8ZNu8Z2eY7zPPRRXxw1IqzO23rSgXX7MqkbCOE7gMgH8CWFrfxbhR6Sx1LEH25POwXVWkluzKn2-).

Statista. 2017. "Number of Homes with Smart Thermostats in North America from 2014 to 2020 (in Millions)." www.statista.com/statistics/625868/homes-with-smart-thermostats-in-north-america/.

Thornburg, T., and J. Jewel. 2015. *Advancing the Adoption of Smart Thermostats in Northern Illinois*. Chicago: Illinois Energy Efficiency Stakeholder Advisory Group.

ilsagfiles.org/SAG_files/Meeting_Materials/2015/6-23-15_Meeting/Smart_Thermostat_SAG_Presentation_6-23-15_Meeting.pdf.

Violette, D., and P. Rathbun. 2014. "Chapter 23: Estimating Net Savings: Common Practices." In *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*. Golden, CO: National Renewable Energy Laboratory. www.nrel.gov/docs/fy14osti/62678.pdf.

Wood, E. 2015. "Are Smart Thermostats the New CFLS?" *Microgrid Knowledge*, October 13. microgridknowledge.com/are-smart-thermostats-the-new-cfls/.

Xu, T., and R. Friedmann. 2016. *A Summary of High-level Issues on the existing ISP Guide (V1.2A, April 2014) and Recommendations for revising the ISP Study Guide*. San Francisco: California Public Utilities Commission. t2wg.cadmusweb.com/Documents/Task_5_-_ISP_Guidance_Document/T2WG_Task5_ISP-Issues-PGESummary-v2_20161222.docx.

Appendix A. Interview Responses

RECENT DEVELOPMENTS

We asked our interviewees an open-ended question about the one or two most important recent developments in the EM&V field. We grouped their responses into seven categories (see table A1).

Table A1. Expert interview responses: most important developments in the field of EM&V

Response	Number of responses
<i>Data.</i> Increased access, availability, timeliness, and granularity of energy consumption data.	8
<i>State policies.</i> California (e.g., AB 802) normalized metered energy consumption (NMEC) as measure of energy savings, and SB 350 doubled energy savings. Also policy developments in Illinois (statewide energy legislation) and New York's Reforming the Energy Vision proceeding.	6
<i>Hardware and software proliferation.</i> New residential hardware at the panel, device, switch, and cord level; reduced cost of meters and sensors; smart meters; metering software; development of new energy software; and data analytics companies.	5
<i>M&V 2.0/advanced M&V.</i> A great data analytic tool; provides good data, more precision; new analysis methods, which do not mask the detail.	3
<i>Technical reference manuals.</i> Wider use of TRMs; evaluation frameworks that include the TRM; improved calibration and consistency of deemed savings values; move toward prospective TRMs, with evaluation results feeding back in to enhance accuracy.	3
<i>Efforts to focus on full market effects.</i> Studying markets to determine effects of past programs and resulting lower average consumption; impacts on market transformation efforts; load forecasting.	2
<i>Changing EM&V methods.</i> Correcting for self-selection, e.g., through nonequivalent comparison groups; integration of holistic analysis, including behavioral aspects into EM&V.	2

The first three categories of EM&V developments, which had the most responses, are external developments that influence the EM&V field. The three categories are also interrelated: new technologies (#3) resulting in the production of more and better energy data (#1) have generated interest and have often led to new state energy efficiency policies (#2). Each has implications for EM&V at several levels, including program planning, implementation, measurement, and evaluation.

CPB

The two most common responses from the EM&V experts interviewed were: (1) they were familiar with the approach as being used in the Northwest; and (2) they were aware that it was an approach discussed/recommended in the CPP's "Draft EM&V Guidance" for energy efficiency. Aside from those two comments (made by at least three experts) the other comments represented a diversity of views, ranging from "It is a 'direct to' net (savings)

approach” to “It is not a net savings approach,” with both of those statements made by experienced individuals in the evaluation community. There were also moderate comments addressing where CPB might be useful, including (1) “CPB is situational.” (2) “It can be very useful. The difficulty is acquiring enough reliable data.” (3) “It is not taking the place of NTG studies,” but CPB (concepts) can be used to ensure that double counting does not occur. Taking all the comments into account, we found considerable diversity: some experts indicated that they were not familiar with CPB approaches, while others viewed CPB as an approach that was essentially used only in the Pacific Northwest, within that region’s somewhat unique set of circumstances, and still others indicated that the clarity and understandability of CPB methods are desirable features.³⁸

CHALLENGES

We also asked our experts an open-ended question about the biggest one or two challenges facing energy efficiency EM&V today. We then grouped their responses into 10 categories (see table A2).

Table A2. Expert interview responses: biggest challenges facing EM&V

Response	Number of responses
Keeping EM&V relevant in a changing world. We need to weigh the value of attribution studies with their costs. Efficiency is "victim of its own success," making it difficult to isolate net impact after decades of doing energy efficiency. This leads to a need to look more at market effects going forward.	4
Using new data sources effectively (AMI, related data processing)	3
Assessing the value of M&V 2.0 and navigating changes that come along with it	3
Making results and language around evaluation more impactful, accessible, timely, and relevant	3
Budget challenges: insufficient EM&V budgets or pressure to reduce EM&V spending	3
Regulatory perceptions: lack of confidence that savings are real or concern that utilities are inflating savings	3
Political changes; specifically, conservative policy trends at federal and state levels that will impact EM&V institutions, focus, and expenditures	2

³⁸ These wide-ranging viewpoints were also found in the US DOE *Uniform Methods Protocols* (2014) in Chapter 23, “Estimating Net Savings: Common Practice.” That chapter sets out a number of approaches that are commonly used to estimate net savings. There was a discussion among commentators and the steering committee on whether CPB approaches were parallel to the other methods presented and should be included in the chapter. CPB methods focus only on ex ante information on market behavior and do not look at ex post information on actions or program participants. As a result, some commentators viewed this as a “deemed” estimation approach. Ultimately, CPB approaches were included because they represent a body of work in the Northwest and were beginning to receive attention elsewhere.

Response	Number of responses
Attribution issues: time lag between participant decisions and evaluation surveys	2
Insufficient evaluation guidance and oversight resulting in reduced confidence in the evaluation results	1
Increasing concentration of the evaluation industry, which negatively impacts quality of work	1

We found a wide range of responses across the interviews, clearly showing diverse concerns in the evaluation expert community. That said, it is possible to discern common themes. We found that the 10 response categories followed two general tracks: (1) specific challenges within EM&V practices, e.g., attribution issues and making the results and language more impactful, and (2) external forces exerting pressure on EM&V and creating challenges, e.g., regulatory perceptions of energy efficiency savings and changes in political landscapes.